

FERNANDO M. S. SILVA FERNANDES
BENEDITO J. COSTA CABRAL
Departamento de Química & CECUL (INIC)
Rua Ernesto de Vasconcelos, Bloco C1-Piso 5
1700 Lisboa,
PORTUGAL



MOLECULAR DYNAMICS BY COMPUTER SIMULATION (*)

The purpose of this paper is to present a short survey of molecular dynamics methods by computer simulation with incidence on general aspects rather than on technicalities.

(*) Presented at the meeting «Química Teórica para a Biotecnologia em Portugal», Estalagem da Boega (Vila Nova de Cerveira), 19-22 July 1987.

1. INTRODUCTION

In the present context, molecular dynamics (MD) and Monte Carlo (MC) methods are understood as computer simulation techniques for studying thermodynamic and structural properties of solids and fluids.

MD methods [1-2] produce trajectories in phase space by numerically solving Newton's equations of motion for the molecules in the sample. Time averages are then calculated over those trajectories.

MC methods [3-4] generate statistical ensembles by giving random displacements to the molecules and accepting, or rejecting, the resulting configurations with a probability proportional to appropriate Boltzmann factors. Ensemble averages are then calculated over the generated configurations.

According to statistical mechanics [5] time and ensemble averages are equal. Therefore MD and MC methods should be equivalent as far as equilibrium properties are concerned. In fact, MC methods can not deal with dynamic phenomena. On the contrary, MD is able to probe both equilibrium and dynamic properties.

MC methods are easier to implement in a computer program for they only require the calculation of the total energy of each configuration while MD also requires the forces on each particle. The computer effort for a comparable statistical accuracy is, however, of the same order of magnitude for MD and MC.

The purpose of this paper is to present a short survey of molecular dynamics methods with greater incidence on general aspects rather than on technicalities.

2. BOUNDARY CONDITIONS

Due to memory and computer time requirements the number, N , of molecules that is possible to follow in time with the present generation of computers is of order 10^2 - 10^4

depending on the complexity of the molecules and the type of computer.

This is a rather small number for the simulation of bulk systems where the number of molecules is of order 10^{23} . Therefore, the use of boundary conditions is inevitable in order to eliminate surface effects.

and the bulk system is approximated by a periodically repeated sample. This is illustrated in Figure 1.

The potential energy and the force on each molecule are calculated by taking into account the periodicity of the system as we shall see.

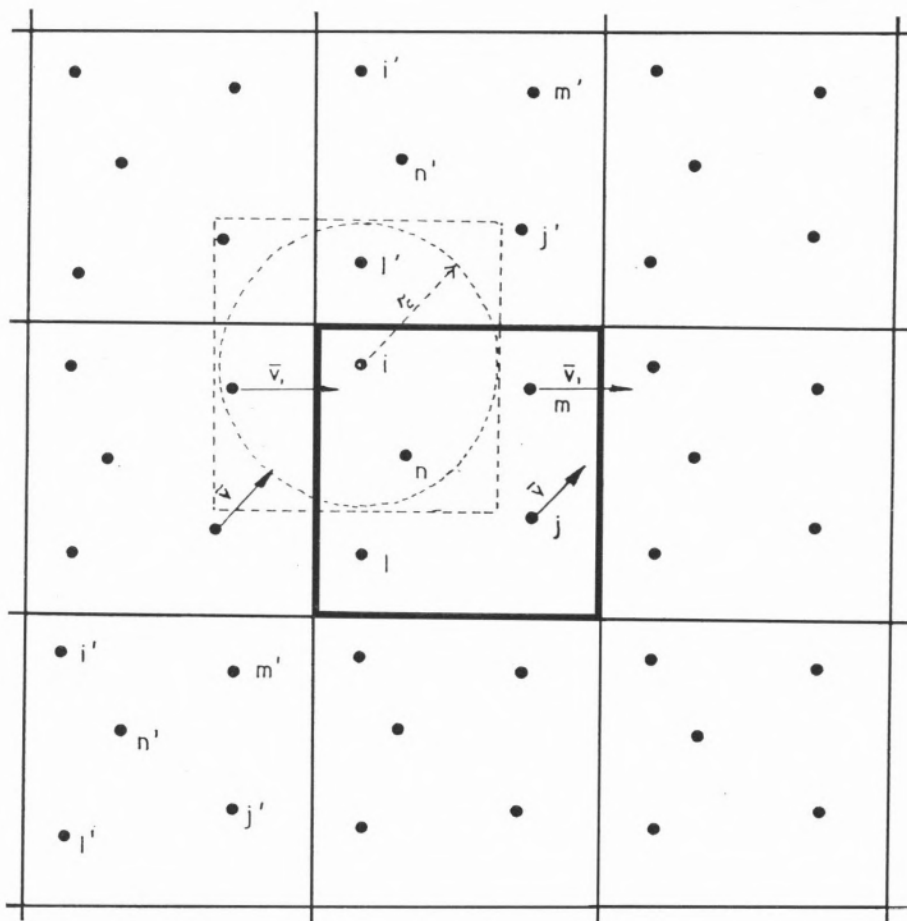


Figure 1

*MD box, its translation and boundary conditions in two dimensions.
(For clarity, only five molecules are shown in each square).*

Cubic boundary conditions are common in most of the calculations: N molecules are enclosed in a cubic box (MD box) whose volume, V , is chosen according to a pre-defined density. The MD box is made to behave as though it were part of an infinite system by surrounding it by periodically repeated images of itself. Thus, all surfaces are eliminated

The density is maintained by assuming that when a molecule leaves the MD box through a wall, an image enters the box, through the opposite wall, with the same velocity.

There are a variety of boundary conditions [6-7]. However, apart from the cube, the most suitable shape seems to be the truncated octahedron.

3. MOLECULAR DYNAMICS AT CONSTANT ENERGY

The following steps are involved in a conventional molecular dynamics program:

a) Assign initial positions, $\mathbf{r}_i(0)$, to the molecules in the MD box.

The positions of a lattice are generally chosen. Alternatively, the positions of a previous run can be taken.

b) Assign initial molecular velocities $\mathbf{v}_i(0)$, such that the total momentum is zero.

It is common practice to assign initial velocities from a Maxwell-Boltzmann distribution [8], but there are other alternatives [9].

c) Calculate the intermolecular potential energy, $U(\mathbf{r}^N)$, and derive the force on each molecule:

$$\mathbf{F}_i = -\nabla_i U(\mathbf{r}^N) \quad (1)$$

The potential energy is usually assumed to be pairwise additive:

$$U(\mathbf{r}^N) = \sum_{i < j} u(r_{ij}) \quad (2)$$

where $u(r_{ij})$ is an effective pair potential [10] and $r_{ij} = |\mathbf{r}_i - \mathbf{r}_j|$.

The calculation of forces and potential energy take into account the boundary conditions. We shall return to that in the next paragraph.

d) Integrate Newton equations of motion for each of the N molecules taking into account the boundary conditions.

There are a variety of numerical algorithms to carry out the integrations, but a very simple and stable one is the Verlet algorithm [11] which may be written [12] in the so-called leapfrog version:

$$\begin{aligned} \mathbf{v}_i(t + \delta t/2) &= \mathbf{v}_i(t - \delta t/2) + \mathbf{F}_i(t) \delta t / m_i \\ \mathbf{r}_i(t + \delta t) &= \mathbf{r}_i(t) + \mathbf{v}_i(t + \delta t/2) \cdot \delta t \\ \mathbf{v}_i(t) &= [\mathbf{v}_i(t + \delta t/2) + \mathbf{v}_i(t - \delta t/2)]/2 \end{aligned} \quad (3)$$

where

$$\mathbf{v}_i(t - \delta t/2) = [\mathbf{r}_i(t) - \mathbf{r}_i(t - \delta t)]/\delta t$$

m_i is the molecular mass and t the integration time-step.

The time-step must be less than the molecular relaxation times and is of order $10^{-16} - 10^{-14}$ seconds depending on the type of molecules.

e) Let the system evolve in time for n_e time-steps where n_e is the number needed to reach equilibrium.

f) Let the system evolve in time for the next n_p time-steps where n_p is the number appropriate to obtain a good statistical accuracy.

g) Calculate time averages over the n_p steps.

The above procedure keeps the volume and the number of molecules constant. Furthermore, once the initial positions and velocities are assigned, the initial total energy of the system is defined:

$$E(0) = 1/2 \sum_i^N m_i v_i^2(0) + U(\mathbf{r}^N(0)) \quad (4)$$

As the system evolves in time without external interferences, then $E(t) = E(0)$. Thus the total energy is conserved along the trajectory in phase space. According to statistical mechanics the time average of any property over that trajectory should be equivalent to the microcanonical ensemble average. In other words, microcanonical ensemble averages can be obtained from the molecular dynamics trajectory.

As the total energy, volume and number of particles are kept constant, the temperature and pressure, for example, fluctuate along the trajectory. The system will reach equilibrium when the averages of fluctuating instantaneous properties are constant in time.

The instantaneous temperature is defined from equipartition theorem [5]:

$$T = \sum_i^N m_i v_i^2 / 3Nk \quad (5)$$

where k is the Boltzmann constant.

The instantaneous pressure is defined from virial theorem [5]:

$$P = (1/3V) \left(\sum_i^N m_i v_i^2 + \sum_i^N \sum_{j>i}^N \mathbf{r}_{ij} \cdot \mathbf{F}_{ij} \right) \quad (6)$$

The algorithm (3) is only appropriate for the molecular centre of mass motion. For polyatomic molecules, rotational and vibrational degrees of freedom must be considered. In many systems of importance, namely hydrocarbons melts, polymers and lipid bilayers, internal rotation and vibration are important in determining equilibrium and dynamical properties.

The inclusion of vibration directly in simulations is not an easy matter. It is essentially a quantum mechanical phenomenon which can not be tackled using classical equations of motion. Herman and Berne [13] described a method of including quantum vibration. On the other hand, the simulation of quantum systems is now possible by means of molecular dynamics of wave packets [14-15] and path integral Monte Carlo methods [16].

The motion of rigid models, or models with internal rotations, may be treated classically. There are a variety of algorithms to describe rotation [17-18]. A standard one is the SHAKE algorithm [19] and its more recent version RATLE [18]. They are based on the Verlet algorithms and allow the integrations to be carried out in cartesian coordinates.

3.1 Calculation of Potential Energy and Forces

The evaluation of potential energy and forces must consider not only the particles in MD box but also their images in order to eliminate surface effects and to approach a bulk system. When the interactions between molecules are short-ranged, that is to say, when they can be neglected after ≈ 3 molecular diameters (the so-called cut-off distance, r_c) the calculation is carried out using the minimum image approximation: given a particle i , the MD box is translated so that is centred on i (see Figure 1). Then, the particle i only interacts explicitly with the particles or images within the sphere of radius r_c centred on particle i . The distance r_c must be less than (or equal to) half of the MD box side length. After the cut-off distance, it is assumed an uniform

distribution of particles and long range corrections are analytically calculated [11].

The minimum image approximation with truncation is largely used in the study of non-polar systems, in particular, of noble gases.

Such an approximation is however unsuitable for ionic or highly polar systems where the electrostatic interactions extend over many molecular diameters. All images have then to be accounted for. The classical method for calculating the electrostatic potential of a system of charges in a periodic cell is the Ewald sum method [20]. However the expression and the computer code [21] are rather involved. Recently Adams and Dubey [22] derived a modified expression for the Ewald sum in terms of an effective pair potential. This enables the calculation to be performed in a very simple and efficient way. The modified Ewald sum is straightforwardly introduced in a conventional program because the calculations are also carried out with the minimum image convention, but now without truncation. Recently Fernandes and Cabral [23] have applied the method to a simulation of molten salts.

3.2 Other Properties

Kinetic and potential energies, temperature and pressure referred to above are first order thermodynamic properties and primary output data from a simulation.

Given an instantaneous property X , its time average over the MD trajectory is denoted by $\langle X \rangle$. For example, the time average of the temperature is:

$$\langle T \rangle = \langle \sum_i^N m_i v_i^2 / 3Nk \rangle \quad (7)$$

Second order properties, such as specific heats and thermal-pressure coefficients, can be calculated from the fluctuations in first order thermodynamic properties. For example, the heat capacity at constant volume is given by:

$$C_V / Nk = 3/2 [1 - 2N (\langle EK^2 \rangle - \langle EK \rangle^2) / 3k^2 T^2]^{-1} \quad (8)$$

where EK is the kinetic energy. They have been thoroughly reviewed by Cheung [24] and by Haile and Graben [25].

Chemical potential, and related properties, can be evaluated by Widom method [26-27]. It basically involves establishing a $m \times m \times m$ grid within the MD cell ($m \approx 10$). At each time step, the potential energy u_i of a test particle is calculated on each of the grid points. Then the chemical potential is given by:

$$\mu = -kT \ln \langle \exp(-\beta u_i) \rangle \quad (9)$$

where $\beta = (kT)^{-1}$.

The pair radial distribution function, $g(r)$, a structural property, is calculated by dividing the configuration space around each molecule into concentric spherical shells of thickness δr . If the average number of neighbours in a shell is $n(r)$ then:

$$g(r) = (V/N)n(r)/4\pi r^2 \delta r \quad (10)$$

The radial distribution function is related to the static structure factor through a Fourier transform [4]. The structure factor, in turn, is obtained from X-ray and neutron scattering. Thus, it is possible to compare the simulated functions with experimental results. On the other hand, the radial distribution function forms the underlying structural basis for first-order thermodynamic properties [28].

Time correlation functions [4] can also be obtained by MD. For instance, the velocity auto-correlation function defined as

$$Z(t) = \langle \mathbf{v}(t) \cdot \mathbf{v}(0) \rangle \quad (11)$$

is an important quantity. It elucidates the diffusive processes in liquids [29]. The Fourier transform of (11) gives the frequency spectrum and the integration gives the self-diffusion coefficient:

$$D = \frac{1}{3} \int_0^\infty Z(t) dt \quad (12)$$

Other transport coefficients, spectroscopic and scattering properties are directly related to

time correlation functions [4,30]. Therefore, they may be quantitatively predicted from MD. Dynamic structure of liquids, which is related to inelastic neutron scattering, is obtained through the space-time van Hove correlation function [4]. Moreover the study of current fluctuations [4], inevitable in order to get a full understanding of the collective modes of a liquid, can also be realized by MD.

4. MOLECULAR DYNAMICS AT CONSTANT TEMPERATURE AND PRESSURE

MD at constant energy has limitations from an experimental standpoint. In fact, as far as experiments are concerned, it is desirable to perform simulations at a constant pre-defined temperature and/or pressure, which is not possible with MD at constant energy.

One way to carry out simulations at these conditions is to integrate the equations of motion subjected to constraints such that the trajectories in phase space correspond to (N, V, T) , (N, P, H) or (N, P, T) statistical ensembles.

There are a variety of methods to maintain temperature at a constant pre-defined value [12,31]. The method we generally follow is the damped force method of Hoover [32] and Evans [33]. It is theoretically founded and easily incorporated into the leapfrog algorithm.

The basic assumption of the method is that each molecule is subjected to the total force.

$$\mathbf{F}_t = \mathbf{F}_i - \alpha \mathbf{p}_i / m_i \quad (13)$$

where \mathbf{F}_i is the force due to the other molecules and $\alpha \mathbf{p}_i / m_i$ is a damping force whose parameter α is calculated by the condition of constant instantaneous temperature $dT/dt = 0$. The motion equations are:

$$\mathbf{p}_i = \mathbf{F}_i - \alpha \mathbf{p}_i / m_i \quad (14)$$

where \mathbf{p}_i is the linear momentum.

The leapfrog algorithm (3) is easily modified to introduce this constraint. The formulae is given by Brown and Clark [12] in a readily implementable form.

To maintain the pressure at a pre-defined value there also are some methods [17,34]. A standard one is the Andersen method [35,36]. Both volume and pressure are allowed to fluctuate, in the later case about a fixed mean. The MD box volume is expanded if the instantaneous internal pressure, P_i , is greater than the desired value P_e and contracted if P_i is less than P_e .

The volume equation of motion is:

$$\ddot{V} = [P_i - P_e]/M \quad (15)$$

where M is the so-called «piston mass».

The motion equations for the particles are:

$$\ddot{\mathbf{R}}_i = \mathbf{F}_i/m + \mathbf{R}_i/3V [\ddot{V} - (2/3)(\dot{V}/V)^2] \quad (16)$$

and

$$\dot{\mathbf{r}}_i = \dot{\mathbf{R}}_i - \mathbf{R}_i \dot{V}/3V \quad (17)$$

$\dot{\mathbf{R}}_i$ is the actual velocity of particle i and $\dot{\mathbf{r}}_i$ its «thermodynamic velocity», that is to say, the velocity related to internal forces and that contribute to thermodynamic properties.

The above equations are also easily incorporated in the leapfrog algorithm. The formulae is given by Brown and Clark [12] and Fincham and Heyes [17].

The piston mass is an adjustable parameter and is chosen so that the fluctuations in the volume do not reach an unphysical frequency or the system does not undergo a catastrophic irreversible expansion or contraction.

5. NONEQUILIBRIUM MOLECULAR DYNAMICS

The MD methods described so far are equilibrium methods. They assume that the system is not subjected to external perturbations

(such as shear forces, temperature gradients, electric fields, etc.) and that it attains equilibrium before the calculation of any properties. Even the calculation of transport coefficients by means of time correlation functions referred to in 3.1., assume the system to be in equilibrium when the correlation functions are evaluated. This is justified by linear response theory [4] which establishes the relation between the linear response of a system to an external perturbation and the properties of the system in equilibrium.

Apart from the self-diffusion coefficient, it is not efficient to calculate transport coefficients by equilibrium MD. The corresponding time correlation functions are difficult to obtain with good accuracy and the computer time involved is often prohibitive [37]. However, it is possible to perform molecular dynamics with the system subjected to external perturbations and to study directly the response of the system. The theory shows [17] that the ratio of the response to the external field gives the transport coefficient of interest. Such methods are known as nonequilibrium molecular dynamics (NEMD) or non-standard molecular dynamics [38-40].

One way of performing NEMD has been devised by Cicotti and Jacucci [41] and is based on the calculation of two trajectories: one with and one without the applied perturbation. The net effect of the perturbation is then obtained by subtraction of the zero-perturbation trajectory from that with the applied perturbation. The method allows the applied perturbations to be sufficiently small so that there is no problem with nonlinear effects. Good statistics are obtained from averaging over a few hundred pairs of trajectories. An important feature of the method is that it produces not merely the steady-state response to an applied perturbation but the average time-dependent response, so that it becomes possible to obtain such properties as the frequency-dependent shear viscosity.

NEMD techniques are not restricted to use for the calculation of transport coefficients but

also can successfully be used to obtain second-order thermodynamic quantities [17] as an alternative to the fluctuation method referred to in 3.1.

New NEMD techniques which have proved to be useful in exploring molecular systems far from equilibrium are presently being investigated [42].

6. FINAL COMMENTS

Computer simulation is a powerful technique to clarify and even to provide new insights into the molecular dynamics of condensed states of matter. It can be used to study either relatively simple problems like the thermodynamic properties of simple systems or complex problems like proteins [43] or rotation-translation coupling [44]. The latter is important, for example, to explain why the physical properties of enantiomers and of their racemic mixture are different. For instance, at room temperature, a racemic mixture of lactic acid is a liquid whereas the individual enantiomers are solid.

The sizes of molecules and the ensembles that may be studied will grow as computer power grows [45]. Intermolecular potentials will be improved and some approximations like the assumption of pairwise additivity will be better understood.

Computer simulation of the molecular dynamics is an invaluable tool for the development of analytical theories and to overcome their inherent mathematical problems.

Our group in Lisbon is involved in the study of simple systems, alcohols, surfaces, and transport properties by MD and MC methods. We welcome the portuguese biochemists and biotechnologists.

ACKNOWLEDGMENTS

One of us (F. Fernandes) is grateful to the Organising Committee of Boega meeting for the kind invitation.

(Received, 14th December 1987)

REFERENCES

- [1] J. B. ALDER, and WAINRIGHT, *J. Chem. Phys.*, **31**, 459 (1959).
- [2] A. RAHMAN, *Phys. Rev.*, **136A**, 405 (1964).
- [3] N. METROPOLIS, A. W. ROSENBLUTH, M. N. ROSENBLUTH, and A. J. TELLER, *J. Chem. Phys.*, **21**, 1087 (1953).
- [4] J. P. HANSEN and I. R. McDONALD, «Theory of Simple Liquids», Academic Press (1976).
- [5] T. L. HILL, «Statistical Mechanics», McGraw Hill (1956).
- [6] D. J. ADAMS, *Information Quarterly*, CCP5, **10**, 30 (1983).
- [7] W. SMITH, *Information Quarterly*, CCP5, **10**, 37 (1983).
- [8] D. W. HEERMANN, «Computer Simulation Methods in Theoretical Physics», Springer-Verlag (1986).
- [9] F. M. S. S. FERNANDES, Ph. Thesis, University of Southampton (1977).
- [10] G. MATTLAND, M. RIGBY, B. SMITH and W. WAKEHAM, «Intermolecular Forces. Their Origin and Determination», Clarendon Press (1981).
- [11] L. VERLET, *Phys. Rev.*, **159**, 98 (1967).
- [12] D. BROWN and J. H. R. CLARKE, *Mol. Phys.*, **51**, 1243 (1984).
- [13] M. F. HERMAN and B. BERNE, *J. Chem. Phys.*, **78**, 4103 (1983).
- [14] W. SMITH and K. SINGER, *Information Quarterly*, CCP5, **13**, 38 (1984).
- [15] W. SMITH and K. SINGER, *Information Quarterly*, CCP5, **17**, 53 (1985).
- [16] W. SMITH and K. SINGER *Information Quarterly*, CCP5, **25**, 18 (1987).
- [17] D. FINCHAM and D. M. HEYES, *Adv. Chem. Phys.*, **63**, 493 (1985).
- [18] H. ANDERSEN *J. Comp. Phys.*, **52**, 24 (1983).
- [19] J. P. RYCHAERT, G. CICOTTI and H. J. C. BERENDSEN, *J. Comp. Phys.*, **23**, 327 (1977).
- [20] P. P. EWALD, *Ann. Physik (Leipzig)*, **64**, 253 (1921).
- [21] W. SMITH, *Information Quarterly*, CCP5, **21**, 37 (1986).
- [22] D. ADAMS and G. DUBEY, *Information Quarterly*, CCP5, **22**, 39 (1986).
- [23] F. M. S. S. FERNANDES, and J. B. CABRAL, *Portugal. Phys.*, **18**, 39 (1987).
- [24] P. S. Y. CHEUNG, *Mol. Phys.*, **33**, 519 (1977).
- [25] J. M. HAILE and H. W. GRABEN, *Mol. Phys.*, **40**, 1433 (1980).
- [26] K. S. SHING and K. E. GUBBINS, *Mol. Phys.*, **46**, 1109 (1982).
- [27] J. G. POWLES, *Chem. Phys. Lett.*, **86**, 335 (1982).
- [28] H. J. M. HANLEY and D. J. EVANS, *Mol. Phys.*, **39**, 1039 (1980).
- [29] A. RAHMAN, *J. Chem. Phys.*, **45**, 2585 (1966).
- [30] D. J. TILDESLEY, Proc. of Nato Advanced Study Institute of Molecular Liquids Dynamics and Interactions, Italy (1983), p. 519.
- [31] D. M. HEYES, *Chem. Phys.*, **82**, 285 (1983).
- [32] W. G. HOOVER, *Physica*, **A118**, 111 (1983).
- [33] D. J. EVANS, *J. Chem. Phys.*, **78**, 3279 (1983).
- [34] H. J. C. BERENDSEN, J. P. M. POSTMA, W. F. VAN GUNSTEREN, A. DiNOLA and J. R. HAAK, *J. Chem. Phys.*, **81**, 3684 (1984).

- [35] H. C. ANDERSEN, *J. Chem. Phys.*, **72**, 2384 (1980).
- [36] J. HAILE and H. GRABEN, *J. Chem. Phys.*, **73**, 2412 (1980).
- [37] E. M. GOSLING, I. R. McDONALD and K. SINGER, *Mol. Phys.*, **26**, 1475 (1973).
- [38] W. G. HOOVER, *Ann. Rev. Phys. Chem.*, **34**, 103 (1983).
- [39] W. G. HOOVER *Phys. Today*, **37**, 44 (1984).
- [40] D. J. EVANS, J. M. HOWARD and H. SIEGFRIED, *Phys. Today*, **37**, 26 (1984).
- [41] CICOTTI and JACUCCI, *Phys. Rev. Lett.*, **35**, 789 (1975).
- [42] W. G. HOOVER, H. A. POSCH, B. L. HOLIAN, M. J. GILLIAN, M. MARECHAL and C. MASSOBRIO, *Molecular Simulation*, **1**, 79 (1987).
- [43] W. VAN GUNSTEREN and H. BERENDSEN, *J. Mol. Biol.*, **176**, 559 (1984).
- [44] M. EVANS and G. EVANS, *Adv. Chem. Phys.*, **63**, 377 (1985).
- [45] D. FINCHAM, *Molecular Simulation*, **1**, 1 (1987).

RESUMO

Dinâmica Molecular por Simulação Computacional

O objectivo deste artigo é o de apresentar uma breve revisão dos métodos de dinâmica molecular por simulação computacional com mais ênfase em aspectos genéricos do que em detalhes técnicos.