Chapter 5

Molecular Dynamics

So far about the electronic structures, from now on, we transfer our attentions to descriptions of the nuclear motion.

As introduced in Chapter 1, molecular dynamics (MD) is a technique which allows us to investigate the statistical and dynamical properties of a real poly-atomic system at the molecular level. When the Born-Oppenheimer approximation is further imposed, this simplifies into reproducing the propagation of the nuclei on their Born-Oppenheimer potential-energy surfaces (PESs). In standard treatment of such propagations, the classical equations of motion governing the microscopic evolution of a many-body entity (composed by nuclei) are often solved numerically, subjecting to the boundary conditions appropriate for the geometry and/or symmetry of the system. Therefore, the method is in principle classical and these nuclei are classical point-like particles in this treatment. We note that there are extensions of this method to the quantum (in the statistical perspective) or semiclassical (in the dynamical perspective) regime, where some of the quantum nuclear effects (QNEs) can be explicitly addressed. In this chapter, to be pediatric, we restrict ourselves to the classical picture that the nuclei are point-like particles for the illustration of the basic underlying principles. Its extension to the descriptions of QNEs will be described in Chapter 6.

Depending on how the inter-atomic interactions are calculated, the MD simulations can be separated into two categories, *i.e.* the traditional classical MD simulations, where the interactions between different nuclei are described by empirical potentials, and the more recent *ab initio* MD simulations, where the interactions between the nuclei are calculated *on-the-fly* using the Hellmann-Feynman theorem after the electronic structures of the system at each specific spatial configuration of the nuclei were obtained quantum mechanically in an *ab initio* manner. Throughout this book, we focus on the *ab initio* methods and consequently the *ab initio* MD, but we note that the principles underlying the propagation of the nuclei as introduced here also apply to those simulations when empirical potentials are used.

Within the framework of *ab initio* MD, it is also worthwhile to note that there is a scheme when the Born-Oppenheimer approximation is not rigorously done but rather an adiabatic separation between the electronic and the nuclear degrees of freedom has been assumed using a special treatment for the movement of the electrons and the nuclei, like the Car-Parrinello (CP) MD [228]. And there is a scheme when the electronic structure optimization is rigorously performed at each spatial configuration of the nuclei along the trajectory, like the Born-Oppenheimer (BO) MD. The former treatment is an approximation to the later one. Due to the lower computation cost which originates from the fact that the self-consistent electronic structure optimization doesn't have to be done at each MD step, but rather a fictitious small mass is assigned to the electronic degrees of freedom which enables an adiabatic separation between the movements of the electrons and the nuclei, the former treatment has been very popular starting from the 1980s till now. The BO MD method only becomes a standard routine after the late 1990s when the improvement of computational algorithm and development of computer power made it feasible to carry out a self-consistent electronic structure optimization at each MD step. Because of this, the development of some methods within the framework of *ab initio* MD, especially those related to the thermostats, has been largely influenced by this historical trend. In later discussions, we will point this out in detail when those concepts which were influenced are introduced.

Our discussions in this chapter are categorized mainly into two parts. Sec. 5.1 presents an introduction to the ideas behind MD. With these concepts, our molecular simulation is able to sample a micro-canonical ensemble and its phase space on a constant-energy shell. In real experiments, however, it is the temperature that is kept constant. Sec. 5.2 then gives us an introduction to some methods to go beyond the simulations of the micro-canonical ensemble. Techniques underlying numerical controls of the molecular temperature (thermostats) are the basis for such simulations. Then, more sophisticated situations when other ensembles, such as the constant-pressure and constant-temperature (isothermal-isobaric) one, must be simulated will also be discussed. With these, in principle, one should be able to do a proper simulation for a system under the situation to be studied and compare the results with experiments.

5.1 Introduction to Molecular Dynamics

The foundation of modern MD technique is the so-called Hamiltonian mechanics, which is often used in descriptions of a micro-canonical ensemble in *Gibbs'* language. This Hamiltonian mechanics is in principle equivalent to the simplest Newton's equation of motion for the description of the nuclear propagations, with the only different that the equations of motion can be derived easily in non-Cartesian coor-

dinates. Because of this advantage, properties in the phase space can be discussed in a easier way in this language. Therefore, although it provides nothing more than what can be obtained by solving the Newton's equation of motion, we prefer using the Hamiltonian mechanics for discussions in this chapter.

Within this framework of the Hamiltonian mechanics, where the Hamiltonian's equation of motion for the nuclei are numerically solved, the simulations provide descriptions for the statistical properties of the system within a certain constant-energy shell in the phase space. There is neither particle exchange, energy exchange between the system under investigation and its environment, nor any change of the system's volume. In other words, most essentially, the total energy (Hamiltonian) is a conserved quantity. Later, extensions of this Hamiltonian mechanics to the non-Hamiltonian ones were proposed, where the total energy of the system is not conserved and other ensembles such as the canonical or isothermal-isobaric ones can be simulated. These different ensembles serve for different purposes in simulations of the real materials. To be pediatric, we start our discussions from the simplest Hamiltonian simulations of the micro-canonical ensemble (or otherwise NVE ensemble, where N stands for the number of the particles, V stands for the volume of the cell under simulation, and E stands for the total energy) and then extend these techniques to the non-Hamiltonian scheme and simulations of other ensembles.

Intuitively, the three key quantities in this MD simulation of a micro-canonical ensemble within the framework of the Hamiltonian mechanics should be the spatial coordinate of each nucleus ($\mathbf{r}_i$), the velocity of each nucleus ($\mathbf{v}_i$), and the force imposed by the quantum glue of the electrons and the classical inter-nuclear Coulomb interactions on each nucleus ($\mathbf{F}_i$) at time zero. These three quantities, together with the Hamiltonian's equation of motion, determine the nuclear evolution of the system and consequently the equilibrium and dynamical properties of the system under investigation. The entity of the interacting nuclei can be viewed as a classical many-body system, with interactions between the nuclei change as the system evolves. For such a many-body entity, it is impossible to solve this equation of motion analytically. Therefore, one needs to resort to numerical integrations for the equation of motion.

5.1.1 The Verlet Algorithm

As mentioned above, the first step in carrying out a numerical integration for the equation of motion is to set up a proper initial state of the system. This includes setting the initial spatial configuration of the nuclei, initial velocities of them, and the initial forces imposed on them. Then the trajectory of the nuclei can be integrated out numerically in principle. Seemingly simple, the integration of such an equation of motion is highly non-trivial due to the fact that a finite time-step must be used. Error accumulation ultimately results in two trajectories that are initially close

diverge exponentially as the dynamic evolves.

Fortunately, the MD simulations differ in principle to the trajectory predictions in classical dynamics such as in the study of the comet's orbital (see e.g. Sec. 4.3 in Ref. [48]). What we are primarily aiming at in the MD simulations is some knowledge on the statistical property of the system whose initial state we already know something about, e.g. the energy. In the language of the *Gibbs' ensemble* (the micro-canonical ensemble in this case), many individual microscopic configurations of a large system lead to the same macroscopic behaviors. This statement implies that it is unnecessary to know the precise propagation of every nuclei in the system in order to predict it macroscopic properties. Rather, a simple average over a large number of states visited during the simulation will give us properly formulated macroscopic observable. As long as our integrator respects the fact that the Newton's equation of motion is time-reversible and the corresponding Hamiltonian dynamics leaves the volume of the phase space explored in the MD simulation unchanged, the energy conservation can be guaranteed if the finite time step between movements are small enough in practical simulations [48]. This indicates that a micro-canonical ensemble is already being simulated. As one of the successful schemes on implementing such a propagation when the total energy and the time-reversibility of the dynamics can be preserved, we take the Verlet algorithm as an example and explain the underlying principles by going through the key equations in below [229].

In this algorithm, the spatial configuration of the system at time $t + \Delta t$ is determined by:

$$\mathbf{r}_i(t + \Delta t) = \mathbf{r}_i(t) + \mathbf{v}_i(t)\Delta t + \frac{\mathbf{F}_i(t)}{2M_i}\Delta t^2 + O(\Delta t^3) + O(\Delta t^4) \tag{5.1}$$

when Taylor expansion is used. Here $O(\Delta t^3)$ ($O(\Delta t^4)$) represents error in order of Δt^3 (Δt^4), which equals $\mathbf{r}'''(t)\Delta t^3/3!$ ($\mathbf{r}''''(t)\Delta t^4/4!$). Further expansions beyond $O(\Delta t^4)$ are neglected. We note that the quantities calculated along the trajectories are only $\mathbf{r}_i$, $\mathbf{v}_i$, and $\mathbf{F}_i$, the terms $O(\Delta t^3)$ and $O(\Delta t^4)$ won't be calculated. But it is important to write them down in Eq. 5.1 in order to show the order of accuracy for $\mathbf{r}_i$ and $\mathbf{v}_i$ calculated along the trajectory.

From Eq. 5.1, it is easy to see that in reverse time the spatial configuration respects:

$$\mathbf{r}_i(t - \Delta t) = \mathbf{r}_i(t) - \mathbf{v}_i(t)\Delta t + \frac{\mathbf{F}_i(t)}{2m_i}\Delta t^2 - O((\Delta t)^3) + O((\Delta t)^4). \tag{5.2}$$

If we sum up Eqs. 5.1 and 5.2, we can have:

$$\mathbf{r}_i(t + \Delta t) = 2\mathbf{r}_i(t) - \mathbf{r}(t - \Delta t) + \frac{\mathbf{F}_i(t)}{m_i}\Delta t^2 + 2O(\Delta t^4). \tag{5.3}$$

The $O(\Delta t^3)$ terms were cancelled out. Therefore, the evolution of the nuclear position obtained from this integrator contains an error in order of Δt^4 (quantities related to the lower order expansions needed in the analytical expression in Eq. 5.3, *i.e.* $\mathbf{r}_i$ and $\mathbf{F}_i$, are all calculated along the trajectory).

To obtain the new velocity, we minus Eqs. 5.1 and 5.2. In so doing, we arrive at:

$$\mathbf{r}_i(t+\Delta t) - \mathbf{r}_i(t-\Delta t) = 2\mathbf{v}_i(t)\Delta t + 2O(\Delta t^3). \tag{5.4}$$

The velocity is then updated by:

$$\mathbf{v}_i(t) = \frac{\mathbf{r}_i(t+\Delta t) - \mathbf{r}_i(t-\Delta t)}{2\Delta t} - O(\Delta t^3)/\Delta t. \tag{5.5}$$

From this equation, it is clear that the velocity contains an error in order of Δt^2.

It is worthwhile to note that Eqs. 5.1 and 5.2 are the basis of the Verlet algorithm, which are obviously time-reversible. Therefore, trajectories obtained from the Verlet algorithm respects the time-reversibility of the Newton's equation of motion. Disregarding the error terms which are not calculated, in this Verlet algorithm, the trajectory is propagated through:

$$\mathbf{r}_i(t+\Delta t) = 2\mathbf{r}_i(t) - \mathbf{r}(t-\Delta t) + \frac{\mathbf{F}_i(t)}{m_i}\Delta t^2 \tag{5.6}$$

and

$$\mathbf{v}_i(t) = \frac{\mathbf{r}_i(t+\Delta t) - \mathbf{r}_i(t-\Delta t)}{2\Delta t}. \tag{5.7}$$

The positions and velocities contain error in order of Δt^4 and Δt^2 respectively.

5.1.2 The Velocity Verlet Algorithm

We note that in Eq. 5.6 the velocity is not used when the new position is generated. There is an equivalent algorithm to the Verlet one in which the velocity is used in the calculation of the new position. This is the so-called velocity Verlet algorithm [230].

In this algorithm, the new position is calculated from:

$$\mathbf{r}_i(t+\Delta t) = \mathbf{r}_i(t) + \mathbf{v}_i(t)\Delta t + \frac{\mathbf{F}_i(t)}{2m_i}\Delta t^2. \tag{5.8}$$

The velocity is updated by:

$$\mathbf{v}_i(t+\Delta t) = \mathbf{v}_i(t) + \frac{\mathbf{F}_i(t+\Delta t) + \mathbf{F}_i(t)}{2m_i}\Delta t. \tag{5.9}$$

We note that a key difference between this algorithm and the Euler scheme, which updates the velocity by:

$$\mathbf{v}_i(t+\Delta t) = \mathbf{v}_i(t) + \frac{\mathbf{F}_i(t)}{m_i}\Delta t, \tag{5.10}$$

is that the new velocity can only be calculated after the new position and its corresponding forces are generated in the velocity Verlet algorithm. This fine numerical difference actually results in the fundamental distinction that the trajectory obtained from the Euler algorithm is not time-reversible and the position contains an error in order of Δt^3. Whereas in the velocity Verlet algorithm the trajectory is time-reservable and the position updated contains an error in order of Δt^4, although this is not clear in Eqs. 5.8 and 5.9.

To prove the equality between the velocity Verlet algorithm and the original Verlet algorithm, we start from the original Verlet algorithm in Eqs. 5.6 and 5.7 and try to deduce the velocity Verlet algorithm in Eqs. 5.8 and 5.9. To this end, we first rewrite Eq. 5.6 into the following form:

$$\mathbf{r}_i(t+\Delta t) = \mathbf{r}_i(t) + \frac{\mathbf{r}_i(t+\Delta t) - \mathbf{r}_i(t-\Delta t)}{2\Delta t}\Delta t + \frac{\mathbf{F}_i(t)}{2m_i}\Delta t^2. \tag{5.11}$$

Then, by replacing the $(\mathbf{r}_i(t+\Delta t) - \mathbf{r}_i(t-\Delta t))/(2\Delta t)$ with $\mathbf{v}_i(t)$ through Eq. 5.7, we arrive at Eq. 5.8. Therefore, the updating of the position is completely equivalent in these two algorithms.

Concerning the deduction of Eq. 5.9, from Eq. 5.8 we see that:

$$\mathbf{v}_i(t) = \frac{\mathbf{r}_i(t+\Delta t) - \mathbf{r}_i(t)}{\Delta t} - \frac{\mathbf{F}_i(t)}{2m_i}\Delta t. \tag{5.12}$$

This equation is equivalent to:

$$\mathbf{v}_i(t+\Delta t) = \frac{\mathbf{r}_i(t+2\Delta t) - \mathbf{r}_i(t+\Delta t)}{\Delta t} - \frac{\mathbf{F}_i(t+\Delta t)}{2m_i}\Delta t, \tag{5.13}$$

which further equals

$$\mathbf{v}_i(t+\Delta t) = \frac{\mathbf{r}_i(t+2\Delta t) - \mathbf{r}_i(t)}{\Delta t} - \frac{\mathbf{r}_i(t+\Delta t) - \mathbf{r}_i(t)}{\Delta t} - \frac{\mathbf{F}_i(t+\Delta t)}{2m_i}\Delta t, \tag{5.14}$$

Using Eq. 5.7, it can be reduced to:

$$\mathbf{v}_i(t+\Delta t) = 2\mathbf{v}_i(t+\Delta t) - \frac{\mathbf{r}_i(t+\Delta t) - \mathbf{r}_i(t)}{\Delta t} - \frac{\mathbf{F}_i(t+\Delta t)}{2m_i}\Delta t, \tag{5.15}$$

which equals

$$\mathbf{v}_i(t+\Delta t) = \frac{\mathbf{r}_i(t+\Delta t) - \mathbf{r}_i(t)}{\Delta t} + \frac{\mathbf{F}_i(t+\Delta t)}{2m_i}\Delta t. \tag{5.16}$$

From Eq. 5.12, we see that $(\mathbf{r}_i(t+\Delta t)-\mathbf{r}_i(t))/\Delta t$ equals $\mathbf{v}_i(t)+\mathbf{F}_i(t)\Delta t/(2m_i)$. Replacing this $(\mathbf{r}_i(t+\Delta t)-\mathbf{r}_i(t))/\Delta t$ term with $\mathbf{v}_i(t)+\mathbf{F}_i(t)\Delta t/(2m_i)$ in Eq. 5.16, one arrives at Eq. 5.9.

The above deduction means that starting from Eqs. 5.6 and 5.7, one easily arrives at Eqs. 5.8 and 5.9. In other words, Eqs. 5.8 and 5.9 propagate the trajectory in the same way as Eqs. 5.6 and 5.7, *i.e.* the velocity Verlet and the Verlet algorithms are completely equivalent. The position (velocity) contains error in order of Δt^4 (Δt^2). They both differ from the Euler algorithm in Eq. 5.10 in the sense that their trajectories are time reversible and preserve higher order of accuracy in terms of Δt.

5.1.3 The Leap Frog Algorithm

The second often used algorithm for the propagation of the MD trajectory, which is also equivalent to the Verlet algorithm, is the so-called Leap Frog algorithm [48, 231]. In this algorithm, the velocities and positions are updated by:

$$\mathbf{v}_i(t+\Delta t/2)=\mathbf{v}_i(t-\Delta t/2)+\frac{\mathbf{F}_i(t)}{m_i}\Delta t, \tag{5.17}$$

and

$$\mathbf{r}_i(t+\Delta t)=\mathbf{r}_i(t)+\mathbf{v}_i(t+\Delta t/2)\Delta t. \tag{5.18}$$

The propagation is equivalent to the Verlet algorithm with the only difference that the velocities are updated at half time step.

For the completeness of this introduction, we use one more paragraph to explain how these equations can be derived from Eq. 5.6 and 5.7. From Eq. 5.7, we have:

$$\mathbf{v}_i(t-\Delta t/2)=\frac{\mathbf{r}_i(t)-\mathbf{r}_i(t-\Delta t)}{\Delta t}, \tag{5.19}$$

and

$$\mathbf{v}_i(t+\Delta t/2)=\frac{\mathbf{r}_i(t+\Delta t)-\mathbf{r}_i(t)}{\Delta t}. \tag{5.20}$$

From Eq. 5.20, it is easy to obtain Eq. 5.18 for the update of the positions. Concerning the deduction of Eq. 5.17, one can rewrite Eq. 5.6 into the following form:

$$\frac{\mathbf{r}_i(t+\Delta t)-\mathbf{r}_i(t)}{\Delta t}=\frac{\mathbf{r}_i(t)-\mathbf{r}_i(t-\Delta t)}{\Delta t}+\frac{\mathbf{F}_i(t)}{m_i}\Delta t. \tag{5.21}$$

Then by inputting Eqs. 5.19 and 5.20 into Eq. 5.21, one gets:

$$\mathbf{v}_i(t+\Delta t/2)=\mathbf{v}_i(t-\Delta t/2)+\frac{\mathbf{F}_i(t)}{m_i}\Delta t, \tag{5.22}$$

which is the same as Eq. 5.17. Therefore the trajectory is equivalent to that of the Verlet algorithm. However, due to the fact that the velocities are not calculated at the same time slides as the positions, the total energy can't be rigorously calculated.

The Verlet, velocity Verlet, and the Leap-Frog algorithms are the often used simple algorithm in standard implementations of the MD method. For most MD simulations, they are sufficiently accurate. From Eq. 5.1, it is easy to see that if higher orders of expansion for the updating of the position in terms of Δt is included, one may use larger time step for the propagation of the trajectory and consequently higher efficiency of the simulations. The predictor-corrector algorithm is such an example [232]. However, these extensions may not respect the time reversal symmetry of the Newton's equation and consequently may not lead to more reliable sampling of the corresponding phase space. In this chapter, we restrict ourselves to the simplest ones introduced and refer the readers interested in more complicated ones to Refs. [48, 233].

5.2 Other Ensembles

When a proper initial state of the system under investigation is set up and the trajectory is generated using the algorithms as introduced above, if the time step is sufficiently small, the total energy E will be conserved within a certain tolerance. On the basis of this, if the *ergodic hypothesis* is further imposed, which means that given an infinite amount of time the trajectory will go through the entire constant energy hypersurface in the phase space, a proper micro-canonical ensemble will be simulated. This micro-canonical ensemble provides accurate description for the statistical/dynamical behavior of the system within a certain constant energy shell in the phase space, when there is neither particle exchange, energy exchange between the system under investigation and its environment, nor any change of the system's volume [234, 235, 236, 237, 238]. Most experiments, however, are performed at situations where these quantities are not conserved. The simplest situation is when the experiment is carried out at a constant temperature T instead of the total energy E, due to the energy exchange between the system under investigation and its environment. In this case, the statistical behavior of the system at a finite temperature must be accurately taken into account, and the states associated with the system in the phase space visited during the time when the experiment is carried out are not within a certain constant energy shell anymore. Canonical, or in other words, the constant volume and constant temperature (NVT) ensemble therefore needs to be simulated [239, 240, 241, 242, 243, 244, 245, 246].

Theoretically, this can be done by introducing a proper thermostat in the simulation, through which the energy exchange between the system under investigation and the environment can be properly described. However, other thermodynamic vari-

ables such as the volume of the simulation cell at finite pressure P may also change due to fluctuations. To describe situations like this, a barostat needs to be further introduced which keeps the tensor of the system constant during the simulations. This ensemble with constant temperature and constant pressure is called a NPT, or isothermal-isobaric ensemble in literature [247, 248, 249]. In ascending order of complexity, other ensembles such as the isoenthalpic-isobaric (NPH) [238, 250, 251, 252], grand-microcanonical (μVL, with μ standing for the chemical potential) [238, 253], etc. also exist which simulate experiments at different situations. In this chapter, we aim at providing an introduction to the principles underlying theoretical simulations of these ensembles. Therefore, we will use the simplest case, *i.e.* the canonical (NVT) ensemble, as an example to explain how these principles work. Its extension to the isothermal-isobaric (NPT) ensemble will also be briefly discussed. For readers interested in those methods for simulating more complicated ensembles, please go to their original papers [238, 238, 250, 251, 252, 253]. We note that in these simulations, since the total energy is not conserved and the equation of motion can not be obtained from the Hamiltonian of the system under simulations, they are often called the non-Hamiltonian MD simulations in literature.

5.2.1 Andersen Thermostat

There are several schemes for the temperature to the controlled. The underlying principle is that at a finite temperature T, the probability of finding a classical particle, e.g. the i^{th} atom in the system, with kinetic energy $\mathbf{p}_i^2/(2m_i)$, which is labelled as $\mathrm{P}(\mathbf{p}_i)$ throughout our discussions, respects the Maxwell-Boltzmann distribution:

$$\mathrm{P}(\mathbf{p}_i) = \left(\frac{\beta}{2\pi m_i}\right)^{\frac{3}{2}} e^{-\beta\frac{\mathbf{p}_i^2}{2m_i}}. \tag{5.23}$$

Here $\beta = 1/(k_{\mathrm{B}}T)$. The simplest implementation of this restriction on simulation of real systems naturally resides in a stochastic collision of nuclei with a thermostat that imposes such a Maxwell-Boltzmann distribution on the velocities. The corresponding thermostat is the so-called Andersen thermostat [239]. When it is used, with a certain initial state, the propagator as introduced in Sec. 5.1 ensures that the system propagates on a certain constant energy shell in the phase space between collisions. The collisions themselves, on the other hand, randomly act on selected particle and reset its new velocity from the Maxwell-Boltzmann distribution function at the desired temperature, which effectively reduce correlation between events along the trajectory for increased sampling efficiency at the cost of accuracy in short-time dynamics. With a reasonable choice of the collision frequency, the canonical ensemble with velocities respecting the Maxwell-Boltzmann distribution can be conveniently simulated. Chapter 6 in Ref. [48] gives details of how this is

implemented, including how the program should be written, which we don't bother. For such details, please refer to their explanation.

Here, we note that although it is efficient and simple to be implemented on imposing a canonical ensemble to be sampled, the stochastic collision results in non-deterministic and discontinuous trajectories. These non-deterministic and discontinuous trajectories lead to all real-time information being lost at collision. Related physical quantities, such as the vibrational frequencies, diffusion coefficients etc., therefore can't be calculated. In addition to this, more seriously, in spite of the fact that a fairly ergodic sampling can often be guaranteed on all degrees of freedom in such a stochastic simulation, when an adiabatic separation between the electronic and ionic degrees of freedom must be assumed, like in Car-Parrinello (CP) molecular dynamics, a breakdown of such adiabatic separation often happens [254, 255]. As mentioned in the introduction of this chapter, *ab initio* MD has been dominated by the CP MD treatment before the late 1990s due to its lower computational cost. Therefore, deterministic methods such as the Nosé-Hoover thermostat which avoid such a numerical problem has been widely adopted since 1980s [240, 241, 242, 243, 244]. It is only after the BO MD gets widely used when stochastic methods such as the Andersen or Langevin thermostat get popular in *ab initio* MD simulations. Because of this historical reason, when one goes through the literatures, one may realize a change of the taste in attempts to control the temperature influenced by this trend. In this chapter, to present a general introduction of these methods and be consistent with this historical trend, we will give a detailed explanation of the principles underlying the Nosé-Hoover thermostat and its related Nosé-Hoover chain. After that, we come back to another stochastic method, *i.e.* the Langevin thermostat, which is widely used in *ab initio* BO MD simulations nowadays. We note that extensions of this method were also designed to be compatible with CP MD recently [254].

5.2.2 Nosé-Hoover Thermostat

As one of the first successful attempts to incorporate deterministic trajectories in the MD simulation at finite temperature, Nosé and Hoover have separately used an extended system with an additional and artificial degree of freedom to control the temperature [240, 241, 242, 243]. The corresponding thermostat is called the Nosé-Hoover thermostat. We note that this idea of extracting a canonical ensemble out of a micro-canonical ensemble on an extended system is the same as how the concept of canonical ensemble is derived from the concept of micro-canonical ensemble in standard textbooks on statistics. In that case, it is required that the degrees of freedom in the thermostat is much larger than that in the canonical ensemble. Here, a very clever choice of the potential to which the added degree of freedom particle is subjected to ensures that one single extra degree of freedom imposes a canonical

ensemble on the system to be studied. In the following, we will show how this works by going through the underlying key equations.

The main assumption of this extended system method is that an additional coordinate s needs to be introduced to the classical N-nuclei system, with the Lagrangian for the extended system

$$L_{\text{Nosé}} = \sum_{i=1}^{N} \frac{m_i}{2} s^2 \dot{\mathbf{r}}_i^2 - V(\mathbf{r}_1, \cdots, \mathbf{r}_N) + \frac{Q}{2}\dot{s}^2 - \frac{g}{\beta}\ln s. \tag{5.24}$$

Here, we use "Nosé" to denote the extended system. Q is the effective mass associated with the additional degree of freedom, g is a parameter to be fixed, and $[g \ln s]/\beta$ is the external potential acting on this extra degree of freedom. The corresponding Hamiltonian of this "Nosé" system can then be written as:

$$H_{\text{Nosé}} = \sum_{i=1}^{N} \frac{1}{2m_i s^2}\mathbf{p}_i^2 + V(\mathbf{r}_1, \cdots, \mathbf{r}_N) + \frac{1}{2Q}p_s^2 + \frac{g}{\beta}\ln s, \tag{5.25}$$

where s, p_s, $\mathbf{r}_i$ and $\mathbf{p}_i$ are variables in the phase space and $\mathbf{p}_i = \partial L_{\text{Nosé}}/\partial \dot{\mathbf{r}}_i = m_i s^2 \dot{\mathbf{r}}_i$.

From this Hamiltonian, the partition function for the micro-canonical (Nosé) ensemble is defined as:

$$Z_{\text{Nosé}} = \frac{1}{N!}\int dp_s ds d\mathbf{p}_1 \cdots d\mathbf{p}_N d\mathbf{r}_1 \cdots d\mathbf{r}_N \delta(H_{\text{Nosé}} - E), \tag{5.26}$$

Now, if we scale the momentum $\mathbf{p}_i$ with the additional degree of freedom by:

$$\mathbf{p'}_i = \mathbf{p}_i/s, \tag{5.27}$$

the partition function in Eq. 5.26 simplifies into:

$$\begin{aligned} Z_{\text{Nosé}} =& \frac{1}{N!}\int dp_s ds d\mathbf{p'}_1 \cdots d\mathbf{p'}_N d\mathbf{r}_1 \cdots \mathbf{r}_N s^{3N} \\ & \delta((\sum_{i=1}^{N} \frac{1}{2m_i}\mathbf{p'}_i^2 + V(\mathbf{r}_1, \cdots, \mathbf{r}_N) + \frac{1}{2Q}p_s^2 + \frac{g}{\beta}\ln s) - E). \end{aligned} \tag{5.28}$$

Further simplification of this equation requires using of the following equation:

$$\delta\left[h(s)\right] = \frac{\delta(s - s_0)}{h'(s_0)}, \tag{5.29}$$

where $h(s)$ is a function of s with a single root for $h(s) = 0$ at s_0 and $h'(s_0)$ is its derivative at this point. Now, if we take:

$$h(s) = (\sum_{i=1}^{N} \frac{1}{2m_i}\mathbf{p'}_i^2 + V(\mathbf{r}_1, \cdots, \mathbf{r}_N) + \frac{1}{2Q}p_s^2 + \frac{g}{\beta}\ln s) - E \tag{5.30}$$

in Eq. 5.28 and make use of Eq. 5.29, Eq. 5.28 can be simplified into:

$$\begin{aligned} Z_{\text{Nosé}} =& \frac{1}{N!}\int dp_s ds d\mathbf{p'}_1\cdots d\mathbf{p'}_N d\mathbf{r}_1\cdots\mathbf{r}_N s^{3N}\frac{\beta s}{g} \\ & \delta\left(s-\exp\left[-\beta\frac{\sum_{i=1}^{N}\frac{1}{2m_i}\mathbf{p'}_i^2+V(\mathbf{r}_1,\cdots,\mathbf{r}_N)+p_s^2/(2Q)-E}{g}\right]\right) \\ =& \frac{1}{N!}\int dp_s ds d\mathbf{p'}_1\cdots d\mathbf{p'}_N d\mathbf{r}_1\cdots\mathbf{r}_N \frac{\beta s^{(3N+1)}}{g} \\ & \delta\left(s-\exp\left[-\beta\frac{\sum_{i=1}^{N}\frac{1}{2m_i}\mathbf{p'}_i^2+V(\mathbf{r}_1,\cdots,\mathbf{r}_N)+p_s^2/(2Q)-E}{g}\right]\right). \end{aligned} \tag{5.31}$$

We integrate out s first. In so doing, we can get the partition function of the "Nosé" system as:

$$\begin{aligned} Z_{\text{Nosé}} =& \frac{1}{N!}\int dp_s d\mathbf{p'}_1\cdots d\mathbf{p'}_N d\mathbf{r}_1\cdots\mathbf{r}_N \frac{\beta}{g} \\ & \exp\left[-\beta(\frac{\sum_{i=1}^{N}\frac{1}{2m_i}\mathbf{p'}_i^2+V(\mathbf{r}_1,\cdots,\mathbf{r}_N)+p_s^2/(2Q)-E}{g})(3N+1)\right] \\ =& \frac{1}{N!}\frac{\beta\exp[E\beta(3N+1)/g]}{g}\int dp_s \exp\left[-\beta\frac{3N+1}{g}\frac{p_s^2}{2Q}\right] \\ & \int d\mathbf{p'}_1\cdots d\mathbf{p'}_N d\mathbf{r}_1\cdots d\mathbf{r}_N \exp\left[-\beta\frac{3N+1}{g}\left(\sum_{i=1}^{N}\frac{1}{2m_i}\mathbf{p'}_i^2+V(\mathbf{r}_1,\cdots,\mathbf{r}_N)\right)\right] \end{aligned} \tag{5.32}$$

This equation indicates that in the extended system with a Hamiltonian as shown in Eq. 5.25, its partition function can be separated into a product of two parts, *i.e.* that of a sub-system with variables $\mathbf{p'}_i$ and $\mathbf{r}_i$ in the phase space and that of a constant C, as:

$$Z_{\text{Nosé}} = C\int d\mathbf{p'}_1\cdots d\mathbf{p'}_N d\mathbf{r}_1\cdots\mathbf{r}_N \exp\left[-\beta\frac{3N+1}{g}H(\mathbf{p'},\mathbf{r})\right]. \tag{5.33}$$

Here,

$$C = \frac{1}{N!}\frac{\beta\exp(E\beta(3N+1)/g)}{g}\int dp_s \exp\left[-\beta\frac{3N+1}{g}\frac{p_s^2}{2Q}\right] \tag{5.34}$$

and

$$H(\mathbf{p'},\mathbf{r}) = \sum_{i=1}^{N}\frac{1}{2m_i}\mathbf{p'}_i^2+V(\mathbf{r}_1,\cdots,\mathbf{r}_N). \tag{5.35}$$

The equation of motion for this extended system, rigorously, is determined by the Hamiltonian in Eq. 5.25 through:

$$\frac{d\mathbf{r}_i}{dt} = \frac{\partial H_{\text{Nosé}}}{\partial \mathbf{p}_i} = \frac{\mathbf{p}_i}{m_i s^2} \tag{5.36a}$$

$$\frac{d\mathbf{p}_i}{dt} = -\frac{\partial H_{\text{Nosé}}}{\partial \mathbf{r}_i} = -\frac{\partial V(\mathbf{r}_1, \cdots, \mathbf{r}_N)}{\partial \mathbf{r}_i} \tag{5.36b}$$

$$\frac{ds}{dt} = \frac{\partial H_{\text{Nosé}}}{\partial s} = \frac{p_s}{Q} \tag{5.36c}$$

$$\frac{dp_s}{dt} = -\frac{\partial H_{\text{Nosé}}}{\partial s} = \left[\sum_{i=1}^{N} \mathbf{p}_i^2/(m_i s^2) - g/\beta\right]/s \tag{5.36d}$$

With the quasi-ergodic hypothesis that the time average along the trajectory equals the ensemble average when the simulation time is sufficiently long, the ensemble average of a certain physical quantity A that depends on $\mathbf{p}'$ and $\mathbf{r}$ in the "Nosé" system equals:

$$\langle A(\mathbf{p}', \mathbf{r})\rangle = \lim_{\tau\to\infty} \frac{1}{\tau} \int_0^\tau A(\mathbf{p}', \mathbf{r}) dt. \tag{5.37}$$

Here $\langle \cdots \rangle$ indicates the micro-canonical ensemble average of the extended (Nosé) system and the right-hand side is the time average over the trajectory generated by Eq. 5.36. If we put in the partition function as simplified in Eq. 5.33 for the "Nosé" system to the left-hand side of Eq. 5.37, we can have:

$$\langle A(\mathbf{p}', \mathbf{r})\rangle = \frac{\int d\mathbf{p}'_1 \cdots d\mathbf{p}'_N d\mathbf{r}_1 \cdots d\mathbf{r}_N A(\mathbf{p}', \mathbf{r}) \exp\left[-\beta \frac{3N+1}{g} H(\mathbf{p}', \mathbf{r})\right]}{\int d\mathbf{p}'_1 \cdots d\mathbf{p}'_N d\mathbf{r}_1 \cdots d\mathbf{r}_N \exp\left[-\beta \frac{3N+1}{g} H(\mathbf{p}', \mathbf{r})\right]}. \tag{5.38}$$

Now if we take $g = 3N + 1$, this further simplifies into:

$$\langle A(\mathbf{p}', \mathbf{r})\rangle = \frac{1/(N!) \int d\mathbf{p}'_1 \cdots d\mathbf{p}'_N d\mathbf{r}_1 \cdots d\mathbf{r}_N A(\mathbf{p}', \mathbf{r}) \exp\left[-\beta H(\mathbf{p}', \mathbf{r})\right]}{1/(N!) \int d\mathbf{p}'_1 \cdots d\mathbf{p}'_N d\mathbf{r}_1 \cdots d\mathbf{r}_N \exp\left[-\beta H(\mathbf{p}', \mathbf{r})\right]}. \tag{5.39}$$

We label the righthand side of the above equation as $\langle A(\mathbf{p}', \mathbf{r})\rangle_{\text{NVT}}$, since it is equivalent to the canonical ensemble average of the operator $A(\mathbf{p}', \mathbf{r})$ in a system with Hamiltonian $H(\mathbf{p}', \mathbf{r})$ at T. In other words,

$$\langle A(\mathbf{p}', \mathbf{r})\rangle = \langle A(\mathbf{p}', \mathbf{r})\rangle_{\text{NVT}}. \tag{5.40}$$

From this equality, it is clear that $\langle A(\mathbf{p}', \mathbf{r})\rangle_{\text{NVT}}$, the canonical ensemble average of $A(\mathbf{p}', \mathbf{r})$ in the sub-system, can be evaluated using the micro-canonical ensemble average of the extended system. Numerical, if the quasi-ergodic hypothesis is further

imposed, by combining Eqs. 5.37 and 5.40 this canonical ensemble average further equals the time average of this physical quantity along the deterministic trajectory of the "Nosé" system, through

$$\langle A(\mathbf{p}', \mathbf{r})\rangle_{\mathrm{NVT}} = \lim_{\tau\to\infty} \frac{1}{\tau} \int_0^\tau A(\mathbf{p}', \mathbf{r}) dt. \tag{5.41}$$

In other words, imagine that the real system we want to investigate is subjected to the Hamiltonian in Eq. 5.35, Eq. 5.40 means that the canonical ensemble of a certain physical quantity in this real system can be evaluated using the micro-canonical ensemble average of an artificial extended system, in which a deterministic trajectory is generated. When the quasi-ergodic hypothesis is imposed on the MD simulation of the micro-canonical extended system, the equality in Eq. 5.41 further means that this canonical ensemble average of the physical quantity in the sub-system is completely equivalent to the time-average of this physical quantity on the deterministic trajectory of the extended system.

Till now, the principle underlying the Nosé-Hoover thermostat is clear. The micro-canonical distribution of the extended ("Nosé") system with an augmented variable s and Hamiltonian as shown in Eq. 5.25 is equivalent to a canonical distribution of the sub-system with variables $\mathbf{p'}_i$ and $\mathbf{r}_i$, where $\mathbf{p'}_i$ is the scaled momentum $\mathbf{p}_i/s$. This feature is obtained by artificially introducing an additional degree of freedom s, which is subjected to an external potential of a special form. And the property that the canonical ensemble can be sampled this way is guaranteed by the above equations. We note that the trajectory in Eq. 5.41 needs to be generated by Eq. 5.36, using an even distribution of the time interval dt.

From Eq. 5.27, Eq. 5.36 and Eq. 5.39, one can also see that the scaled momentum $\mathbf{p'}_i$ is the real momentum of the canonical system under investigation and the momentum $\mathbf{p}_i$ before scaling is only for the artificial extended "Nosé" system. Because of this difference, from now on, we call the variables related to $\mathbf{p'}_i$ the real variables and those related to $\mathbf{p}_i$ the virtual ones. Since Eq. 5.36 is the equation of motion from which the trajectory is determined, the ensemble average in Eq. 5.39 must follow an even distribution of the virtual time (t). Following this rigorous route that the trajectory is generated for the extended system with an even distribution of the virtual time and the propagation generated from Eq. 5.36 is on the virtual variables, while the ensemble average is done for the sub-system with real momentum $\mathbf{p'}_i$, there exists some obvious inconsistency which is cumbersome to handle.

To avoid such an inconsistency in calculating the trajectory and doing the ensemble averages in different manners, we can rewrite the equations of motion in Eq. 5.36 by scaling the key variables in the virtual system to the real one through: $\mathbf{r}_i{}' = \mathbf{r}_i$, $\mathbf{p}_i{}' = \mathbf{p}_i/s$, $s' = s$, $p_s{}' = p_s/s$, and $dt' = dt/s$. In terms of these real

variables, the equation of motion in Eq. 5.36 can be rewritten as:

$$\frac{d\mathbf{r}_i{}'}{dt'} = s\frac{d\mathbf{r}_i}{dt} = s\frac{\mathbf{p}_i}{m_i s^2} = \frac{\mathbf{p}_i}{m_i s} = \frac{\mathbf{p}_i{}'}{m_i} \tag{5.42a}$$

$$\begin{aligned}\frac{d\mathbf{p}_i{}'}{dt'} =& s\frac{d(\mathbf{p}_i/s)}{dt} = \frac{d\mathbf{p}_i}{dt} - \frac{1}{s}\frac{ds}{dt}\mathbf{p}_i \\ =& -\frac{\partial V(\mathbf{r}_1, \cdots, \mathbf{r}_N)}{\partial \mathbf{r}_i} - \frac{ds}{dt}\mathbf{p}_i{}' \\ =& -\frac{\partial V(\mathbf{r}_1{}', \cdots, \mathbf{r}_N{}')}{\partial \mathbf{r}_i{}'} - \frac{s'p_s{}'}{Q}\mathbf{p}_i{}'\end{aligned} \tag{5.42b}$$

$$\frac{ds'}{dt'} = s\frac{ds}{dt} = \frac{sp_s}{Q} = \frac{s'^2 p_s{}'}{Q} \tag{5.42c}$$

$$\begin{aligned}\frac{dp_s{}'}{dt'} =& s\frac{d(p_s/s)}{dt} = \frac{dp_s}{dt} - \frac{1}{s}\frac{ds}{dt}p_s \\ =& \left[\sum_{i=1}^{N} \mathbf{p}'^2_i/m_i - g/\beta\right]/s' - s'p'^2_s/Q.\end{aligned} \tag{5.42d}$$

The deterministic trajectory of a system involving variables $\mathbf{r}_i{}' = \mathbf{r}_i$, $\mathbf{p}_i{}' = \mathbf{p}_i/s$, $s' = s$, $p_s{}' = p_s/s$, and $dt' = dt/s$ can be determined from Eq. 5.42. Along this trajectory, the conserved quantity is:

$$H'_{\text{Nosé}} = \sum_{i=1}^{N} \frac{1}{2m_i}\mathbf{p}'^2_i + V(\mathbf{r}'_1, \cdots, \mathbf{r}'_N) + \frac{s'^2 p'^2_s}{2Q} + \frac{g}{\beta}\ln s'. \tag{5.43}$$

But we note that this is not the Hamiltonian of the system since the equations of motion can't be obtained from it. The MD simulation performed in this manner, accordingly, is called non-Hamiltonian molecular dynamics in literature [246].

Now, we can generate a trajectory using Eq. 5.42 which satisfies the equation of motion defined by the Hamiltonian in Eq. 5.25. Accordingly, the expectation value of a physical quantity $A(\mathbf{p}', \mathbf{r}')$ should be evaluated with it. However, we note that the application of the quasi-ergodic hypothesis as shown in Eq. 5.41, which relates the ensemble average of the extended micro-canonical ensemble to the real canonical subsystem, is achieved by using the trajectory generated with Eq. 5.36 for the extended system, where an even distribution of dt is imposed. In order for this quasi-ergodic hypothesis to be applied to the new trajectory of the extended system, where an even distribution of dt' is used, a mathematical transformation and justification must be made, which we will show in the following paragraphs.

Along the new trajectory, the time average of a certain physical quantity, *e.g.* $\mathbf{A}(\mathbf{p}(t')/s(t'), \mathbf{r}(t'))$, can be evaluated by:

$$\lim_{\tau' \to \infty} \frac{1}{\tau'} \int_0^{\tau'} \mathbf{A}(\mathbf{p}(t')/s(t'), \mathbf{r}(t'))dt' \tag{5.44}$$

according to the quasi-ergodic hypothesis, where an even distribution on dt' is imposed. Since $dt' = dt/s$ and the trajectory is determined for the real sub-system, the virtual time interval dt varies in each time step during the MD simulation. The τ' in Eq. 5.44 is related to the virtual time τ in Eq. 5.37 through:

$$\tau' = \int_0^\tau \frac{1}{s(t)} dt. \tag{5.45}$$

We note that till now, along the trajectory generated by Eq. 5.42, the time average of the physical quantity as given in Eq. 5.44 is not related to the ensemble average of any kind. In order to relate it to an ensemble average of some kind, we can use Eq. 5.45 and rewrite Eq. 5.44 in terms of the trajectory in the virtual time as:

$$\frac{1}{\tau'} \lim_{\tau' \to \infty} \int_0^{\tau'} \mathbf{A}(\mathbf{p}(t')/s(t'), \mathbf{r}(t'))dt' \tag{5.46a}$$

$$= \frac{\tau}{\tau'} \frac{1}{\tau} \lim_{\tau \to \infty} \int_0^{\tau} \frac{\mathbf{A}(\mathbf{p}(t)/s(t), \mathbf{r}(t))}{s(t)} dt \tag{5.46b}$$

$$= \frac{\frac{1}{\tau} \lim_{\tau \to \infty} \int_0^{\tau} \frac{\mathbf{A}(\mathbf{p}(t)/s(t), \mathbf{r}(t))}{s(t)} dt}{\tau'/\tau} \tag{5.46c}$$

$$= \frac{\frac{1}{\tau} \lim_{\tau \to \infty} \int_0^{\tau} \frac{\mathbf{A}(\mathbf{p}(t)/s(t), \mathbf{r}(t))}{s(t)} dt}{\frac{1}{\tau} \lim_{\tau \to \infty} \int_0^{\tau} \frac{1}{s(t)} dt} \tag{5.46d}$$

.

Here $\frac{1}{\tau} \lim_{\tau \to \infty} \int_0^\tau \cdots dt$ is the time average of a physical quantity along the trajectory in the virtual time (that generated by Eq. 5.36).

Then, making use of Eq. 5.37, these quantities related to the time average of the virtual trajectory can be further simplifies into:

$$\frac{\left\langle \frac{\mathbf{A}(\mathbf{p}(t)/s(t), \mathbf{r}(t))}{s(t)} \right\rangle}{\left\langle \frac{1}{s(t)} \right\rangle}, \tag{5.47}$$

where $\langle \cdots \rangle$ denotes the ensemble average of the extended (Nosé) system. Now, we can make use of the partitioning function in Eq. 5.31 again, and further simplify

Eq. 5.47 into:

$$\begin{aligned}
\frac{\left\langle \frac{\mathbf{A}(\mathbf{p}(t)/s(t),\mathbf{r}(t))}{s(t)} \right\rangle}{\left\langle \frac{1}{s(t} \right\rangle} =& \frac{\left\{ \frac{\int d\mathbf{p'}_1 \cdots d\mathbf{p'}_N d\mathbf{r}_1 \cdots \mathbf{r}_N A(\mathbf{p'},\mathbf{r}) \exp\left[-\beta \frac{3N}{g} H(\mathbf{p'},\mathbf{r})\right]}{\int d\mathbf{p'}_1 \cdots d\mathbf{p'}_N d\mathbf{r}_1 \cdots \mathbf{r}_N \exp\left[-\beta \frac{3N+1}{g} H(\mathbf{p'},\mathbf{r})\right]} \right\}}{\left\{ \frac{\int d\mathbf{p'}_1 \cdots d\mathbf{p'}_N d\mathbf{r}_1 \cdots \mathbf{r}_N \exp\left[-\beta \frac{3N}{g} H(\mathbf{p'},\mathbf{r})\right]}{\int d\mathbf{p'}_1 \cdots d\mathbf{p'}_N d\mathbf{r}_1 \cdots \mathbf{r}_N \exp\left[-\beta \frac{3N+1}{g} H(\mathbf{p'},\mathbf{r})\right]} \right\}} \\
=& \frac{\int d\mathbf{p'}_1 \cdots d\mathbf{p'}_N d\mathbf{r}_1 \cdots \mathbf{r}_N A(\mathbf{p'},\mathbf{r}) \exp\left[-\beta \frac{3N}{g} H(\mathbf{p'},\mathbf{r})\right]}{\int d\mathbf{p'}_1 \cdots d\mathbf{p'}_N d\mathbf{r}_1 \cdots \mathbf{r}_N \exp\left[-\beta \frac{3N}{g} H(\mathbf{p'},\mathbf{r})\right]} \\
=& \frac{\frac{1}{N!} \int d\mathbf{p'}_1 \cdots d\mathbf{p'}_N d\mathbf{r}_1 \cdots \mathbf{r}_N A(\mathbf{p'},\mathbf{r}) \exp\left[-\beta \frac{3N}{g} H(\mathbf{p'},\mathbf{r})\right]}{\frac{1}{N!} \int d\mathbf{p'}_1 \cdots d\mathbf{p'}_N d\mathbf{r}_1 \cdots \mathbf{r}_N \exp\left[-\beta \frac{3N}{g} H(\mathbf{p'},\mathbf{r})\right]}
\end{aligned} \tag{5.48}$$

From this equation, it is clear that if we take $g = 3N$ this quantity equals the ensemble average of a canonical system simulated by Eq. 5.42, which we denote as $\langle \cdots \rangle_{\mathrm{c}}$. And by linking Eqs. 5.44, 5.46, 5.47, with 5.48, we can get:

$$\frac{1}{\tau'} \lim_{\tau' \to \infty} \int_0^{\tau'} \mathbf{A}(\mathbf{p}(t')/s(t'), \mathbf{r}(t'))dt' = \langle A(\mathbf{p'},\mathbf{r}) \rangle_{\mathrm{c}} \tag{5.49}$$

It is worth to note that in Eq. 5.46 the quantity to be integrated has a denominator $s(t)$, which cancels one $s(t)$ in the $s(t)^{3N+1}$ in Eq. 5.31. It is this tiny difference that results in the fundamentally different manners for the canonical ensemble to be sampled by the trajectories from Eq. 5.36 and Eq. 5.42. In the former case, the "Nosé" thermostat simulates a canonical ensemble using trajectory generated with an evenly distributed virtual time and the parameter g needs to be set as $3N + 1$. While in the latter case, the refined form of the "Nosé" thermostat simulates a canonical ensemble using trajectory generated with an evenly distributed real time and the parameter g is set as $3N$. In most of the current implementations of this thermostat, the latter form is used.

We note that in real implementations of Eq. 5.42, it can be further simplified by replacing the variable s' by ξ' and p'_s by $p'_{\xi'}$, where $\xi' = \ln s'$ and $p'_{\xi'} = s' p_{s'}$. In so

doing, Eq. 5.42 can be rewritten as:

$$\frac{d\mathbf{r}_i{}'}{dt'} = \frac{\mathbf{p}_i{}'}{m_i} \tag{5.50a}$$

$$\frac{d\mathbf{p}_i{}'}{dt'} = -\frac{\partial V(\mathbf{r}_1{}', \cdots, \mathbf{r}_N{}')}{\partial \mathbf{r}_i{}'} - \frac{p'_{\xi'}}{Q}\mathbf{p}_i{}' \tag{5.50b}$$

$$\frac{d\xi'}{dt'} = \frac{p'_{\xi'}}{Q} \tag{5.50c}$$

$$\frac{dp'_{\xi'}}{dt'} = \sum_{i=1}^{N} \mathbf{p}'^2_i/m_i - g/\beta. \tag{5.50d}$$

Here, Q is still a parameter we need to set in our simulations which is associated with the mass of the particle in the extra degree of freedom. From Eq. 5.50 (b), it is clear that $p'_{\xi'}$ is associated with the friction coefficient in controlling the temperature. Since the variables $\mathbf{r}_i{}'$, $\mathbf{p}_i{}'$, dt' are the real variables, g equals $3N$. The energy which is conserved (Eq. 5.43) is now reformed to:

$$H'_{\text{Nosé}} = \sum_{i=1}^{N} \frac{\mathbf{p}'^2_i}{2m_i} + V(\mathbf{r}'_1, \cdots, \mathbf{r}'_N) + \frac{p'^2_{\xi'}}{2Q} + \frac{g}{\beta}\xi'. \tag{5.51}$$

Again, we note that this is not a real Hamiltonian since the equations of motion can't be generated from it. It is a quantity which should be conserved during the MD simulation.

5.2.3 Nosé-Hoover Chain

As simple as it it, the Nosé-Hoover thermostat succeeded in sampling the canonical ensemble in a series of systems. There are, however, many situations when the quasi-ergodic hypothesis fails. A one-dimensional system when the external potential is a harmonic oscillator is such an example. In this case, the Andersen thermostat still works well on sampling the Maxwell-Boltzmann distribution of the velocities. The Nosé-Hoover thermostat, however, only samples a micro-canonical ensemble [242]. Now, it is realized that the reason for this breakdown is due to the fact that in the deduction of the equations from Eq. 5.28 to Eq. 5.33, it is essential to identify all the conserved quantities [246].

In the 1990s, a scheme in which the Nosé-Hoover thermostat is coupled to other thermostats to take into account of these additional conservation laws has been proposed by Martyna and Tuckerman *et al.* to get over this deficiency [245, 246]. In this improved version of the Nosé-Hoover thermostat, a system of N nuclei is

coupled to a chain of thermostat (say the chain number is M) through:

$$\frac{d\mathbf{r}_i{}'}{dt'} = \frac{\mathbf{p}_i{}'}{m_i} \tag{5.52a}$$

$$\frac{d\mathbf{p}_i{}'}{dt'} = -\frac{\partial V(\mathbf{r}_1{}', \cdots, \mathbf{r}_N{}')}{\partial \mathbf{r}_i{}'} - \frac{p'_{\xi'_1}}{Q_1}\mathbf{p}_i{}' \tag{5.52b}$$

$$\frac{d\xi'_1}{dt'} = \frac{p'_{\xi'_1}}{Q_1} \tag{5.52c}$$

$$\frac{dp'_{\xi'_1}}{dt'} = \left[\sum_{i=1}^{N} \frac{\mathbf{p}_i'^2}{m_i} - g/\beta\right] - \frac{p'_{\xi'_2}}{Q_2} p'_{\xi'_1} \tag{5.52d}$$

$$\frac{d\xi'_j}{dt'} = \frac{p'_{\xi'_j}}{Q_j} \quad \cdots\cdots\cdots\cdots\cdots\cdots 2 \le j \le M-1 \tag{5.52e}$$

$$\frac{dp'_{\xi'_j}}{dt'} = \left[\frac{p'^2_{\xi'_{j-1}}}{Q_{j-1}} - 1/\beta\right] - \frac{p'_{\xi'_{j+1}}}{Q_{j+1}} p'_{\xi'_j} \tag{5.52f}$$

$$\frac{d\xi'_M}{dt'} = \frac{p'_{\xi'_M}}{Q_M} \tag{5.52g}$$

$$\frac{dp'_{\xi'_M}}{dt'} = \left[\frac{p'^2_{\xi'_{M-1}}}{Q_{M-1}} - 1/\beta\right], \tag{5.52h}$$

where $g = 3N$. The conserved total energy out of these equations of motion is:

$$H'_{\mathrm{NHC}} = \sum_{i=1}^{N} \frac{\mathbf{p}_i'^2}{2m_i} + V(\mathbf{r}'_1, \cdots, \mathbf{r}'_N) + \sum_{j=1}^{M} \frac{p'^2_{\xi'_j}}{2Q_j} + \frac{g}{\beta}\xi'_1 + \sum_{j=2}^{M} \frac{1}{\beta}\xi'_j. \tag{5.53}$$

Now, it has been shown that the Nosé-Hoover chain is very efficient in sampling ergodically the statistic property of many systems. It is also of practical use that dynamic properties such as the vibrational frequency can be extracted from their trajectories, although a rigorous justification of such applications is absent. As mentioned in the end of Sec. 5.2.1, after the BO MD method become a standard routine in practical *ab initio* MD simulations since the late 1990s, other stochastic methods such as the Andersen and Langevin thermostats also became widely used. In Sec. 5.2.1, we have introduced the Andersen thermostat. In the following, we will give a brief introduction to the method of controlling temperature using Langevin dynamics.

5.2.4 Langevin Thermostat

The Langevin dynamics is one simple and efficient scheme to control the temperature of a canonical ensemble, which is often used in *ab initio* MD simulations nowadays.

The equations of motion for the particles are:

$$\dot{\mathbf{r}}_i = \mathbf{p}_i/m_i \tag{5.54a}$$
$$\dot{\mathbf{p}}_i = \mathbf{F}_i - \gamma\mathbf{p}_i + \mathbf{R}_i \tag{5.54b}$$

Here, $\mathbf{r}_i$ and $\mathbf{p}_i$ are the position and velocity of the i^{th} particle in the simulation. $\mathbf{F}_i$ is the force imposed by the other particles explicitly treated in the simulation. The two parameters which are essential to control the temperature in this scheme are γ and $\mathbf{R}_i$. The first one (γ) represents a viscous damping due to fictitious "heat bath" particles, while the second one ($\mathbf{R}_i$) represents the effect of collisions with these particles. A balance between them, which decelerates or accelerates the particles to be simulated respectively, keeps the average kinetic energy and consequently the temperature of the system constant.

In MD simulations where all atoms are explicitly treated, these "heat bath" particles are purely fictitious, originating from the thermal fluctuations. In computer simulation of large molecules in solvent composed by smaller ones (but these solvent molecules won't be explicitly simulated), this "heat bath" represents thermal effects originating from real collisions between the smaller molecules in the solvent which are implicitly treated and the large ones which are explicitly treated under simulation. Therefore, in the limiting case for very large friction coefficient, the Brownian movement of the large molecules in solvent can also be described by this equation, which simplifies into:

$$\dot{\mathbf{r}}_i = \gamma^{-1} m_i^{-1} \left[\mathbf{F}_i + \mathbf{R}_i\right] \tag{5.55}$$

because $\dot{\mathbf{p}}_i = 0$. In this book, we focus on the MD simulations where all nuclei are explicitly treated. Therefore the Brownian movement of the large molecules in solvent is unrelated to the main problem we want to solve. But it is worthwhile to note that the principles underlying this dynamics also apply to these problems.

There are four properties for the stochastic forces $\mathbf{R}_i$ to comply with in order to control the temperature in simulations using Langevin dynamics, *i.e.* i) they are uncorrelated with $\mathbf{r}_i$ and $\mathbf{p}_i$, ii) the autocorrelation function of $\mathbf{R}_i(t)$ has a form of δ-function, iii) their mean values are zero, and iv) their distribution is Gaussian-like and the unit variance of this Gaussian function is $\sqrt{2k_{\text{B}}T\gamma m/\Delta t}$. For a clarification of these four properties, we note that the first one is easy to understand since it just implies that the random force imposed by the collisions of the nuclei and the fictitious "heat bath" particles is uncorrelated with the position and momentum of the nuclei. The second point needs some explanation. It originates from the fact that the timescale of the collisions between the particles under simulations and the fictitious "heat bath" ones is much smaller than that of the ionic movement. Therefore, the autocorrelation function of the random force $\mathbf{R}_i$ reads:

$$\langle \mathbf{R}_i(t)\mathbf{R}_i(t')\rangle = 2k_{\text{B}}T\gamma m\delta(t-t'). \tag{5.56}$$

These t and t' only take values with integer time-step along the trajectory. Besides this, it is also assumed that during a MD movement, there are many collisions between the particles under simulation and the "heat bath". Therefore, one can impose a Gaussian distribution on $\mathbf{R}_i$ based on the central limit theorem [256]. Then the Stokes-Einstein relation for the diffusion coefficient can be used to show that the unit variance associated with this Gaussian distribution is $\sqrt{2k_{\mathrm{B}}T\gamma m/\Delta t}$ and its mean value is zero.

With this, the principles underlying the choice of $\mathbf{R}_i$ for the simulation of a canonical ensemble using the Langevin dynamics is clear. We note that this choice of $\mathbf{R}_i$ limits this thermostat to the equilibrium simulations only. For its extension to the non-equilibrium state simulations, one may refer to Ref. [257]. The last point which needs to be explained for the simulation of a canonical ensemble using Langevin dynamics is that the equations of motion in Eq. 5.54 are equivalent to a Fokker-Planck equation in description of the phase space probability density distribution function $\rho(\mathbf{r}_1, \cdots, \mathbf{r}_N, \mathbf{p}_1, \cdots, \mathbf{p}_N)$ of the form:

$$\frac{\partial \rho}{\partial t} + \sum_{i=1}^{N} \left[\frac{\mathbf{p}_i}{m_i} \cdot \nabla_{\mathbf{r}_i} \rho + \mathbf{F}_i \cdot \nabla_{\mathbf{p}_i} \rho \right] = \gamma \sum_{i=1}^{N} \nabla_{\mathbf{p}_i} \cdot \left[\mathbf{p}_i \rho + m_i k_{\mathrm{B}} T \nabla_{\mathbf{p}_i} \rho \right], \tag{5.57}$$

which has the canonical phase space probability distribution function as a stationary solution. Therefore, this method is justified for the simulation of the canonical ensemble. From these principles, it is clear that the choice of γ is the key issue in this simulation, which actually determines a compromise between the statistical sampling efficiency of the preservation of the accuracy in short-term dynamics. As a stochastic method, one should not expect the dynamics to be comparable with those from the simulations using deterministic trajectories, like those of Nosé-Hoover thermostat simulations. But the stochastic feature normally guaranteed the ergodicity to be sampled in a more efficient way.

In the end, it is worthwhile to note that although traditional stochastic thermostats often cause breakdown of the adiabatic separation in the CP MD simulations, a recent development of the Langevin dynamics actually showed that using correlated noise, it is possible to tune the coupling of a stochastic thermostat with the various degrees of freedom for the nuclear motion [254]. This development not only enables the using of the Langevin dynamics in the CP MD simulations, it also improves the canonical sampling efficiency because this thermostat can be tailored in a predictable and controlled fashion. For more details about this recent development, please refer to Ref. [254].

5.2.5 Constant Pressure, Andersen and Parrinello-Rahman Barostats

In simulations of a NVT ensemble, a non-flexible unit cell is used. Therefore, although the thermal fluctuation of the nuclei can be properly described using thermostats, the fluctuation of the simulation cell is completely absent. This fluctuation of the unit cell is a consequence of the mechanical contact of the system with its environment, which is inevitable in most experiments. Therefore, similar to the control of temperatures using thermostats in the NVT simulations, it is also highly desirable that a barostat can be introduced to account for this fluctuation of the unit cell, so that the environment under simulation is as close to real experiments as possible.

Based on this consideration, there has been many methods proposed to control this quantity in the MD simulations, from the earliest Andersen scheme [239, 258], Parrinello-Rahman scheme [247, 248], to the more recent work using Langevin dynamics [256, 259]. The corresponding simulations are called NPT, or isobaric-isothermal simulations in literature. In this section, we will use the Andersen and Parrinello-Rahman schemes as examples to illustrate how this is done. Principles underlying the scheme using Langevin dynamics please refer to Refs. [256, 259].

In the Andersen scheme [239], a constant pressure simulation method for a fluid system is proposed by introducing the instantaneous volume of the simulation cell as an environmental variable in the effective Lagrangian. Due to the isotropic feature of the system (fluid) to be simulated, a cubic cell is used with volume Ω. The side of this cubic cell equals $\Omega^{1/3}$. Then, the Cartesian coordinates of the nuclei ($\mathbf{r}_i$) can be scaled with respect to the instantaneous size of the simulation cell by through $\mathbf{x}_i = \mathbf{r}_i/(\Omega^{1/3})$. In terms of these scaled variables and the instantaneous Ω, the Lagrangian can be written as:

$$L_{\text{Andersen}} = \sum_{i=1}^{N} \frac{m_i}{2} \Omega^{2/3} \dot{\mathbf{x}}_i^2 - \sum_{i>j} V(\Omega^{1/3} \mathbf{x}_{i,j}) + \frac{M}{2} \dot{\Omega}^2 - P_0 \Omega. \tag{5.58}$$

Here, M is a parameter which can be viewed as an effective inertia for the expansion and contraction of the volume. P_0 is the external potential. $\mathbf{x}_{i,j}$ is the scaled distance (with no unit) between the i^{th} atom and the j^{th} atom in the contracted or expanded cell.

There is a physical interpretation of the $P_0\Omega$ term in Eq. 5.58, in the original words of Andersen: "*Suppose the fluid to be simulated is in a container of variable volume. The fluid can be compressed by a piston. Thus, Ω, whose value is the volume, is the coordinate of the piston, $P_0\Omega$ is a PV potential derived from an external pressure P_0 acting on the piston, and M is the mass of the piston. The piston is not of the usual cylindrical type that expands or contracts the system along only one*

direction; instead, a change in Ω *causes an isotropic expansion or contraction.*" But we note that this interpretation is not entirely consistent with the Lagrangian in Eq. 5.58, since from $\mathbf{x}_i = \mathbf{r}_i/(\Omega^{1/3})$, one easily derives

$$\dot{\mathbf{r}}_i = \Omega^{1/3}\dot{\mathbf{x}}_i + \frac{1}{3}\Omega^{-2/3}\dot{\Omega}\mathbf{x}_i. \tag{5.59}$$

This equation means that the kinetic energy of the atoms should also contain contributions from $\dot{\Omega}$, which is absent in Eq. 5.58. In spite so, we acknowledge that Eq. 5.58 still presents a well-defined Lagrangian.

Now we analyze the dynamics generated from this Lagrangian. The momentum conjugate to $\mathbf{x}_i$ is denoted as π_i, which equals

$$\pi_i = \frac{\partial L_{\text{Andersen}}}{\partial \dot{\mathbf{x}}_i} = m_i\Omega^{2/3}\dot{\mathbf{x}}_i. \tag{5.60}$$

The momentum conjugate to Ω is denoted as Π, which equals

$$\Pi = \frac{\partial L_{\text{Andersen}}}{\partial \dot{\Omega}} = M\dot{\Omega}. \tag{5.61}$$

The Hamiltonian corresponding to the Lagrangian in Eq. 5.58 is then written as:

$$\begin{aligned} &H_{\text{Andersen}}(\mathbf{x}_1, \cdots, \mathbf{x}_N, \pi_1, \cdots, \pi_N, \Omega, \Pi) = \sum_{i=1}^{N} \dot{\mathbf{x}}_i \cdot \pi_i + \Omega\Pi - L_{\text{Andersen}} \\ &= (2\Omega^{2/3})^{-1} \sum_{i=1}^{N} \frac{1}{m_i}\pi_i \cdot \pi_i + \sum_{i>j} V(\Omega^{1/3}\mathbf{x}_{i,j}) + (2M)^{-1}\Pi^2 + P_0\Omega \end{aligned} \tag{5.62}$$

With this Hamiltonian, one can derive the equations of motion as:

$$\begin{aligned} \frac{d\mathbf{x}_i}{dt} =& \frac{H_{\text{Andersen}}}{\partial \pi_i} = \frac{\pi_i}{m_i\Omega^{2/3}} \\ \frac{d\pi_i}{dt} =& -\frac{H_{\text{Andersen}}}{\partial \mathbf{x}_i} = -\Omega^{1/3} \sum_{i>j} \frac{\mathbf{x}_{i,j}}{|\mathbf{x}_{i,j}|} \frac{\partial V(\Omega^{1/3}\mathbf{x}_{i,j})}{\partial \mathbf{x}_i} \\ \frac{d\Omega}{dt} =& \frac{H_{\text{Andersen}}}{\partial \Pi} = \frac{\Pi}{M} \\ \frac{d\Pi}{dt} =& \frac{H_{\text{Andersen}}}{\partial \Omega} \\ =& -(3\Omega)^{-1} \left(-2(2\Omega^{2/3})^{-1} \sum_{i=1}^{N} \frac{1}{m_i}\pi_i \cdot \pi_i + \Omega^{1/3} \sum_{i<j} \mathbf{x}_{i,j} \frac{\partial V(\Omega^{1/3}\mathbf{x}_{i,j})}{\partial \mathbf{x}_i} + 3P_0\Omega \right) \end{aligned}$$

With these equations of motion, one can use the MD method as introduced in Sec. 5.1 for the micro-canonical system to a trajectory for $\mathbf{x}_1, \cdots, \mathbf{x}_N$, $\pi_1, \cdots, \pi_N$,

Ω, and Π. We note that the volume is not a conserved quantity in real space. However, the scaled extended system with variables $\mathbf{x}_1, \cdots, \mathbf{x}_N$, $\pi_1, \cdots, \pi_N$, Ω, and Π can still be regarded as a micro-canonical ensemble in which the volume is unity since the coordinates $\mathbf{x}_i$ are all scaled to lie within a dimensionless unit cell. Under the ergodic hypothesis, the time average over this trajectory is equivalent to the ensemble average of the micro-canonical extended system. Therefore, for any function of $\mathbf{x}_1, \cdots, \mathbf{x}_N$, $\pi_1, \cdots, \pi_N$, Ω, and Π, the time average over this trajectory will be regards as equivalent to the ensemble average of this quantity on the extended system. And the NPT ensemble average of the real sub-system can be naturally evaluated using this trajectory too. We note that this method is one of the first methods which employed an extended Lagrangian in a MD simulation. Later development of many ensemble simulation methods, *e.g.* the Nosé-Hoover thermostat, has benefitted a lot from this idea.

One shortcoming of the Andersen barostat is that an isotropic change of the simulation cell must be assumed. This assumption is reasonable in simulations of fluids. However, when it comes to solids, especially in the case when the unit cell is not cubic, in addition to the volume, one must also allow the shape of the simulation cell to change. This generalization is essential in studies of structural phase transitions. Based on this consideration, Parrinello and Rahman soon designed a scheme which allows the change of the simulation cell's shape in MD simulations [247, 248]. In this method, instead of using only one variable Ω, nine variables associated with three vectors in the Cartesian coordinate ($\mathbf{a}$, $\mathbf{b}$, and $\mathbf{c}$) are used which allow both the shape and the volume of the simulation cell to be changed. These nine variables are $\mathbf{A}_{i,j}$, where i (and j) goes from 1 to 3. $\mathbf{A}$ is the matrix representation of ($\mathbf{a}$, $\mathbf{b}$, $\mathbf{c}$). The volume (Ω) of the simulation cell equals $\mathbf{a} \cdot (\mathbf{b} \times \mathbf{c})$. The Lagrangian for this extended system is:

$$L_{\mathrm{PR}} = \frac{1}{2} \sum_{i=1}^{N} m_i (\mathbf{A}\dot{\mathbf{x}}_i)^{\mathrm{T}} \mathbf{A}\dot{\mathbf{x}}_i - \sum_{i>j} V(\mathbf{A}^{-1}\mathbf{x}_{i,j}\mathbf{A}) + \frac{M}{2} \sum_{i,j=1}^{3} \dot{\mathbf{A}}_{i,j} - P_0\Omega, \qquad (5.63)$$

Using the same trick as we have applied to the Andersen barostat (and Nosé-Hoover thermostat), one can derive the equations of motion for $\mathbf{x}_i$ and $\mathbf{A}_{i,j}$. We don't go through this deduction here, for such details, please refer to Refs. [247, 248].

5.3 Examples for Practical Simulations in Real Poly-Atomic Systems

Once understanding the principles of MD as introduced above, one can easily carry out a practical MD simulation using a computer code which is accessible in either a commercial or open-source way. Taking one of the most often used setting, an *ab*

initio Born-Oppenheimer MD simulation using a Nosé-Hoover Chain thermostat for the NVT ensemble as an example, in Fig. 5.1, we show one snapshot of a system composed by 32 water molecules and one extra proton. This simulation is based on the implementation of the MD method in the VASP code [260, 261, 262, 263, 264]. As mentioned in the introduction, the nuclei are treated as classical point-like particles. Therefore, the ball-and-stick model is still used for the representation of the nuclei in this snapshot. The electronic structures, however, are computed "on-the-fly" in an *ab initio* manner as the system evolves. This is reflected by the silver contour representing the density distribution of the electrons at this snapshot, which describes the bond making and breaking processes during the MD simulation in a seamless manner.

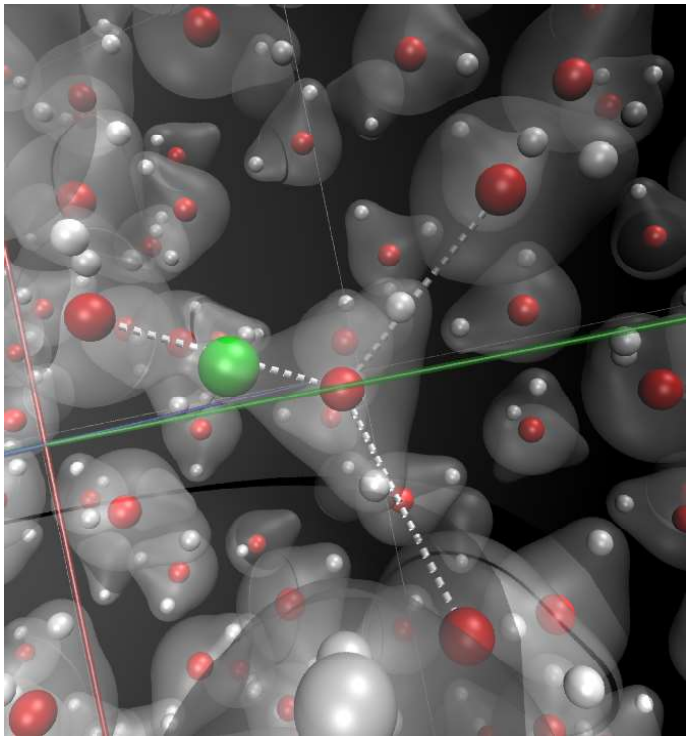

Figure 5.1: Illustration on how an *ab initio* MD simulation is like. In this figure, we show one snapshot of an *ab initio* MD simulation on a system composed by 32 water molecules and one extra proton. The extra proton is denoted by the sphere in green. The oxygen nuclei and other hydrogen nuclei are shown by spheres in red and white respectively. The silver contour means the electron density distribution, which is computed in an *ab initio* manner "on-the-fly" as the system evolves.

From such snapshots and statistical (as well as dynamical) analysis on them, one can get quite some information concerning the evolution of the electronic structures

and associated nuclear motion at the molecular level in such systems. For example, one can analyze the electron charge transfer between the extra proton and the bulk water solution explicitly using the electronic structure of such a snapshot. Specifically, this can be done by calculating the electron density distribution of the system composed by the water molecules and the extra proton first, and then that for each of them (solely the water molecules and solely the proton) separately. Taking the difference between the first one and the sum over the latter two, one obtains how the electron charge density is redistributed by the interactions between the extra proton and the water solution in a very clean manner. In Fig. 5.2, we show how this electron charge transfer is like in the neighborhood of the excess proton. Besides the electron accumulation and depletion on the water molecules and hydrogen bond associated with the extra proton, obvious electron charge redistribution exists in the neighboring region, indicating that the hydrogen bonds strength in the second shell are also significantly influenced by the presence of this extra proton.

The above analysis also indicates that the behaviors of the hydrogen bonds in the neighboring shells and the excess proton are correlated. Because of this, for a better understanding of the proton transport mechanism in liquid water, one can investigate in more detail the hydrogen bond dynamics in the presence of this excess proton. As a matter of fact, this issue has been intensively studied from both the experimental and theoretical perspectives in the last two decades [264, 265, 266, 267, 268, 269, 270, 271, 272, 273, 274]. Now, it has been widely accepted that the behavior of this extra proton is basically determined by the dynamics of the hydrogen bonds within the neighboring shells [264, 265, 266, 271]. This story can be understood pictorially by taking some sequential snapshots of the MD simulation, as shown in Fig. 5.3. In these snapshots, the pivot oxygen (O_p) is defined as the oxygen nucleus which owns three hydrogens simultaneously, and O_s is defined as the oxygen nucleus which shares the most active proton with the pivot one. The pair (hydrogen bond) between O_p and O_s is called "the special pair". In the MD simulation, the pivot oxygen normally stays on one site, *i.e.* the pivot oxygen identity doesn't change, peaceful for a certain time interval ($\sim$ 1 ps). During this time, the spacial pair switches identity between the hydrogen bonds in which the pivot oxygen is involved frequently, which is often called the special pair dancing in literature [267, 268, 269, 270, 271]. The coordination number of O_s is normally 4 during this dancing process (the first three panels in Fig. 5.3). Upon going from the third panel to the forth one, the coordination number of O_s reduces from 4 to 3. Then proton transport happens (the fifth panel). After the proton transport the most active proton begins to exchange sites about a new pivot oxygen (the last two panels). In short, the proton transport happens through the well-known Grotthuss mechanism [275, 276], with the hydrogen bond dynamics determining the inter-conversion between different hydration states of the excess proton.

The second example we want to show briefly here for the practical simulations

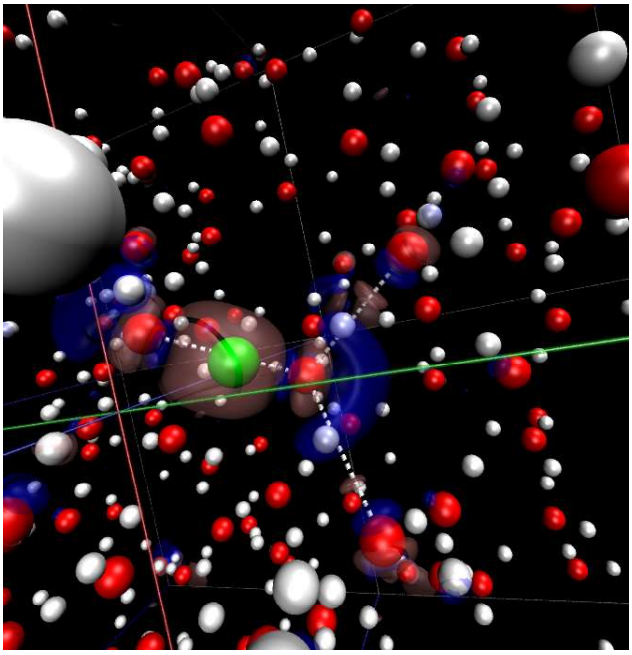

Figure 5.2: Illustration on how the electron charge transfer is analyzed in an *ab initio* MD simulation. The same simulation cell and notation of the nuclei are used as in Fig. 5.1. Contour in brown means electron charge accumulation and contour in blue means electron charge depletion. This analysis shows that an extraordinarily strong hydrogen bond is associated with the extra proton. Besides this, significant electron charge redistribution also exists on the neighboring hydrogen bonds, enhancing most of them. Therefore, in order for a reasonable characterization of the transfer process of this extra proton, besides the dynamics of this extra proton, the hydrogen bond dynamics of the "shell" is also essential.

using the *ab initio* MD method is on the phase diagram of condensed matter under extreme conditions, e.g. high pressures. It is well-known that experimental realization of pressures above 100 GPa is extremely difficulty. For theoreticians, however, this is readily doable since an *ab initio* MD simulation using a NPT ensemble easily reproduces the atomic level evolution of the system under investigation, if i) the electronic structures are accurate, ii) the simulation cell is large enough, and iii) the simulation time is long enough. As a matter of fact, this simulation technique has already been widely used in studies of the high pressure phase diagrams for a series of systems [277, 278, 279, 280, 281]. One of the most prominent example is the high pressure phase diagram of hydrogen, as shown in Fig. 5.4. Below 300 GPa, it is well-known that three regions exist, consisting of the molecular solid, molecular liquid, and the atomic liquid. The dissociation line between the molecular liquid and the atomic liquid as shown in the figure was just determined by the *ab initio*

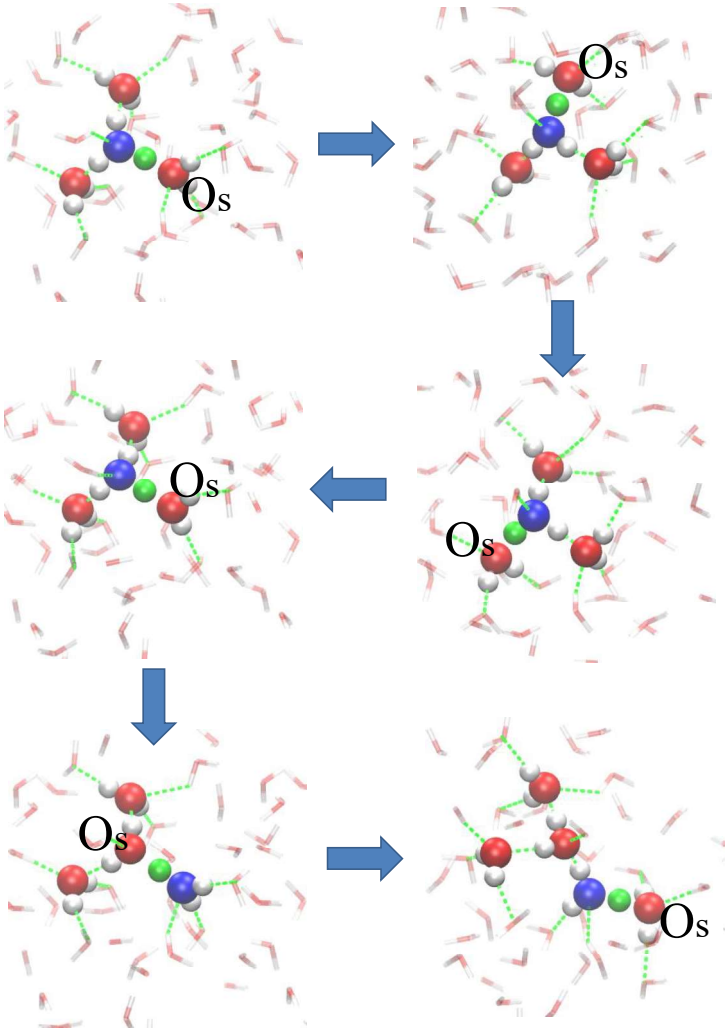

Figure 5.3: Illustration on the hydrogen bond dynamics related to the proton transport in liquid water obtained from an *ab initio* MD simulation using classical nuclei. Red spheres are oxygen nuclei and white spheres are hydrogen nuclei. The excess proton is in green and the pivot oxygen (O_p) is in blue. Hydrogen bonds (except for the most active one) are indicated by green dashed lines. O_s is the oxygen which shares the most active proton with the pivot oxygen. The first three panels show that the most active proton switches identity between the three hydrogen bonds related to O_p. Going from the third to the forth, the coordination number of O_s reduces from 4 to 3. Then proton transport happens upon going from the forth to the fifth panel. After the proton transport the most active proton begins to exchange sites about a new pivot oxygen (the last two panels).

MD simulations in Refs. [277, 278].

We note that the accuracy of the electronic structures obtained from the present *ab initio* method, which is affordable for the MD simulations, is still a problem which determines the accuracy of the theoretical predictions of this phase diagram [282, 283, 284, 285]. As a matter of fact, this is a problem which is vital for the accuracy of the MD simulations in general. It is also why when chemical bond breaking and forming happen, non-reactive force field based method must be abandoned in descriptions of the inter-atomic interactions and *ab initio* MD method prevails in recent years. Therefore, as the last reminder of this chapter, when the MD simulation technique is used to interpret the experiments or make predictions, one must pay special care to this accuracy of the inter-atomic interactions, for all spatial configurations visited during the simulation.

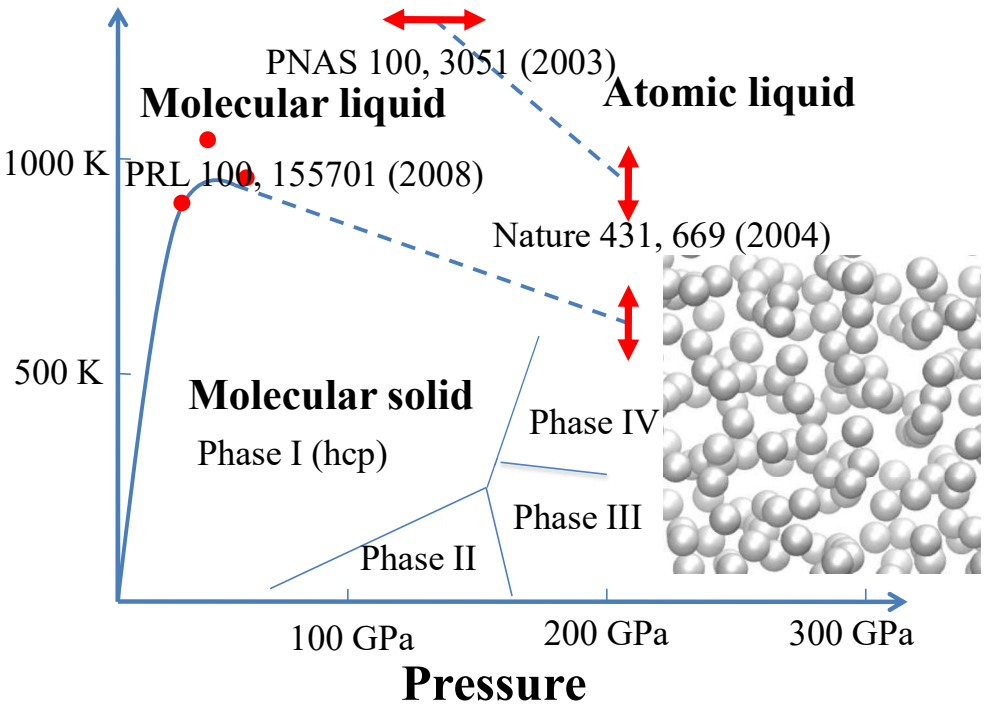

Figure 5.4: Illustration on hydrogen phase diagram under pressure. The dissociation line between the molecular liquid phase and the atomic liquid phase was obtained using the *ab initio* MD simulations, in Refs. [277, 278].