

Molecular Dynamics simulations of Biomolecules

Adolfo Bastida

Universidad de Murcia (Spain)



Máster Universitario en Bioinformática

Modelado Molecular

1 Solvents

- ## 2 Periodic boundary conditions
- Surface effects
 - Periodic boundary conditions
 - Long-range interactions

- ## 3 Molecular Dynamics
- Hamilton's equations
 - Finite difference methods
 - Ensembles
 - Tests
 - Static properties
 - Dynamic properties



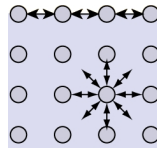
Solvents

- Most of chemical reactions take place in solution
- Good solvent $\left\{ \begin{array}{l} \text{Inert} \\ \text{Able to solve all reactants} \\ \text{Adequate ebullition point} \\ \text{Easy to separate} \end{array} \right.$
- Inert $\neq$ neutral $\Rightarrow \left\{ \begin{array}{l} \text{Energy fluxes} \\ \text{Intermolecular interactions} \end{array} \right.$
- Quantum-mechanical simulations $\Rightarrow$ computationally forbidden.
- Molecular Dynamics



Periodic boundary conditions

- Cubic box $\Rightarrow$ Surface effects $\Rightarrow$ Surface tension



- 
 $5^3=125$ molecules

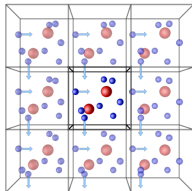
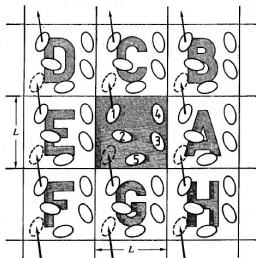
{	Surface	$25+(25-5)+(25-9)+(25-9)$ $+(25-13)+(25-16)=98$ (78.4 %)
	Bulk	$125-98=27$ (21.6 %)
- 
 $6^3=216$ molecules

{	Surface	$36+(36-6)+(36-11)+(36-11)$ $+(36-16)+(36-20)=152$ (70.4 %)
	Bulk	$216-152=64$ (29.6 %)
- 
 $10^3=1000$ mol.

{	Surface	$100+(100-10)+(100-19)+(100-19)$ $+(100-28)+(100-36)=488$ (48.8 %)
	Bulk	$1000-488=512$ (51.2 %)



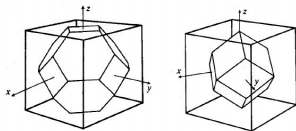
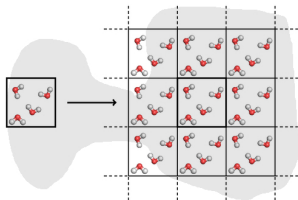
Periodic boundary conditions



- Central box $\Rightarrow$ infinite lattice
- Periodic image in each box moves in the same way
- If a molecule leaves a box, one of its images will enter through the opposite face
- The number density in all boxes is constant
- Only $(\mathbf{r}_i, \mathbf{p}_i)$ of the molecules in the central box are stored



Periodic boundary conditions

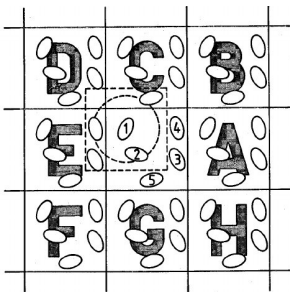


- Special conditions { Phase transitions
Supercritical fluids
- Test $\Rightarrow$ increase box size
- Other boxes $\Rightarrow$ space filling polyhedra

{ Rhombic dodecahedron
Truncated octahedron



Periodic boundary conditions

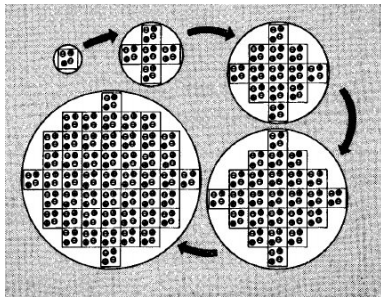


- Infinite number of interactions (potential energy and forces)
- Minimum Image Convention $\Rightarrow$ each molecule interacts with the closest periodic image of the other $N - 1$ molecules
- Define a box centered in molecule 1 $\Rightarrow$ interacts with 2, 3_E, 4_E, and 5_C
- Pairwise-additive interactions $\Rightarrow \frac{1}{2} N(N - 1)$ terms



Long-range interactions

Ewald summation



X. Quian and T. Schlick, *Journal of Chemical Physics* 116, 5971 (2002)

$$V = \frac{1}{2} \sum_{\mathbf{n}} \sum'_{i,j} \frac{q_i q_j}{|\mathbf{r}_{ij} + \mathbf{n}|}$$

$$\mathbf{n} = (n_x L, n_y L, n_z L)$$

$$\sum' \Rightarrow i \equiv j, \mathbf{n} \neq 0$$

$$\begin{aligned} \gamma^{zz}(\epsilon_s = 1) = & \frac{1}{2} \sum_{i=1}^N \sum_{j=1}^N \left(\sum_{|\mathbf{n}|=0}^{\infty} z_i z_j \frac{\text{erfc}(\kappa |\mathbf{r}_{ij} + \mathbf{n}|)}{|\mathbf{r}_{ij} + \mathbf{n}|} \right. \\ & + (1/\pi L^3) \sum_{\mathbf{k} \neq 0} z_i z_j (4\pi^2/k^2) \exp(-k^2/4\kappa^2) \cos(\mathbf{k} \cdot \mathbf{r}_{ij}) \Big) \\ & - (\kappa/\pi^{1/2}) \sum_{i=1}^N z_i^2 + (2\pi/3L^3) \left| \sum_{i=1}^N z_i \mathbf{r}_i \right|^2. \end{aligned}$$



Hamilton's equations

- Newton's second law

$$\mathbf{F}_i = m_i \mathbf{a}_i = m_i \frac{d^2 \mathbf{r}_i}{dt^2} \Rightarrow \text{Second order differential equation}$$

- Hamilton's equations

$$\left. \begin{array}{l} \dot{\mathbf{r}}_i = \mathbf{p}_i / m_i \\ \dot{\mathbf{p}}_i = \mathbf{F}_i \end{array} \right\} \Rightarrow \text{Coupled first order differential equations}$$

- Equivalent formulations $\Rightarrow \mathbf{F}_i = \dot{\mathbf{p}}_i = m_i \ddot{\mathbf{r}}_i = m_i \mathbf{a}_i$

- Forces $\Rightarrow \mathbf{F}_i = -\nabla_{\mathbf{r}_i}^2 V = -\left(\frac{\partial V}{\partial x_i}, \frac{\partial V}{\partial y_i}, \frac{\partial V}{\partial z_i}\right)$



Verlet algorithm

$$\begin{aligned}
 \mathbf{r}(t) &= \mathbf{r}(t_0) + \underbrace{\dot{\mathbf{r}}(t_0)}_{\mathbf{v}(t_0)} (t - t_0) + \frac{1}{2} \underbrace{\ddot{\mathbf{r}}(t_0)}_{\mathbf{a}(t_0)} (t - t_0)^2 + \cancel{O((t - t_0)^3)} \\
 &= \mathbf{r}(t_0) + \mathbf{v}(t_0) (t - t_0) + \frac{1}{2} \mathbf{a}(t_0) (t - t_0)^2
 \end{aligned} \tag{1}$$

$$\mathbf{r}(t \pm \delta) = \mathbf{r}(t_0) + \mathbf{v}(t_0) (t \pm \delta - t_0) + \frac{1}{2} \mathbf{a}(t_0) (t \pm \delta - t_0)^2 \tag{2}$$

$$\begin{aligned}
 (1), (2) \Rightarrow \mathbf{r}(t+\delta) + \mathbf{r}(t-\delta) - 2\mathbf{r}(t) &= \cancel{\mathbf{r}(t_0)} + \cancel{\mathbf{v}(t_0)(t+\delta-t_0)} + \frac{1}{2} \mathbf{a}(t_0)(t+\delta-t_0)^2 \\
 &\quad + \cancel{\mathbf{r}(t_0)} + \cancel{\mathbf{v}(t_0)(t-\delta-t_0)} + \frac{1}{2} \mathbf{a}(t_0)(t-\delta-t_0)^2 \\
 &\quad - \cancel{2\mathbf{r}(t_0)} - \cancel{2\mathbf{v}(t_0)(t-t_0)} - \mathbf{a}(t_0)(t-t_0)^2 \\
 &= \frac{1}{2} \mathbf{a}(t_0) \left(\cancel{(t-t_0)^2} + \delta^2 + \cancel{2(t-t_0)\delta} \right) \\
 &\quad + \frac{1}{2} \mathbf{a}(t_0) \left(\cancel{(t-t_0)^2} + \delta^2 - \cancel{2(t-t_0)\delta} \right) \\
 &\quad - \cancel{\mathbf{a}(t_0)(t-t_0)^2} = \mathbf{a}(t_0) \delta^2
 \end{aligned}$$



Verlet algorithm

$$\mathbf{r}(t + \delta) = 2 \mathbf{r}(t) - \mathbf{r}(t - \delta) + \mathbf{a}(t_0) \delta^2$$

$$\boxed{\mathbf{r}(t_0 + \delta) = 2 \mathbf{r}(t_0) - \mathbf{r}(t_0 - \delta) + \mathbf{a}(t_0) \delta^2} \Leftarrow \text{Verlet}$$

$$(1) \Rightarrow \dot{\mathbf{r}}(t) = \mathbf{v}(t) = \mathbf{v}(t_0) + \mathbf{a}(t_0) (t - t_0) \quad (3)$$

$$\begin{aligned} \mathbf{r}(t+\delta) - \mathbf{r}(t-\delta) &= \cancel{\mathbf{r}(t_0)} + \mathbf{v}(t_0) (\cancel{t} + \delta - \cancel{t_0}) + \frac{1}{2} \mathbf{a}(t_0) (t + \delta - t_0)^2 \\ &\quad - \cancel{\mathbf{r}(t_0)} - \mathbf{v}(t_0) (\cancel{t} - \delta - \cancel{t_0}) - \frac{1}{2} \mathbf{a}(t_0) (t - \delta - t_0)^2 \\ &= 2 \mathbf{v}(t_0) \delta + \frac{1}{2} \mathbf{a}(t_0) \left(\cancel{(t-t_0)^2} + \delta^2 + 2(t-t_0)\delta - \cancel{(t-t_0)^2} - \delta^2 + 2(t-t_0)\delta \right) \\ &= 2 \mathbf{v}(t_0) \delta + 2 \mathbf{a}(t_0) (t - t_0) \delta = 2 \mathbf{v}(t) \delta \\ \mathbf{v}(t) &= \frac{\mathbf{r}(t+\delta) - \mathbf{r}(t-\delta)}{2 \delta} \end{aligned}$$

$$\boxed{\mathbf{v}(t_0) = \frac{\mathbf{r}(t_0+\delta) - \mathbf{r}(t_0-\delta)}{2 \delta}} \Leftarrow \text{Verlet}$$



Leap-frog algorithm

$$(3) \text{ with } t \rightarrow t_0 + \frac{\delta}{2} \Rightarrow \mathbf{v}(t_0 + \frac{\delta}{2}) = \mathbf{v}(t_0) + \mathbf{a}(t_0) (t_0 + \frac{\delta}{2} - t_0) = \mathbf{v}(t_0) + \frac{1}{2} \mathbf{a}(t_0) \delta \quad (4)$$

$$\begin{aligned} (2) \text{ with } t \rightarrow t_0 \Rightarrow \mathbf{r}(t_0 + \delta) &= \mathbf{r}(t_0) + \mathbf{v}(t_0) \delta + \frac{1}{2} \mathbf{a}(t_0) \delta^2 \\ &= \mathbf{r}(t_0) + \delta \left(\mathbf{v}(t_0) + \frac{1}{2} \mathbf{a}(t_0) \delta \right) \\ &= \mathbf{r}(t_0) + \mathbf{v}(t_0 + \frac{\delta}{2}) \delta \\ (4) \end{aligned}$$

$$\boxed{\mathbf{r}(t_0 + \delta) = \mathbf{r}(t_0) + \mathbf{v}(t_0 + \frac{\delta}{2}) \delta} \Leftarrow \text{Leap-frog}$$

$$(4) \Rightarrow \left. \begin{aligned} \mathbf{v}(t_0 + \frac{\delta}{2}) &= \mathbf{v}(t_0) + \frac{1}{2} \mathbf{a}(t_0) \delta \\ \mathbf{v}(t_0 - \frac{\delta}{2}) &= \mathbf{v}(t_0) - \frac{1}{2} \mathbf{a}(t_0) \delta \end{aligned} \right\} \mathbf{v}(t_0 + \frac{\delta}{2}) - \mathbf{v}(t_0 - \frac{\delta}{2}) = \mathbf{a}(t_0) \delta$$

$$\boxed{\mathbf{v}(t_0 + \frac{\delta}{2}) = \mathbf{v}(t_0 - \frac{\delta}{2}) + \mathbf{a}(t_0) \delta} \Leftarrow \text{Leap-frog}$$



Other algorithms

$$\mathbf{r}(t_0 + \delta) = \mathbf{r}(t_0) + \mathbf{v}(t_0) \delta + \frac{1}{2} \mathbf{a}(t_0) \delta^2$$

$$\mathbf{v}(t_0 + \frac{\delta}{2}) = \mathbf{v}(t_0) + \frac{1}{2} \mathbf{a}(t_0) \delta$$

$$\mathbf{v}(t_0 + \delta) = \mathbf{v}(t_0 + \frac{\delta}{2}) + \frac{1}{2} \mathbf{a}(t_0 + \delta) \delta$$

⇐ Velocity Verlet

$$\mathbf{r}(t_0 + \delta) = \mathbf{r}(t_0) + \mathbf{v}(t_0) \delta + \frac{2}{3} \mathbf{a}(t_0) \delta^2 - \frac{1}{6} \mathbf{a}(t_0 - \delta) \delta^2$$

$$\mathbf{v}(t_0 + \delta) = \mathbf{v}(t_0) + \frac{1}{3} \mathbf{a}(t_0 + \delta) \delta + \frac{5}{6} \mathbf{a}(t_0) \delta - \frac{1}{6} \mathbf{a}(t_0 - \delta) \delta$$

⇐ Beeman

Verlet, Leap-frog, velocity Verlet, Beeman, ... ⇒ Similar global errors



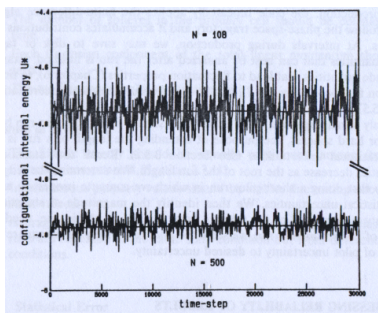
References

- L. Verlet. *Computer “experiments” on classical fluids. I. Thermodynamical properties of Lennard-Jones molecules.* *Phys. Rev.* **159**, 98-103 (1967).
- W.C. Swope, H.C. Andersen, P.H. Berens and K.R. Wilson. *Computer simulation method for the calculation of equilibrium constants for the formation of physical clusters of molecules: Application to small water clusters.* *J. Chem. Phys.* **76**, 637-649 (1982).
- W.H. Press, S.A. Teukolsky, W.T. Vetterling and B.P. Flannery. *Numerical recipes in Fortran 77.* *Cambridge University Press.* 1992. Chapter 16.7.



NVE ensemble

- NVE ensemble $\Rightarrow \begin{cases} N/V \rightarrow \text{density} \\ E \text{ constant} \rightarrow \text{Hamilton's equations} \end{cases}$



$$E = \frac{p^2}{2m} + V(\mathbf{r})$$

$$\dot{E} = \frac{2 \mathbf{p} \cdot \dot{\mathbf{p}}}{2m} + \dot{V}(\mathbf{r}) = \frac{\mathbf{p} \cdot \dot{\mathbf{p}}}{m} + \nabla_{\mathbf{r}} V \cdot \dot{\mathbf{r}}$$

$$= \frac{\mathbf{p}}{m} \cdot (\dot{\mathbf{p}} + \nabla_{\mathbf{r}} V) = 0 \Rightarrow \dot{\mathbf{p}} = -\nabla_{\mathbf{r}} V$$



NVT ensemble

- Equilibrium $\Rightarrow$ degree of freedom $\Rightarrow \frac{1}{2} k_B T$
- Temperature $\Rightarrow K = \sum_{i=1}^N \frac{1}{2} m_i \mathbf{v}_i^2 = 3N \left(\frac{1}{2} k_B T \right) \Rightarrow T = \frac{2K}{3N k_B}$
- Scaling

$$K = \sum_{i=1}^N \frac{1}{2} m_i \mathbf{v}_i^2 = 3N \left(\frac{1}{2} k_B T \right) \quad \left\{ \begin{array}{l} \frac{\sum_i m_i \mathbf{v}_i^2}{\sum_i m_i \mathbf{v}_i'^2} = \frac{T}{T_f} \rightarrow \sum_i m_i \mathbf{v}_i^2 = \sum_i m_i \mathbf{v}_i'^2 \frac{T}{T_f} \end{array} \right.$$

$$\mathbf{v}_i^2 = \mathbf{v}_i'^2 \frac{T}{T_f}$$

$$\mathbf{v}_i' = \frac{T_f}{T} \mathbf{v}_i$$



NVT ensemble

- Coupling to a external bath [Berendsen 1984]

$$\mathbf{v}_i' = \lambda \mathbf{v}_i \rightarrow \lambda = \sqrt{1 + \frac{\Delta t}{\tau} \left(\frac{T_f}{T} - 1 \right)}$$

- Other thermostats (Nose-Hoover,...)
- Equilibration (*NVT*) - Production (*NVE*)
- Other emsembles (*NPT*,...)



Conservation principles

- Mass $\Rightarrow M = \sum_i m_i$
- Energy $\Rightarrow E = \sum_i \frac{p_i^2}{2m_i} + V(\mathbf{r}_i)$
- Linear momentum $\Rightarrow \sum_i \mathbf{p}_i = 0$
- Angular momentum $\Rightarrow \sum_i \mathbf{r}_i \times \mathbf{p}_i = 0$

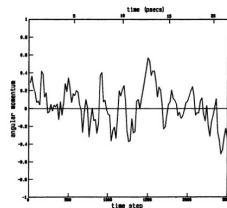


FIGURE 2.15 Fluctuations in the y -component of the angular momentum $\mathcal{Q}_y(t) = \sum_i (x_i p_{iz} - z_i p_{ix})$, computed during a molecular dynamics simulation of 108 Lennard-Jones atoms at $\rho r^3 = 0.7$, $kT/\epsilon = 1$. Units are σ for position and $(m/\epsilon)^{1/2}$ for velocity. Values are plotted at intervals of 20 integration time-steps; using potential parameters for argon, 20 time-steps = 0.172 psec.



Maxwell-Boltzmann distribution

- Velocity distribution

$$\rho(v_x) = \sqrt{\frac{m}{2\pi k_B T}} e^{-\frac{1}{2}mv_x^2/k_B T} \Leftarrow \int_{-\infty}^{\infty} \rho(v_x) dv_x = 1$$

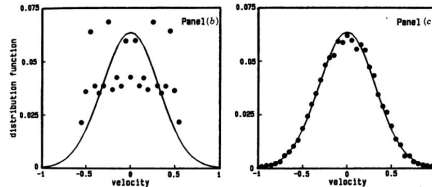


FIGURE 3.9 Approach to equilibrium for the 256-hard-sphere simulation described in Figure 3.7. (a) Same as Figure 3.8(a). (b) The velocity distribution $f(v_x)dv_x$ averaged over the first 5000 collisions of the run (points) compared to the Maxwell distribution (line) at the same kinetic energy. (c) The velocity distribution $f(v_x)dv_x$ averaged over collisions 6000–11,000 of the run (points) compared to the Maxwell distribution (line).



Static properties

- Time average

$$\langle A \rangle = \lim_{\tau \rightarrow \infty} \frac{1}{\tau} \int_0^{\tau} A(t) dt \approx \frac{1}{M} \sum_{k=1}^M A(t_k)$$

- Fluctuations

$$\delta A = A - \langle A \rangle$$

$$\begin{aligned} \langle (\delta A)^2 \rangle &= \langle (A - \langle A \rangle)^2 \rangle \\ &= \langle A^2 - 2A\langle A \rangle + \langle A \rangle^2 \rangle \\ &= \langle A^2 \rangle - \langle A \rangle^2 \end{aligned}$$



Static properties

- Pressure

$$P = \rho k_B T + \frac{\langle W \rangle}{V}$$

$$W = \frac{1}{3} \sum_i^N \mathbf{r}_i \cdot \mathbf{F}_i \rightarrow \text{internal virial}$$

- Heat capacity $\Rightarrow C_V = \left(\frac{\partial E}{\partial T}\right)_V \rightarrow C_V^* = \left(1 - \frac{2}{3NT^{*2}} \langle (\delta K^*)^2 \rangle\right)^{-1}$
- Adiabatic compressibility $\Rightarrow \kappa = -\frac{1}{V} \left(\frac{\partial V}{\partial P}\right)_S$

$$\kappa^* = \left(7P^* - \frac{16}{3}\rho^*T^* - 8\rho^*U^* - \frac{N}{\rho^*}T^*\langle(\delta P^*)^2\rangle\right)^{-1}$$

- Thermal pressure coefficient $\Rightarrow \gamma = \left(\frac{\partial P}{\partial T}\right)_V$

$$\gamma^* = \frac{2}{3} C_V^* \left(\rho^* - \frac{1}{T^{*2}} \langle \delta K^* \delta P^* \rangle\right)$$



Static properties

- Radial distribution function $g(r)$

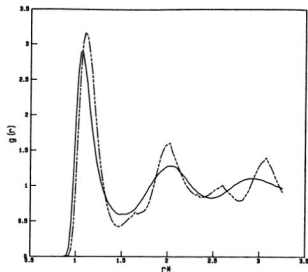


FIGURE 6.22 The radial distribution function $g(r)$ may often be used to distinguish a fluid state (continuous line) from a metastable solidlike state (broken line). These curves are from simulations that used the Lennard-Jones potential with 256 atoms at a density of $\rho\sigma^3 = 0.9$. The fluid state was at $kT/\epsilon = 1.087$, while the metastable state was at $kT/\epsilon = 0.80$. See Figure 5.7 for the location of these states on the Lennard-Jones phase diagram.

$$g(r) = \frac{\rho(r)}{\rho}$$

$$r \in (R, R + \delta) \rightarrow V = \frac{4}{3} \pi ((R + \delta)^3 - R^3)$$

$$\rho(R) = \frac{n(R)}{V}$$



Dynamic properties

- Time correlation functions

$$\langle A(t_0) B(t_0 + t) \rangle = \langle A(0) B(t) \rangle = \lim_{\tau \rightarrow \infty} \frac{1}{\tau} \int_0^\tau A(t_0) B(t_0 + t) dt_0$$

$$\simeq \frac{1}{M} \sum_{k=1}^M A(t_k) B(t_k + t)$$

- Limits $\langle A(0) B(t) \rangle$

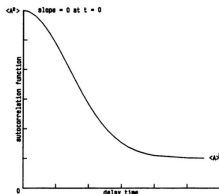


FIGURE 7.2 An autocorrelation function $\langle A(0)A(t) \rangle$ is initially equal to $\langle A^2 \rangle$, but as the delay time grows, it decays, initially with zero slope, to $\langle A \rangle^2$.

$$t = 0 \Rightarrow \langle AB \rangle \simeq \frac{1}{M} \sum_{k=1}^M A(t_k) B(t_k)$$

$$t \rightarrow \infty \Rightarrow \langle A \rangle \langle B \rangle \text{ (uncorrelated)}$$



Dynamic properties

- Transport coefficients

Flux = -coefficient $\times$ gradient

Fick's law $\Rightarrow N\dot{x} = -D \frac{\partial N}{\partial x} \begin{cases} N(x, t) \rightarrow \text{number atoms/unit volumen} \\ \dot{x}(x, t) \rightarrow \text{local velocity} \end{cases}$

TABLE 7.1 Generalized Einstein and Green-Kubo Formulas for Self-Diffusion Coefficient, Shear, and Longitudinal Viscosity

$$K = \lim_{t \rightarrow \infty} \langle [A(t) - A(0)]^2 \rangle / 2t = \int_0^\infty d\tau \langle \dot{A}(\tau) \dot{A}(0) \rangle$$

K	$A(t)$	$\dot{A}(t)$
Self-diffusion coefficient D	$x_i(t)$	$\dot{x}_i(t)$
Shear viscosity $\eta \, V kT$	$m \sum_i^N \dot{x}_i(t) y_i(t)$	$m \sum_i^N \dot{x}_i(t) \dot{y}_i(t) + \sum_{i < j}^N y_{ij}(t) F_{ijx}(t)$
Longitudinal viscosity ^a $(\frac{4}{3}\eta + \zeta) V kT$	$m \sum_i^N \dot{x}_i(t) x_i(t)$	$m \sum_i^N \dot{x}_i(t) \dot{x}_i(t) + \sum_{i < j}^N x_{ij}(t) F_{ijx}(t) - PV$

^a ζ is the bulk viscosity.