

# Computer Simulation of Liquids

M. P. ALLEN

*H. H. Wills Physics Laboratory  
University of Bristol*

and

D. J. TILDESLEY

*Department of Chemistry  
Imperial College of Science, Technology and Medicine, London*

CLARENDON PRESS • OXFORD

(2.45) we have zero, and on the right we have  $\sigma^2(E)$  at constant- $NVT$  and a combination of thermodynamic derivatives which turn out to equal  $(\partial E/\partial \beta) = -k_B T^2 C_V$ , where  $C_V$  is the constant-volume heat capacity.

## 2.4 Simple thermodynamic averages

A consequence of the equivalence of ensembles is that, provided a suitable phase function can be identified in each case, the basic thermodynamic properties of a model system may be calculated as averages in any convenient ensemble. Accordingly, we give in this section expressions for common thermodynamic quantities, omitting the subscripts which identify particular ensembles. These functions are usually derivatives of one of the characteristic thermodynamic functions  $\Psi_{\text{ens}}$ . Examples are  $P = -(\partial A/\partial V)_{NT}$  and  $\beta = (1/k_B T) = (1/k_B)(\partial S/\partial E)_{NV}$ .

The kinetic, potential, and total internal energies may be calculated using the phase functions of eqns (1.1)–(1.3):

$$E = \langle \mathcal{H} \rangle = \langle \mathcal{K} \rangle + \langle \mathcal{V} \rangle. \quad (2.46)$$

The kinetic energy is a sum of contributions from individual particle momenta, while evaluation of the potential contribution involves summing over all pairs, triplets etc. of molecules, depending upon the complexity of the function as discussed in Chapter 1.

The temperature and pressure may be calculated using the virial theorem, which we write in the form of 'generalized equipartition' [Münster 1969]:

$$\langle p_k \partial \mathcal{H} / \partial p_k \rangle = k_B T \quad (2.47a)$$

$$\langle q_k \partial \mathcal{H} / \partial q_k \rangle = k_B T \quad (2.47b)$$

for any generalized coordinate  $q_k$  or momentum  $p_k$ . These equations are examples of the general form  $\langle \mathcal{A} \partial \mathcal{H} / \partial q_k \rangle = k_B T \langle \partial \mathcal{A} / \partial q_k \rangle$  which may be easily derived in the canonical ensemble. They are valid (to  $\mathcal{O}(N^{-1})$ ) in any ensemble.

Equation (2.47a) is particularly simple when the momenta appear as squared terms in the Hamiltonian. For example, in the atomic case, we may sum up  $3N$  terms of the form  $p_{i\alpha}^2/m_i$ , to obtain

$$\left\langle \sum_{i=1}^N |\mathbf{p}_i|^2/m_i \right\rangle = 2 \langle \mathcal{K} \rangle = 3Nk_B T. \quad (2.48)$$

This is the familiar equipartition principle: an average energy of  $k_B T/2$  per degree of freedom. It is convenient to define an instantaneous 'kinetic temperature' function

$$\mathcal{T} = 2\mathcal{K}/3Nk_B = \frac{1}{3Nk_B} \sum_{i=1}^N |\mathbf{p}_i|^2/m_i \quad (2.49)$$

whose average is equal to  $T$ . Obviously, this is not a unique definition. For a system of rigid molecules, described in terms of centre-of-mass positions and velocities together with orientational variables, the angular velocities may also appear in the definition of  $\mathcal{T}$ . Alternatively, it may be useful to define separate 'translational' and 'rotational' temperatures each of which, when averaged, gives  $T$ . In eqn (2.47a) it is assumed that the independent degrees of freedom have been identified and assigned generalized coordinates  $q_k$  and momenta  $p_k$ . For a system of  $N$  atoms, subject to internal molecular constraints, the number of degrees of freedom will be  $3N - N_c$  where  $N_c$  is the total number of independent internal constraints (fixed bond lengths and angles) defined in the molecular model. Then, we must replace eqn (2.49) by

$$\mathcal{T} = \frac{2\mathcal{K}}{(3N - N_c)k_B} = \frac{1}{(3N - N_c)k_B} \sum_{i=1}^N |\mathbf{p}_i|^2/m_i. \quad (2.50)$$

We must also include in  $N_c$  any additional global constraints on the ensemble. For example, in the 'molecular dynamics' constant- $NVEP$  ensemble, we must include the three extra constraints on centre of mass motion.

The pressure may be calculated via eqn (2.47b). If we choose Cartesian coordinates, and use Hamilton's equations of motion (see Chapter 3), it is easy to see that each coordinate derivative in eqn (2.47b) is the negative of a component of the force  $\mathbf{f}_i$  on some molecule  $i$ , and we may write, summing over  $N$  molecules,

$$-\frac{1}{3} \left\langle \sum_{i=1}^N \mathbf{r}_i \cdot \nabla_{\mathbf{r}_i} \mathcal{V} \right\rangle = \frac{1}{3} \left\langle \sum_{i=1}^N \mathbf{r}_i \cdot \mathbf{f}_i^{\text{tot}} \right\rangle = -Nk_B T. \quad (2.51)$$

We have used the symbol  $\mathbf{f}_i^{\text{tot}}$  because this represents the sum of intermolecular forces and external forces. The latter are related to the external pressure, as can be seen by considering the effect of the container walls on the system:

$$\frac{1}{3} \left\langle \sum_{i=1}^N \mathbf{r}_i \cdot \mathbf{f}_i^{\text{ext}} \right\rangle = -PV. \quad (2.52)$$

If we define the 'internal virial'  $\mathcal{W}$

$$-\frac{1}{3} \sum_{i=1}^N \mathbf{r}_i \cdot \nabla_{\mathbf{r}_i} \mathcal{V} = \frac{1}{3} \sum_{i=1}^N \mathbf{r}_i \cdot \mathbf{f}_i = \mathcal{W} \quad (2.53)$$

where now we restrict attention to intermolecular forces, then

$$PV = Nk_B T + \langle \mathcal{W} \rangle. \quad (2.54)$$

This suggests that we define an instantaneous 'pressure' function [Cheung 1977]

$$\mathcal{P} = \rho k_B \mathcal{T} + \mathcal{W}/V = \mathcal{P}^{\text{id}} + \mathcal{P}^{\text{ex}} \quad (2.55)$$

whose average is simply  $P$ . Again, this definition is not unique; apart from the different ways of defining  $\mathcal{W}$ , which we shall see below, it may be most convenient (say in a constant temperature ensemble) to use

$$\mathcal{P}' = \rho k_B T + \mathcal{W}/V = \langle \mathcal{P}^{\text{id}} \rangle + \mathcal{P}^{\text{ex}} \quad (2.56)$$

instead. Both  $\mathcal{P}$  and  $\mathcal{P}'$  give  $P$  when averaged, but their fluctuations in any ensemble will, in general, be different. It should be noted that the above derivation is not really valid for the infinite periodic systems used in computer simulations: there are no container walls and no external forces. Nonetheless, the result is the same [Erpenbeck and Wood 1977].

For pairwise interactions,  $\mathcal{W}$  is more conveniently expressed in a form which is explicitly independent of the origin of coordinates. This is done by writing  $\mathbf{f}_i$  as the sum of forces  $\mathbf{f}_{ij}$  on atom  $i$  due to atom  $j$

$$\sum_i \mathbf{r}_i \cdot \mathbf{f}_i = \sum_i \sum_{j \neq i} \mathbf{r}_i \cdot \mathbf{f}_{ij} = \frac{1}{2} \sum_i \sum_{j \neq i} \mathbf{r}_i \cdot \mathbf{f}_{ij} + \mathbf{r}_j \cdot \mathbf{f}_{ji} \quad (2.57)$$

The second equality follows because the indices  $i$  and  $j$  are equivalent. Newton's third law  $\mathbf{f}_{ji} = -\mathbf{f}_{ij}$  is then used to switch the force indices

$$\sum_i \mathbf{r}_i \cdot \mathbf{f}_i = \frac{1}{2} \sum_i \sum_{j \neq i} \mathbf{r}_{ij} \cdot \mathbf{f}_{ij} = \sum_i \sum_{j > i} \mathbf{r}_{ij} \cdot \mathbf{f}_{ij} \quad (2.58)$$

where  $\mathbf{r}_{ij} = \mathbf{r}_i - \mathbf{r}_j$ , and the final form of the summation is usually more convenient. It is essential to use the  $\mathbf{r}_{ij} \cdot \mathbf{f}_{ij}$  form in a simulation that employs periodic boundary conditions. So we have at last

$$\begin{aligned} \mathcal{W} &= \frac{1}{2} \sum_i \sum_{j > i} \mathbf{r}_{ij} \cdot \mathbf{f}_{ij} = -\frac{1}{2} \sum_i \sum_{j > i} \mathbf{r}_{ij} \cdot \nabla_{\mathbf{r}_{ij}} v(\mathbf{r}_{ij}) \\ &= -\frac{1}{2} \sum_i \sum_{j > i} w(r_{ij}) \end{aligned} \quad (2.59)$$

where the intermolecular pair virial function  $w(r)$  is

$$w(r) = r \frac{dv(r)}{dr} \quad (2.60)$$

Like  $\mathcal{V}$ ,  $\mathcal{W}$  is limited by the range of the interactions, and hence  $\langle \mathcal{W} \rangle$  should be a well-behaved, ensemble-independent function in most cases.

For molecular fluids we may write

$$\begin{aligned} \mathcal{W} &= \frac{1}{2} \sum_i \sum_{j > i} \mathbf{r}_{ij} \cdot \mathbf{f}_{ij} = -\frac{1}{2} \sum_i \sum_{j > i} \mathbf{r}_{ij} \cdot (\nabla_{\mathbf{r}_{ij}} \mathcal{V})_{\Omega_i \Omega_j} \\ &= -\frac{1}{2} \sum_i \sum_{j > i} w(r_{ij}) \end{aligned} \quad (2.61)$$

where  $\mathbf{r}_{ij}$  is the vector between the molecular centres. Here we have made it

clear that the pair virial is defined as a position derivative at constant orientation of the molecules

$$w(r_{ij}) = r_{ij} \left( \frac{\partial v(r_{ij}, \Omega_i, \Omega_j)}{\partial r_{ij}} \right)_{\Omega_i \Omega_j} \quad (2.62)$$

The pressure function  $\mathcal{P}$  is defined through eqn (2.55) as before. For interaction site models, we may treat the system as a set of atoms, and use eqns (2.59), (2.60), with the summations taken over distinct pairs of sites  $ia$  and  $jb$  (compare eqn (1.12)). When doing this, however, it is important to include all intramolecular contributions (forces along the bonds for example) in the sum. Alternatively, the molecular definition, eqns (2.61), (2.62) is still valid. In this case, for computational purposes, eqn (2.62) may be rewritten in the form

$$w(r_{ij}) = \sum_a \sum_b \frac{w_{ab}(r_{ab})}{r_{ab}^2} (\mathbf{r}_{ab} \cdot \mathbf{r}_{ij}) \quad (2.63)$$

where  $\mathbf{r}_{ab} = \mathbf{r}_{ia} - \mathbf{r}_{jb}$  is the vector between the sites and  $w_{ab}(r_{ab})$  is the site-site pair virial function. This is equivalent to expressing  $\mathbf{f}_{ij}$  in eqn (2.61) as the sum of all the site-site forces acting between the molecules. Whether the atomic or molecular definition of the virial is adopted, the ensemble average  $\langle \mathcal{W} \rangle$  and hence  $\langle \mathcal{P} \rangle = P$  should be unaffected.

Quantities such as  $\langle N \rangle$  and  $\langle V \rangle$  are easily evaluated in the simulation of ensembles in which these quantities vary, and derived functions such as the enthalpy are easily calculated from the above.

Now we turn to the question of evaluating entropy-related ('statistical') quantities such as the Gibbs and Helmholtz functions, the chemical potential  $\mu$ , and the entropy itself. A direct approach is to conduct a simulation of the grand canonical ensemble, in which  $\mu$ , or a related quantity, is specified. It must be said at the outset that there are some technical difficulties associated with grand canonical ensemble simulations, and we return to this in Chapter 4. There are also difficulties in obtaining these functions in the other common ensembles, since they are related directly to the partition function  $Q$ , not to its derivatives. To calculate  $Q$  would mean summing over all the states of the system. It might seem that we could use the formula

$$\exp(A^{\text{ex}}/k_B T) = Q_{NVT}^{\text{ex}-1} = \langle \exp(\mathcal{V}/k_B T) \rangle_{NVT} \quad (2.64)$$

to estimate the excess statistical properties, but, in practice, the distribution  $\rho_{NVT}$  will be very sharply peaked around the largest values of  $\exp(-\mathcal{V}/k_B T)$ , i.e. where  $\exp(\mathcal{V}/k_B T)$  is comparatively small. Consequently, any simulation technique that samples according to the equilibrium distribution will be bound to give a poor estimate of  $A$  by this route. Special sampling techniques have been developed to evaluate averages of this type [Valleau and Torrie 1977] and we return to this in Chapter 7. It is comparatively easy to obtain free energy differences for a given system at two different temperatures by integrating the internal energy along a line of constant density:

$$\left(\frac{A}{Nk_B T}\right)_2 - \left(\frac{A}{Nk_B T}\right)_1 = \int_{\beta_1}^{\beta_2} \left(\frac{E}{Nk_B T}\right) \frac{d\beta}{\beta} = - \int_{T_1}^{T_2} \left(\frac{E}{Nk_B T}\right) \frac{dT}{T}. \quad (2.65)$$

Alternatively, integration of the pressure along an isotherm may be used:

$$\left(\frac{A}{Nk_B T}\right)_2 - \left(\frac{A}{Nk_B T}\right)_1 = \int_{\rho_1}^{\rho_2} \left(\frac{PV}{Nk_B T}\right) \frac{d\rho}{\rho} = - \int_{V_1}^{V_2} \left(\frac{PV}{Nk_B T}\right) \frac{dV}{V}. \quad (2.66)$$

To use these expressions, it is necessary to calculate ensemble averages at state points along a reversible thermodynamic path. To calculate absolute free energies and entropies, it is necessary to extend the thermodynamic integration far enough to reach a state point whose properties can be calculated essentially exactly. In general, these calculations may be expensive, since accurate thermodynamic information is required for many closely spaced state points.

One fairly direct, and widely applicable, method for calculating  $\mu$  is based on the thermodynamic identities

$$\exp(-\mu/k_B T) = Q_{N+1}/Q_N = Q_N/Q_{N-1} \quad (2.67)$$

valid at large  $N$  for both the constant- $NVT$  and constant- $NPT$  ensembles. From these equations, we can obtain expressions for the chemical potential in terms of a kind of ensemble average [Widom 1963, 1982]. If we define the excess chemical potential  $\mu^{ex} = \mu - \mu^{id}$  then we can write

$$\mu^{ex} = -k_B T \ln \langle \exp(-\mathcal{V}_{test}/k_B T) \rangle \quad (2.68)$$

where  $\mathcal{V}_{test}$  is the potential energy change which would result from the addition of a particle (at random) to the system. This is the 'test particle insertion' method of estimating  $\mu$ . Eqn 2.68 also applies in the constant- $\mu VT$  ensemble [Henderson 1983]. A slightly different formula applies for constant  $NVE$  because of the kinetic temperature fluctuations [Frenkel 1986]

$$\mu^{ex} = -k_B \langle \mathcal{T} \rangle \ln[\langle \mathcal{T} \rangle^{-3/2} \langle \mathcal{T}^{3/2} \exp(-\mathcal{V}_{test}/k_B \mathcal{T}) \rangle] \quad (2.69a)$$

where  $\mathcal{T}$  is the instantaneous kinetic temperature. Similarly for the constant- $NPT$  ensemble, it is necessary to include the fluctuations in the volume  $V$  [Shing and Chung 1987]

$$\mu^{ex} = -k_B T \ln[\langle V \rangle^{-1} \langle V \exp(-\mathcal{V}_{test}/k_B T) \rangle]. \quad (2.69b)$$

In all these cases the 'test particle', the  $(N+1)$ th, is not actually inserted: it is a 'ghost', i.e. the  $N$  real particles are not affected by its presence. There is an alternative formula which applies to the removal of a test particle (selected at random) from the system [Powles, Evans, and Quirke 1982]. This 'test particle' is not actually removed: it is a real particle and continues to interact normally with its neighbours. In practice, this technique does not give an accurate estimate of  $\mu^{ex}$ , and for hard spheres (for example) it is completely unworkable [Rowlinson and Widom 1982]. We defer a detailed discussion of the applicability of these methods and more advanced techniques until Chapter 7.