TU/e EINDHOVEN UNIVERSITY OF TECHNOLOGY

Monte Carlo simulations of fluid systems with waterlike molecules

Citation for published version (APA):
Bol, W. (1979). *Monte Carlo simulations of fluid systems with waterlike molecules*. [Phd Thesis 1 (Research TU/e / Graduation TU/e), Chemical Engineering and Chemistry]. Technische Hogeschool Eindhoven. https://doi.org/10.6100/IR101722

DOI:
10.6100/IR101722

Document status and date:
Published: 01/01/1979

Document Version:
Publisher's PDF, also known as Version of Record (includes final page, issue and volume numbers)

Please check the document version of this publication:

• A submitted manuscript is the version of the article upon submission and before peer-review. There can be important differences between the submitted version and the official published version of record. People interested in the research are advised to contact the author for the final version of the publication, or visit the DOI to the publisher's website.
• The final author version and the galley proof are versions of the publication after peer review.
• The final published version features the final layout of the paper including the volume, issue and page numbers.

Link to publication

General rights
Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

• Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
• You may not further distribute the material or use it for any profit-making activity or commercial gain
• You may freely distribute the URL identifying the publication in the public portal.

If the publication is distributed under the terms of Article 25fa of the Dutch Copyright Act, indicated by the "Taverne" license above, please follow below link for the End User Agreement:

www.tue.nl/taverne

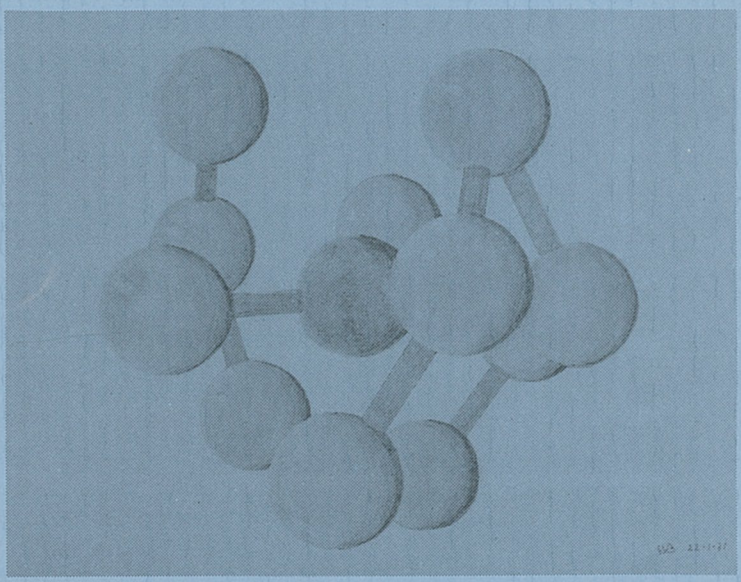
Take down policy
If you believe that this document breaches copyright please contact us at:

openaccess@tue.nl

providing details and we will investigate your claim.

Download date: 04. Oct. 2023

MONTE CARLO SIMULATIONS OF FLUID SYSTEMS WITH WATERLIKE MOLECULES



W. BOL

MONTE CARLO SIMULATIONS OF FLUID SYSTEMS WITH WATERLIKE MOLECULES

PROEFSCHRIFT

TER VERKRIJGING VAN DE GRAAD VAN DOCTOR IN DE TECHNISCHE WETENSCHAPPEN AAN DE TECHNISCHE HOGESCHOOL EINDHOVEN, OP GEZAG VAN DE RECTOR MAGNIFICUS, PROF. IR. J. ERKELENS, VOOR EEN COMMISSIE AANGEWEZEN DOOR HET COLLEGE VAN DEKANEN IN HET OPENBAAR TE VERDEDIGEN OP DINSDAG 18 SEPTEMBER 1979 TE 16.00 UUR

DOOR

WILLEM BOL

GEBOREN TE ROTTERDAM-CHARLOIS

DRUK: WIBRO HELMOND

DIT PROEFSCHRIFT IS GOEDGEKEURD DOOR DE PROMOTOREN

PROF.DR. C.L. VAN PANTHALEON VAN ECK

EN PROF.DR. J.A.A. KETELAAR

Aan Adrie
Dik en Lisette
Arie Klaas en Nellie
Leonard
Timon
Vincent

CONTENTS

1. INTRODUCTION

The substances that constitute the material world present themselves in many forms. These forms can be classified into the gases, the solids and the liquids.
This classification is closely connected with intuitive experience and moreover it appears to make scientific sense.

It is one of the major objectives of physical chemistry to explain the observable properties of the materials surrounding us in terms of the molecular picture created by physicists and chemists. For gases, crystalline solids and liquids the explanation appears to go along quite different paths. In the case of gases and crystalline solids the starting point is an extremely simplified model.
In the simplified gas (the ideal gas), the particles perform their thermal motion within some volume, interaction between them is supposed to be absent. There is complete disorder both in the location and in the motion of the particles. There is only interaction between the particles and the walls of the vessel, resulting in elastic collisions. The ideal gas can be treated theoretically quite satisfactorily with a randomized geometry. The properties of many (gaseous) physical systems approach closely the properties calculated with this theory. Deviations of the properties of physical systems from the theoretical values can be treated as perturbations due to interaction between the particles, for instance by taking into account the volume of the particles.

In the case of crystalline solids the reference state is the 'ideal crystal', an arrangement of a great number of elementary cells with molecules (atoms, ions) arranged in a simple lattice. In this case too, many physical systems correspond closely to the simplified model, notably systems with high density and low temperature. Deviations from that picture are treated in terms of crystal 'imperfections', lattice vibrations, vacancies and interstitial particles.

In the case of liquids the situation is different.
An anthropomorphic description on a molecular scale of a liquid in

the sense of an 'ideal liquid' does not exist. In a negative way it can be stated that both the complete disorder of the perfect gas and the long range order of the crystalline state are absent although there exists certainly some short range order.
The theoretical treatment should therefore account for the short range order but also contain a random aspect.
For the theoretical treatment of liquids on this basis several approaches exist : the cell model theory, the cluster theory and the computer simulation approach.
The cell model theory in its simplest form describes in fact a crystalline state. The volume is divided into cells, which are arranged in a simple lattice. Each cell contains one molecule. The starting point for the calculation of the free energy and the other thermodynamic properties is the concept of the 'free volume', that is the fraction of the cell volume in which the molecule apparently can move unhampered by its neighbours.
Especially this concept has been exposed to severe criticism in the previous years. For instance Hildebrand et al. [1.1] state :
"Theories based upon this concept of free volume cannot be made consistent with all thermodynamic properties of a liquid, although some workers continue to try. It is not a physical parameter and cannot be a thermodynamic one".
Hildebrands statement may seem slightly exaggerated, since it is likely that at normal density there will be some space to move for the particles of some configuration. In the course of our computer simulations, however, we could prove (appendix 5) that the free volume as defined according to the cell model cannot be related to the partition function in the way the theory suggests. The deviation is so important that it is impossible to consider the free volume theory as an approximate one that could be adapted to physical reality by certain adjustments such as an alternative definition of the free volume.
In accordance with the opinion of Hildebrand and other workers, we conclude that the free volume theory of liquids should be abandoned altogether.

Stillinger, [1.14] page 45 - 70, reviews more sophisticated lattice theories, having not the disadvantage of the free volume theory.
For waterlike potentials, however, Stillinger expects the methods to be not very efficient.

The 'cluster' approach starts from the ideal gas theory. Here clusters of two, three, four, etc. particles are taken into account.
The effect is considered to be a perturbation of the simpler situation. This leads to a polynomial, the virial expansion of Kamerlingh Onnes [1.2]

$$\frac{p}{kT} = \rho + B_2\rho^2 + B_3\rho^3 + \ldots\ldots \qquad (1.1)$$

The B's are functions of temperature only. In simple cases the lower B's can be calculated from the intermolecular potentials, B_2 in a two particle system, B_3 in a three particle system and so on.
This theory is conceptionally sound, but applies only to relatively low densities. At higher densities the concept of isolated clusters fails. In more sophisticated variants of the cluster theory duplets, triplets, quadruplets and so on are treated without the assumption of isolation from other molecules. As a rule these variants are brought into the form of integral equations. These integral equations can be solved numerically when an intermolecular potential is given. The result is a distribution function.
Important members of this group are the Yvon-Born-Green theory [1.3], the Percus-Yevick theory [1.4] and the hypernetted chain theory [1.5]. The distribution functions obtained can be used to evaluate macroscopic properties of the liquid with the Kirkwood-Buff theory [1.6]. The Kirkwood-Buff equations are exact. The distribution functions are approximated, but can be verified with X-ray and neutron diffraction. These sophisticated cluster theories play an important role in the physical chemistry of liquids. Rather complicated systems can be dealt with. A drawback of these theories is that for waterlike molecules it is not possible to solve exactly the three dimensional integral equations obtained. The results that have been reported rest on simplifying statistical mechanical approximations and thus inevitably convey incertainty. (Ben-Naim [1.15] and Stillinger [1.14] p. 71).
In that case the computer simulation method is certainly more promising. The computer simulation method is the third of the approaches mentioned above. Here a greater number of molecules, e.g. one hundred is taken into consideration and their mutual interactions are accounted for . Periodic boundary conditions are introduced in order to eliminate sur-

face effects. An advantage of the method is that to some extent an anthropomorphic character is present. A disadvantage is that long calculation times are needed.

The method has reached a certain maturity already. Barker et al. [1.7] calculated the properties of argon from the best intermolecular potential they could obtain . The result was that excellent agreement with experiment was obtained for the thermodynamic properties of solid, liquid and gaseous argon, for the lattice spacing of crystalline argon, the cohesive energy at 0 K, thermal expansion, phonon dispersion in crystals, viscosity and other transport properties of the gas.

At first agreement with experiment was poor in the case of the internal energy and of the elastic constant of crystalline argon. However, here the experimental values were shown to be wrong after all.

Good results have also been obtained in molten salts with simple ions such as molten KCl (Michielsen [1.16]).

At present for these liquids with relatively simple interaction between the particles it is a matter of consideration whether a property needed for the solving of some problem will be measured or calculated.

For more complicated substances like water the job has not been finished as yet [1.8]. It proves to be very difficult to find a reliable algorithm that accounts for all details in the interaction between the molecules. Nevertheless there is no reason why these difficulties will not be overcome some day and water and any other liquid will be calculable just like argon or KCl.

However, there is a quantitative restriction. The results of the calculations on one liquid do not apply to any other. Since there is a tremendous quantity of different liquids that is of interest and since each calculation demands much computer time it is not likely that all problems of the physical chemistry of liquids will be solved with computer simulations in the near future.

Therefore, it will be wise to combine the accurate calculation of some key problems with appropriate calculations of a more or less approximate character.

To this purpose Barker and Henderson [1.9] have developed a perturbation theory. In this approach some simple model potential is used as a reference and refinements of the potential are considered as a perturbation. Thermodynamic properties are evaluated as a function of a perturbation parameter.

From a technical point of view the above mentioned approach is one of computer simulation of the behaviour of a system of particles.
In current practice there are two methods.
One is the method of molecular dynamics [1.13].
In this case a number of molecules is located in three dimensional space and an initial velocity is given to each of them. Subsequently the motion of the molecules due to inertia and interaction forces is calculated with the algorithms of classical mechanics. The result is a system of which energy is constant. When proper conditions are inserted, volume or pressure are kept constant too, so that the system can be characterised to be of the NVE respectively NPE type. In molecular dynamics the detailed positions of the molecules as a function of time are obtained, which enables the calculation of energy and pressure as well as many dynamic properties.
A method for the evaluation of the partition function (and the free energy and the entropy) on the basis of molecular dynamics has not been developed.
The second is the Monte Carlo calculation technique of Metropolis, Rosenbluth, Rosenbluth, Teller and Teller [1.10]
In this case a random set of configurations is generated, configurations that pertain to a preset temperature. A great number of trials for small random steps is made, which steps are accepted or not, depending on the energy change the step would involve. In that way a set of configurations for a NVT- (or NPT-) system is obtained. In appendix 2 we describe briefly this method of Metropolis et al. Averaging over the configurations obtained, leads to the equilibrium properties of the system. Free energy can also be evaluated, albeit not after a simple averaging procedure.
Monte Carlo yields no dynamic properties of the system since the time scale is absent.

In the present work we restrict ourselves to Monte Carlo simulations with simple potentials, the hard sphere potential in chapter 2 and a tetrahedral potential with a hard core in chapter 3.
For the evaluation of the free energy we used the method of 'multistage sampling' of Valleau et al. [1.11]. This method will be discussed in chapter 2. In that chapter we will compare the free energy of a hard-spheres system from the multistage sampling method,

applied to a system with overlapping spheres, with the 'exact' value from the formula of Carnahan and Starling [1.12]. The agreement is excellent. This, together with the fact that the method is conceptually sound makes that we consider it to be reliable.

In chapter 3 we will calculate the equilibrium properties of a system with molecules having a hard core and a short-range interaction with tetrahedral symmetry. These molecules will be called 'tetrahedral molecules' in the next chapters.

Next we will calculate the consequence for the system when the molecular properties are changed in the sense that the short-range interaction is made polar. The resulting interaction can be considered to approach the interaction between water molecules to some extent. Thus we can speak of 'waterlike' molecules.

Finally in chapter 4 we will compare the properties of a system with 'waterlike' molecules with the properties of physical water.

2. CALCULATION OF EQUILIBRIUM PROPERTIES

2.1. The Monte Carlo Method

In mathematics different calculation schemes are indicated by the name 'Monte Carlo'. In the simulation of liquid systems one special scheme is generally adopted and indicated as 'the' Monte Carlo method. This scheme was developed by Metropolis et al. [1.10] and will be discussed in appendix 2. The simulations deal with a number of particles located in space. If the location (and eventually the orientation) of each of the N particles is fixed within a volume V we will speak of a configuration of N particles in the volume V. The assembly of all different configurations that are possible is called the configuration space (which is dependent of the magnitude of N and V).
With the calculation scheme of Metropolis et al. the probability of any configuration to be found during the simulation is proportional to the relevant Boltzmann factor. This is very convenient for the statistical treatment of the data obtained.
In this section we will deal with some more or less technical details, that is to say the periodic boundary conditions, the choice of N (the number of molecules), the intermolecular potential energy and the initial configuration. These details must be established before the computer simulation can start.

Periodic boundary conditions

Computer simulations demand a great deal of machine time. Therefore it is necessary to apply the calculations to a very restricted number of molecules. This restriction can lead to difficulties since unavoidably an important fraction of the molecules would be located at the surface of any tiny set of molecules (in a cluster of 100 molecules about half this number is situated at the surface). Moreover during the simulation 'evaporation' would occur.
For these reasons periodic boundary conditions are introduced. The molecules are placed in a cubic box and this arrangement is repeated by translation (indefinitely) in all directions. In that way the system is surrounded by an infinite number of identical

boxes of which 26 are in contact with the original one. Consequently all N molecules in the box are surrounded by other molecules in all directions.
During the simulation a molecule can eventually leave the box. The periodic boundary conditions ensure, that at the same time an identical molecule enters the box at the opposite side. In this way the number of molecules in the box is kept constant.

The number of molecules

The number of molecules N in the box has to be chosen with some care. If it is chosen too high the calculation time, being about proportional to N^2, would become prohibitively long. When on the other hand N is too low, the results would become poorly comparable with macroscopic quantities.
In section 2.3 we will see that for a hard-spheres fluid the value of the free energy is reasonably good in accordance with what it should be, even for the value of N as low as 44.

Another thing that has to be envisaged in the choice of the number of molecules in the system is the possibility of crystallisation due to the periodicity imposed by the boundary conditions. Especially when the number of molecules is an exact multiple of a third power, subdivision of the box can occur and the probability of a regular configuration will be enhanced.
In the calculations of chapter 3 we have chosen the number N = 91, being far enough from 81 (27 cells of 3), 88 and 96 (8 cells of 11 or 12).
Moreover 91 is the product of the numbers 7 and 13, the symbols of good and bad luck, both indispensable in Monte Carlo.

The intermolecular potential

The potential energy function of the system which is to be simulated depends on the relative positions of the molecules.
If the system consists of two molecules, the potential energy depends only on the positions of those two :

$$U = E_2 (q_1, q_2) \qquad (2.1)$$

U is the energy of the system and E_2 is the contribution of one pair of molecules. q_i denotes the position of molecule i in geometric space.
For rigid rotors for instance q would consist of six coordinates, three for the orientation and three for the location.
In a system with three molecules the potential energy can not always be calculated from the pair potential of the three pairs, but an extra term appears, the three-body contribution E_3 :

$$U = \sum_{k<l} \sum E_2 (q_k, q_l) + E_3(q_1, q_2, q_3) \qquad (2.2)$$

For an N-body system we could write :

$$U = \sum E_2 + \sum E_3 + \sum E_4 + ... + E_N \qquad (2.3)$$

Often this series can be replaced with only restricted loss of precision by a slightly modified two-body interaction, the effective pair-potential.
Only when close accordance between simulation and physical data is desired a three-body interaction must be added. This has been performed by Barker et al. [1.7] for the calculations regarding noble gases. They report that in this case the three-body contributions appeared to be relatively small.
Almost all other computer simulations reported in literature have been performed with (effective) pair potentials. The present work too deals merely with two-body interactions, since the aim of the work is restricted to the investigation of the properties of systems with certain simple potentials.

The initial configuration

When N molecules are situated randomly in a certain volume, the potential energy will often be extremely high, due to the overlap of molecules. This will be especially the case when the density is not very low. Therefore an initial set of steps is necessary in order to obtain a configuration with lower energy. This is a very fast process as can be seen in fig. 2.1 where the results of a Monte Carlo simulation that we have performed with Lennard Jones particles have been plotted.

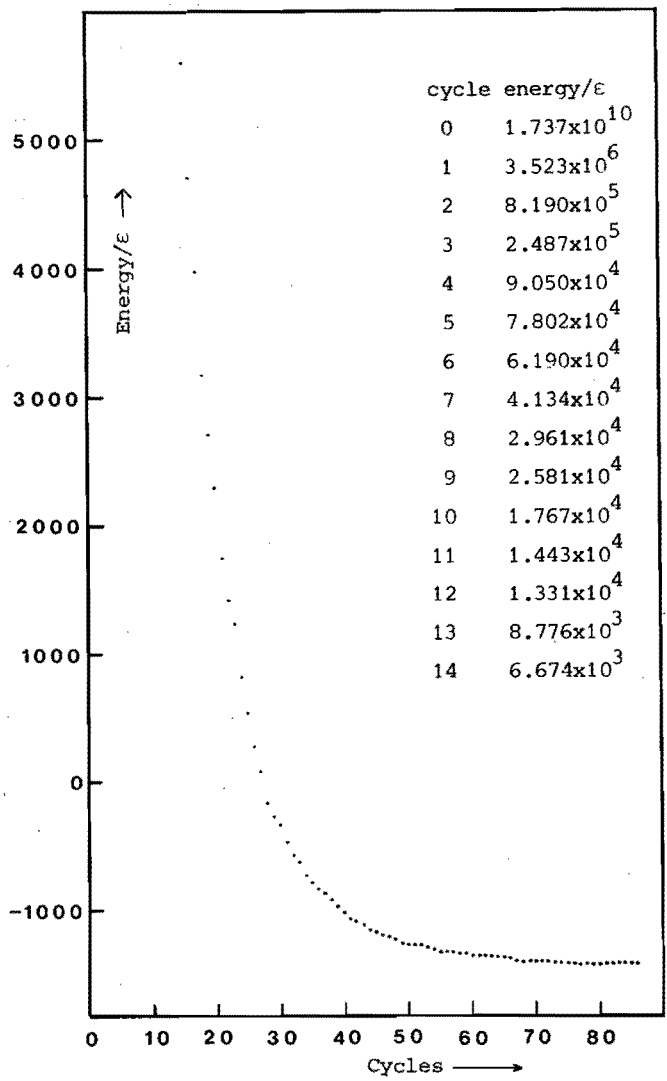


Fig. 2.1 Energy of a system with 271 Lennard-Jones particles as a function of the number of Monte Carlo cycles passed. In the initial situation the particles were randomly distributed. A Monte Carlo cycle consists of one proposed step for each of the 271 particles.
The very first cycles cannot be plotted, they are listed. Energy in units of ε, that is the energy parameter in the Lennard-Jones potential (2.4). The density ($=\sigma^3 N/V$) = 0.75 and the temperature = $\varepsilon/2k$, where k is Boltzmann's constant.

Even as many as 271 particles prove to be properly arranged in about 50 cycles, that is 50 x 271 steps.
The Lennard Jones potential is :

$$E_2 = 4\varepsilon\left[\left(\frac{\sigma}{r_{kl}}\right)^{12} - \left(\frac{\sigma}{r_{kl}}\right)^{6}\right] \qquad (2.4)$$

where r_{kl} is the internuclear distance of two particles k and l, σ is the diameter of the particles and ε is the (absolute) value of energy in the minimum of the potential curve.

The example of fig. 2.1 is a case of a pair potential that is finite for all values of r_{kl}. In the case of hard spheres where the potential energy is infinite if r_{kl} is below some limiting value σ, initialisation demands for an ad hoc potential.
We have used for this purpose a potential that has been deprived of all unnecessary complications:

$$E_2 = \left(\frac{\sigma}{r_{kl}}\right)^2 \text{ if } r_{kl} < \sigma \text{ and } E_2 = 0 \text{ if } r_{kl} \geq \sigma$$

As soon as energy had fallen to zero the initial procedure was discontinued.
Many authors report to have avoided this initialisation problem by starting with molecules, situated on a regular lattice. In that case there is some risk that the system will remain in a region of configuration space close to a crystalline configuration (which is not necessarily the original one). Since initialisation after random distribution appears to be extremely simple, it is not necessary to run that risk.

The Computer

All calculations have been performed on the Burroughs 7700 computer of this University.
This machine is a relatively fast one. The central processor time needed for one Monte Carlo step with 91 tetrahedral particles, including random displacement of one particle, calculation of the energy, decision whether the displacement is accepted and storage of the relevant data for statistics proved to be 0.006 seconds.

2.2. Evaluation of the Equilibrium Properties

When a run of Monte Carlo cycles is executed, the average value of a number of properties is calculated. Important properties are the potential energy, the distance distribution function, the pressure and the free energy.
The total potential energy is calculated for each configuration and the mean value is obtained as an average over all configurations. Likewise the distance distribution function is the average of the frequency with which the different distances occur in the configurations.
The pressure can be calculated from the virial function :

$$\text{Vir} = -\tfrac{1}{2} \left\langle \sum r_{kl} \cdot F_{kl} \right\rangle \tag{2.5}$$

which is easily evaluated since in all configurations all r_{kl} and F_{kl} are known. (The intermolecular force $F = -\nabla \cdot E_2$ for pairwise interaction).
The interaction of the molecules with the wall of the vessel is not included in the virial because of the periodic boundary conditions. In this case the pressure is related to the virial with the equation :

$$\frac{PV}{NkT} = 1 - \frac{2}{3}\frac{\text{Vir}}{NkT} \tag{2.6}$$

(See Hirschfelder, Curtiss and Bird [2.4] page 135).
For molecules with a hard core $\nabla \cdot E_2$ does not exist everywhere. In that case the pressure can be evaluated from the distribution function.
The algorithm that is available for rigid spheres (see Hansen and Mc Donald [2.1] chapter 3.2) will be extended in appendix 4 to the tetrahedral molecules we will introduce in chapter 3.
The evaluation of the free energy is more complicated. Hansen and Verlet [2.2] calculated for a system with Lennard Jones particles the value of the pressure as a function of volume at constant temperature and evaluated from these data a value for

the excess free energy with the equation :

$$\frac{A'}{NkT} = \int_{\rho=0}^{\rho} \left(\frac{P}{\rho kT} - 1\right)\frac{d\rho}{\rho}$$

$\rho = N/V$ and A' is the excess free energy, that is the difference between the free energy of the system and of the ideal gas with the same temperature and the same ρ.
Valleau and Card have developed a method that requires slightly less computer time, the method of multistage sampling [1.11].
The starting point is again an ideal gas and the effect of non-zero pair interaction is introduced as a perturbation.
In the present work we have adopted the latter method.
The method can be described as follows :
Starting point is the partition function of a NVT system in classical mechanics. Since it is possible to discriminate between potential and kinetic energy in classical systems the partition function is given bij (A1.3) which can be written slightly simplified as :

$$Q = Q_i \cdot Z'_N$$

Q_i is the partition function of the ideal gas,
Z'_N is the excess configuration integral.
For spherical particles is :

$$Z'_N = \frac{1}{V^N} \cdot \int \exp(-U/kT)\ dx_1 \ldots\ldots dx_N \qquad (2.7a)$$

x_i represents the coordinates of the location of particle i.
For non-spherical particles is :

$$Z'_N = \frac{1}{V^N \cdot (8\pi^2)^N} \cdot \int \exp(-U/kT)\ dq_1 \ldots\ldots dq_N \qquad (2.7b)$$

q_i represents the coordinates of location and orientation of the particle i.

When the potential energy is not the same for all configurations we can define a fraction of the configuration space C(U) pertaining to some energy U.

There are two different cases.
In the first place, if U is a continuous function of the coordinates, C(U) can be defined (for spherical particles) with the formula :

$$C(U)\ dU \equiv \frac{1}{V^N} \cdot \int \xi(U, dU)\ dx_1 \ldots\ldots dx_N \qquad (2.8)$$

$\xi(U, dU) = 1$ for those configurations where the potential energy is between U and U+dU. $\xi(U, dU) = 0$ in all other cases.
Secondly for hard spheres and for the other potentials we will deal with in the present work, the potential energy is a discontinuous function of the coordinates and the system can adopt only a restricted number of energy levels.
Each energy level can be labeled with an integer number j.
The properties of the system that are related to the energy level can be denoted with an indexed symbol. For instance the fraction of the configuration space with the potential energy U_j can be denoted with the symbol C_j.
For spherical particles is :

$$C_j \equiv \frac{1}{V^N} \cdot \int \xi(U, j) dx_1 \ldots\ldots dx_N \qquad (2.9a)$$

For non-spherical particles is :

$$C_j \equiv \frac{1}{V^N \cdot (8\pi^2)^N} \cdot \int \xi(U, j) dq_1 \cdot \ldots\ldots dq_N \qquad (2.9b)$$

In both cases is $\xi(U, j) = 1$ if the potential energy equals U_j and $\xi(U, j) = 0$ if not.
Summation over all possible levels gives :

$$\sum C_j = 1 \qquad (2.10)$$

The excess configuration integral becomes now :

$$Z'_N = \sum C_j \cdot \exp(-U_j/kT) \qquad (2.11)$$

If a Boltzmann distribution exists the probability of the system to be at level j is proportional to $C_j \cdot \exp(-U_j/kT)$.

We will indicate this probability with the symbol P_j.
The sum of all probabilities is unity :

$$\sum P_j = 1 \qquad (2.12)$$

So we can state :

$$P_j \equiv C_j \cdot \exp(-U_j/kT)/Z'_N \qquad (2.13)$$

The multistage sampling method implies a perturbation analysis along the following line:

Perturbation parameter is λ :

$$\lambda \equiv -\varepsilon/kT \qquad (2.14)$$

ε being some reference value for the energy, for instance the energy parameter appearing in the relevant formula for the pair potential.

Furthermore we introduce the factor

$$\phi_j \equiv U_j/\varepsilon \qquad (2.15)$$

With Monte Carlo simulation the apparent probability PA_j of a certain number of energy levels is obtained. This value is an estimate for the real probability (see appendix 2) :

$$PA_j \approx P_j \qquad (2.16)$$

and $$PA_j \approx P_j = C_j \exp(\lambda \cdot \phi_j) / Z'_N \qquad (2.17)$$

Which follows from (2.13), (2.14), (2.15) and (2.16).

If $\lambda = 0$ then can be concluded from (2.10), (2.12) and (2.13) that $Z'_N = 1.0$ and $C_j = P_j$ for any j and by consequence that $C_j \approx PA_j$.
So all values C_j could be estimated from a sufficiently long Monte Carlo run with $\lambda = 0$.
Unfortunately the length of a run is limited.
When N is about 100, the computing time is about 0.006 sec per step. So runs of 10^7 steps would require about 17 hours. If this

is taken as an upper limit, values of C_j below 10^{-7} are not accessible, although these values certainly are of great interest. Therefore a further Monte Carlo run is executed with λ slightly above the previous value. Again a set of values PA_j is obtained and if λ is chosen correctly, then a part of the set overlaps with the previous one.

$\mathcal{E}'_N$ can now be evaluated with (2.17) since $\mathcal{E}'_N$ is not a function of j and in the overlapping domain C_j is known from the previous run. $\mathcal{E}'_N$ being evaluated for the present value of λ, the range of C_j can be extended outside the overlapping domain, λ is increased again and so on until sufficient values of C_j are obtained.

Then all thermodynamic properties, free energy etc. can be calculated, including pressure (if a range of volumes is evaluated in order to calculate $\frac{\partial A}{\partial V}$).

Pressure however has also been obtained from each Monte Carlo run with the formula (2.6) or the formulae from appendix 4.

The results should agree.

2.3. Application to Systems with Overlapping Spheres and with Hard Spheres

The excess free energy A' (the difference in free energy with the ideal gas) of a hard-spheres fluid has been studied extensively elsewhere. See [1.12] and the references therein.

A relevant result of these studies is the equation of Carnahan and Starling :

$$\frac{A'}{NkT} = \frac{4y - 3y^2}{(1-y)^2} \qquad (2.18)$$

where $y = \frac{\pi}{6} \cdot \frac{N\sigma^3}{V}$

The Carnahan and Starling equation is a very close approximation of the accurate virial expansion evaluated by Ree and Hoover [2.5].

When we calculate the same free energy with multistage sampling the result can be used to evaluate the precision of the method.

In order to be able to apply the method in a fluid with overlapping spherical particles we define a pair potential by :

$$\left.\begin{aligned} E_2 &= 0 \qquad \text{if} \quad r_{kl} > \sigma \\ E_2 &= \varepsilon \qquad \text{if} \quad r_{kl} \leq \sigma \end{aligned}\right\} \tag{2.19}$$

r_{kl} being the distance between the centres of the two spheres.

The energy parameter ε has a positive value.
If the parameter $\lambda = -\varepsilon/kT$ equals zero, the excess configuration integral is unity (ideal gas) and if $\lambda = -\infty$ the system is identical with a system with hard spheres.
The quantity ϕ_j of (2.15) equals j for this potential, j being the number of overlapping pairs in the system.

$$\phi_j = j \tag{2.20}$$

We have performed a set of Monte Carlo simulations according to the algorithm of Metropolis et al. (see appendix 2 and [1.10]) for various numbers of molecules N, and various values of the parameter λ.
The volume V has been chosen in such a way that the number density $D = N\sigma^3/V$ was 0.6 in all cases.
The calculations are very simple when $\lambda = 0$ and N is small.
As an example in fig. 2.2 a histogram is given of the results for 82816 random configurations with $N = 4$ and $\lambda = 0$.
In most cases there appear to be 3 or 4 overlapping pairs.
A small number of configurations -54- appears to be without overlap.
This number is important because these are the only configurations that are allowed in the corresponding hard-spheres case.
The fraction $54/82816 = 0.00065$ is therefore an estimate for the fraction of configuration space C_o, that is accessible for the hard-spheres fluid.
Since the number of 54 is rather small from the point of view of statistics, this estimate is rather inaccurate. A better estimate could be obtained by repeating the calculation a number of times. In this case this would be possible because the calculation of 82816 configurations took only one minute on the local computer.

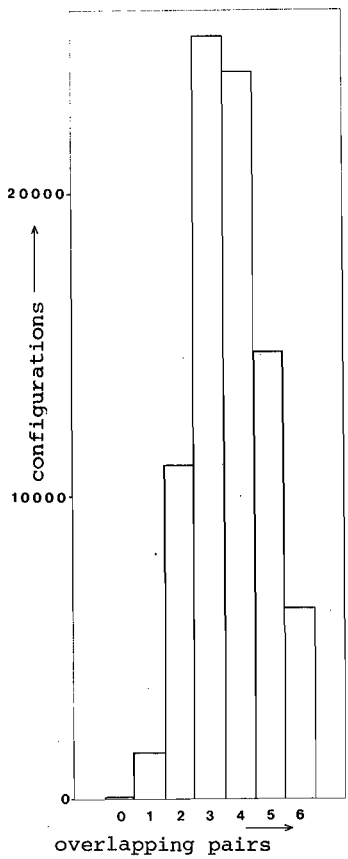


Fig. 2.2 The number of configurations with each of the seven possibilities for the number of overlapping pairs of spheres, as obtained in a Monte Carlo run. N = 4, λ = 0, the total number of configurations is 82816. (See also table 2.1)

More satisfactory, however, is the application of the multistage sampling method as described in section 2.2.
In table 2.1 the results of a second one-minute Monte Carlo run with $\lambda = -3$ are given together with the original run.
From the first run the value of $\ln C_2 = -2.017$ can be considered to be a good approximation from the point of view of statistics. Therefore the value of the logarithm of the excess configuration integral Z'_N can be calculated as the difference between the values of column 3 and column 5. In this case $\ln Z'_N = -2.017-4.234 = -6.260$. By consequence $\ln C_0 = -7.412$ and $C_0 = 0.000604$.

Table 2.1 The results of two Monte Carlo runs with 4 molecules, with $\lambda = 0$ and -3 respectively.

$\phi_{j(=j)}$	Number of configurations with $\lambda = 0$	$\ln PA_j \approx \ln C_j$	Number of configurations with $\lambda = -3$	$\ln PA_j - \lambda.j \approx \ln C_j - \ln Z'_N$
0	54	- 7.3	25390	- 1.152
1	1547	- 3.98	39873	+ 2.299
2	11017	- 2.017	13860	+ 4.243
3	25162	- 1.191	1188	+ 4.79
4	23981	- 1.239	41	+ 4.4
5	14764	- 1.724	0	
6	6291	- 2.578	0	

For a larger value of N, in this case N = 44, the procedure has been displayed graphically in fig. 2.3 and 2.4. In fig. 2.3 the quantity $\ln PA_j - \lambda.j$ is plotted as a function of j.

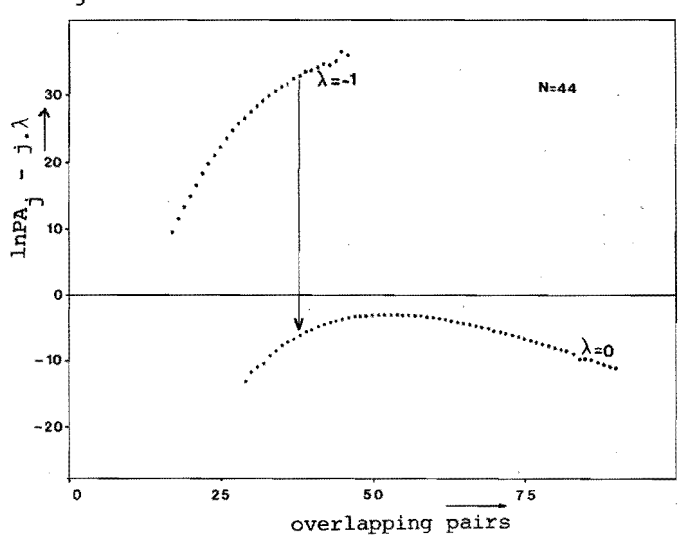


Fig. 2.3 The quantity $\ln PA_j - j.\lambda \approx \ln C_j - \ln Z'_N$ for two Monte Carlo runs. The two curves are parallel. Shifting the upper curve vertically downwards can make the two curves coincide. The magnitude of the shift is an estimation for the excess configuration integral Z'_N for $\lambda = -1$.

For $\lambda = 0$ the dots can be considered to be approximations of $\ln C_j$.
For $\lambda \neq 0$ is :

$$\ln PA_j \approx \ln C_j + \lambda.j - \ln Z'_N \qquad (2.21)$$

So the ordinate of the $\lambda = -1$ curve can be considered to be an estimate of the quantity $\ln C_j - \ln Z'_N$.

The term $\ln Z'_N$ has to be evaluated in such a way that the upper set of dots, shifted downwards in fig. 2.3 over a distance $\ln Z'_N$, will give a best fit with the other set.
The evaluation of $\ln Z'_N$ can be performed algebraically as follows :

Suppose we have two sets of values, resulting from Monte Carlo simulation with $\lambda = \lambda_1$ and $\lambda = \lambda_2$. The value of $\ln Z'_N(\lambda_1)$ is known either for $\lambda_1 = 0$ or as the result of a previous cycle. $\ln Z'_N(\lambda_2)$ is to be evaluated.
Now we define a function γ of the real variable j in a way that for integer values of j applies :

$$\gamma(j) = \ln C_j \qquad (2.22)$$

From the apparent probabilities obtained in the two Monte Carlo simulations we can evaluate a number of approximations for the derivative of γ independent of the unknown $Z'_N(\lambda_2)$.

$$\frac{d\gamma}{dj} \approx \ln PA_{j+1} - \ln PA_j - \lambda \qquad (2.23)$$

with a standard least squares method [2.3] $d\gamma/dj$ is approximated by a polynomial (we used to adopt a polynomial of degree 4) that fits best to the data of both simulations.
The polynomial $x_1 + x_2.j + x_3.j^2 + x_4.j^3 + x_5.j^4$ is integrated analytically. The integration constant x_0 can be evaluated with the equations (2.11), (2.14), (2.15), (2.20) and (2.22).

$$x_0 = \ln Z'_N(\lambda_1) - \ln \sum \exp(x_1.j + ... + \frac{1}{5} x_5.j^5 + j.\lambda_1) \qquad (2.24)$$

(summation over integer values of j).

In the same way we now find the unknown $\ln Z'_N(\lambda_2)$:

$$\ln Z'_N(\lambda_2) = x_0 + \ln \sum \exp(x_1.j + ... + \frac{1}{5} x_5.j^5 + j.\lambda_2) \qquad (2.25)$$

After the operation with the two runs of fig. 2.3 (with $\lambda = 0$ and $\lambda = -1$) we have not reached the desired value of C_0. The lower limit of the domain covered is $j = 17$. In order to extend the domain until $j = 0$ it proved to be necessary to repeat the procedure with $\lambda = -2, -3, -4$ and -6.
The final result, $\ln C_0$ is about -89, can be read from fig. 2.4.

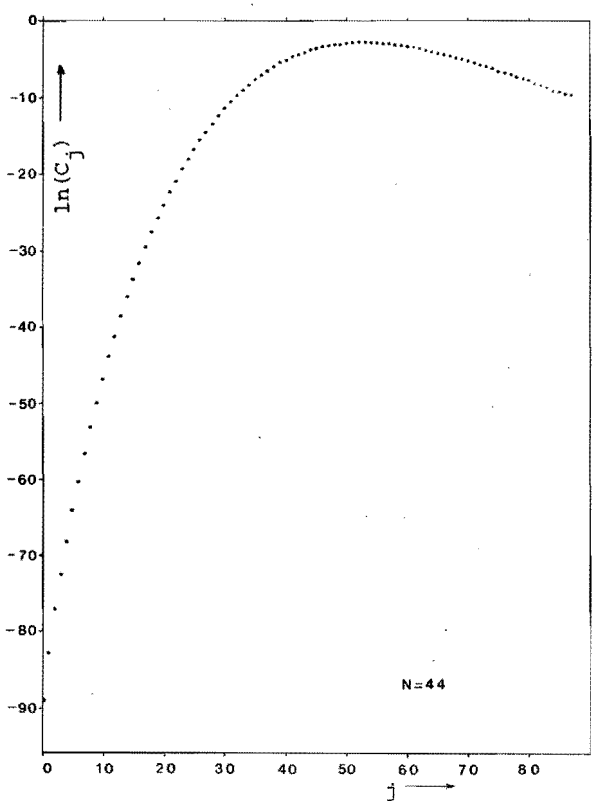


Fig. 2.4 The result of the fitting procedure. The quantity C_j that is the fraction of all possible configurations that gives rise to j overlapping pairs of molecules, is plotted - on a logarithmic scale - as a function of j.
The curve is composed from the results of six Monte Carlo runs, with $\lambda = 0.0, -1.0, -2.0, -3.0, -4.0$ and -6.0 respectively.

Since in a hard-spheres fluid ($\lambda = -\infty$) the potential energy is $+\infty$ if there is any overlap between spheres and zero if not, of the summation of (2.11) will be only one term left: $\ln Z'_N = \ln C_0$.
By consequence the excess free energy is:

$$A' = -kT.\ln C_0 \qquad (2.26)$$

In the above mentioned case the value of A'/NkT becomes 89/44 = 2.02. This value can be compared with the value one can calculate with the formula (2.18) of Carnahan and Starling. That is A'/NkT = 2.042 for the present density. The agreement is very good.

Table 2.2 The result of the calculation of configuration space in a fluid of hard spheres with density = 0.6. The calculation is performed with the Monte Carlo method as discussed in chapter 2. (For N=2 the result is obtained analytically and for N=1 is trivial). The last column is calculated with formula (A 3.9) and (A 3.13).

N	Total number of Monte Carlo steps	$\ln C_0$	$-\frac{\ln C_0}{N}$	$\frac{A'}{NkT}$ calculated
1	-	0.0	0.0	1.149
2	-	-2.5893	1.295	1.495
3	0.18×10^6	-5.616	1.872	1.651
4	0.16×10^6	-7.412	1.853	1.739
5	0.15×10^6	-9.402	1.880	1.798
8	0.12×10^6	-14.587	1.823	1.888
20	1.63×10^6	-39.807	1.990	1.981
27	1.37×10^6	-53.772	1.992	1.997
44	2.95×10^6	-88.992	2.023	2.014
64	1.61×10^6	-131.203	2.050	2.023
91	1.77×10^6	-185.356	2.037	2.029
172	1.67×10^6	-351.919	2.046	2.035
∞	-	-	-	2.04208

In table 2.2 the values of A'/NkT, results of a set of multistage sampling experiments, are given for N = 3, 4, 5, 8 up to 172. (the value for N = 2 is evaluated analytically and for N = 1 it is trivial).
From the table it is apparent that for low values of N the configuration volume as calculated with multistage sampling is slightly larger than would be calculated directly from the Carnahan and Starling equation. This can be ascribed mainly to the 'communal effect' of appendix 3. When the correction of appendix 3 is applied to the Carnahan and Starling formula (last column of table 2.2) the agreement with the multistage sampling results is very close, even for small N.

For larger values of N (64 and up) the correction is not very important. It seems even that the value from multistage sampling is closer to the uncorrected value than to the corrected one. Possibly other systematic errors exist that cancel the communal effect. Therefore since we deal in chapter 3 only with systems with 91 molecules, we decided on using only the uncorrected Carnahan and Starling equation henceforth.

Finally we have to make two remarks, one about the machine time required for the multistage sampling calculations and one about the accuracy.
The machine time required for a single Monte Carlo run depends strongly on the absolute value of the parameter λ.
When $\lambda = 0$ every proposed step is allowed and consequently the configuration space is sampled rather efficiently.
Otherwise the probability of accepting an unfavorable move is proportional to $1/\exp(|\lambda|)$. If for instance $\lambda = -6$ that is about 1/400. Therefore the system will stick a long time to a favorable position. The machine time, which is proportional to the number of attempted moves, is therefore greatly determined by the simulation with the highest absolute value of λ. Once this simulation is finished with M_1 attempted moves, one can start a next simulation with lower absolute value of λ and a lower number of moves M_2. If the above effect of sticking in favorable situations would

be the principal factor that limits the accuracy, a rational choice of M_2 would be :

$$M_2 = M_1/\exp(|\lambda_1 - \lambda_2|)$$

In fact we have chosen a series of λ climbing up with the unity (λ = -6, -5, -4, -3, -2, -1 and 0) with $M_{i+1} \approx \frac{1}{2}M_i$ and a minimum of 10^5 steps.
In this way is $\sum_{i=2}^{i_{max}} M_i \approx M_1$ and the multistage sampling method demands for an extra machine time of the same magnitude as the longest run.

Concerning the accuracy of the method we will make the following comments.
The data of table 2.2 for N > 20 suggest that the accuracy of the calculation of the free energy from Monte Carlo calculations would be better than one percent.
It is not easy to verify this precision in a straightforward way from the statistical errors that one would expect in the Monte Carlo simulations. Intuitively, however, we felt the need for some verification. Therefore we have repeated one of the sets of Monte Carlo runs three times with different initial configurations in order to estimate repeatability.

Table 2.3 Excess free energy for hard spheres: $A'/NkT = -\ln C_o/N$ as calculated from 4 different Monte Carlo simulations. The number of particles, N is 91.

Number of single steps	$\ln C_o$	$-\frac{\ln C_o}{N}$
1.77×10^6	-185.356	2.0369
1.77×10^6	-186.539	2.0499
1.77×10^6	-184.729	2.0300
1.77×10^6	-185.522	2.0387
Mean	-185.54	2.039

The results are given in table 2.3. The standard deviation in $\ln C_0$ calculated from the deviation of the four values obtained, appears to be $\sigma = 0.75$. That is 0.4 % of the mean value of -185.54.
The repeatability is certainly good, notwithstanding the strong correlation existing in Monte Carlo calculation between the properties of subsequent configurations - especially when λ is high- and notwithstanding the fact that no less than eight curves had to be fitted, the end of the last curve giving the desired value.

We can conclude from table 2.3 that the mean value for $-\ln C_0/N$ is: 2.039 ± 0.004 for $N = 91$.
This value is to be compared with 2.042 from the Carnahan and Starling approximation.
From a statistical point of view it can now be stated that the above hypothesis " ..... the value from multistage sampling is closer to the uncorrected value than to the corrected one" is in accordance with the results of table 2.3. (The chance of this statement being false is about 2.5 % against 97.5 % of it being true).

Because of these results and because the method is conceptually sound we consider the method of multistage sampling to be a reliable one for the evaluation of the partition function.

2.4. Extension to Other Potentials

Once the configuration space has been divided into fractions C_j for a specific potential E_2 and all C_j have been evaluated this result can be extended to some new potential E_2'.

In principle the procedure for achievement of this extension is as follows.
A Monte Carlo run is performed with the potential E_2. One of the most abundant energy levels that appears is indicated with the integer j. For each of the configurations obtained the energy is calculated with the new potential E_2' (without consequence for the acceptance of proposed steps). In the assembly of alternative energies obtained we suppose that discrete energy levels appear. If not, the energies obtained are joined together in a number of

quasi levels. Say that we have found that a fraction α_{ji} of the configurations with energy level j has obtained the energy of level i on the basis of potential E_2'. Then we can state that $\alpha_{ji} \cdot C_j$ is the fraction of the configuration space where the energy level is j on the basis of E_2 and the energy level is i on the basis of the potential E_2'.

When a second Monte Carlo run is performed with the potential E_2' with such a temperature that a representative sample of the configurations with the energy level i is obtained, a fraction α_{ij} of these configurations will be found to lead to the energy of level j on the basis of the alternative potential E_2.
Since the overlapping fraction of the configuration space is the same in both cases we can conclude that

$$\alpha_{ji} \cdot C_j = \alpha_{ij} \cdot C_i' \tag{2.27}$$

Where C_i' is the corresponding fraction of the configuration space when it is divided on the basis of the potential E_2'.
From equation (2.27) the value of C_i' can be calculated.

If it proves to be difficult to find an overlapping region, one can perform a Monte Carlo simulation with an intermediate potential $(1-\lambda)\, E_2 + \lambda E_2'$. Subsequently the other fractions of the configuration space (as far as covered by the relevant Monte Carlo runs) can be calculated and if necessary the range can be extended with multistage sampling.

After evaluation of C_i' the configuration integral can be calculated on the basis of equation (2.11).

When E_2 is the hard-spheres potential there is only one energy level. When moreover the potential E_2' comprises a hard core that is greater than or equal to the hard sphere of E_2 then is $\alpha_{ij} = 1$ since all configurations obtained with E_2' must lead to zero energy with E_2. In that case the above mentioned second Monte Carlo run can be omitted.

The evaluation of the fractions of the configuration space for the new potential of chapter 3 will be performed in accordance with the above procedure.

3. NON-SPHERICALLY SYMMETRIC MOLECULES

3.1. General Considerations

Computer simulation of liquids with non-spherical molecules has been reported in the literature by several authors. In many cases the object has been to simulate liquid water.
Examples of computer simulations with more complicated molecules are presented in the work of Curro and of Ryckaert et al. [3.5]. They obtained promising results with the simulation of n-alkanes.

The reason why many others have concentrated on water is obvious, water is an important liquid, not only in physical chemistry. Complete succes as in the case of noble gases is not attained as yet because of the very complicated character of this liquid. Rahman and Stillinger have performed a number of calculations using semi-empirical potentials [3.1]. The first calculations were made with a potential designed by Ben-Naim and Stillinger [3.14], later with an improved version, the ST2 potential. Both potentials are pair interactions between rigid molecules. Six adjustable parameters are involved.

As mentioned above the results are not exactly in accordance with the physical data, but the agreement is certainly satisfactory, both for equilibrium properties (distribution function) as for dynamic properties (diffusion etc.).
The ST2 potential is an effective pair-potential in the sense that the mean effect of multi-particle interaction is included. Consequently such properties that are related to the properties of isolated pairs of molecules, are rather poorly represented. This is the case for the second virial coefficient of the gas phase [3.13]. On the other hand a many-particle effect such as the polarisation of the molecules is represented only as a mean value. Fluctuations of the polarisation and change of the polarisation with temperature are absent.
At the moment the ST2 potential must be considered to be the best pair potential that is available for the simulation of water. Nevertheless there is one detail in this potential that should be improved, that is the so called switching function.

In order to avoid a catastrophe when two point charges would coincide, the coulomb interaction is gradually 'switched off' when the distance between the centres of two molecules decreases. Consequently at short distance two neighbours interact as spherical molecules, which is non physical and causes an artifact in the distribution function [3.15]. Perhaps this may be of minor importance, but an elegant alternative exists [3.16].

Other important potentials are obtained with the quantum-mechanical approach. Popkie et al. and later Matsuoka et al. designed pair potentials for water with 9 or more parameters [3.2], which parameters were adjusted in a way that a best fit was obtained with the data of quantum-mechanical calculations of water dimers. Owicki and Scheraga and Lie et al. have made simulations with one of the Matsuoka potentials [3.3]. The radial distribution agrees well with the experimental data. This is an important criterion since the equilibrium thermodynamic properties are closely related to the distribution function [1.6].

Less satisfacory is the fact that the second virial coefficient of the gas is rather poorly in agreement with the physical data [3.17]. Contrary to expectation it seems that the potential that is the result of quantum-mechanical calculation of dimers gives better results for systems of many molecules than for clusters of two molecules.

Stillinger and David [3.4] have recently developed a potential which also represents the polarisation of the molecules. This aspect of the multi-particle interaction between water molecules is believed to be very important. Verification may be expected soon.

Thus further refinement of the intermolecular potential will result gradually in better agreement between calculated and experimental properties of liquid water.
However, in order to be able to discuss (and to understand) the relation between molecular properties and the macroscopic properties of liquids, it can be useful also to investigate the properties of systems with 'simple' model molecules. In that way it is possible to evaluate to what extent simple models are relevant for the understanding of the properties of liquids.

In the present work we will deal with such a simple model of a waterlike molecule, based on an idea of Ben-Naim [3.6]. The molecule consists of a hard sphere with additionally a short range pair interaction between neighbouring molecules. This interaction is of tetrahedral symmetry. Therefore we will speak of 'tetrahedral molecules'. Long range pair interaction between molecules is absent. In a later development (section 3.3.) we will evaluate the consequence of the introduction of a polar character into the molecules, still without long range interaction. The polar molecules can be considered to have some characteristics of the watermolecule.

Tetrahedral symmetry has been realised by assigning four vectors to the hard sphere, vectors pointing into the proper direction (fig. 3.1). When the molecules are close to each other, a bond exists only when both molecules have one of their vectors close to the line connecting the centres of the molecules (fig. 3.2.).

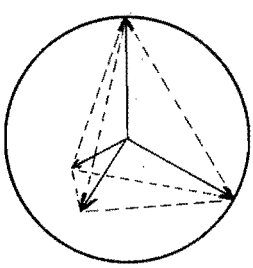


Fig. 3.1. 'Molecule' consisting of a hard sphere with additionally four vectors which are pointing to the corners of a tetrahedron.

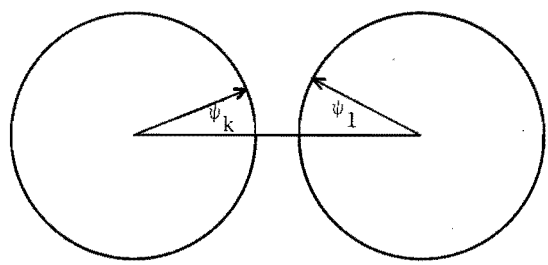


Fig. 3.2. Two molecules with vectors making angles ψ_k and ψ_l with the line connecting the centres. A bond exists when the internuclear distance and ψ_k and ψ_l are below preset values.

Thus the potential energy of interaction is:

$$\left.\begin{aligned} E_2 &= +\infty \qquad \text{if } r_{kl} < \sigma \\ E_2 &= \varepsilon \qquad \text{if } \sigma \leq r_{kl} < f.\sigma \\ &\qquad \underline{\text{and}}\ |\psi_k| < \psi_{max}\ \underline{\text{and}}\ |\psi_l| < \psi_{max} \\ E_2 &= 0 \text{ in all other cases} \end{aligned}\right\} \qquad (3.1)$$

The factor f will be taken as 1.10 and the angle ψ_{max} as 27^{o}. (The value of ψ_{max} has been chosen as large as possible without introducing the possibility of two molecules to be bonded to the same vector of a third one)
The energy ε is negative.

3.2. Simulations with Tetrahedral Molecules

3.2.1. Outline

With the above mentioned pair potential Monte Carlo simulations have been performed for systems with 91 molecules in a cubic box with periodic boundary conditions. For the volume of the box four different values have been chosen in such a way that the density D ($=N\sigma^3/V$) was 0.6667, 0.6000, 0.5455 and 0.5000 respectively. In that way the corresponding volumes increase proportionally to 9 : 10 : 11 : 12.

λ \ D	0.6667	0.6000	0.5454	0.5000
0	0.75 - 0.0	0.60 - 0.0	0.60 - 0.0	0.60 - 0.0
2	0.73 - 0.27	1.09 - 0.23	0.73 - 1.18	0.55 - 0.27
3	0.91 - 0.23	1.41 - 0.34	0.91 - 0.43	1.09 - 0.46
4	1.82 - 0.70	3.64 - 1.01	1.82 - 1.09	2.00 - 1.06
5	3.64 - 1.52	5.46 - 1.60	3.64 - 1.76	3.64 - 1.52
6	9.67 - 3.22	10.01 - 2.73	9.37 - 2.73	9.02 - 2.73

Table 3.1. The number of single steps in the 24 Monte Carlo simulations at different densities and at different values of the parameter λ ($= -\varepsilon/kT$).
The numbers are presented in units of 10^6 steps. In each case two numbers are given. The first one is the total number of steps. The second one is the number of initial steps that has not been counted for statistical purposes.

These rather low density values have been chosen in order to obtain information in the domain that is relevant because the density of physical water is close to these values.
For the parameter $\lambda = -\varepsilon/kT$ six values have been chosen: λ = 0, 2, 3, 4, 5 and 6. Higher values of λ are not practical since the approach towards equilibrium conditions becomes very slow in those cases.
In table 3.1 the number of steps performed in each of the 24 cases is reported.

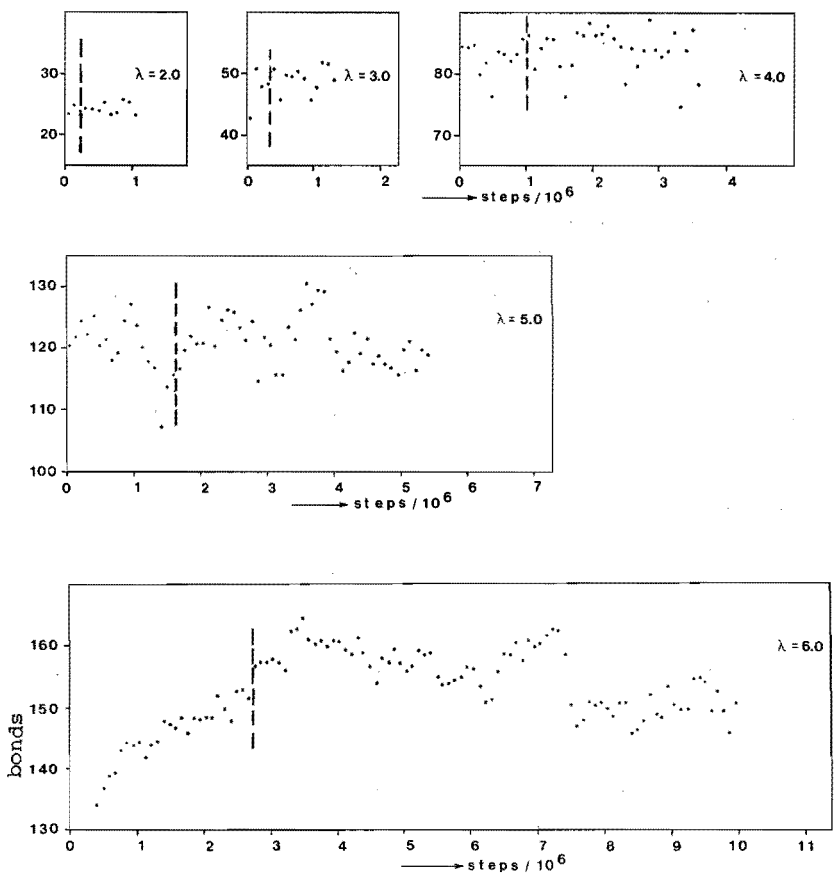


Fig. 3.3. The number of bonds as found in the course of five Monte Carlo simulations. N = 91 molecules. Density = 0.600. λ is mentioned inside each frame.
Each dot in the figures gives the mean number of bonds in 91000 successive steps. The vertical broken lines give the number of steps, below which the results have been omitted for calculation of the mean energy, pressure, etc.

In all calculations of mean values and in other statistical operations a number of initial steps has been discarded (except for the case $\lambda = 0$). The choice of the number of initial steps to be omitted is a rather arbitrary one.
In fig. 3.3 the number of bonds has been given as a function of the progress of the simulation with the indication of the limit of discrimination. We believe in all cases the choice has been made on the safe side. Only in the cases that $\lambda = 6$ the situation settles extremely slowly and the limit of discrimination should certainly not be chosen at a smaller number of steps than indicated.

3.2.2. Energy

We are dealing with a simple pair potential which has been defined in such a way (eq. 3.1) that the potential energy of the system is exclusively composed of the contributions of those pairs of molecules which have been specified as 'bonded' molecules. For all other pairs the energy is zero. Accordingly the potential energy of the system is simply proportional to the number of bonds.

If some numerical value is given to ε, each value of the parameter λ in table 3.1, except $\lambda = 0$, corresponds with a well defined temperature.
The value $\lambda = 0$ can only occur when $\varepsilon = 0$. The system is a hard-spheres fluid in that case.

Table 3.2 gives the mean number of bonds we found in the 24 computer simulations. In the case $\lambda = 0$ the bonds are registered in conformity with the other cases, namely if both the distance and the orientation of two neighbours are within the specified limits.
For $\lambda = 0$ the finding of a bond has no influence on the acceptance or rejection of a proposed step and by consequence has no influence on the structure of the liquids: there is only a purely 'administrative' existence of bonds.

Table 3.2 shows that the number of bonds increases with increasing density. That is understandable, since when the density increases the number of neighbours and the chance for making a bond increase accordingly.

However, for $\lambda = 5$ and 6 the situation is different. There appears a maximum at $D = 0.6$.
Apparently this is a structural effect. The tetrahedral molecule exhibits a preference for a tetrahedral surrounding. If at some density many configurations would exist in which the majority of the molecules has tetrahedral coordination, compression would tend to increase the mean coordination number and thus would tend to decrease the possibility of arrangements with four bonds per molecule. It must be mentioned that this effect is not specific for particles with a preference for tetrahedral coordination.
In the case of preference for hexahedral or octahedral coordination compression above a certain density would also diminish the possibility of the realisation of bonds. However, probably the effect will be weaker because of the smaller difference between the density that is most favorable for bonding and the density of a closely packed system (with twelve-coordination).

λ \ D	0.6667	0.6000	0.5454	0.5000
0	4.94	4.16	3.30	2.89
2	27.82	24.29	20.69	18.03
3	55.9	49.1	44.4	41.4
4	89.2	83.8	77.7	70.2
5	121.0	121.1	115.8	108.1
6	150.1	155.2	145.5	139.8

Table 3.2 Mean number of bonds, present in the 24 Monte Carlo simulations at different densities and different values of the parameter λ. The maximum possible number of bonds equals $2N$ (=182).

3.2.3. Distribution Functions and Coordination Numbers

During each simulation a great number of distances between pairs of molecules is calculated. It is easy to use these data for quantitative analysis and keep administration of the distances lying between r and r + Δr, for all values of r between 0 and half of the edge of the cubic volume of the system.
When a sufficiently large number of cases has been administrated, the mean number of particles, found at a distance between r and r + Δr from any reference particle is calculated. This mean number is identified with the product G(r) . Δr.

The magnitude of G(r) depends on three functions, firstly the magnitude of the volume between r and r + Δr, which is $4\pi r^2 \Delta r$, secondly the magnitude of the mean density of the system, N/V and thirdly the apparent local deviation from that density. Only the third function contains new information about the system that is studied. Therefore we will eliminate the others in order to get the new function :

$$g(r) \equiv G(r)/(4\pi r^2 . N/V) \qquad (3.2)$$

This new function is called the radial distribution function.

In fig. 3.4 sixteen radial distribution functions have been plotted. In all cases no distances $r < \sigma$ can exist. Accordingly the curve begins with $R = r/\sigma = 1.0$.
For $\lambda = 0$ the curves give the situation in a hard-spheres fluid with neighbours situated mainly between R = 1.0 and R = 1.25*).

*) It may seem puzzling why spheres that do not attract each other exhibit a g(r) above unity for low values of r. Apparently this is a geometric effect. Each hard sphere has an excluded volume of $\pi . \sigma^3/6$ where the centre of a second sphere cannot be located. When there are two spheres, the excluded volume for a third one is two times that value if the two spheres are more than 2σ apart, but less if they are closer to each other (due to the overlap of the two excluded volumes).
So when the distance between two spheres decreases, the remaining volume for a third increases. Consequently the probability of a short distance between two spheres increases on the introduction of a third one. When more spheres are introduced the effect is still reinforced.

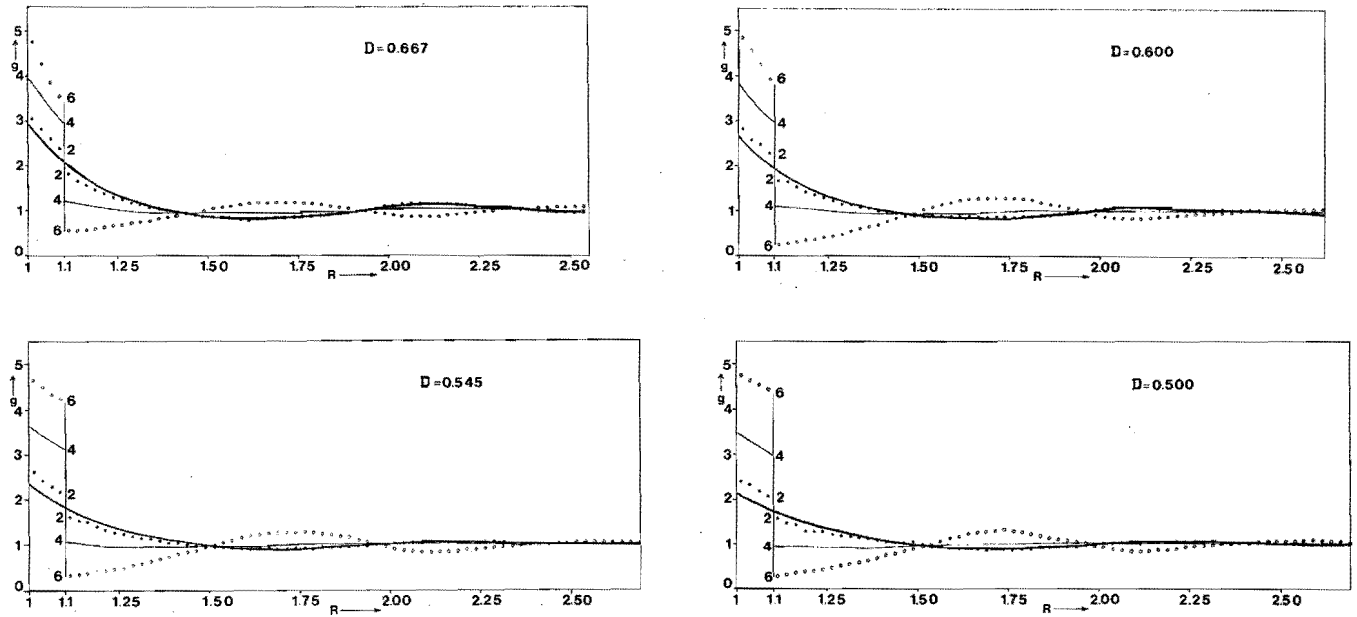


Fig. 3.4. Radial distribution function g, plotted as a function of R = r/σ. The curves result from sixteen Monte Carlo runs, with four different values of the density D = Nσ³/V and four values of the parameter λ = − ε/kT. The value of the density is indicated at the top of each of the four frames. Except for λ = 0 (which value pertains to the drawn line without discontinuity at R = 1.1, the distribution function of the hard spheres fluid) the value of λ is given next to each curve at the discontinuity at R = 1.1.

Furthermore there is a maximum between 2.0 and 2.25 and a minimum at about 1.7. Maximum and minimum both tend to move to a lower value of R with increasing density. Moreover the height of the maximum and the depth of the minimum increase with density.
When λ increases, a sharply limited maximum arises between R = 1.0 and 1.1. For λ = 6 there is a minimum just beyond 1.1 and at about 2.1 and a maximum at about 1.7. For λ = 2 and 4 the curves are intermediate between the curves of λ = 0 and λ = 6.
The difference between λ = 6 and the hard spheres is striking. The order of the maxima and minima have completely interchanged. A curious consequence of this interchange is that for λ = 4 the curve is almost straight above 1.1.

The existence of points of intersection at R = 1.5, 2.0 and 2.3 suggests that there are two different types of second neighbours. Apparently the second neighbours in a hard-spheres system are situated at a longer distance than the second neighbours in a strongly interlocked structure. These short-distance second neighbours are attached with a bond to the same molecule. This can easily be verified since we know all details of all configurations.
Consequently in all cases can be evaluated what the distance is between two molecules, both bound to a third one. In this way a special distribution function can be calculated with only that type of distances. In fig. 3.5 that special distribution function is compared with g(r). For λ = 6 the peak at R = 1.7 proves to be caused by bonded second neighbours. As can be expected the effect is almost absent in the case λ = 4.
The tetrahedral symmetry of the molecule demands that the most probable distance for second neighbours is $\frac{2}{3}\sqrt{6}$ times the mean length of a bond.
Accordingly the most probable distance becomes 1.715 (the mean length is 1.05 σ). Detailed analysis of the curve in fig. 3.5 proves that the peak occurs at slightly smaller R. This is an artifact of the transformation of G(r) into g(r) in accordance with eq. (3.2). Division by r^2 results in a shift of the maximum. When G(r) itself is analysed the maximum is situated correctly at the calculated position. This applies for all four densities.
Considering fig. 3.4 it is evident that all g(r) show a discontinuity at $r = f \cdot \sigma$ except when λ = 0.

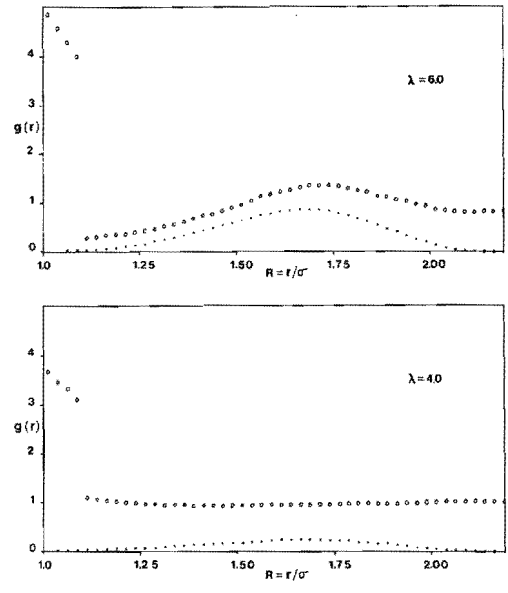


Fig. 3.5. Contribution of bonded second neighbours to the radial distribution function g(r). Circles : all neighbours. Dots : second neighbours. Density = 0.600. λ is given in the upper right.

It will be clear that this discontinuity arises from the fact that the possibility of making a bond exists for molecule-pairs with a distance between σ and f . σ and for pairs with larger intermolecular distance not.
The probability of the occurrence of pairs at a distance r is proportional to the product of the Boltzmann factor $\exp(-E_2/kT)$ (which shows a discontinuity at r = f . σ) and the magnitude of the corresponding fraction of the configuration space (which is a continuous function of r). Accordingly the function y(r) should not show a discontinuity any more. y(r) defined as:

$$\begin{aligned} y(r) &\equiv g_1(r)/\exp(-\varepsilon/kT) + g_2(r) \quad (\text{if } \sigma \leq r < f.\sigma) \\ y(r) &\equiv g(r) \quad (\text{if } r \geq f.\sigma) \end{aligned} \tag{3.3}$$

In these equations $g_1(r)$ is the distribution function pertaining to all those pairs which have formed a bond and $g_2(r)$ is the contribution to the distribution function of all other pairs (for which the interaction energy is zero).

In fig. 3.6, the function y(r) has been plotted for a couple of simulation experiments. The discontinuity at R = 1.1 appears to have been removed, as expected.

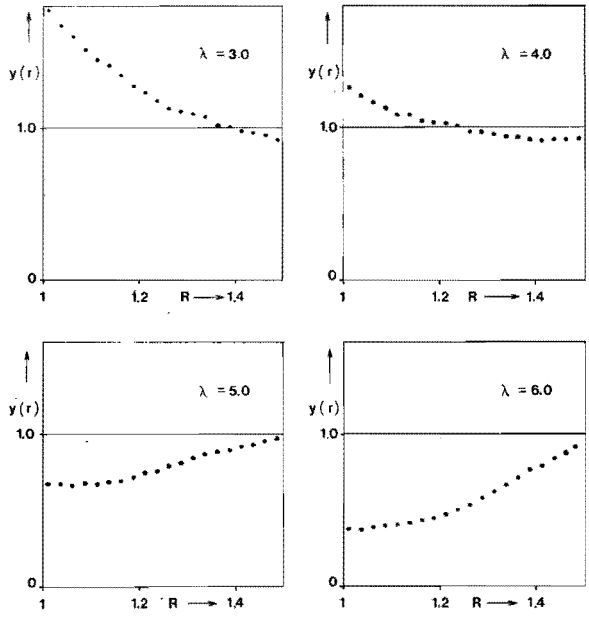


Fig. 3.6. The function y(r) as defined with equation (3.3) as a function of R = r/σ.
There is no significant indication of the existence of a discontinuity at R = 1.1.
The density is 0.600 and λ is printed inside each frame.

Since y(r) is in some respect more transparant than g(r) we will use y(r) in the discussion of the next few pages.

With the data we have generated it is possible to obtain all statistical data on the relative position of pairs we could need. The relative orientation of two molecules, however, depends on five independent parameters. At the same time the distance is relevant. Consequently the relative position of two molecules is a six dimensional function. In the literature no satisfactory method to display this function has been proposed

We will restrict ourselves to a simplification which can easily be generated by our computer program. In our program we deal with a potential in which two angles, ψ_k and ψ_l play a role (formula 3.1). For each pair with a low internuclear distance, we have to check whether ψ_k and ψ_l are below ψ_{max} in order to know whether a bond exists or not.

For orientational statistics we have extended the procedure to all pairs, irrespective the internuclear distance.
In that way we get three types of pairs :

a. pairs for which both ψ_k and ψ_l are below the limiting value
b. pairs for which either ψ_k or ψ_l is below the limiting value
c. pairs for which ψ_k and ψ_l are both outside the limit.

For each of the three types a radial distribution can be made exactly as described above.
We have done so, we have performed the operations of the formulae (3.2) and (3.3) and obtained three new functions, $q_1(r)$, $q_2(r)$ and $q_3(r)$, the designation 1, 2 and 3 being in accordance with the three types of pairs in the above list.
The sum of the three functions equals $y(r)$

$$q_1(r) + q_2(r) + q_3(r) = y(r) \qquad (3.4)$$

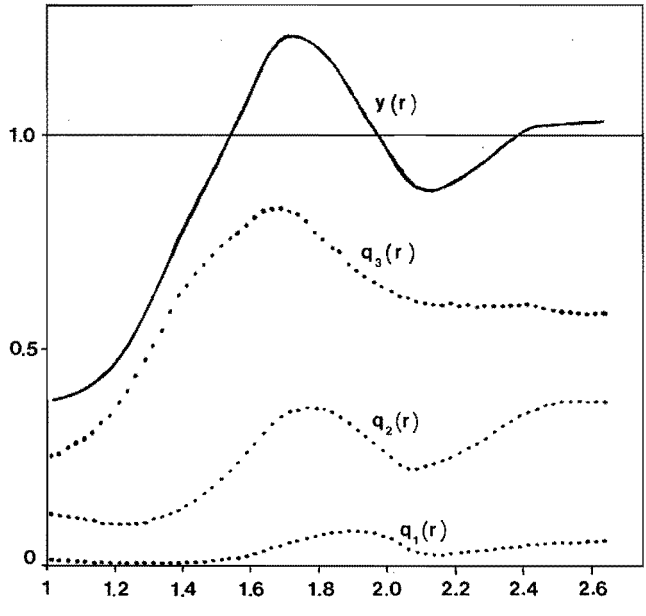


Fig. 3.7. The distribution function y(r) split up into three functions with different orientational characteristics :
q_1 both relevant angles are within the limits
q_2 only one angle within the limits
q_3 none of the angles within the limits.
The density is 0.600 and λ is 6.0.

In fig. 3.7 the functions are plotted for a special case, namely D = 0.600 and λ = 6.0. It can be seen that the q-curves are similar to y(r).

This is an obvious result since when there are many molecules at a certain distance, the probability of finding molecules of a certain orientation will be rather high. It is, however, the deviation from the mean pattern that would give new information about structure. Therefore it is better to transform the q's into a new function by division by the value that would exist if the orientation would be random.

That is easily achieved, since in case of random orientation the function would simply be the product of y(r) and a probability factor.

If β is the probability of a random vector to have an angle below ψ_{max} with one of the four vectors of our tetrahedral molecule, then the probability of two vectors to be correct is β^2, of one vector to be correct is $2\beta(1-\beta)$ and of no vector correct is $(1-\beta)^2$.

Accordingly three new functions have been defined as follows:

$$\omega_1(r) \equiv q_1(r)/(y(r).\beta^2) \tag{3.5}$$

$$\omega_2(r) \equiv q_2(r)/(y(r).2\beta(1-\beta)) \tag{3.6}$$

$$\omega_3(r) \equiv q_3(r)/(y(r).(1-\beta)^2) \tag{3.7}$$

The value of β is a function of ψ_{max}, which is easily evaluated analytically to be:

$$\beta = 2(1-\cos(\psi_{max})) \tag{3.8}$$

In fig. 3.8, 3.9 and 3.10 the functions ω_1, ω_2 and ω_3 are plotted for eight computer simulations.

The case λ = 0 is not mentioned since ω_1, ω_2 and ω_3 equal unity for all r, apart from inevitable statistical fluctuations.

For λ = 2 a slight deviation from unity appears which deviation increases when λ increases.

All curves of ω_1 and ω_2 (fig. 3.8 and fig. 3.9) exhibit a minimum at R = 1.4. In the case of larger λ there is a maximum in ω_1 at R = 1.9 and a maximum in ω_2 at R = 1.8.

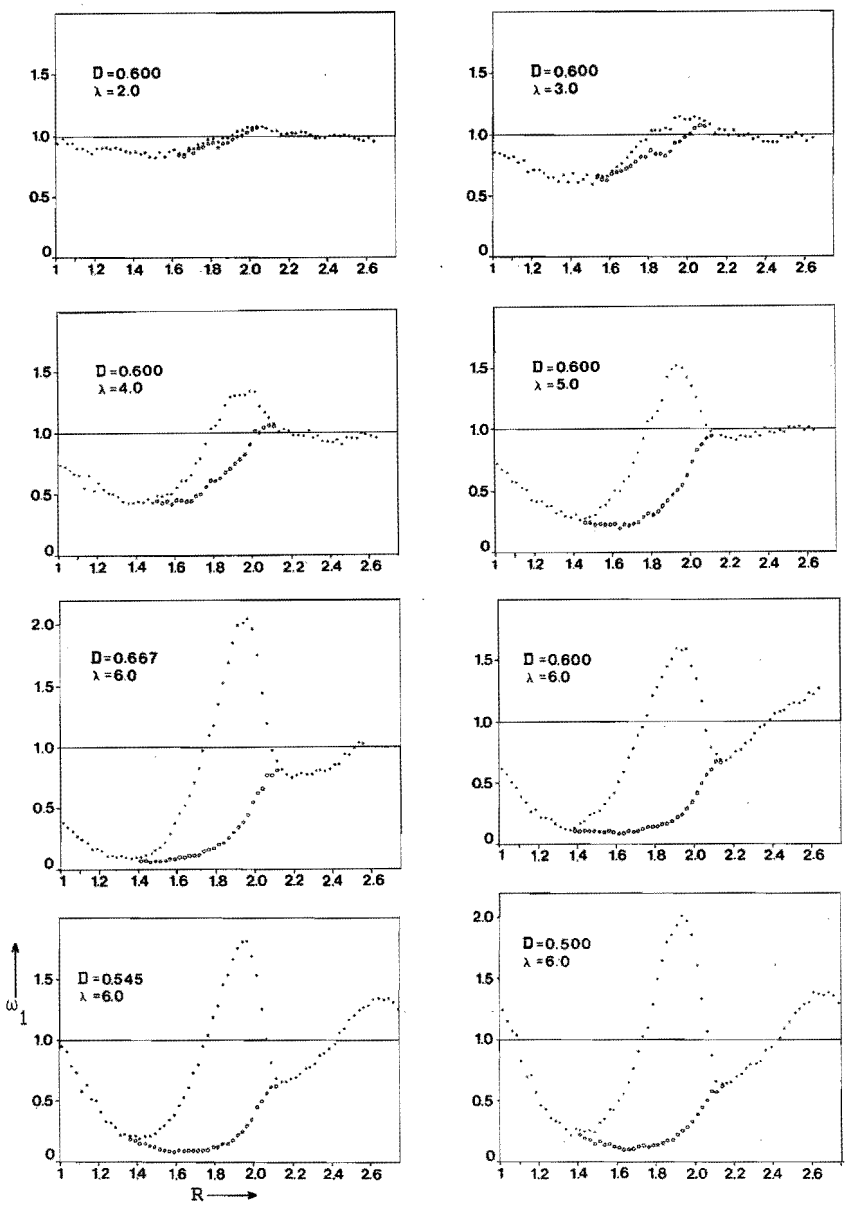


Fig. 3.8. First orientational distribution function ω_1. This is the relative number of pairs that satisfy the orientational condition of the potential (formula 3.1). The number as found in the Monte Carlo simulations has been divided by the number that one would expect in the case of random orientation. Moreover the procedure of eq. (3.3) is applied. Circles give the same function if second neighbours are omitted.

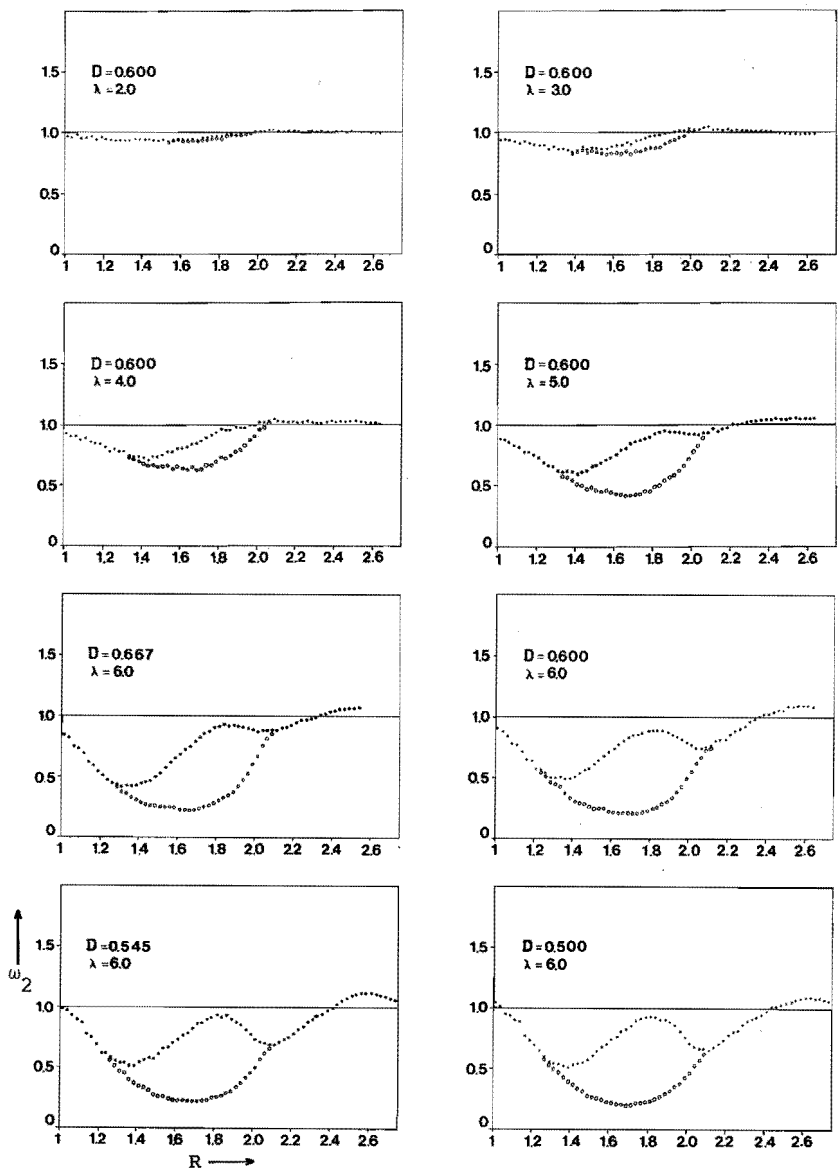


Fig. 3.9. Second orientational distributionfunction ω_2. This is the relative number of pairs of which only one molecule satisfies the orientational condition of the potential (formula 3.1).
Like in fig. 3.8. the number is relative to the number one would expect in case of random orientation. Circles give the same function if second neighbours are omitted.

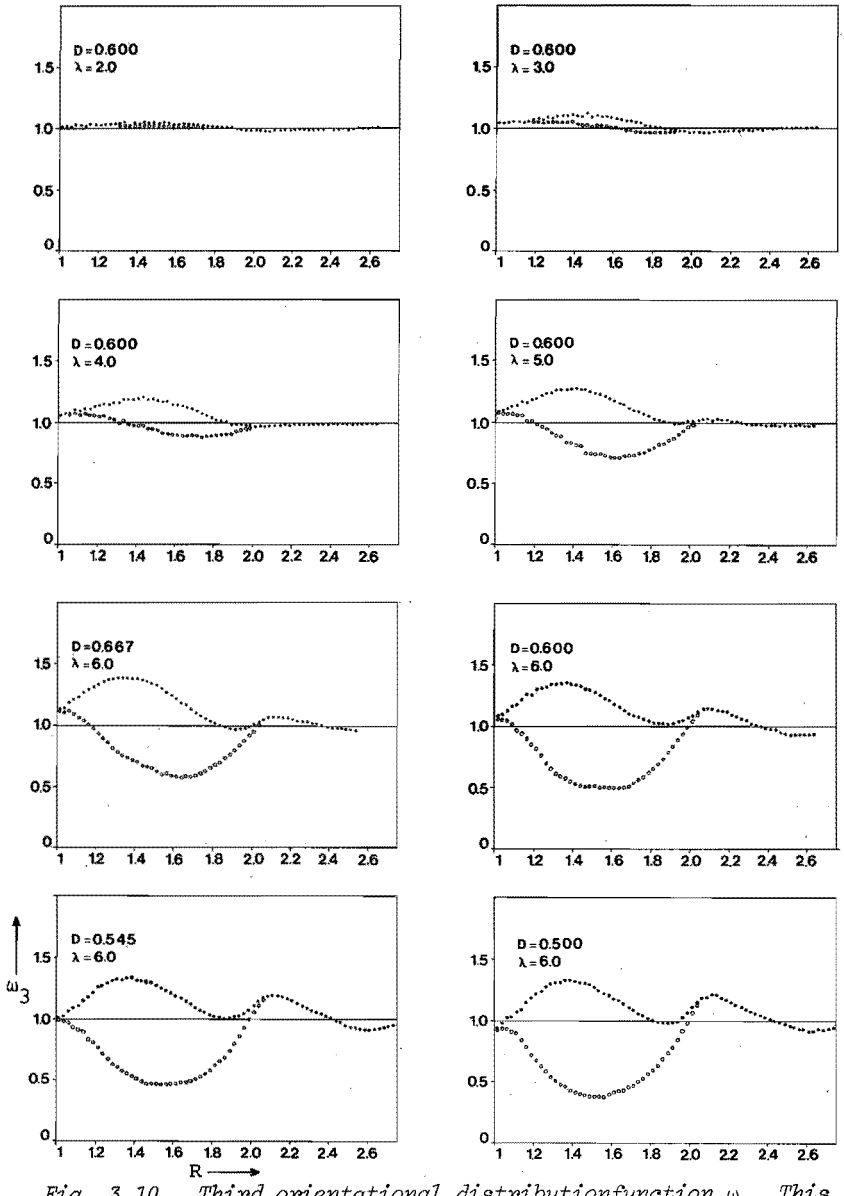


Fig. 3.10. Third orientational distributionfunction ω_3. This is the relative number of pairs of which none of the molecules satisfies the orientational condition of the potential (formula 3.1).

Like in the previous figures the number is relative to the number one would expect in case of random orientation. Circles give the same function if second neighbours are omitted.

ω_3 has the opposite character, a maximum at R = 1.4 and a minimum at R = 1.8. This opposite behaviour originates from eq. 3.4, which demands that $\Sigma q_i/y$ must equal unity.
The maximum in ω_1 and ω_2 at R $\tilde{\gtrsim}$ 1.8 originates from second neighbours. If the pairs that have a bond with the same third molecule are left out, the peak disappears entirely. In fig. 3.11 this situation is illustrated. It is clear that when the angle ABC is large, it is rather trivial that the chance is large for one or two of the vectors of A and C to be sufficiently in line with AC. Accordingly, if the angle ABC is large the probability of A and C to be assigned to the functions ω_1 or ω_2 is also relatively large. The maximum therefore exceeds slightly the value R = 1.71, the most probable distance of second neighbours. When considering second neighbours only we find a maximum for the first group (ω_1, two angles below ψ_{max}) at R = 1.90 and for the second group (ω_2, one angle below ψ_{max}) at R = 1.78.
The minimum in ω_1 and ω_2 at R = 1.4 (or in the curves without second neighbours at R = 1.7) can easily be understood. The bonded neighbours, although they are made invisible in the figures by the operation of formula 3.3, are present in reality and they block an important part of the space near the vector concerned. Therefore eventually molecules at a distance below R = 2.1 will be forced more or less to the space remote from the four tetrahedral vectors. By consequence ω_3 will be large in that region and ω_1 and ω_2 will be small.

Above R = 2.1 there exists no such direct relationship between pairs as mentioned above. For λ up to 5.0 the curves appear to deviate not significantly from unity in that region.
Orientational correlations for R > 2.1 appear to be very weak. For λ = 6.0, however, there appears to be marked deviation from unity at high values of R. Apparently the orientational correlation between remote pairs becomes strong with increasing number of bonds. (it should be kept in mind that long range forces between pairs are not present.) In this respect there is a marked difference between λ = 5.0 and λ = 6.0.
The curves suggest that there is much detail at R > 2.6. This region, however, can be explored only in larger systems.

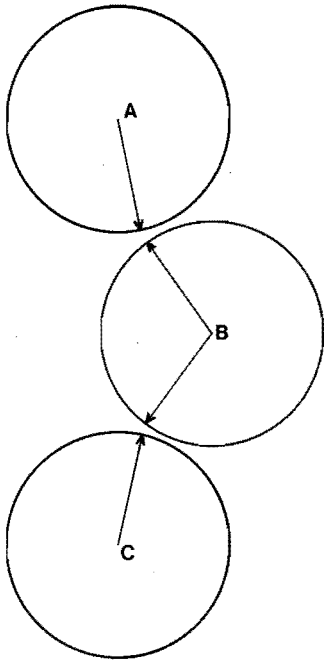


Fig. 3.11. Second neighbours with orientational condition fulfilled. The molecules A and C are second neighbours in the sense that both have a bond with B. The bonding vectors of A and C are orientated in a way that their angle with the line connecting the centres is less than ψ_{max}, *which is not the case for all second neighbours. A pair of molecules like A and C will be included in the first orientational function* ω_1.

In the above discussion of the minimum at R = 1.4 (or 1.7) we mentioned the division of space around a molecule into a part close to the vectors of the tetrahedral molecule and a part corresponding to the remaining space. Now imagine the existence of four identical cones which have an angle at their top which equals $2\psi_{max}$.

The four cones are tetrahedrally arranged with their tops at the centre of the reference molecule. Molecules whose centre is inside one of the cones, exhibit at least one angle $\psi < \psi_{max}$. Using this representation we can derive from our curves an alternative way of arranging the facts: we define a set of two functions, $t_1(r)$ and $t_2(r)$, which represent radial distribution functions of the molecules with their centre inside and outside the cones respectively.

The relation between t and q is:

$$\left.\begin{aligned} t_1(r) &\equiv q_1(r) + \tfrac{1}{2}q_2(r) \\ t_2(r) &\equiv q_3(r) + \tfrac{3}{2}q_2(r) \end{aligned}\right\} \qquad (3.9)$$

In fig. 3.12 is plotted the function $\Theta_1(r) \equiv t_1(r)/y(r)$

At the same time the figure shows $\Theta_2(r) \equiv t_2(r)/y(r)$, which equals $1.0 - \Theta_1(r)$.

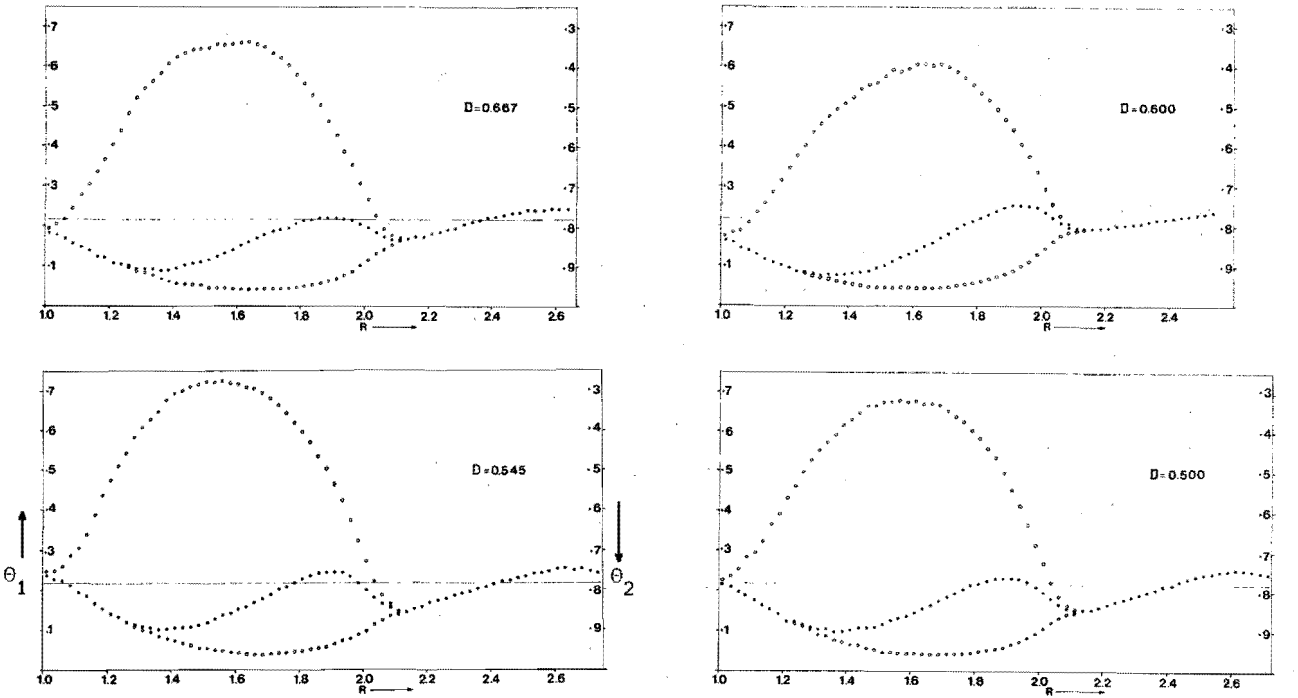


Fig. 3.12. Discrimination of the molecules, surrounding a reference molecule, into two groups. One fraction, θ_1 (black points, left scale) is the fraction lying close to one of the four bonding directions of the reference molecule, irrespective of the other bonding conditions. The other fraction, θ_2 (black points, right scale) is the fraction outside the bonding region. The circles illustrate the situation if second neighbours are ignored. The lower curve is θ_1 minus second neighbours (left scale), the upper curve is θ_2 minus second neighbours (right scale). The horizontal line gives θ in the case of random orientation. Horizontal axis : $R = r/\sigma$, vertical axis : fraction of the total number of molecules at the pertinent distance, after elimination of the discontinuity with (3.3).

Coordination Numbers

Finally we can make a few remarks on the well known concept of the coordination number, which is closely related to the distribution function.

We will discuss two alternatives. Firstly we consider the "all in" coordination number, including both bonded and non-bonded neighbours at a distance between σ and $f \cdot \sigma$ from a reference molecule. In table 3.3.1 the mean value of these coordination numbers is given for the 24 Monte Carlo simulations we are dealing with. Additionally has been mentioned the fraction (in per cent) of the coordinating neighbours having a bond with the reference molecule.

As can be expected the coordination number increases with density and with the parameter λ. The percentage of bonded neighbours is low for $\lambda = 0$ and is independent of the density in that case. For larger values of λ the percentage generally decreases with increasing density. For $\lambda = 6$ the latter decrease is very slow when going from $D = 0.5$ up to $D = 0.6$, but from there to $D = 0.667$ the percentage falls rather significantly.

λ \ D	0.667		0.600		0.545		0.500	
0	2.310	4.75	1.875	4.75	1.575	4.75	1.357	4.75
2	2.51	24.3	2.09	25.6	1.79	25.4	1.55	25.6
3	2.80	43.8	2.38	45.3	2.09	46.8	1.88	48.3
4	3.15	62.1	2.80	65.8	2.54	67.2	2.24	68.8
5	3.48	76.4	3.21	82.9	3.01	84.5	2.73	86.9
6	3.81	86.7	3.65	93.5	3.39	94.2	3.24	94.8

Table 3.3.1 The mean coordination number of the molecules in the 24 Monte Carlo simulations, as mentioned in table 3.1. In each case there are two numbers, left: the coordination number and right: the percentage of coordinating neighbours that is bonded.

From this point of view we can state that the maximum in the number of bonds that has been observed in table 3.2 for $\lambda = 6$ and $D = 0.6$ is caused by the fact that on increasing the density, the fraction of the neighbours that are bonded falls rapidly. On the other hand on decreasing the density the coordination number decreases with a minor increase of the fraction of bonded neighbours. In other words for $D = 0.6$ the geometrical conditions for making bonds are favorable. On compression some neighbours are forced into the non-bonding regions around the molecules. On expansion the distance between the molecules becomes too long for bonding.

Secondly there is an alternative approach. When we restrict ourselves to the bonded neighbours there are five types of molecules namely those with 0-, 1-, 2-, 3- and 4- bonded neighbours.
In the course of the simulations we counted the frequency of each of these groups, which frequencies (in percents of the total) are presented in table 3.3.2.
The data are closely related to the total number of bonds of table 3.2. In the present table we can see that the evolution from a low to a high number of bonds takes place progressively with λ.
Close examination of the data shows that for small λ the distribution resembles closely the binomial distribution as if we were dealing with completely independent events.
For larger value of λ the distribution deviates slightly from the binomial in the sense that the fraction of non, singly and fourfold bonded molecules in the model is slightly larger than would be expected from the binomial formula.
As an illustration we repeated the data of $\lambda = 6$ from table 3.3.2 in table 3.3.3 together with the fractions that would result from a binomial distribution with the same total number of bonds.

D	λ	Number of bonded neighbours :				
		0	1	2	3	4
0.500	0	93.8	6.0	0.1	0.0	0
	2	65.9	29.0	4.7	0.4	0.0
	3	36.0	41.5	18.2	3.9	0.3
	4	14.3	35.5	34.1	13.9	2.2
	5	2.7	15.1	34.5	34.7	13.1
	6	0.5	4.6	18.6	38.7	37.5
0.545	0	92.9	6.9	0.2	0.0	0.0
	2	61.7	31.7	6.1	0.5	0.0
	3	32.7	42.0	20.4	4.6	0.4
	4	11.7	32.1	34.1	17.8	4.2
	5	2.2	13.0	31.0	35.8	17.9
	6	0.4	3.3	15.4	38.1	42.9
0.600	0	91.2	8.5	0.3	0.0	0
	2	56.4	34.6	8.1	0.8	0.0
	3	28.5	41.9	23.4	5.7	0.6
	4	9.1	28.5	36.4	21.2	4.8
	5	1.7	11.1	27.8	38.1	21.3
	6	0.2	1.7	9.8	33.5	54.8
0.667	0	89.6	10.0	0.4	0.0	0
	2	50.6	37.9	10.2	1.3	0.1
	3	23.8	39.4	27.2	8.4	1.2
	4	7.0	26.0	35.8	24.7	6.5
	5	1.4	10.2	29.2	39.4	19.8
	6	0.3	2.6	12.5	36.5	48.2

Table 3.3.2 The fractions of molecules (in per cent) with 0, 1, 2, 3 and 4 bonded neighbours respectively.

D	λ	Number of bonded neighbours : 0	1	2	3	4
0.500	6	0.5 0.29	4.6 3.83	18.6 19.03	38.7 42.03	37.5 34.81
0.545	6	0.4 0.16	3.3 2.58	15.4 15.42	38.1 40.99	42.9 40.85
0.600	6	0.2 0.05	1.7 1.09	9.8 9.46	33.5 36.52	54.8 52.88
0.667	6	0.3 0.09	2.6 1.78	12.5 12.54	36.5 39.33	48.2 46.26

Table 3.3.3 The fractions of molecules (in per cents) with 0, 1, 2, 3 and 4 bonded neighbours respectively.
Above: the data from the Monte Carlo simulations as in table 3.3.2
Below: the binomial distribution.

3.2.4 Pressure

The pressure can be evaluated from Monte Carlo calculations in two ways.
In the first place when the free energy is calculated,which will be the case in the next section, the pressure can be evaluated as the derivative of the free energy with respect to the volume at constant temperature.
Secondly the pressure can be evaluated from the virial. In this section we will report the results of the second possibility.

In appendix 4 we developed formula (A 4.6) and (A 4.7) for the calculation of PV/NkT, from the distribution-function data. These data are obtained in the Monte Carlo simulations and are discussed in section 3.2.3.

We have applied (A 4.6) to the results of the 24 Monte Carlo simulations, previously mentioned (see table 3.1), and have obtained PV/NkT in 24 cases.
The results are plotted in figure 3.13.

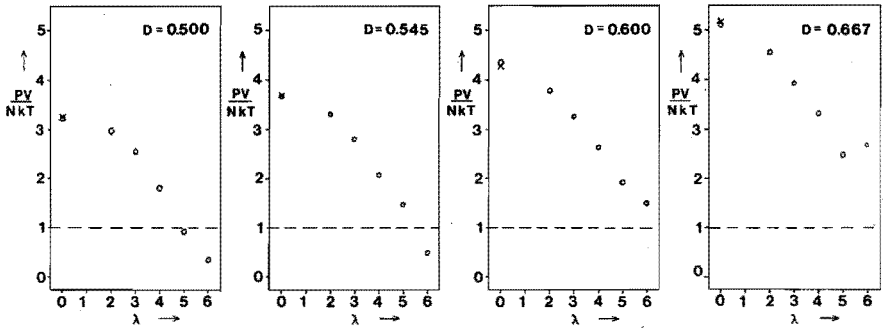


Fig. 3.13. PV/NkT as a function of λ (= − ε/kT). The circles indicate the values calculated with (A 4.6) from the data of the 24 Monte Carlo simulations of table 3.1. The crosses indicate the value calculated with the formula of Carnahan and Starling for hard spheres (3.10). The density is indicated inside each frame.

For $\lambda = 0$ the system is identical to the hard-spheres fluid of the same density. The Carnahan and Starling formula for PV/NkT is :

$$\frac{PV}{NkT} = \frac{1 + y + y^2 - y^3}{(1 - y)^3} \tag{3.10}$$

$y = D \cdot \pi/6$.

The values resulting from (3.10) are indicated in figure 3.13 with a cross. The accordance with the data from (A 4.6) is good.

The pressure can be evaluated from PV/NkT. For that purpose it is necessary to introduce reduced variables defined by :

$$\left.\begin{aligned} P^* &= P\sigma^3/|\varepsilon| \\ V^* &= V/(N\sigma^3) \\ T^* &= kT/|\varepsilon| \end{aligned}\right\} \tag{3.11}$$

(See [1.9])

The evaluation of these reduced quantities from $D(=N\sigma^3/V)$, $\lambda(=-\varepsilon/kT)$ and PV/NkT can be performed with the aid of the equations :

$$P^* = \frac{PV}{NkT} \cdot \frac{D}{|\lambda|} \tag{3.13}$$

$$V^* = 1/D \tag{3.14}$$

$$T^* = 1/|\lambda| \tag{3.15}$$

In table 3.4 the resulting values of P* are given.

	D	0.667	0.600	0.545	0.500
λ	V* / T*	1.500	1.667	1.833	2.000
2	0.500	1.51	1.13	0.91	0.75
3	0.333	0.87	0.65	0.51	0.43
4	0.250	0.55	0.39	0.28	0.23
5	0.200	0.33	0.23	0.16	0.09
6	0.167	0.30	0.15	0.05	0.03

Table 3.4 The reduced pressure P* as a function of the volume and the temperature.

3.2.5. Free Energy

3.2.5.1. Pairwise Combination of Monte Carlo Simulations

The free energy can be calculated when the partition function is known ($A = - kT\ln Q$), and the partition function can be calculated when the details of the configuration space are known.

The starting point for investigating the configuration space in a system with tetrahedral molecules with a hard core is the hard-spheres fluid. Each configuration that can exist for the tetrahedral molecules, also exists in the hard spheres case as far as the location of the centres of the molecules is concerned. Likewise the reverse holds : each configuration of a hard spheres fluid can exist for tetrahedral molecules with a hard core as well.

Once the density is given, the magnitude of the configuration space of a hard spheres system is fixed and can be calculated with the Carnahan and Starling algorithm (eq. (2.18) of section 2.3).
When the possibility of bonding is introduced, the configuration space is divided into parts corresponding to a different number of bonds and by consequence different energies. Since for tetrahedral molecules the number of bonds in a N-molecules system must lie between 0 and 2N, the configuration space is divided into 2N + 1 parts. We denote the relative magnitude of each part with C_j, j being the number of bonds present. Just as we have done in section 2.2.

The analogon of equation (2.9b) is now :

$$C_j \equiv \frac{1}{\mathcal{Z}'_{HS}} \int \xi\,(j)\,dq_i \qquad (3.17)$$

$\mathcal{Z}'_{HS}$ is the excess configuration integral of the corresponding hard spheres ystem. (See appendix 1 equation (A 1.9)).
ξ (j) = 1 if the number of bonds equals j and ξ (j) = 0 if not.

In analogy to section 2.2,the equations (2.10) up to (2.16) apply with the exception that in (2.11) and (2.13) $\mathcal{Z}''_N$ must be read in stead of $\mathcal{Z}'_N$. ($\mathcal{Z}''_N$ as defined in (A 1.9)).

Moreover equation (2.20) applies : $\phi_j = j$, since all bonds have the same contribution to the energy. The analogon of (2.17) becomes :

$$PA_j \overset{\sim}{\sim} Pj = C_j \exp(\lambda \cdot j)/\mathcal{Z}''_N(\lambda) \qquad (3.18)$$

From a set of Monte Carlo simulations we can evaluate $\mathcal{Z}''_N$ for all values of λ with the method of pairwise combination as described in chapter 2 section 3. In that way a function γ (j) is defined and $d\gamma/dj$ is estimated from the apparent probabilities in the simulations.
Finally $\ln\mathcal{Z}''_N$ is evaluated as an integration constant (equations (2.22) up to (2.25)).
We made these calculations for the 24 simulations mentioned in section 3.2.1 (table 3.1). The resulting values of $\mathcal{Z}''_N(\lambda)$ are given in table 3.5.1.

Table 3.5.1 The value of the logarithm of the excess configuration integral, $\ln(\mathcal{Z}''_N(\lambda))$ as calculated with the aid of (2.25) from the pairwise combination of Monte Carlo simulations with the same density and different values of λ.

λ_1	λ_2	λ \ D	0.6667	0.6000	0.5455	0.5000
0	2	2	+ 27.22	+ 23.60	+ 19.20	+ 16.74
2	3	3	+ 67.96	+ 59.16	+ 50.55	+ 45.30
3	4	4	+ 140.93	+ 124.69	+ 110.24	+ 100.49
4	5	5	+ 246.25	+ 227.61	+ 208.30	+ 190.67
5	6	6	+ 382.80	+ 366.39	+ 340.22	+ 316.34

Likewise, with the help of the integrated equation for γ :

$$\ln(C_j) = \gamma(j) = x_0 + x_1 \cdot j + \tfrac{1}{2} x_2 \cdot j^2 \ldots + \tfrac{1}{5} x_5 \cdot j^5 \qquad (3.19)$$

C_j can be evaluated for the range of j-values, that is relevant for the pair of Monte Carlo simulations under consideration. The result is given in table 3.5.2.

It is interesting to compare the present result with what should be expected from the simple lattice theory [3.18]. As will be described in appendix 5, Gosling and Singer [A.6] have developed a method for the calculation of the free energy on the basis of the simple lattice theory, starting with the acceptance ratio of the moves in a Monte Carlo simulation.
In this line of thought one starts with a configuration of N particles in a volume V. The volume is divided into N Voronoi polyhedrons (with one particle in each polyhedron). When one particle is moved leaving all others at their places, the motion is hindered by the surrounding particles. There appears to be only a small 'free volume' v_f.
v_f is a fraction of the volume of the Voronoi polyhedron (which is about V/N).

In a Monte Carlo simulation a great number of moves within a preset volume v_m is proposed (v_m is a fraction of V). A fraction ϕ_t of the proposed moves is allowed.

The free volume is in accordance with Gosling and Singer :

$$v_f = \phi_t \cdot v_m$$

Identifying this free volume with v_f in the lattice theory of liquids (see Barker [3.18] page 30) then the excess partition function with respect to the ideal gas is :

$$Q'_t = \frac{(\phi_t \cdot v_m)^N}{V^N}$$

Table 3.5.2 The logarithm of the fraction of the configuration space where the number of bonds is j, ($\ln C_j$), as a function of j for the four densities investigated.

j \ D	0.667	0.600	0.545	0.500
0	-5.18	-4.29	-3.34	-2.99
5	-1.70	-1.79	-2.12	-2.34
10	-4.12	-5.11	-6.68	-7.60
15	-9.37	-11.20	-13.97	-15.68
20	-16.59	-19.23	-23.14	-25.73
25	-25.34	-28.81	-33.76	-36.91
30	-35.38	-39.68	-45.46	-49.36
35	-46.50	-51.72	-58.33	-62.86
40	-58.48	-64.77	-72.21	-77.30
45	-71.42	-78.73	-87.02	-92.62
50	-85.19	-93.51	-102.70	-108.75
55	-99.73	-109.09	-119.10	-125.85
60	-114.93	-125.48	-136.38	-143.80
65	-130.87	-142.62	-154.40	-162.58
70	-147.54	-160.47	-173.13	-182.14
75	-164.95	-179.02	-192.54	-202.48
80	-183.11	-198.23	-212.58	-223.24
85	-202.02	-218.09	-232.97	-244.73
90	-221.70	-238.65	-254.02	-266.88
95	-242.13	-259.77	-275.71	-289.68
100	-263.37	-281.53	-298.06	-313.14
105	-285.41	-303.94	-321.10	-337.26
110	-308.25	-326.99	-344.85	-362.06
115	-331.90	-350.71	-369.36	-387.28
120	-356.37	-375.11	-394.54	-413.54
125	-381.51	-400.23	-420.18	-440.16
130	-407.35	-426.08	-446.78	-467.88
135	-434.04	-452.62	-474.27	-496.52
140	-461.60	-479.88	-502.73	-526.15
145	-490.11	-507.86	-532.24	-556.86
150	-519.66	-536.61	-562.91	-588.74
155	-550.34	-566.17	-594.86	-
160	-582.29	-596.65	-	-
165	-	-628.21	-	-

Since we deal also with rotational steps we must perform the same procedure for the angular motion. When the maximum rotation for each of the three axis is α, the result is :

$$Q'_r = \frac{(\phi_r \cdot 2\,(1-\cos\alpha) \cdot 2\alpha)^N}{(8\,\pi^2)^N}$$

ϕ_r is the acceptance ratio for rotational steps.

The excess partition function becomes :

$$Q' = Q'_t \cdot Q'_r$$

As an example we will consider one simulation of chapter 3, and it is the simulation with D = 0.6, λ = 6. In that case we have fixed $v_m = (0.1\sigma)^3$ and α = 0.35. The resulting acceptance ratios were : ϕ_t = 0.30 and ϕ_r = 0.27. Furthermore N = 91 and $V = \frac{N\sigma^3}{D}$.
With these data we calculate $\ln Q'$ = -1832.
From the tables 2.3 and 3.5.2 we know that this quantity should be :

$$\ln Q' = -566 + (-2.042 \times 91) = -752.$$

The discrepancy is to be attributed to the fact that the free volume theory is based on the assumption that starting from one basic configuration (almost) all relevant configurations can be realised by shifting the particles inside the free volume.
Apparently that is not true and we conclude that if one would prefer this simple lattice theory, one should conceive of a tremendous number of 'basic configurations'.

That tremendous number can be evaluated with the multistage sampling method and in the above-mentioned case it appears to be :

$$\exp(-752 + 1832) = 10^{469} \text{ (that is about } 10^{5N}).$$

3.2.5.2. Accuracy

There are two approaches for estimating the accuracy of the results of the previous section: the evaluation of the repeatability and the check with pressure as mentioned in section 2.2.
Beginning with the latter, we can compare 10 of the 24 values of the pressure, evaluated in section 3.2.4 with the pressure that can be calculated from free energy with the formula :

$$P = - \left(\frac{\partial A}{\partial V}\right)_T \qquad (3.20)$$

The free energy A is evaluated from the results of table 3.5.1 with the aid of the formulae (A 1.3b) and (A 1.9).
The two alternative ways of calculating the pressure of the system are essentially independent since the determination with the aid of (3.20) involves a combination of many Monte Carlo simulations, whereas the calculation of the pressure with the virial according to appendix 4 involves the data from one single simulation.
The greater part of the information, that is used for the calculation of the pressure from free energy, is absent when the pressure is calculated with the virial.

From (3.20), (A 1.3b), (A 1.9) and the Carnahan and Starling formula (3.10) we can derive that :

$$\frac{PV}{NkT} = \frac{V}{N}\left(\frac{\partial \ln Z''_N}{\partial V}\right)_T + \frac{1 + y + y^2 - y^3}{(1-y)^3} \qquad (3.21)$$

(where $y = D \cdot \pi/6$).

We can approximate the derivative of Z''_N from the results of table 3.5.1 by stating :

$$V \cdot \left(\frac{\partial \ln Z''_N}{\partial V}\right)_T \approx \left(\frac{\Delta \ln Z''_N}{D\left[\frac{1}{D_2} - \frac{1}{D_1}\right]}\right)_\lambda \qquad (3.22)$$

(Note that $\left(\frac{1}{D_2} - \frac{1}{D_1}\right)$ equals $\left(\frac{1}{D_3} - \frac{1}{D_2}\right)$ and $\left(\frac{1}{D_4} - \frac{1}{D_3}\right)$)

In (3.22) $\Delta \ln Z''_N$, which is a function of density, is approximated For a given value of λ by fitting to a polynomial of the third degree:

$$\Delta \ln Z''_N(D_2) = \nu_1 . \ln Z''_N(D_1) + \nu_2 . \ln Z''_N(D_2) + \nu_3 . \ln Z''_N(D_3) + \nu_4 . \ln Z''_N(D_4) \qquad (3.23)$$

and

$$\Delta \ln Z''_N(D_3) = -\nu_4 . \ln Z''_N(D_1) - \nu_3 . \ln Z''_N(D_2) - \nu_2 . \ln Z''_N(D_3) - \nu_1 . \ln Z''_N(D_4) \qquad (3.24)$$

with $\nu_1 = -2/6$, $\nu_2 = -3/6$, $\nu_3 = 6/6$ and $\nu_4 = -1/6$, whereas D_1, D_2, D_3 and D_4 are the four densities, mentioned in table 3.5.1 and $Z''_N(D_1)$ etc. are the values of Z''_N of the same table with the relevant value of λ and D.

The quantities $\Delta \ln Z''_N(D_1)$ and $\Delta \ln Z''_N(D_4)$ are not approximated in this way since the differential of a curve can only be evaluated with some accuracy when there are data available on both sides of the reference point.
In fig. 3.14 we have plotted the ten resulting values of PV/NkT in comparison with those from section 3.2.4.

For $\lambda = 0$, $\ln Z''_N$ is always zero, without deviation. Therefore and because of the cumulative character of the problem we assume the inaccuracy to be proportional to λ.

If we calculate now the root mean square of the ten values of $\Delta p/\lambda$, the result is : $\sigma_p/\lambda = 0.029$ (Δp is the difference between pV/NkT from the free energy and from the virial and σ_p is the standard deviation of $\frac{PV}{NkT}$).

This root mean square deviation can be used to estimate the inaccuracy σ_z of $\ln Z''_N$ of table 3.5.1 as follows :

ΔP depends on the inaccuracy of both the virial and $\ln Z''_N$. We consider the worst case, namely that the difference is entirely to be ascribed to inaccuracy of $\ln Z''_N$.
Furthermore we assume the inaccuracy in $\ln Z''_N$, σ_z to be proportional to λ and independent of the density.

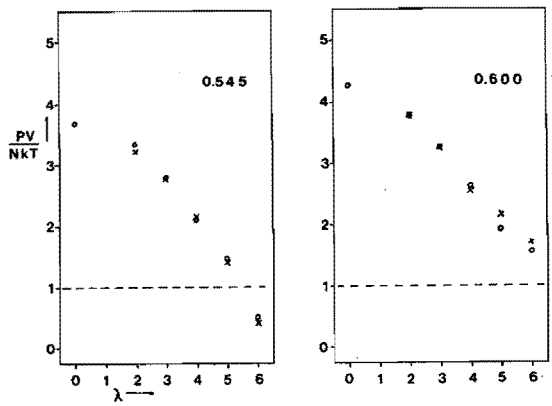


Fig. 3.14. Comparison of PV/NkT as calculated from the virial (circles) and from the free energy (crosslets). The density is indicated inside each frame.

Now we can calculate σ_z with (3.21), (3.22) and (3.23) :

$$\sigma_z = \sigma_p \cdot \frac{ND\left(\frac{1}{D_2} - \frac{1}{D_1}\right)}{\sqrt{\nu_1^2 + \nu_2^2 + \nu_3^2 + \nu_4^2}} \qquad (3.25)$$

The result is : $\sigma_z = 0.22\ \lambda$.

Alternatively we estimate the same inaccuracy from the repeatability. For that purpose we divided the data from each of the 24 Monte Carlo simulations into two groups.
One group consists of half the original number of Monte Carlo steps and by consequence the standard deviation in the resulting values of $\ln Z_N''$ is increased by a factor $\sqrt{2}$.
For the standard deviation of the difference between the corresponding $\ln Z_N''$ of the two groups again a factor $\sqrt{2}$ must be introduced.
From the root-mean square of the 20 resulting differences we could calculate the second approximation of σ_z :

$$\sigma_z = \frac{1}{\sqrt{2} \cdot \sqrt{2}} \times 0.56\ \lambda = 0.28\ \lambda \qquad (3.26)$$

We can conclude that the inaccuracy of $\ln Z_N''$ is relatively small. For instance if $\lambda = 6$ the magnitude of the inaccuracy in $\ln Z_N''$ is about 0.5 % and if $\lambda = 2$ about 2.5 %.

The repeatability procedure mentioned above yields not only data on the accuracy of $\ln\Xi''_N$, but in principle it can yield data on the accuracy of all other quantities involved.
For the quantity $\ln C_j$ the root mean square deviation appears to be of the same magnitude as for $\ln\Xi''_N$:

$$\sigma_c = 0.28\ \lambda$$

This can be understood since a relation like (2.11) suggests that there may be a strong relation between the inaccuracy of C_j and Ξ'_N (or Ξ''_N).

For the mean number of bonds as mentioned in table 3.2 the repeatability calculation has given the standard deviation :

$$\sigma_B = 0.22\ \lambda \qquad (3.27)$$

3.2.5.3. Fitting of the Data with an Analytical Expression.

Before dealing with the thermodynamic consequence of the results of table 3.5.2 we have fitted the data with analytical functions of the number of bonds j and the density D. That is necessary in order to make interpolations for calculation at other temperatures and densities than those of the Monte Carlo simulations and to make the evaluation of derivatives possible. The results of table 3.5.2 are composed of a set of 20 different functions of the type (2.22):

$$\gamma(j) = \ln C_j$$

In all instances where one is obliged to jump over from one curve to another there is a small kink, enough to make differentiation difficult.
The results of table 3.5.2 are those values that correspond as closely as possible to the evidence of the computer experiment, and fitting will strain somewhat this result. Therefore we should find an algorithm that represents this 'experimental evidence' as accurately as possible.
Representation of the curves with four polynomials that fit from j = 0 up to 182 (= 2N) or 167 (the highest number of j found) for the 4 densities, is not attractive. In order to get a reasonably

good fit it proved to be necessary to increase the degree of the polynomial to twelve. Apart from the fact that this is not very elegant, it is almost impossible to unify the four curves to a smooth function of density at constant j. So we must find an algorithm that fits better than a polynomial.

A major improvement is obtained when the data are compared with the binomial formula that applies for independent events. Division of all C_j by the binomial factor :

$$\binom{2N}{j} \equiv \frac{(2N)!}{j!\,(2N-j)!} \tag{3.28}$$

yields a new function that is only slightly curved (see fig. 3.15). We define the new function γ of the real variable j in such a way that for integer values of j :

$$\gamma(j) = \ln C_j - \ln\binom{2N}{j} \tag{3.29}$$

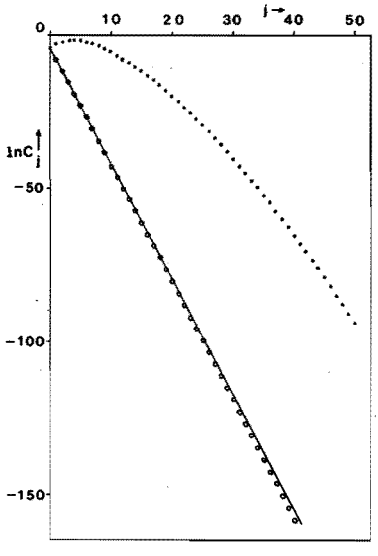


Fig. 3.15. The dots represent $\ln C_j$ as calculated in section 3.2.5.1. The circles represent the values of $\ln C_j - \ln\binom{2N}{j}$. If C_j would satisfy exactly the binomial distribution, the circles would coincide with the full line.

Since the more direct information from Monte Carlo simulations yields the derivative of this curve we have assembled in fig. 3.16 the apparent values of these derivatives :

$$\frac{d\gamma}{dj} \approx \ln PA_{j+1} - \ln PA_j + \ln\left(\frac{j+1}{2N-j}\right) - \lambda \qquad (3.30)$$

PA_j represents the apparent probability as obtained from a Monte Carlo simulation.

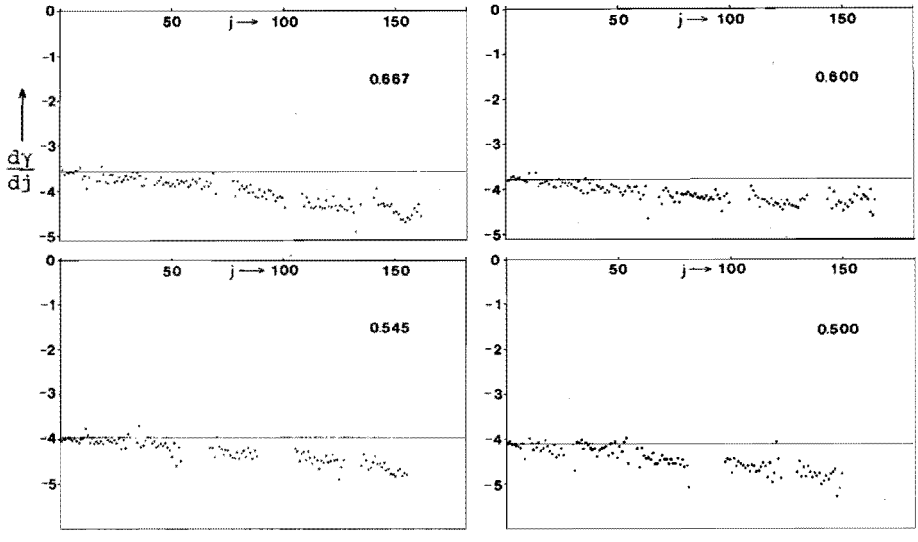


Fig. 3.16. The derivative of the function γ, estimated from the Monte Carlo simulations after equation (3.30) (dots). The horizontal line gives the corresponding derivative of the binomial distribution. The density is printed inside each frame.

As can be seen from fig. 3.16 the function is only slightly curved and by consequence the equation should be rather simple. We used the formula :

$$\gamma(D,j) = x_0 + x_1.\Delta j + x_2.\Delta j^2 + x_3.\Delta j^p \qquad (3.31)$$

Where the coefficients x_i are functions of density. The parameter Δj is the difference between the present number of bonds j and the mean number of 'bonds' in a hard-spheres system j_m :

$$\Delta j = j - j_m \qquad (3.32)$$

The first and second coefficient of (3.31), x_0 and x_1 can be evaluated rather accurately from the requirement that for $\lambda = 0$ the system is identical with the corresponding hard spheres system in the following way.

From the distribution function of a hard spheres fluid the mean number of neighbours with the distance within the limits of bonding (eq. 3.1) can be obtained. In our case we made six separate Monte Carlo simulations with 91 hard spheres and the densities 0.462, 0.500, 0.545, 0.600, 0.667 and 0.750 respectively. We counted the mean number ν of neighbours with a distance from the reference particle lying between σ and $f \cdot \sigma$. The resulting data were fitted into the polynomial :

$$\nu = a_0 + a_1.D + a_2.D^2 + a_3.D^3 + a_4.D^4 \qquad (3.33)$$

The coefficients, a_0 and a_1 are evaluated analytically, whereas a_2, a_3 and a_4 are fitted with least squares [2.3] to the Monte Carlo data.
In table 3.6 the results are given.

Table 3.6. The numerical values of the coefficients of eq (3.33) for the calculation of the mean numbers of neighbours with distance between σ and 1.1σ in a hard-spheres fluid.

a_0	0.0
a_1	$+ 1.3865 = \frac{4\pi}{3} (1.1^3-1)$
a_2	+ 3.080
a_3	- 3.571
a_4	+ 5.441

We can take the view that the orientation is completely random since we are dealing with hard spheres. Accordingly the chance for one of two neighbours to meet the orientational condition can be evaluated analytically :

$$\beta = 2(1 - \cos\psi_{max}) \qquad (3.34)$$

By consequence the chance of two neighbours to be both oriented correctly is β^2.

In the system there are $\frac{1}{2}N.\nu$ neighbours and the mean number of bonds is:

$$j_m = \tfrac{1}{2}N.\nu.\beta^2 \tag{3.35}$$

If this is identified with the number of bonds pertaining to the binomial distribution:

$$BC_j \equiv \binom{2N}{j} . (1-\alpha)^{2N-j}.\alpha^j \tag{3.36}$$

then α should fulfill the equality $j_m = 2N\alpha$ and by consequence:

$$\alpha = \tfrac{1}{4}\nu\beta^2 \tag{3.37}$$

If

$$x_0 = j_m . \ln\alpha + (2N-j_m)\ln(1-\alpha) \tag{3.38}$$

and

$$x_1 = \ln(\alpha/(1-\alpha)) \tag{3.39}$$

the requirement is fulfilled that for $\lambda = 0$ and by consequence Δj is small, equation (3.31) is in accordance with (3.36), the binomial distribution.

The binomial distribution applies for independent stochastic events. If $\lambda = 0$ the orientation of the molecules is random and the number of bonds is so small that the probability of a bond between neighbours is almost independent of the existence of others.
The algorithm of (3.33) up to (3.39) leads to a set of values for the quantities of ν, j_m, α, x_0 and x_1, which are gathered in table 3.7.

When the data of the Monte Carlo simulations with tetrahedral molecules would be in accordance with the binomial distribution all dots in fig. 3.16 would coincide with the horizontal line shown (apart from inevitable statistical fluctuations).

Table 3.7 The value of the parameter ν, j_m, α and the coefficients x_0 and x_1 as a function of the density, calculated with the equations (3.33) up to (3.39)

D	ν	j_m	α	x_0	x_1
0.500	1.357	2.934	0.01612	- 15.021	- 4.1114
0.545	1.575	3.405	0.01871	- 16.920	- 3.9599
0.600	1.875	4.053	0.02227	- 19.428	- 3.7820
0.667	2.310	4.994	0.02744	- 22.884	- 3.5678

Apparently this is not the case. The presence of bonds appears to decrease the possibility to form new bonds. This effect is accounted for with the last two terms of (3.31).
The coefficient x_2 determines the inclination of the curves through the dots of fig. 3.16 at low and intermediate Δj.
x_2 is a function of density and can be represented by :

$$x_2 = b_1 + b_2 D + b' D^2 \qquad (3.40)$$

It appears to be possible to eliminate the coefficient b' by the proper choice of the power p of parameter Δj in the last term of (3.31).
In fig. 3.17 it is shown that b' changes rapidly with p. When p = 4.50 the value of b' is zero, whereas at the same time the corresponding residue of the least squares procedure is at its minimum.

This proves that the curvature of the curves through the dots of fig. 3.16 is correctly accounted for when p = 4.5. The result is that the variable quantity b' is eliminated by a proper choice of another one, p, with a resultant improvement of the fit.

The corresponding values of b_1 and b_2 are mentioned in table 3.8.
The above-mentioned exponent p being fixed, x_3 determines the curvature of the curves in fig. 3.16 for large Δj. The curves are bending differently for the four densities. For D = 0.667 and D = 0.500 the curves bend downwards and for D = 0.600 the curve bends upwards.

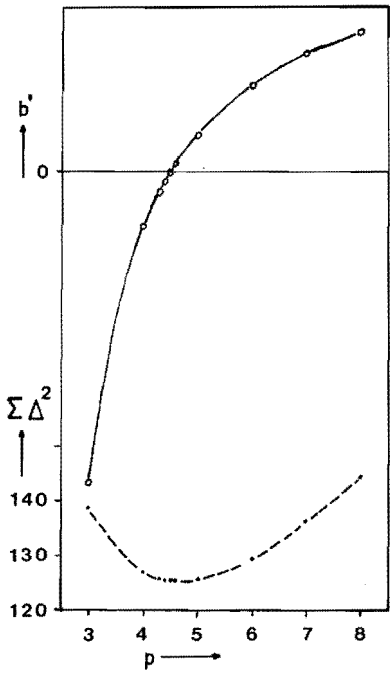


Fig. 3.17. The coefficient b' (upper curve) and the least squares residue (lower curve) resulting from a set of least squares calculations with variable exponent p of eq. (3.31). Best fit is obtained when p=4.50 and b'=0

Apparently the possibility of increasing the number of bonds in a highly bonded system is greater for D = 0.600 than for the other densities. This effect is caused by the previously discussed effect, arising when the density is favourable for the establishment of structures with more or less tetrahedral coordination of the molecules. (For instance a diamond structure would lead to D = 0.56, just like the tridimyte or ice I structure. The ice polymorphs would lead to higher values of D. Ice II : D = 0.70 and ice III : D = 0.68).
By consequence the experiments show that the formula for the evaluation of x_3 must be designed in such a way that x_3 is negative for high and low density and positive for intermediate density.
We obtained the best results with the formula:

$$x_3 = b_3 + b_4.D + b_5 \sqrt{(D-b_6)^2 + b_7^2} \qquad (3.41)$$

which represents an oblique hyperbola (fig. 3.18).
Of all curves we considered this one yielded the lowest value of the least squares residue :

parabola	→ residue	=	170.0
gaussian	→ "	=	145.2
hyperbola with vertical axis	→ "	=	148.8
oblique hyperbola	→ "	=	125.3

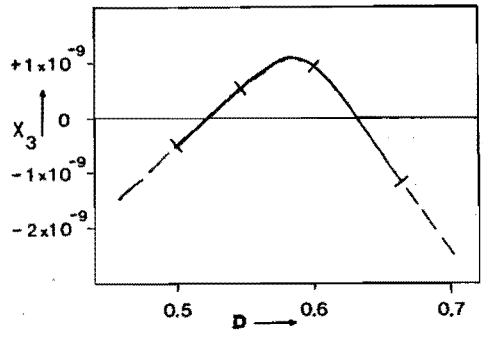


Fig. 3.18. The coefficient x_3 as a function of the density (calculated with the aid of eq. (3.41) and the numerical values of table 3.8). The marks indicate the densities of the Monte Carlo simulations.

Equation (3.41) has five adjustable coefficients whereas there are only four different values of D. One can fix b_7 and adjust the coefficients b_1 up to b_6 with a least squares program [3.7] in such a way that γ from (3.31) is fitted best to the values of $\ln C_j$ as found in section 3.2.5.1.

The value of b_7 can be chosen between 0.0 and 0.08 without the slightest effect on the result. Therefore b_7 has to be chosen on the basis of other evidence. The pressure appears to be very sensitive for the value of b_7. When $b_7 = 0.0$, there is a discontinuity at $D = b_6$ and by consequence there is a jump in the pressure at that point. And by consequence a steep fall of pressure when b_7 is low. When b_7 is high the value of the pressure at high density and low temperature, especially for $D = 0.667$ and $\lambda = 6.0$, becomes very high. Since we know from the virial (section 3.2.4) what value the pressure should have approximately we can conclude that b_7 should not be too high. Finally we decided on the value of $b_7 = 0.02$, which is an acceptable compromise. The resulting values of b_1 up to b_7 are given in table 3.8.

Table 3.8. The value of the coefficients of eq. (3.40) and (3.41) as obtained from least squares fitting.

b_1	$- 9.8844 \times 10^{-4}$
b_2	$- 2.4250 \times 10^{-3}$
b'	0.0
b_3	$+ 5.7040 \times 10^{-9}$
b_4	$- 6.8174 \times 10^{-9}$
b_5	$- 30.4404 \times 10^{-9}$
b_6	+ 0.5896
b_7	+ 0.02
p	+ 4.50

Finally we will give a summary of the formulae for the evaluation of the free energy as we have developed in this section :

$$A = -kT\ln Q$$

$$Q = Q_i \cdot \Xi'_{HS} \cdot \Xi''_N$$

Q_i = the partition function of the ideal gas.

$$\ln \Xi'_{HS} = N \frac{3y^2 - 4y}{(1 - y)^2}$$

$$y = D \cdot \pi/6$$

$$D = N\sigma^3/V$$

$$\Xi''_N = \sum_{j=0}^{2N} C_j \exp(-j\,\varepsilon/kT)$$

$$C_j = \binom{2N}{j} \exp(\gamma)$$

$$\gamma = x_0 + x_1 \Delta j + x_2 \Delta j^2 + x_3 \Delta j^{4.5}$$

$$\Delta j = j - \tfrac{1}{2} N \nu \beta^2$$

$$\beta = 2(1 - \cos\psi_{max}) \quad , \quad \psi_{max} = 27^{o}$$

$$\nu = a_0 + a_1 D + a_2 D^2 + a_3 D^3 + a_4 D^4$$

a_0, a_1, a_2, a_3, a_4 see table 3.6

$$x_0 = 2N\,(\alpha \ln\alpha + (1-\alpha)\ln(1-\alpha))$$

$$x_1 = \ln(\alpha/(1-\alpha))$$

$$x_2 = b_1 + b_2 D$$

$$x_3 = b_3 + b_4 D + b_5 \sqrt{(D-b_6)^2 + b_7^2}$$

b_1, ..., b_7 see table 3.8

3.2.5.4. Review of the Thermodynamic Properties of a System with Tetrahedral Molecules.

With the algorithm of the previous section the fraction of configuration space C_j can be calculated for any value of the number of bonds j and those values of the density lying within the interval covered by the original Monte Carlo simulations. Accordingly the equilibrium thermodynamic properties can be derived.

For instance the partition function is calculated with the combination of (A1.3b), (A1.7b) and (A1.9)

$$Q = Q_i \cdot Z'_{HS} \cdot Z''_N \qquad (3.42)$$

In this way the partition function is the product of three factors:

Firstly Z''_N, the contribution of bonding, which can be evaluated from C_j :

$$Z''_N = \sum C_j \cdot \exp(-j.\varepsilon/kT) \qquad (3.43)$$

Secondly Z'_{HS}, the factor originating from the fact that all configurations with overlapping molecules are excluded. This factor is given by the Carnahan and Starling equation :

$$\ln Z'_{HS} = N \cdot \left(\frac{3y^2 - 4y}{(1-y)^2} \right)$$

$(y = D.\pi/6)$

Finally Q_i is the partition function of the ideal gas.
Likewise for the other thermodynamic properties we can speak of a bonding contribution, a hard-spheres contribution and an ideal-gas contribution. For the energy for instance is the bonding contribution $\varepsilon.j$, the hard-spheres contribution is zero and the ideal-gas contribution is 3NkT.

Just as in section 3.2.4 we will express the thermodynamic variables in reduced form, i.e. in dimensionless units.
The reduction formulae are :

$$(3.11)\quad \begin{cases} P^* = P \cdot \sigma^3/|\varepsilon| \\ V^* = V/(N \cdot \sigma^3) \\ T^* = k \cdot T/|\varepsilon| \end{cases}$$

and

$$\left.\begin{aligned} U^* &= U/(N \cdot |\varepsilon|) \\ A^* &= A/(N \cdot |\varepsilon|) \\ S^* &= S/(N \cdot k) \end{aligned}\right\} \qquad (3.44)$$

The effect of the reduction of the variables is that some results become independent of the value of ε, σ and N. In other cases, however, there remains a slight effect of the value of these parameters on the resulting thermodynamic properties.
Particularly the calculation of the ideal-gas contribution to some variables (the partition function, the free energy en the entropy), involves the non-reduced volume and temperature. For that reason it is necessary to attach some value to the bonding energy ε and to the diameter σ of the molecules.

In accordance with X-ray evidence for physical water [3.19] we have adopted $\sigma = 2.7 \times 10^{-10}$ m. The energy of the hydrogen bond in physical water is estimated differently by different authors [3.20]. We adopted $\varepsilon = 1800 \cdot k$ (k is the Boltzmann constant) which corresponds with 15 kJ/mol. With this choice of ε we achieved a reasonable accordance between the model and physical water as far as the temperature of the inversion of the thermal expansion is concerned. (see chapter 4).

The effect of the fixation of the value of σ and ε is that the reduced variables can be related with the non-reduced ones. For better understanding of the figures of this and the next sections we have tabulated in table 3.9.1, 3.9.2 and 3.9.3 some examples of this relationship.

Table 3.9.1 The reduced temperature compared with the experimental scales. The value of T* correspond with the values in the figures 3.19 up to 3.30 and 4.2 up to 4.5.

T*	Kelvin	centigrades
0.133	240	-33
0.150	270	- 3
0.152	273	0
0.163	293	20
0.167	300	27
0.196	353	80
0.200	360	87
0.250	450	177
0.252	453	180
0.330	593	320
0.333	600	327
0.500	900	627

Table 3.9.2 The relation between some values of the reduced volume and the molar volume.

V*	$cm^3 mol^{-1}$
1.4	16.60
1.6	18.97
1.8	21.34
2.0	23.71

Table 3.9.3 The relation between some values of the reduced pressure and the pressure in atm.

p*	atm.
0.0	0
0.1	1263
0.2	2525
0.3	3788
0.4	5050
0.5	6313

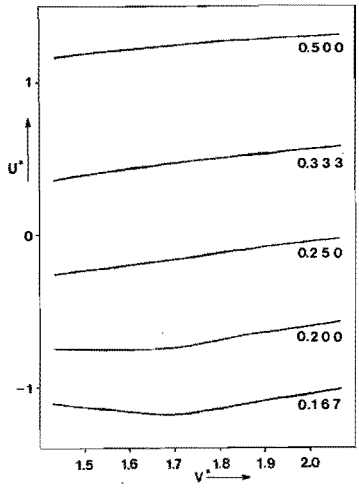


Fig. 3.19. The internal energy of a system with tetrahedral molecules, plotted as a function of the volume. The temperature is indicated next to the curves. All quantities are in reduced vorm.

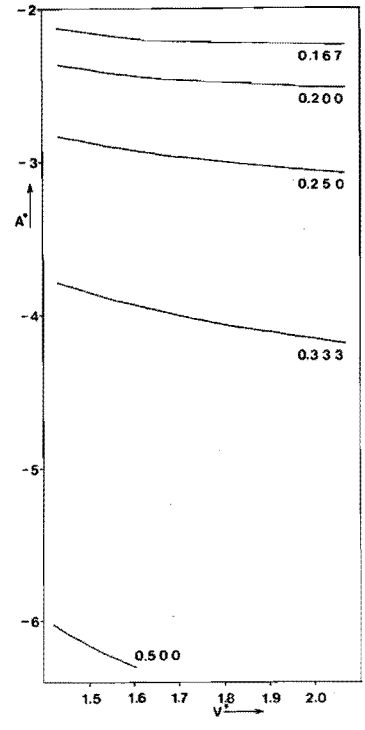


Fig. 3.20. The free energy as a function of the volume. The temperature is indicated next to the curves. All quantities are in reduced form.

With the approximation of section 3.2.5.3 we have evaluated the energy, the free energy, the entropy and the equation of state of the present system with tetrahedral molecules.
The results have been plotted in the figures 3.19, 3.20, 3.21 and 3.22.

Although the figures are self evident we will discuss just a few details. The energy exhibits a minimum at the lowest temperature ($T^* = 0.167$). This is due to the minimum in the potential energy which is discussed in section 3.2.2.

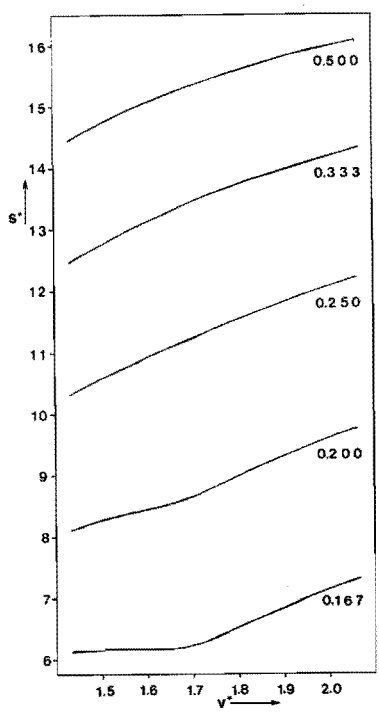


Fig. 3.21. The entropy as a function of the volume. The temperature is indicated next to the curves. All quantities are in reduced form.

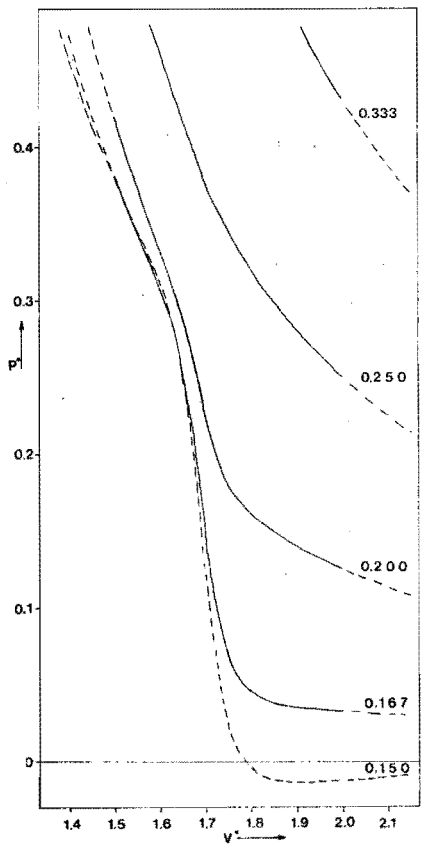


Fig. 3.22. P - V* curves for different temperatures. Broken curves are extrapolations outside the region of the original Monte Carlo simulations. T* is indicated next to the curves.*

Another noteworthy detail is the fact that the entropy has a horizontal section from about V* = 1.50 up to 1.63 at T* = 0.167. Close examination of the numeric material shows that there is even a weak maximum at V* = 1.57 and a minimum at V* = 1.62. That means that $\left(\frac{\partial S}{\partial V}\right)_T$ changes from positive to negative and back.

Since $\left(\frac{\partial S}{\partial V}\right)_T = \left(\frac{\partial P}{\partial T}\right)_V$ (3.45)

there is inversion of expansive properties in that region.

It can also be observed in the P V T diagram, fig. 3.22, that the isotherms of T* = 0.167 and of T* = 0.150 intersect. When at a constant pressure P* = 0.30 the temperature is decreased from T* = 0.20 to T* = 0.167 the volume decreases as usual. However, when the temperature is decreased still further until T* = 0.15 the volume appears to increase. This implies an inversion of the thermal expansion.

The equation of state, however, is very sensitive to the details of the approximation procedure in section 3.2.5.3. Still we are convinced that the inversion of the thermal expansion observed is not an artifact. During our investigations we have applied various alternative approximations for the coefficient x_3 (see section 3.2.5.3). Although the equation of state was distinctly different in all cases the inversion of the thermal expansion was always present at the same value of P* and V* (see fig. 3.23).

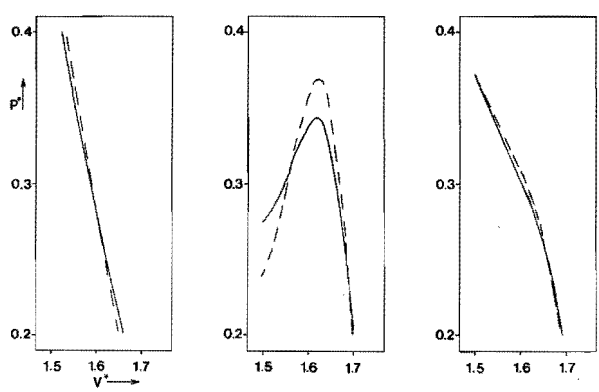


Fig. 3.23. P - V* curves for two temperatures (T*=0.167 : full line and T*=0.15 : broken line) calculated with alternative algorithms. It should be noted that, in all cases the lines intersect, even when the shape of the curves is strongly deviating, as in the middle frame where (artificially) two liquid phases occur. Left: parabolic approximation for coefficient x_3, middle: gaussian approximation and right: hyperbolic approximation (non-oblique).*

3.3 . Polar Molecules

The tetrahedral molecule model, as described in the previous chapter can be transformed into a polar model.
To that purpose a sign is associated with each of the four vectors of the molecule. Two positive and two negative, see fig.3.24.
The conditions for the formation of a bond are not changed, but the sign of the bonding energy depends on the sign of the two vectors involved.
The pair potential between two polar molecules k and l respectively now becomes :

$$\left.\begin{array}{lll} E_2 = +\infty & & \text{if } r_{kl} < \sigma \\ E_2 = -\text{sign}(k)\,\text{sign}(l) \cdot \varepsilon & & \text{if } \sigma \leq r_{kl} < f \cdot \sigma \\ & & \text{and } |\psi_k| < \psi_{max} \\ & & \text{and } |\psi_l| < \psi_{max} \\ E_2 = 0 \quad \text{in all other cases} & & \end{array}\right\} \quad (3.46)$$

Sign (k) and sign (l) are equal to +1 or -1, depending on the sign associated with the vector involved.
It should be noted that this potential accounts only for short range interactions. Long range interaction as exists for physical molecules [3.8] does not occur since for $r_{kl} > 1.1\sigma$ there is no interaction whatsoever between our molecules.

The above-mentioned transformation has two consequences.
Firstly the symmetry of the molecule decreases and by consequence the multiplicity factor η in the formula for the partition function (Appendix 1 formula (A 1.3b) decreases likewise. As mentioned in appendix 1 the value of η in the symmetrical model is 12, whereas in the polar model $\eta = 2$.

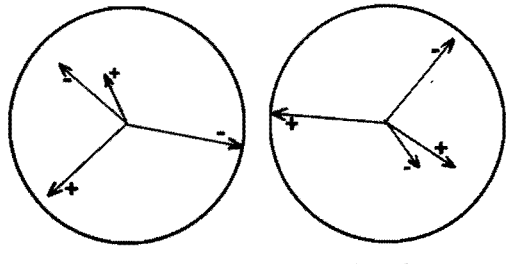


Fig. 3.24 Two 'polar' molecules.

By consequence in an N-particle system the number of physically distinguishable configurations is increased by a factor 6^N on transformation from non polar to polar.
Secondly the (potential) energy is changed on transformation When in a system with symmetric molecules the number of bonds is j, then the energy is $j \cdot \varepsilon$. After transformation there is for each bond the possibility of an energy contribution of $+\varepsilon$ or $-\varepsilon$.
The total energy of the system is between $j \cdot \varepsilon$ and $-j \cdot \varepsilon$.

When for instance a system consists of two non-polar molecules with one bond, a single configuration results in 36 new configurations after transformation into polar molecules.
Exactly 50 percent of the new configurations has an energy $+\varepsilon$, the other 50 percent has an energy of $-\varepsilon$. If in a system with more bonds we suppose the chances for the bonding energies of all bonds to be positive or negative are equal, the distribution of the possibilities is given by a binomial distribution. In an N-particle system with j bonds the transformation increases the configuration space with a factor 6^N. Of these possible configurations a fraction

$$fr = \binom{j}{q} \cdot \frac{1}{2^j} \tag{3.47}$$

has q 'antibonds' with energy $-\varepsilon$ and j-q bonds with energy ε.

The energy of the system is in that case $(j - 2q)\varepsilon$. Accordingly we will use $i \equiv j - 2q$ as the integer that indicates the energy level.

With the help of eq. (3.47) we can calculate now the distribution of the configuration space with respect to the energy of the system.

We will start with the 2N+1 compartments of the configuration space in the non-polar case, as evaluated in the previous sections. For all possible values of q, C_j as obtained with the approximation of section 3.2.5.3, is multiplied by the relevant fraction (3.47). The resulting values are properly combined for summation as shown in fig. 3.25.

j \ i	....	-3	-2	-1	0	+1	+2	+3	....
0					C_0				
1				$\frac{1}{2}\cdot C_1$		$\frac{1}{2}\cdot C_1$			
2			$\binom{2}{2}\cdot\frac{1}{2^2}\cdot C_2$		$\binom{2}{1}\cdot\frac{1}{2^2}\cdot C_2$		$\binom{2}{0}\cdot\frac{1}{2^2}\cdot C_2$		
3		$\binom{3}{3}\cdot\frac{1}{2^3}\cdot C_3$		$\binom{3}{2}\cdot\frac{1}{2^3}\cdot C_3$		$\binom{3}{1}\cdot\frac{1}{2^3}\cdot C_3$		$\binom{3}{0}\cdot\frac{1}{2^3}\cdot C_3$	
4	....		$\binom{4}{3}\cdot\frac{1}{2^4}\cdot C_4$		$\binom{4}{2}\cdot\frac{1}{2^4}\cdot C_4$		$\binom{4}{1}\cdot\frac{1}{2^4}\cdot C_4$		....
5		$\binom{5}{4}\cdot\frac{1}{2^5}\cdot C_5$		$\binom{5}{3}\cdot\frac{1}{2^5}\cdot C_5$		$\binom{5}{2}\cdot\frac{1}{2^5}\cdot C_5$		$\binom{5}{1}\cdot\frac{1}{2^5}\cdot C_5$	
⋮	....	....	....	....	....	....	....	....	....

Fig. 3.25. Survey of the transformation procedure. The fractions of the configuration space C_j are divided into smaller fractions, which are placed in rows. The polar equivalence CP_i is obtained by summation of the columns.

The result is :

$$CP_i = \sum_{q=0}^{N-i/2} C_j \cdot \binom{j}{q} \cdot \frac{1}{2^j} \qquad (3.48)$$

where $j = i + 2q$ and q is integer.
With this equation we performed the transformation from non polar into polar.

In fig. 3.26 the effect of the transformation on the partition of the configuration space can be seen.
In table 3.10 some values of CP_i are given. This table is to be compared with table 3.5.2. However, it should be kept in mind that the data of 3.5.2 have been obtained directly from the Monte Carlo simulations, whereas the data of 3.10 are calculated from the values of C_j as obtained with the approximation of section 3.2.5.3.

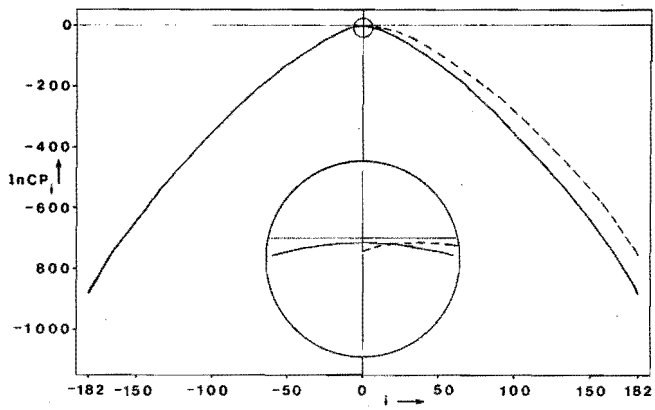


Fig. 3.26. The effect of the transformation of a system (with N=91) from non-polar to polar molecules. The broken curve gives the logarithm of the fraction of the configuration space in the non-polar case (lnC_j), the full curve gives the corresponding $lnCP_i$ for the polar case. The central part of the figure, magnified ten times, is plotted in the lower circle.

It is easy to verify that for the larger values of j (above about twenty) in the summation of (3.48) the one term with q = 0 outweighs all others, so that i = j and :

$$CP_i \approx \frac{C_j}{2^j} \tag{3.49}$$

From (3.18) can be concluded that for non-polar molecules

$$\ln P_{j+1} - \ln P_j = \ln C_{j+1} - \ln C_j + \lambda$$

For the maximum probabilities is $P_j \approx P_{j+1}$ and by consequence

$$\ln C_{j+1} - \ln C_j \approx -\lambda \tag{3.50}$$

Likewise for polar molecules :

$$\ln CP_{i+1} - \ln CP_i \approx -\lambda' \tag{3.51}$$

If now i = j, the combination of (3.49), (3.50) and (3.51) results in :

$$\lambda' \approx \lambda + \ln 2 \tag{3.52}$$

Table 3.10 $\ln CP_i$, the logarithm of the fraction of configuration space at energy level i. For negative i is $CP_i = CP_{-i}$.

i \ D	0.6667	0.600	0.545	0.500
0	- 1.70	- 1.59	- 1.49	- 1.41
10	- 10.56	- 11.93	- 13.15	- 14.26
20	- 29.99	- 33.42	- 36.37	- 38.95
30	- 55.47	- 60.98	- 65.67	- 69.73
40	- 85.42	- 92.99	- 99.40	- 104.93
50	- 119.09	- 128.67	- 136.78	- 143.77
60	- 156.04	- 167.55	- 177.35	- 185.81
70	- 196.03	- 209.36	- 220.84	- 230.79
80	- 238.92	- 253.92	- 267.08	- 278.54
90	- 284.66	- 301.12	- 315.95	- 329.00
100	- 333.23	- 350.89	- 367.42	- 382.14
110	- 384.71	- 403.23	- 421.49	- 438.01
120	- 439.23	- 458.18	- 478.22	- 496.70
130	- 496.99	- 515.84	- 537.73	- 558.39
140	- 558.30	- 576.42	- 600.24	- 623.34
150	- 623.64	- 640.26	- 666.12	- 691.99
160	- 693.81	- 708.00	- 736.05	- 765.08
170	- 770.31	- 781.01	- 811.42	- 844.08
180	- 857.81	- 863.77	- 896.74	- 933.90
182	- 878.70	- 883.54	- 917.05	- 954.81

In order to verify this we calculated the mean number of bonds for the polar case with $\lambda' = \lambda + \ln 2$ and compared the results obtained with the results of a calculation for non-polar molecules and λ. The comparison is given in table 3.12. The accordance appears to be very good.

Evidently the most probable configurations, occuring in a system with non-polar molecules and parameter λ correspond with the most probable configurations in a system with polar molecules and parameter $\lambda' = \lambda + \ln 2$.

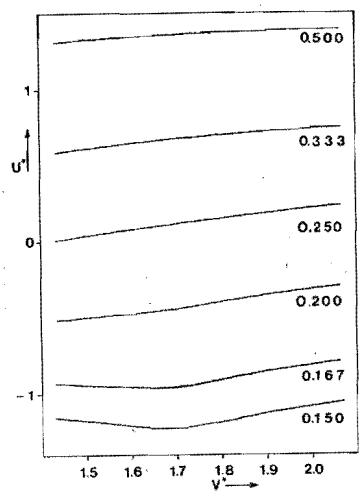


Fig. 3.27. The energy as a function of the volume for a system with polar molecules. The temperature is indicated next to each curve. All quantities are in reduced form.

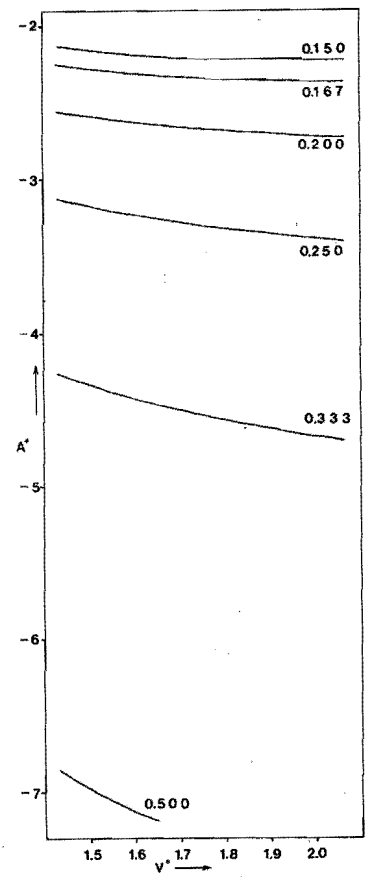


Fig. 3.28. The free energy as a function of the volume (polar molecules). The temperature is indicated next to each curve . All quantities are in reduced form.

Consequently some results reported in chapter 3.2 for non-polar molecules apply also for polar molecules when λ is increased with ln2. This is the case for the mean number of bonds, as mentioned above, the distribution functions and the configuration integral Z_N'' ,but not for the thermodynamic functions, U*, A*, S* and P*, since the ideal gas contribution is different for λ and λ'. Therefore we have evaluated the variables U*, A*, S* and P* using the values of CP_i mentioned before. The resulting functions have been plotted in the figures 3.27 up to 3.30. These figures should be compared with 3.19 up to 3.22.

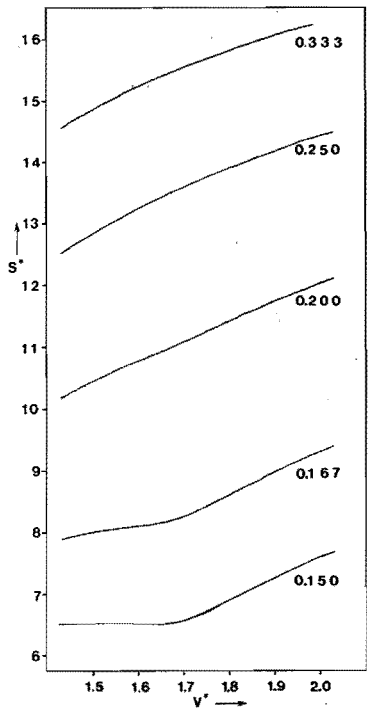


Fig. 3.29. The entropy as a function of the volume (polar molecules). The temperature is indicated next to the curves. All quantities are in reduced form.

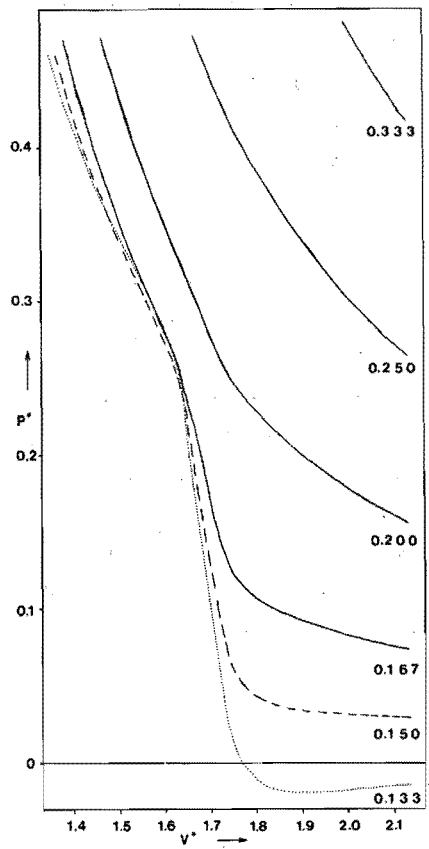


Fig. 3.30. P-V*diagram of a system with polar molecules. T* is indicated next to the curves.*

Table 3.12 The mean number of bonds in a 91 particles system with polar molecules. In brackets the difference with the corresponding non-polar case with $\lambda = \lambda' - \ln 2$.

λ' \ D	0.667	0.600	0.545	0.500
2.693	28.2(-0.1618)	23.9(-0.1334)	20.8(-0.1132)	18.4(-0.0982)
3.693	55.0(-0.0486)	48.9(-0.0417)	43.8(-0.0364)	39.7(-0.0323)
4.693	89.2(-0.0120)	84.0(-0.0109)	77.7(-0.0099)	72.1(-0.0091)
5.693	122.4(-0.0024)	122.3(-0.0024)	115.7(-0.0022)	108.8(-0.0021)
6.693	148.4(-0.0004)	152.9(-0.0004)	147.6(-0.0004)	140.4(-0.0004)

The present analytical transformation of the partition of the configuration space of a system with non-polar molecules into the same partition for polar molecules is a very efficient procedure. It demands about 0.5 minute of computer time. That is about 1/16000 of the time a complete set of Monte Carlo simulations necessary for the evaluation of CP_i with multistage sampling would demand.

At this point the reader may have still one reservation: the transformation operation, notably eq. (3.47) is based upon the supposition that the chances for the bonding energies of the bonds in a system to be positive or negative are independent of each other. It is clear that especially in systems with many bonds this cannot be strictly the case.
Fortunately there is evidence that the effect of interdependency between the bonds is very small.
Pauling [3.9] stated in 1935 that ice should exhibit a residual entropy at 0 K. He estimated this entropy to be given by the equation

$$S = Nk\ln(6/4) \tag{3.53}$$

At the same time Giauque et al. [3.10] confirmed this experimentally. The Pauling formula would lead to a residual entropy of 0.806 cal/K.mol, Giauque and Stout found by accurate measurements of the specific heat of ice a value of 0.82 ± 0.015 cal/K.mol.
This supports the suppositions underlying Paulings' formula, which suppositions are exactly the same as those we used as a starting point for the above discussed transformation.
(Since in a crystalline substance at 0 K only one energy level exists it can easily be verified that Paulings formula (3.53) results from the combination of (3.48) (with i = 2N) and the factor 6^N originating from the symmetry of the water molecule).

Later evidence is even more convincing. Onsager and Dupuis [3.11] stated in 1960 that the above argument of Pauling is not full proof. They put forward the 'interlacing effect' that increases the number of possible configurations, which might be compensated by some ordering of the bonds within the ice crystal at 0 K. So there could

be an unknown degree of order without discrepancy between (3.53) and the experimental data. Thus Onsager and Dupuis emphasized the importance of the correct evaluation of the interlacing effect.

The interlacing effect originates from the existence of rings of bonded molecules. The magnitude of the fraction with $q = 0$ as calculated with (3.47) would increase with a factor $(1 + 1/3^n)$ as a result of the presence of one n-membered ring in the system.
In an ice I crystal with N molecules a number of 2N six-membered rings is present. These rings are strongly connected, which makes the problem to be an intricate mathematical puzzle.
J.F. Nagle [3.12] succeeded in evaluating the effect in 1966. His conclusion is that the interlacing effect is small indeed, the residual entropy is raised by only one percent.

Consequently the above-mentioned transformation procedure can safely be used. The more so as in liquids the number of bonds is smaller than in ice which results in a less interlaced structure.

If necessary the partition of the configuration space can also be transformed with the method of section 2.4 into the corresponding partition for an other pair potential. It should be noted, however, that this method will demand a considerable amount of computer time albeit less than a complete new set of simulations would require.

4. DISCUSSION

4.1. Comparison with Hard Spheres and with a Square-Well Fluid

In the previous chapter we have dealt with a system of molecules with tetrahedral interaction and we have evaluated a part of the configuration space of such a system with the method of multistage sampling. The configuration space appears to be evaluated with reasonable accuracy, sufficient for the evaluation of the free energy, the entropy, etc. Besides the pressure, the energy and radial distribution functions are obtained which makes the picture of the equilibrium thermodynamic properties complete in the region of temperatures and densities investigated.

We can compare the resulting equation of state of the present system with a hard spheres system on the one hand and a system with spherical square-well molecules on the other.
The spherical square-well molecules are defined by the pair potential :

$$\left.\begin{array}{lll} E_2 = +\infty & \text{if} & r < \sigma \\ E_2 = \varepsilon & \text{if} & \sigma \leq r < f \cdot \sigma \\ E_2 = 0 & \text{if} & r \geq f \cdot \sigma \end{array}\right\} \qquad (4.1)$$

In accordance with eq. (3.1) is $f = 1.1$ and ε is negative.

From a set of special Monte Carlo simulations of hard-spheres systems with high density ($D = 0.75$, 1.00 and 1.087) we evaluated the coefficients, necessary for the estimation of PV/NkT with the help of the second order perturbation theory of Barker and Henderson [4.1]. In fig. 4.1 the resulting values are plotted for $\lambda = 2.0$. The value of PV/NkT appears to be appreciably lower than in the case of hard spheres.

As can be seen in fig. 4.1 the restriction of the attraction between pairs of molecules to small regions (tetrahedrally arranged according to (3.1)) decreases strongly the effect of the attraction between molecules. As could be expected.
Increasing λ for the tetrahedral model from 2 to 4 causes a decrease of PV/NkT. Above $\lambda = 4$ the situation becomes more complicated.

That is caused by the geometric effect that tetrahedrally surrounded molecules are more probable at certain densities (D is about 0.6) than at other densities. (The effect is discussed in chapter 3. It appears to affect the number of bonds, the partition of the configuration space and the thermal expansion at constant pressure).

The conclusion that can be drawn from fig. 4.1 is that the volume of a liquid with a considerable tetrahedral interaction between the molecules tends to be 50 - 80 % larger than the volume of the corresponding liquid with spherical molecules.

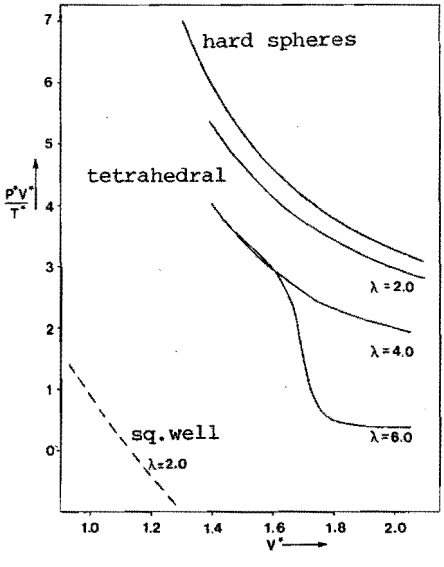


*Fig. 4.1 The parameter P*V*/T* as a function of the volume. The upper curve gives the data of the hard spheres fluid as calculated with the formula of Carnahan and Starling (3.16). The lowest curve pertains to a system with molecules with spherically symmetric interaction (square well molecules). The other curves are for the non-polar model. The relevant values of the parameter λ≡1/T* is indicated next to each curve.*

4.2. Comparison with Physical Water

4.2.1 The Equation of State

After comparison of the non-polar tetrahedral model with the hard-spheres model and with the spherical square-well model we can compare the properties of the model with the properties of a physical liquid. Different liquids deserve consideration because of the tetrahedral properties of the molecules or atoms : water, molten germanium, silicon, silica etc.

For comparison with the model we have chosen water in accordance with the starting point of this work. This choice implies that the model system that is considered for the comparison with the physical liquid must be the polar system of section 3.3.

In the next section we will pay attention to some molecular models for physical water that have been proposed in the literature on the basis of spectroscopic evidence.
In the present section we will consider the equilibrium thermodynamic properties and the equation of state. For the polar model these properties are plotted in the figures 3.27 up to 3.30 of the previous chapter.
For physical water the properties are taken from Dorsey [4.2], from a steam table [4.3] and from Eisenberg and Kauzmann [4.4] page 67 and 100. The reduction of the variables has been performed with the same value of σ and ε as in section 3.2.5.4, namely $\sigma = 2.7 \times 10^{-10}$ m and $\varepsilon = -1800$ k . See also the tables 3.9.1, 3.9.2 and 3.9.3.
In figure 4.2 up to 4.5 U*, A*, S* and P* have been plotted as a function of V* and T*, both for physical water and for the model system.

In some respects there is accordance between water and the model, in other respects there appears to be an important difference. That is exactly what should be expected,since the model molecule has some analogy with the physical water molecule but differs from it considerably in other respects.
For instance for the entropy the accordance is satisfactory. The entropy of water seems to be mainly determined by the entropy calculated for a hard-spheres system, corrected for the effect of preference for tetrahedral surrounding and for the polar character of the molecules without long range effect. The latter correction, for the polar character, results mainly in a shift in temperature (See eq. (3.52)). Another detail that is present both in the model and in the physical liquid is the anomaly in the equation of state at lower temperatures. This anomaly is found when a preference for tetrahedral surrounding is introduced in a hard-spheres system. The effect is conserved after transformation to polar molecules.
In physical water the coefficient of thermal expansion changes sign

at T = 277 K (T* = 0.154) whereas the inherent minimum in the isothermal compressibility (See Eisenberg and Kauzmann [4.4] p. 184) is is at T* = 0.177, reduced pressure being about zero. In the model the inversion occurs at higher pressure. With P* = 0.3 thermal expansion changes sign at T* = 0.150 and the minimum in isothermal compressibility is at T* = 0.185 (see fig. 4.6 and 4.7). As is discussed in chapter 3 the anomaly is caused by the fact that at a reduced volume of about 1.66 the configurations with preponderantly tetrahedral surrounding of the molecules are more probable than at a smaller volume, whereas at a larger volume the intermolecular distances become too long for bonding.

Since this is a geometric effect and in the physical liquid geometry must be essentially the same as in the model, we believe that the configurations we have obtained in the computer simulations will have some resemblance with the physical liquid. For instance at the situation of anomaly in the equation of state (T* = 0.15 and P* = 0.3) we observed that in the model about 85 % of the maximum possible number of bonds is realised. It is likely that physical water at the temperature of zero thermal expansion will also be in a configuration with a rather high fraction of the hydrogen bonds intact.

There is a clear discrepancy between model and physical water as far as the energy, the pressure and the heat capacity are concerned. The energy as well as the pressure and the heat capacity of the model system are too high. That is probably due to the fact that spherically symmetric forces (dispersion forces) between the physical molecules exist which are absent in the model.

At least for the energy and the pressure can be verified that the introduction of this type of interaction will have an effect of the right magnitude. When a spherically symmetric attraction is introduced as a perturbation of our model we can suppose that the first order effect will be a change in the energy and the pressure without changing the distribution function.
For evaluation of the effect we suppose the pair potential (3.46) to be changed by adding a perturbation potential :

$$\left.\begin{aligned} E_p &= C \cdot \left(\frac{\sigma}{r}\right)^6 \quad \text{if} \quad r > f \cdot \sigma \\ E_p &= C \cdot \left(\frac{1}{f}\right)^6 \quad \text{if} \quad r \leq f \cdot \sigma \end{aligned}\right\} \qquad (4.2)$$

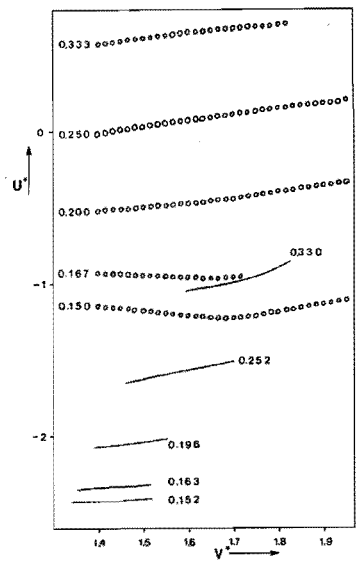


Fig. 4.2 The internal energy of physical water as a function of the volume (full lines). The corresponding values for the polar model are indicated with circles. The latter data are identical with those of fig. 3.27. The (reduced) temperature T is indicated next to each curve.*

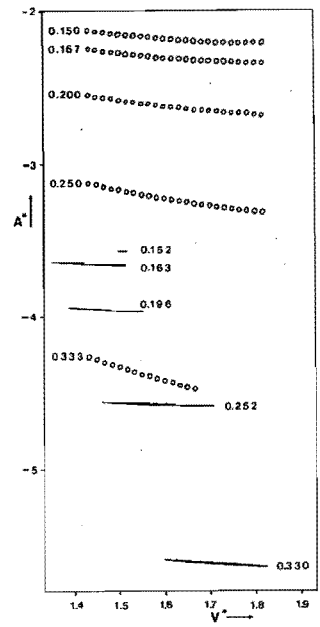


Fig. 4.3 The free energy A of physical water as a function of the volume (full lines). The corresponding values for the polar model are indicated with circles. The latter data are identical with those of fig. 3.28. The (reduced) temperature T* is indicated next to each curve.*

In (4.2) is f = 1.1 and the perturbation parameter C is negative.
The excess energy and pressure become now :

$$U^*_{excess} = \tfrac{1}{2} N \int_{\sigma'}^{\infty} 4\pi r^2 \cdot \rho \cdot g(r) \cdot E_p dr \qquad (4.3)$$

$$P^*_{excess} = -\frac{1}{3V} \cdot \tfrac{1}{2} N \int_{f.\sigma}^{\infty} 4\pi r^2 \cdot \rho \cdot g(r) \cdot r \cdot \left[\frac{dE_p}{dr}\right] dr \qquad (4.4)$$

We have calculated these variables for D = 0.6 and $\lambda' = \lambda + \ln 2 = 6.69$ in accordance with (3.52). The value of C in (4.2) is chosen in accordance with the Slater-Kirkwood formula (see Eisenberg and Kauzmann [4.4] pag. 44). We have taken the value of g(r) for D = 0.6 and $\lambda = 6.0$ of section 3.2.3.

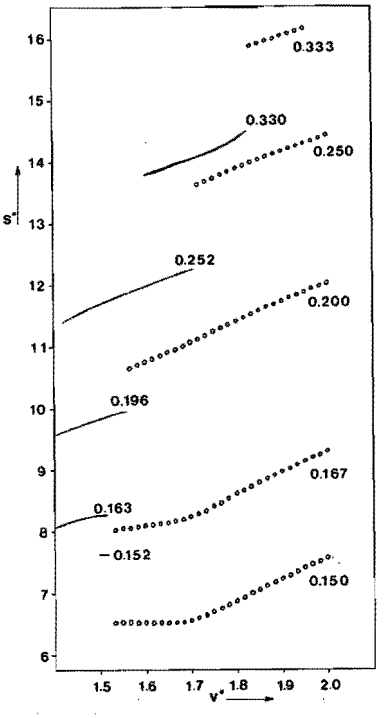


Fig. 4.4 The entropy of physical water as a function of the volume (full lines). The circles give the corresponding values for the polar model. The latter data are identical with those of fig. 3.29. The (reduced) temperature T is indicated next to each curve.*

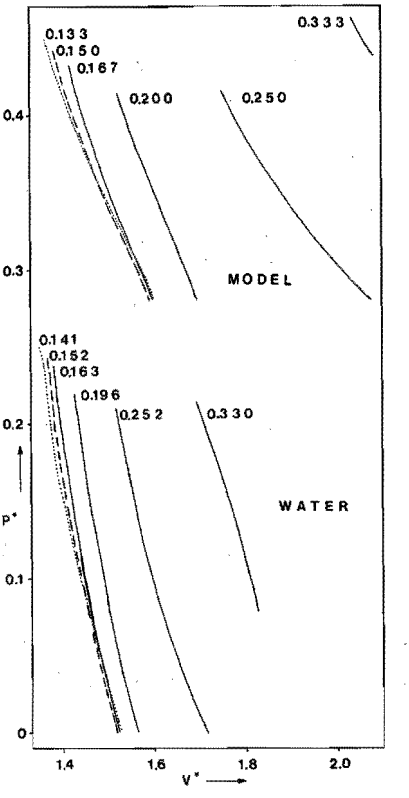


Fig. 4.5 The equation of state of physical water as compared with the (polar) model. The upper set of curves pertains to the model. They are identical with the curves of fig. 3.30 (only the upper part of the curves has been drawn). The lower set of curves (below P=0.3) represents physical water. T* is indicated next to each curve.*

The resulting values were : $U^*_{excess} = -1.138$

$P^*_{excess} = -0.563$

This result is satisfactory since the magnitude is of the right order.

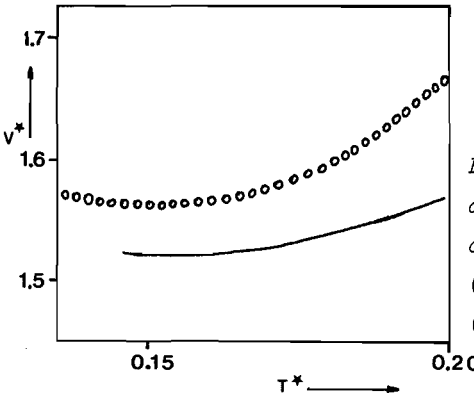


Fig. 4.6 The reduced volume V as a function of T* for the model system with polar molecules (circles) and for physical water (full line).*

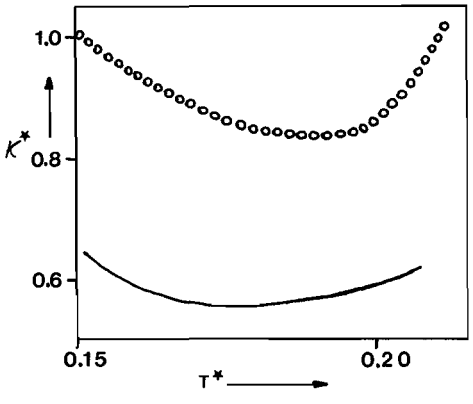


Fig. 4.7 The isothermal compressibility $\kappa^* = -\frac{1}{V^*}\left(\frac{dV^*}{dP^*}\right)_{T^*}$ *as a function of the temperature for the model system with polar molecules (circles) and for physical water (full line).*

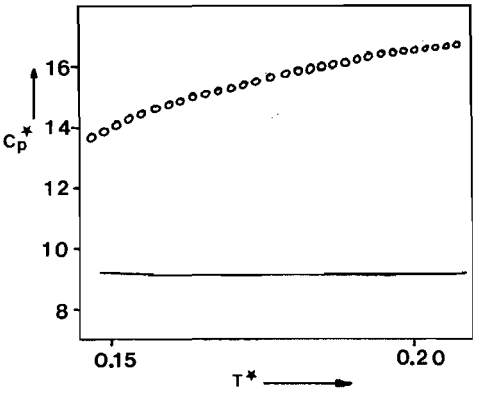


Fig. 4.8 The specific heat at constant pressure $C_p^* = \left(\frac{\partial H^*}{\partial T^*}\right)_{P^*}$ *as a function of T* for the model system with polar molecules (circles) and for physical water (full line)*

4.2.2 Spectroscopic Data

Numerous investigations of the spectroscopy of water have been conducted. A review has been given by Eisenberg and Kauzmann [4.4] page 228.
The interpretation of the spectroscopic data of water appeared to result in different conclusions with regard to such properties as association, the number of bonded and of free OH-groups etc.
Unlike the situation in our computermodel the bonds between water molecules are not all of the same strength and consequently the impact of bonding on the spectroscopic properties is rather complicated. Every statement about numbers of bonds, etc., is a simplification of the real situation.

Four papers are particularly of interest for the comparison of the model with physical water.
Haggis et al. [4.5] have estimated the degree of bonding from dielectric measurements at a wavelength of 1 - 10 cm. They found a fraction of 91 % of the maximum possible bonds in water of 273 K.
Buijs and Choppin [4.6] have studied the infrared absorption of water in the 1.1 to 1.3 μ region. From the results they estimated the degree of bonding to be 52 % at 279 K.
At the corresponding temperature and a reduced pressure of 0.3 in our model about 85 % of all bonds are realised. This is between the estimates of Haggis et al. and of Buijs and Choppin.
Walrafen [4.7] has studied the raman intensities at 152 to 175 cm^{-1}. He concluded that 62 % of the watermolecules is fourfold bonded at 273 K and about 30 % at 316 K. We can compare these values with 55 % and 21 % in the model (table 3.3.2).
Finally, Stevenson [4.8] has studied the ultraviolet absorption of water in the region of 0.175 to 0.195 μ. From this absorption he calculated that 0.12 % of the watermolecules is non-bonded at 296 K and 0.9 % at 365 K. This is to be compared with about 0.7 % and 6 % for the model (interpolated in table 3.3.2).
Compared with the results of Walrafen and of Stevenson the degree of bonding is rather low. Possibly the bonding energy (ε = 1800 k) should have been chosen about 10 % higher.

As a conclusion we can state that in many respects there is a qualitative accordance between the properties of physical water and the model. In some respects the accordance is even rather good.

In this way we have obtained an idea how far the properties of water are determined by the preference of the molecules for tetrahedral coordination.

4.3 Structure and Order

Many authors have discussed the molecular situation in liquid water in terms of 'structure' or 'order' and we can ask ourselves whether we can contribute something to that discussion after the simulation experiments of chapter 3.

The theories are of two types.

Firstly the 'continuum' theories. An important representative is the theory of Pople and Lennard Jones [4.9]. The authors start with fourfold bonding of each water molecule with the possibility of bending the direction of the bonds away from tetrahedral arrangement. They could calculate with this theory a radial distribution function matching rather well with the experimental data (of Morgan and Warren [4.10]).Other authors that have advocated this continuum model are mentioned in reference [4.19].

Secondly others have specified some structure that might predominantly exist in liquid water. Sometimes in a one phase model with some random aspect and sometimes in a two phase model.
As early as in 1892 Roentgen [4.11] has proposed a two phase model, namely 'icebergs' floating in a 'normal' liquid of higher density. We must keep in mind that the structure of ice was not known at that time so Roentgens' theory did not have the pretentions of the recent ones. Later Hall [4.12] has proposed a mixture of ice with a liquid of closely packed water molecules. Davis and Litowitz [4.13] brought forward a model with ice and a second phase of graphitic structure. Probably the best known theory that started with two phases is the 'flickering cluster' theory of Frank and Wen [4.14]. On the basis of cooperative effects in hydrogen bonding they advocated the existence of short living clusters of ice like structure in a liquid of

non-bonded molecules. Némethy and Scheraga [4.15] have further developed this theory, specifying the magnitude of the clusters and the number of water molecules that might be fourfold, threefold, twofold, singly and non-bonded. This theory is certainly the best developed of the two phase theories.
Samoilov [4.16] has proposed a single phase model of ice with part of the cavities filled with interstitial water molecules. This model has been further developed by Narten, Danford and Levy [4.17] on the basis of their X-ray analysis of liquid water.
There is experimental evidence against the two phase model and against the models with an appreciable amount of ice I.
Mysels [4.18] proved that the light scattering, that must be expected in the case of two phases of different density, is absent, at least at room temperature. (In supercritical conditions, however, (the region of critical opalescence) clusters of the type Frank and Wen proposed will certainly exist.)

Secondly, water behaves as a normal liquid when it is undercooled below the freezing point. That proves that water contains no significant amount of crystallites of ice I, at least not more than other liquids contain crystallites of the corresponding type.
Apart from this experimental evidence we feel that it is not very likely that freely moving molecules would restrict themselves to one or two 'structures' as is the case in crystalline substances. The situation in a liquid is basically different from the situation in a crystal. In the crystalline state a vibrating particle returns to its original lattice site even after a prolonged time. In a liquid a particle performs a brownian motion without reference to any specified site. After a long time it will be possible to decide whether the motion of a particle has a vibrational or a brownian character. After a short time, however, it will be impossible to make that distinction.

Analogous to the dynamic difference is the configurational difference between liquid and crystal. In a crystal long-range order exists which is absent in a liquid. On the short range the coordination of a particle may be hardly different in the two states.
Consequently we agree with the theory of Pople and Lennard Jones [4.9]. Our model of chapter 3 is well in accordance with this theory.

Accordingly we consider the liquid phase as dynamic. In the course of time there is one configuration after another in an endless succession of a tremendous number of different configurations. What we mean with 'A tremendous number of different configurations' has been specified in section 3.2.5.1 with the theory of Gosling and Singer [A .6] on the basis of the acceptance ratios during Monte Carlo simulations. The number of possible basic configurations in a system with tetrahedral molecules at $T^* = 0.167$ and $V^* = 1.67$ is estimated to be about 10^{5N}.
(These different or 'basic' configurations can be identified with what Eisenberg and Kauzmann [4.4] define as 'vibrationally-averaged structures'.)

Consequently the description of the 'structure' of a liquid is perforce restricted to the presentation of mean quantities or functions such as distribution functions and correlation functions.
Only at short distance it may make sense to specify details about order or specific interactions (bonds).

4.4 Future Developments

It is possible to evaluate the thermodynamic properties of a model liquid with the Monte Carlo method of Metropolis et al., combined with the multistage sampling method of Valleau et al.
In the present work we discussed the effect of two details on the thermodynamic properties of a liquid i.e. the preference for tetrahedral coordination and polar properties.

In further developments the effect of any detail of the intermolecular potential on the thermodynamic properties of a liquid can be taken into consideration. In principle this can be continued until finally every detail of the thermodynamic properties of some physical liquid is understood.

The method is not restricted to the application to pure liquids. Also mixtures can be taken into consideration. With the same objective.

The restriction of the method will be that the more complex the problem is, the more computer time will be demanded.

APPENDICES

THE PARTITION FUNCTION IN CLASSICAL MECHANICS

The classical partition function of a N V T system is :

$$Q = \frac{1}{N!h^{3N}} \int \exp\left(-\frac{H(p,q)}{kT}\right) dq_1 \ldots dq_N \cdot dp_1 \ldots dp_N \qquad (A\ 1.1)$$

(See Hill [A.1] page 118).

H(p,q) is the energy. It is a starting point of classical mechanics that it is separable into two terms, the kinetic and the potential energy :

$$H(p,q) = \frac{1}{2m}(p_1^2 + \ldots . p_N^2) + U(q_1, \ldots . , q_N) \qquad (A\ 1.2)$$

p_i is the momentum of a particle i and
q_i defines the position of the particle i in the geometric space.
Consequently the integral of (A 1.1) is separable which results for spherical particles in :

$$Q = \frac{z_N}{N!\Lambda^{3N}} \qquad (A\ 1.3a)$$

And for non-spherical particles :

$$Q = \frac{z_N}{N!\Lambda^{3N}} \cdot \frac{1}{\eta^N . R_A^N . R_B^N . R_C^N} \qquad (A\ 1.3b)$$

See Hill [A.1] p.165.

$$\Lambda = \frac{h}{\sqrt{2\pi mkT}} \qquad (A\ 1.4)$$

is the result of the integration of the translational part of the kinetic energy.

$$R_A = \frac{h}{\sqrt{2\pi I_A kT}} \qquad (A\ 1.5)$$

is the result of the integration over the rotation coordinate

around the principal axis A of the molecule . R_B and R_C are similar quantities for the axis B and C.

$$Z_N = \int \exp\left(\frac{-U(q_1 \ldots\ldots q_N)}{kT}\right) dq_1 \ldots\ldots dq_N \qquad \text{(A 1.6)}$$

is called the configuration integral.

η is a multiplicity factor. When a symmetric molecule is rotated, there are several orientations that are indistinguishable from others. So on integration of the classical integral over all orientations Z_N is overestimated with regard to physical systems with a factor which depends on symmetry. For a regular tetrahedron is $\eta = 12$, for a molecule with the symmetry of H_2O (point group C_{2v}) is $\eta = 2$. (See Hill [A.1] p. 166).

For those systems where the potential energy is zero for all configurations (ideal gases and equivalent systems), Z_N can be evaluated analytically :

For spherical particles : $\qquad Z_N = V^N \qquad \text{(A 1.7a)}$

For non-spherical particles : $\qquad Z_N = V^N . (8\pi^2)^N \qquad \text{(A 1.7b)}$

For those systems where the potential energy is not for all configurations equal to zero we can define an excess configuration integral :

$$Z'_N \equiv Z_N / V^N \qquad \text{(A 1.8)}$$

for spherical particles.

In chapter 3 we will use still an other excess configuration integral for non-spherical molecules with a hard core :

$$Z''_N \equiv Z_N / \{Z'_{HS} . V^N . (8\pi^2)^N\} \qquad \text{(A 1.9)}$$

Z'_{HS} is the excess configuration integral of a hard-spheres system, it is equal to the quantity C_o of a system with overlapping spheres as treated in section 2.3.

The method of multistage sampling as we have used in the chapters 2 and 3 is in fact a method for evaluating Z'_N or Z''_N .

THE SAMPLING METHOD OF METROPOLIS ET AL.

In 1953 Metropolis et al. [A.3] described an efficient algorithm for the calculation of ensemble averages with (at the time new) fast computers. The method has been generally adopted and next to the method of Molecular Dynamics this method has been used for the computer simulation of many liquids.
If we start with N particles in a volume V (having previously defined the pair interaction between two particles) it is easy to calculate the potential energy U from the positions of the particles by summation of the pair contributions.
This can be done for a great number of randomly chosen configurations ('Monte Carlo' method). For statistics each configuration should be given the weight exp(-U/kT).

So the equilibrium value of some quantity of interest would be:

$$< F > \approx \frac{\Sigma \; F.\exp(-U/kT)}{\Sigma \quad \exp(-U/kT)} \qquad \text{(A 2.1)}$$

This method is not practical for close packed systems such as liquids because in randomly chosen configurations overlapping molecules are very probable and by consequence U becomes generally high and exp(-U/kT) becomes very small.
Only a few configurations will contribute perceptible to the result. For the other configurations the calculation is done in vain.

Therefore Metropolis et al. designed a modified Monte Carlo scheme. This scheme, which is generally accepted, is called 'The' Monte Carlo method in the jargon of computer simulation of liquids.

Instead of choosing configurations randomly, then weighing them with exp(-U/kT), they choose configurations with a probability proportional to exp(-U/kT) and weigh them evenly.

So the equilibrium value of the quantity F becomes after M steps:

$$< F > \approx \frac{\Sigma \; F}{M} \qquad \text{(A 2.2)}$$

The procedure starts with some configuration (preferably a not too improbable one). Then a small change is made, for instance one of the N particles is moved over a small (random) distance. The change of energy ΔU, caused by the move is calculated. If the move would bring the system to a state of lower energy, the move is allowed and the particle is put in its new position. If, however, $\Delta U > 0$ the move is allowed with probability $\exp(-\Delta U/kt)$; i.e. a random number between 0 and 1 is chosen and if it is below $\exp(-\Delta U/kT)$ the particle is moved to its new position. Otherwise it is returned to its old position.
Then the procedure is repeated with another particle and so on.

It is clear that any configuration can be reached and that the configurations with low value of the energy become more probable.

That the probability of a configuration to occur in such a chain is proportional to $\exp(-U/kT)$ can be proven as follows.
For simplicity we suppose that there are only a finite number of configurations and that ν_r is the number of times the system is in configuration r after a very long chain of configurations. Let P_{rs} be the probability that configuration r changes into configuration s in one proposed move.
Now $P_{rs} = P_{sr}$ because any random move is as probable as its reverse.
If we assume that $U_r > U_s$ the number of times the configuration r changes into s is $\nu_r \cdot P_{rs}$ and the number of times the opposite take place is

$$\nu_s \cdot P_{sr} \cdot \exp(-(U_r - U_s)/kT)$$

If the number of changes in one direction would be unequal to the number in the opposite direction, ν_r (or ν_s) would increase relatively faster with the length of the chain than ν_s (or ν_r). This would go on until equilibrium is reached with :

$$\frac{\nu_r}{\exp(-U_r/kT)} = \frac{\nu_s}{\exp(-U_s/kT)} \qquad \text{(A 2.3)}$$

For two configurations that are more than one step apart it can easily be derived that the corresponding version of (A 2.3) is valid.

In other words: The probability of any configuration to be found in a chain constructed with the Metropolis algorithm, is proportional to exp(- U/kT).

It is convenient to choose the single moves of such a magnitude that the probability of a rejection is of the same order as the probability of acceptance.

If the particle is non spherical the move of a particle must imply a rotation and a translation, which can of course also be separated into two moves :

one move that is restricted to a rotation and the next time a move that is only a translation.

COMMUNAL EFFECT IN PERIODIC CELLS

Monte Carlo simulations must be executed with systems with rather a low number of particles; the order of magnitude is mostly 100. This leads to underestimation of the partition function since if a number of systems with N molecules in a volume V would be combined to a greater one, multiplication of the partition functions would ignore all configurations with a different number of molecules in each of the composing volumina.
If for instance two systems of N molecules each in a volume V are combined to a system of 2N molecules in a volume 2V at the same temperature T, one could start from the supposition that each state of one system can be combined with each state of the other.
That would lead to:

$$Q_{2N} = Q_N \cdot Q_N \qquad \text{(A 3.1)}$$

However, the molecules are free to move from one part of the volume 2V to the other, so states with N-1 molecules in one part and N+1 in the other should be envisaged as well.
And so on.
By consequence :

$$Q_{2N} = \sum_{K=N}^{-N} Q_{N-K} \cdot Q_{N+K} \qquad \text{(A 3.2)}$$

(The effect originates from the fact that the whole volume is available to all molecules. We will describe it therefore as "communal effect".)

For convenience we introduce a correction factor:

$$\Phi \equiv \frac{\sum_{K=N}^{-N} Q_{N-K} \cdot Q_{N+K}}{Q_N^{\,2}} \qquad \text{(A 3.3)}$$

This combined with (A 1.3) from appendix 1 gives :

$$\Phi = \sum_{K=N}^{-N} \frac{N!}{(N-K)!} \cdot \frac{N!}{(N+K)!} \cdot \frac{\mathcal{Z}_{N-K}}{\mathcal{Z}_N} \cdot \frac{\mathcal{Z}_{N+K}}{\mathcal{Z}_N} \qquad \text{(A 3.4)}$$

It is convenient to introduce now two new factors :

$$\Psi \equiv \sum_{k=N}^{-N} \frac{(N!)^2}{(N-K)! \,.\, (N+K)!} \qquad \text{(A 3.5)}$$

$$\theta \equiv \Phi / \Psi \qquad \text{(A 3.6)}$$

Ψ can be evaluated with the Stirling formula :

$$M! \approx (M/e)^M \,.\, \sqrt{2\pi M}$$

which is a good approximation for high values of M (M > 10).

Thus, since $\sum_{K=N}^{-N} \frac{(2N)!}{(N-K)! \,.\, (N+K)!} = 2^{2N}$,

we obtain :

$$\Psi = (N!)^2 \,.\, 2^{2N} / (2N)! = \sqrt{\pi N} \qquad \text{(A 3.7)}$$

After doubling the system the logarithm of the partition function becomes :

$$\ln Q_{2N} = 2\ln Q_N + \ln(\theta_1 . \sqrt{\pi . N}) \qquad \text{(A 3.8)}$$

A second doubling gives :

$$\ln Q_{4N} = 2\ln Q_{2N} + \ln(\theta_2 . \sqrt{\pi . 2N})$$

which leads to :

$$\frac{\ln Q_{4N}}{4N} = \frac{\ln Q_N}{N} + \frac{\ln\sqrt{\pi . N}}{N} (\tfrac{1}{2} + \tfrac{1}{4}) + \frac{(\tfrac{1}{2}\ln\theta_1 + \tfrac{1}{4}\ln\theta_2)}{N} + \tfrac{1}{4}\frac{\ln\sqrt{2}}{N}$$

Repeating the doubling until macroscopic dimensions gives :

$$\frac{\ln Q_N}{N} = \frac{\ln Q_M}{M} - \frac{\ln\sqrt{\pi.N}}{N} - \frac{(\frac{1}{2}\ln\theta_1 + \frac{1}{4}\ln\theta_2 + ...)}{N} - \frac{\ln\sqrt{2}}{N} \qquad \text{(A 3.9)}$$

Where M is a macroscopic number, far exceeding N.
If θ would be a constant (which will appear to be a good approximation for hard spheres), (A 3.9) would become :

$$\frac{\ln Q_N}{N} = \frac{\ln Q_M}{M} - \frac{\ln(\theta\ .\sqrt{2\pi.N})}{N} \qquad \text{(A 3.10)}$$

Q_M is the partition function of a macroscopic system for which the equation of Carnahan and Starling [A.2] can be used :

$$\frac{\ln Q_M}{M} - \frac{\ln Q_{iM}}{M} = \frac{4y - 3y^2}{(1 - y)^2} \qquad \text{(A 3.11)}$$

Q_{iM} is the partition function of an ideal gas with M particles in a volume $V \cdot M/N$ at the temperature T.
Furthermore it can easily be derived that for an ideal gas θ equals unity. So we get with (A 3.10) :

$$\frac{\ln Q_{iM}}{M} - \frac{\ln Q_{iN}}{N} = \frac{\ln\sqrt{2\pi.N}}{N} \qquad \text{(A 3.12)}$$

Now from (A 3.10), (A 3.11) and (A 3.12) we obtain :

$$\ln Q_N - \ln Q_{iN} = -N \cdot \frac{4y - 3y^2}{(1 - y)^2} - \ln\theta \qquad \text{(A 3.13)}$$

In words :
The excess partition function of an N-particle system is to be corrected with a factor θ with regard to the equation of Carnahan and Starling. θ is defined by (A 3.6), (A 3.5) and (A 3.3).

When we approximate B in formula (A 3.4) with the equation of Carnahan and Starling, θ can be evaluated numerically for different values of N. The result is that θ is approximately constant for $N > 8$. (This result is used already for equation (A 3.10)).

In table A 3.1 is mentioned the value of θ for a few values of N.

Table A 3.1: The correction factor θ for a hard-spheres system as a function of N. Density = $N.\sigma^3/V = 0.60$.

N	θ
1	0.5000
2	0.3767
4	0.3037
8	0.2913
16	0.2930
32	0.2941
64	0.2947
128	0.2950
256	0.2952
512	0.2952

CALCULATION OF PRESSURE FROM DISTRIBUTION FUNCTION

In a system of N molecules in a volume V the pressure is the result of the interaction of the molecules with the walls of the vessel.
When the molecules exhibit no intermolecular forces, we have a perfect gas and the pressure is given by the formula:

$$P = \frac{NkT}{V} \qquad \text{(A 4.1)}$$

When intermolecular forces do exist, deviation from this value occurs. Repulsive forces attribute positively to the pressure, attractive forces negatively.
We will start with the equation of state, formula (2.6) :

$$\frac{PV}{NkT} - 1 = -\frac{2}{3} \cdot \frac{Vir}{NkT}$$

(with $Vir = -\frac{1}{2}\langle \sum r_{kl} \cdot F_{kl} \rangle$).

If only pairwise interaction exists, as is the case in our models, the equation of state can be rewritten as :

$$\frac{PV}{NkT} - 1 = -\frac{2\pi N}{3kTV} \int_0^{\infty} g(r) \cdot r^3 \cdot \frac{dE_2}{dr}\, dr \qquad \text{(A 4.2)}$$

(See Hill [A.1] p. 305.)

The symbols have the same meaning as in chapter 2 and 3, E_2 being the pair potential and g(r) the radial distribution function. The radial distribution function comprises the information about the occurrence of pairs at each value of r. Since only pairwise interaction exists, that is sufficient. For hard spheres the formula is not applicable in this form since $\frac{dE_2}{dr}$ does not exist everywhere. In this case the potential is replaced by a continuous function, e.g. $E_2 = -\varepsilon\{\frac{\sigma}{r}\}^n$ and the limit for $n = \infty$ is taken (See Hansen and McDonald [A.4] , page 52).

In accordance with chapter 3, ε has a negative value

Another complication is that E_2 depends on the orientation of the molecules we deal with in chapter 3. (formula 3.3)

In that case we have to distinguish between those pairs of molecules that are oriented favourably for making bonds and those pairs that are not.

For the last group we can replace E_2 by $-\varepsilon(\frac{\sigma}{r})^n$ as for the hard spheres, for the other group we should replace E_2 by a more complicated form :

$$E_2' = -\varepsilon(\frac{\sigma}{r})^n + \frac{\varepsilon}{1 + (\frac{r}{f\sigma})^n} \qquad \text{(A 4.3)}$$

and again take the limit for $n = \infty$.

We divide the distribution function into two parts along the lines of section 3.2.3. :

$$g(r) = q_1(r) \cdot \exp(-\frac{E_2'}{kT}) + \{q_2(r) + q_3(r)\} \cdot \exp(-\frac{E_2}{kT}) \qquad \text{(A 4.4)}$$

q_1 refers to the favourably oriented pairs, q_2 and q_3 to the others.

Combination of (A 4.2) with (A 4.4) gives :

$$\frac{PV}{NkT} - 1 = -\frac{2\pi N}{3V}\int_0^\infty q_1(r) \cdot r^3 \cdot \exp(-\frac{E_2'}{kT}) \cdot \frac{d(E_2'/kT)}{dr}\, dr$$

$$-\frac{2\pi N}{3V}\int_0^\infty \{q_2(r) + q_3(r)\} \cdot r^3 \cdot \exp(-\frac{E_2}{kT}) \cdot \frac{d(E_2/kT)}{dr}\, dr \qquad \text{(A 4.5)}$$

We will consider the two integrals of (A 4.5) seperately.

If $E_2 = -\varepsilon(\frac{\sigma}{r})^n$ is inserted in the second integral and n is very large, the product $\exp(-\frac{E_2}{kT}) \cdot \frac{d(E_2/kT)}{dr}$ is very small in most cases. Only when r equals about σ, the product differs appreciable from zero. So

$$\int_0^\infty \{q_2(r) + q_3(r)\} \cdot r^3 \cdot \exp(-\frac{E_2}{kT}) \cdot \frac{d(E_2/kT)}{dr}\, dr \quad \text{becomes:}$$

$$= \{q_2(\sigma) + q_3(\sigma)\} \cdot \sigma^3 \int_{\sigma^-}^{\sigma^+} \exp(-\frac{E_2}{kT}) \cdot \frac{d(E_2/kT)}{dr} \, dr$$

$$= \{q_2(\sigma) + q_3(\sigma)\} \cdot \sigma^3 \int_{\infty}^{0} \exp(-\frac{E_2}{kT}) \, d(E_2/kT)$$

$$= - \{q_2(\sigma) + q_3(\sigma)\} \cdot \sigma^3$$

σ^- and σ^+ are the value of r just below and just above r = σ respectively. (See fig. A.1)

When (A 4.3) is inserted in the first integral of equation (A 4.5) and n is very high, the product $\exp(E_2'/kT \cdot d(E_2'/kT)/dr$ is very small except when r equals about σ or f.σ.

If we now introduce $A = \frac{-\varepsilon(\sigma/r)^n}{kT}$ and $B = \frac{\varepsilon}{kT} \cdot \frac{1}{1 + (\frac{r}{f\sigma})^n}$

the first integral of (A 4.5) becomes :

$$\int_0^{\infty} q_1(r) \cdot r^3 \exp(-A-B) \cdot \frac{d(A+B)}{dr} \, dr$$

$$= \int_0^{\infty} q_1(r) \cdot r^3 \exp(-B) \cdot \exp(-A) \frac{dA}{dr} \, dr +$$

$$\int_0^{\infty} q_1(r) \cdot r^3 \exp(-A) \cdot \exp(-B) \frac{dB}{dr} \, dr$$

which becomes for a high value of n:

$$q_1(\sigma) \cdot \sigma^3 \cdot \exp(-\frac{\varepsilon}{kT}) \int_{\infty}^{0} \exp(-A) dA + q_1(f\sigma) \cdot (f\sigma)^3 \int_{\frac{\varepsilon}{kT}}^{0} \exp(-B) dB =$$

$$= -q_1(\sigma) \cdot \sigma^3 \cdot \exp(-\frac{\varepsilon}{kT}) + q_1(f\sigma) \cdot f^3\sigma^3 \cdot \{-1+\exp(-\frac{\varepsilon}{kT})\}$$

Moreover, in Monte Carlo simulations the right hand side of the last equation should be multiplied by a factor N/(N-1) in order to correct for the fact that the centre of mass is not fixed (see Hansen and McDonald [A.4] p.52).

The equation of state becomes:

$$\frac{PV}{NkT} - 1 = \frac{N}{N-1} \cdot D \cdot \frac{2\pi}{3} \left\{ q_1(\sigma) \exp(-\varepsilon/kT) + q_2(\sigma) + q_3(\sigma) - q_1(f\sigma) \cdot \{ \exp(-\varepsilon/kT) - 1 \} \cdot f^3 \right\} \qquad \text{(A 4.6)}$$

($D = N\sigma^3/V$ = density)

It is possible to bring this equation in a simplified form with help of the following relations:

$$\begin{aligned} g(\sigma^+) &= q_1(\sigma) \exp(-\tfrac{\varepsilon}{kT}) + q_2(\sigma) + q_3(\sigma) \\ g(\sigma^-) &= 0 \\ g(f\sigma^+) &= q_1(f\sigma) + q_2(f\sigma) + q_3(f\sigma) \\ g(f\sigma^-) &= q_1(f\sigma) \exp(-\tfrac{\varepsilon}{kT}) + q_2(f\sigma) + q_3(f\sigma) \end{aligned}$$

In which $g(f\sigma^+)$ means: the value of the(discontinuous) function g just above the discontinuity of $r = f\sigma$ and $g(f\sigma^-)$ the value of the function g just below the same discontinuity. (See fig. A.1) We can now rewrite (A 4.6)

$$\frac{PV}{NkT} - 1 = \frac{N}{N-1} \cdot D \cdot \frac{2\pi}{3} \left\{ \{ g(\sigma^+) - g(\sigma^-) \} + f^3 \{ g(f\sigma^+) - (g(f\sigma^-)) \} \right\} \qquad \text{(A 4.7)}$$

In this way the pressure can be calculated directly from the discontinuities in g(r) at $r = \sigma$ and $r = f\sigma$.

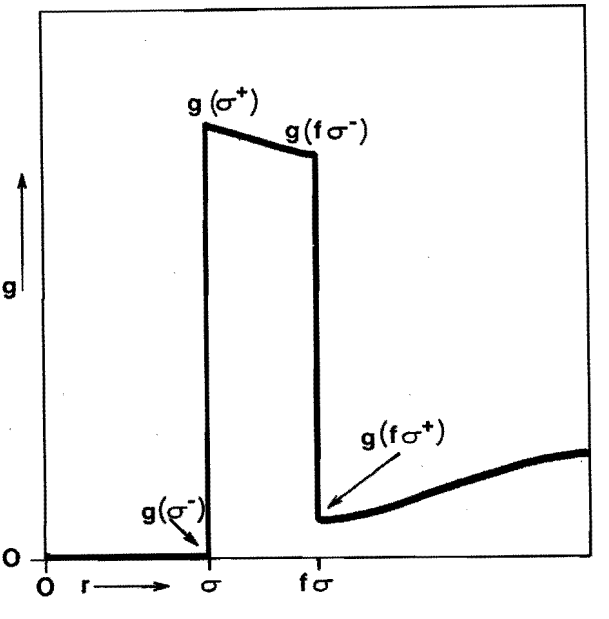


Fig. A.1 The radial distribution function g(r) with indication of the symbols as used in formula A 4.7.

FREE VOLUME

In 1972 Gosling and Singer announced a method for the estimation of the free energy from the acceptance ratio of steps in a Monte Carlo simulation [A.6].
As has been described in appendix 2, in the Metropolis Monte Carlo method small moves are proposed in order to sample a range of configurations.
If the move would lead to a decrease of energy the move is accepted. If it would lead to an increase of energy the move is sometimes accepted, sometimes not. The probability of acceptance is determined in a way that the sampling after a great number of moves will fulfill the Boltzman partition between configurations of different energy (see appendix 2).
Generally, the acceptance ratio (that is the number of accepted moves divided by the total number proposed) is considered to be a trivial by-product of the procedure. Gosling and Singer, however, used the acceptance ratio to calculate a "free volume" for the particles since acceptance of a move can be considered to prove that the proposed place of the particle is free.
Identification of this free volume with the v_f of the well known cell theory of liquids [A.5] leads to the possibility to determine the free energy and all thermodynamic properties of the system.
They applied the method to a Lennard Jones (12-6) potential with parameters corresponding to the argon atom. The method proved to yield excellent accordance with the experimental free energy of the physical argon liquid.

In 1973 Valleau and Whittington [A.7] proved the method to be unsound. The agreement between calculation and experiment is due to the fact that two important errors appear to cancel each other.

The first error is due to the fact that in the cell theory of liquids Gosling and Singer have used the lowest possible energy of the system as a reference level.
The Metropolis sampling, however, has a reference level the energy that is actually calculated in the present configuration.

This error will lead to an overestimation of the partition function.

The second error has a quite different character.
Gosling and Singer assumed that the following relation should exist:

$$\langle a_N \rangle = \langle a_1 \rangle^N \qquad \text{(A 5.1)}$$

That means that the probability of acceptation after simultaneously moving all N particles to any position in the volume V (which value for hard spheres would lead to the correct partition function) equals the N-th power of the probability of acceptation after moving a single particle. The right hand side of (A 5.1) would lead to to a good approximation if one basic configuration would exist from which all other relevant configurations can be realised by the displacement of the particles inside their own free volume.

That the hypothesis (A 5.1) is false can best be illustrated with the case of a hard-spheres fluid. In that case energy is the same in all relevant configurations and by consequence the above-mentioned first error will not confuse the argument.
The chance of any configuration of N hard spheres in a volume V to be accepted is simply the chance of finding a configuration without any overlap of spheres. The chance for the first of the N spheres placed in the volume is 1.0, adding the second will give a chance slightly below 1.0. This chance becomes less and less until for the last sphere the chance of finding a place in a volume filled with N-1 spheres is rather small (if the density is not too low). The total chance is the product of the N factors mentioned. The chance of moving a sphere to some place between the N-1 others approximates the last (the lowest) of the N factors. The N-th power of that chance is considerably lower than the product, just mentioned. This leads to underestimation of configuration volume, so that the two effects could cancel.

It is interesting to verify the argument with a Monte Carlo calculation. To that end we made a Monte Carlo experiment with 91 hard spheres at the density $N\sigma^3/V = 0.60$. The number of (attempted) moves was 910000 of which 5465 appeared to be accepted. There was no restriction to the magnitude of the move in the sense that the new

location of the particle to be moved was chosen randomly within the entire volume V. By consequence the acceptance ratio gives an estimation of the highest possible value of the free volume v_f:

$$\frac{v_f}{V} = \frac{5465}{910000} = 0.00600 \text{ and } \ln(v_f/V) = -5.12$$

That would lead to $\ln Q'/N = -5.12$.
Q' is the excess partition function of the system compared with the ideal gas.
(The standard deviation of this result can be approximated from the repeatability. It proved to be $\pm$ 0.02).
The result mentioned is to be compared with the value $\ln Q/N = -2.042$, calculated with the formula of Carnahan and Starling (chapter 2 section 3).

The discrepancy between the two approaches amounts to $\exp(N\ (-2.042 + 5.12)) = 10^{122}$ for a system with 91 particles. This proves the statement of Valleau and Whittington, that this error would lead to a serious underestimation of the partition function, to be correct.

The discrepancy is so important that we must conclude that the cell theory of liquids in its simplest form should be abandoned. It is possible to modify the cell theory in such a way that the discrepancy is small. It would not be practical to build such a modified theory with more than one basic configurations, because in that case the number of basic configurations would become as large as 10^{122}. However, one should leave the concept of one particle per cell and take into consideration cells with more than one particle. Computer simulation with periodic boundary conditions presents such an alternative method. In chapter 2, table 2.2, we can see which number of particles would suffice.

In this way the computer simulation can be considered to be a modification of the cell model, however, without a closed analytical solution.

REFERENCES

Chapter 1.

1.1 J.H. Hildebrand, J.M. Prausnitz & R.L. Scott, "Regular and Related Solutions", Van Nostrand, New York, 1970.

1.2. H. Kamerlingh Onnes, Comm.Phys.Lab. Leiden, 1901, no. 71. K. Singer, ed. "Statistical Mechanics", A Specialist Periodic Report, The Chem. Soc., London, 1973.

1.3 M. Born & H.S. Green, "A General Kinetic Theory of Liquids", Cambridge, London, 1949.

1.4 J.K. Percus & G.J. Yevick, Phys. Rev. 110, 1 (1958).

1.5 J.M.J. van Leeuwen, J. Groeneveld & J. de Boer, Physica 25, 792 (1959).

1.6 J.G. Kirkwood & F.P. Buff, J.Chem.Phys. 19, 774 (1951).

1.7 J.A. Barker, R.A. Fisher & R.O. Watts, Mol.Phys. 21, 657 (1971). J.A. Barker, R.O. Watts, J.K. Lee, T.P. Schafer & Y.T. Lee, J.Chem.Phys. 61, 3081 (1974).

1.8 G.C. Lie, E. Clementi & M. Yoshimine, J.Chem.Phys. 64, 2314 (1976).
I.R. McDonald & M.L. Klein, J.Chem.Phys. 68, 4875 (1978).
F.H. Stillinger & C.W. David, J.Chem.Phys. 69, 1473 (1978).

1.9 J.A. Barker & D. Henderson, Rev.Mod.Phys. 48, 638 (1976).
J.A. Barker & D. Henderson, Acc.Chem.Res. 4, 303 (1971).

1.10 N. Metropolis, A.W. Rosenbluth, M.N. Rosenbluth, A.H. Teller & E. Teller, J.Chem.Phys. 21, 1087 (1953).

1.11 J.P. Valleau & D.N. Card, J.Chem.Phys. 57, 5457 (1972).

1.12 N.F. Carnahan & K.E. Starling, J.Chem.Phys. 51, 635 (1969).

1.13 B.J. Alder & T.E. Wainwright, J.Chem.Phys. 31, 459 (1959).
A. Rahman & F.H. Stillinger, J.Chem.Phys. 55, 3336 (1971).

1.14 F.H. Stillinger, Adv.Chem.Phys. 31, 1 (1975).

1.15 A. Ben-Naim, J.Chem.Phys. 52, 5531 (1970).

1.16 J. Michielsen, P. Woerlee, F. v.d.Graaf & J.A.A. Ketelaar, J.C.S. Faraday II, 71, 1730 (1975).

Chapter 2

2.1 J.P. Hansen & I.R. McDonald, "Theory of Simple Liquids", Academic Press, London, 1976.

2.2 J.P. Hansen & L. Verlet, Phys. Rev. 184, 151 (1969).

2.3 R.T. Birge & J.W. Weinberg, Rev.Mod.Phys. 19, 298 (1947). The method has been programmed in Algol by Donkersloot in: M.C.A. Donkersloot, "Algemene Foutberekening ...", report 760401 Eindhoven University of Technology (in dutch, available on request).

2.4 J.O. Hirschfelder, C.F. Curtiss & R.B. Bird, "Molecular Theory of Gases and Liquids", Wiley, New York 1964.

2.5 F.H. Ree & H.G. Hoover, J.Chem.Phys. 40, 939 (1964).

Chapter 3

3.1 A. Rahman & F.H. Stillinger, J.Chem.Phys. 55, 3336 (1971).
F.H. Stillinger & A. Rahman, J.Chem.Phys. 57, 1281 (1972).
A. Rahman & F.H. Stillinger, J.Am.Chem.Soc. 95, 7943 (1973).
F.H. Stillinger & A. Rahman, J.Chem.Phys. 60, 1545 (1974).
A. Rahman & F.H. Stillinger, Phys.Rev.A. 10, 368 (1974).
F.H. Stillinger & A. Rahman, J.Chem.Phys. 61, 4973 (1974).
R.O. Watts, Mol.Phys. 28, 1069 (1974).

3.2 H. Popkie, H. Kistenmacher & E. Clementi, J.Chem.Phys. 59, 1325 (1973).
O. Matsuoka, E. Clementi & M. Yoshimine, J.Chem.Phys. 64, 1351 (1976)

3.3 G.C. Lie, E. Clementi & M. Yoshimine, J.Chem.Phys. 64, 2314 (1976).
G.C. Lie & E. Clementi, J.Chem.Phys. 64, 5308 (1976).
J.C. Owicki & H.A. Scheraga, J.Am.Chem.Soc. 99, 7403 (1977).

3.4 F.H. Stillinger & C.W. David, J.Chem.Phys. 69, 1473 (1978).

3.5 J.G. Curro, J.Chem.Phys. 61, 1203 (1974).
J.P. Ryckaert, G.Ciccotti & H.J.C. Berendsen, J.Comp.Phys. 23, 327 (1977)

3.6 A. Ben-Naim, "Water and Aqueous Solutions", Plenum Press, New York, 1974, p. 283.

3.7 G.A.M. Eilers, J.L. de Jong, R. Kool & H.N. Linssen, "Miniquad", COSOR Note R 75-22, Eindhoven University of Technology, Department of Mathematics. (Available on Request).

3.8 A.J.C. Ladd, Mol.Phys. 36, 463 (1978).

3.9 L. Pauling, J.Am.Chem.Soc. 57, 2680 (1935).

3.10 W.F. Giauque & M. Ashley, Phys. Rev. 43, 81 (1933).
W.F. Giauque & J.W. Stout, J.Am.Chem.Soc. 58, 1144 (1936).

3.11 L. Onsager & M. Dupuis, Rc.Scu.int.Fis. 'Enrico Fermi' 10, 294 (1960).

3.12 J.F. Nagle, J.Math.Phys. 7, 1484 (1966).

3.13 L.L. Shipman & H.A. Scheraga, J.Phys.Chem. 78, 909 (1974).

3.14 A. Ben-Naim & F.H. Stillinger, "Structure and Transport Processes in Water and Aqueous Solutions", ed. R.A. Horne, Interscience, 1972.

3.15 W. Bol, Report CECAM Workshop on the Molecular Dynamics of Water, Orsay, 1972, p. 51.

3.16 G.A. v.d. Velde, Report CECAM Workshop on the Molecular Dynamics of Water, Orsay, 1972, p. 38.

3.17 I.R. McDonald & M.L. Klein, J.Chem.Phys. 68, 4875 (1978).

3.18 J.A. Barker, "Lattice Theories of the Liquid State", Pergamon Press, Oxford, 1963, page 29.

3.19 H.A. Narten, "X-ray Diffraction Data on Liquid Water", ORNL-4578, July 1970, Oak Ridge National Laboratory, Oak Ridge U.S.A.
W. Bol, J.Appl.Cryst. 1, 234 (1968).

3.20 D. Eisenberg & W. Kauzmann, "The Structure and Properties of Water", Clarendon Press, Oxford 1969, p. 179.

Chapter 4

4.1 J.A. Barker & D. Henderson, Rev.Mod.Phys. 48, 638 (1976).

4.2 N.E. Dorsey, "Properties of Ordinary Water Substance", Hafner, New York, 1968.

4.3 C.W. Burnham, J.R. Hollowy, N.F. Davis, J.A. Hunter, W.S. Gillam & L. Leiserson, "Thermodynamic Properties of Water", R. and D. Progress Report No 414, March 1969, U.S. Department of the Interior, U.S.Gov.Printing Office, Washington D.C.

4.4 D. Eisenberg & W. Kauzmann, "The Structure and Properties of Water", Clarendon Press, Oxford 1969.

4.5 G.H. Haggis, J.B. Hasted & T.J. Buchanan, J.Chem.Phys. 20, 1452 (1952).

4.6 K. Buijs & G.R. Choppin , J.Chem.Phys. 39, 2035 (1963).
This paper has been criticised by:
D.F. Hornig, J.Chem.Phys. 40, 3119 (1964).

4.7 G.E. Walrafen, J.Chem.Phys. 44, 1546 (1966).

4.8 D.P. Stevenson, J.Phys.Chem. 69, 2145 (1965).

4.9 J.A. Pople & J. Lennard Jones, Proc.Roy.Soc.A., 205, 163 (1951).

4.10 J. Morgan & B.E. Warren, J.Chem.Phys. 6, 666 (1938).

4.11 W.C. Roentgen, Ann.Phys. 45, 91 (1892).

4.12 L. Hall, Phys.Rev. 73, 775 (1948).

4.13 C.M. Davis & T.A. Litovitz, J.Chem.Phys. 42, 2563 (1965).

4.14 H.S. Frank & W.Y. Wen, Disc.Far.Soc. 24, 133 (1957).

4.15 G. Némethy & H.A. Scheraga, J.Chem.Phys. 36, 3382 (1962).

4.16 O.J. Samoilow, "Die Struktur Wässriger Elektrolytlösungen und die Hydratation von Ionen", Teubner, Leipzig, 1961.

4.17 A.H. Narten, M.D. Danford & H.A. Levy, Disc.Far.Soc. 43, 97 (1967).

4.18 K.J. Mysels, J.Am.Chem.Soc. 86, 3503 (1964).

4.19 T.T. Wall & D.F. Hornig, J.Chem.Phys. 43, 2079 (1965).
M. Falk & T.A. Ford, Can.J.Chem. 44, 1699 (1966).
J. Schiffer & D.F. Hornig, J.Chem.Phys. 49, 4150 (1968).
W. Kauzmann, in A.Alfsen & A.J.Berteaud, Eds., "L'Eau et les Systèmes Biologiques", Centre National de la Recherche Scientifique, Paris, 1976.

Appendices

A.1 T.L. Hill, "An Introduction to Statistical Thermodynamics", Addison-Wesley, London, 1962.

A.2 N.F. Carnahan & K.E. Starling, J.Chem.Phys. 51, 635 (1969).

A.3 N. Metropolis, A.W. Rosenbluth, M.N. Rosenbluth, A.H. Teller, & E. Teller, J.Chem.Phys. 21, 1087 (1953).

A.4 J.P. Hansen & I.R. McDonald, "Theory of Simple Liquids", Academic Press, London, 1976.

A.5 S. Glasstone, "Theoretical Chemistry", Van Nostrand, New York, 1944, Page 465.
R.H. Fowler & E.A. Guggenheim, "Statistical Thermodynamics", Cambridge University Press, Cambridge, 1939, page 331.
J.H. Barker, "Lattice Theories of the Liquid State", Pergamon Press, Oxford, 1963, page 30, 127.

A.6 E.M. Gosling & K. Singer, Pure Appl. Chem., 22, 303 (1970).

A.7 J.P. Valleau & S.G. Whittington, J.C.S. Faraday II, 69, 1004 (1973).

SUMMARY

The present thesis is an account of the results of computer simulation of the properties of liquids. The calculations have been based on simplified models. The results obtained provide an estimate of the equilibrium thermodynamic properties of the model-liquids chosen, without implication of suppositions concerning the structure of the liquid.
The energy, the pressure and the distribution functions have been obtained directly from the 'Monte Carlo' method of Metropolis et al. . The Free Energy is evaluated indirectly with the aid of the multistage sampling method of Valleau et al.

Firstly the method has been applied to a 'hard spheres' fluid. It appears that the free energy of the hard-spheres fluid can be evaluated very accurately. A standard deviation of only 0.2 % can be obtained with only a limited number of moves (about 7.10^6).
Secondly a system with rigid spherical non-polar molecules has been studied in which each molecule can have a bond with up to four neighbours. The bonds thus formed are directed tetrahedrally in space. In this model long range interactions are absent. A system containing 91 of such molecules in a cubic box with periodic boundary conditions has been investigated.

The equation of state, the energy, the free energy, the entropy and the distribution functions (including a number of orientational functions) have been calculated.

In this case too the free energy is evaluated rather accurately. The standard deviation of the excess free energy can be estimated to be about 0.5 %.

Furthermore an analytical method has been developed for the transformation of the above-mentioned properties of the system with non-polar molecules into the properties of the corresponding system with polar molecules (still without long range interaction).

The properties of the latter system can be compared with the physical

properties of water. The water molecule is polar and has preferential directions for making hydrogen bonds that deviate only slightly from the tetrahedral directions.

In water long range interaction between the molecules is present as well as a spherically symmetric attraction. Moreover the molecules are not rigid.

Evidently the model and the physical liquid are not identical, but in some aspects the model and the real liquid are comparable.
The entropy calculated for the model closely corresponds with the entropy of liquid water.
Furthermore the anomaly in the thermal expansion at constant pressure and the anomaly in the temperature coefficient of isothermal compressibility (well-known properties of real water) can be observed in the results of the model calculations at a corresponding volume but at a significantly different pressure.

So it is justified to attribute the anomalous expansion properties of water to the preference of the molecules for tetrahedral coordination.
There appears to be a considerable difference between the model and the physical liquid as concerns the energy, the free energy, the specific heat and the region of the pressure where expansion is anomalous. However, it can be shown that the introduction of spherically symmetric interaction of the correct order of magnitude between the molecules would substantially reduce the differences observed.

During the computer simulation calculations many details of the model system are accessible to close examination. The opportunity has been used for checking the classical lattice theory of liquids, based on the conception of a 'vibrationally-averaged structure' of the liquid involving a certain 'free volume' for the molecule to move in. The calculation of the magnitude of the free volume of the molecule in the models discussed leads to the conclusion that the lattice theory would lead to a discrepancy of 10^{5N} for a N-particle system.

Consequently it makes no sense to consider 'a' structure of a liquid. Instead one should conceive of a tremendous number of equivalent structures. It is senseless to speculate about geometric arrangements that involve more than a few molecules (which is general practice in the literature). This conclusion is in accordance with the results of many physical experiments viz. x-ray diffraction. The latter experiments yield information about the short range order, but they leave questions regarding long range arrangements unanswered.

SAMENVATTING

In dit proefschrift worden de resultaten beschreven van de computersimulatie van vloeistoffen. De berekeningen zijn beperkt tot vereenvoudigde modellen. De verkregen resultaten zijn de evenwichts thermodynamische eigenschappen van de gekozen modelvloeistof. Er wordt daarbij van tevoren geen veronderstelling over de structuur van de vloeistof gemaakt.

De energie, de druk en de distributiefunctie worden direct verkregen met de 'Monte Carlo' methode van Metropolis e.a.. De vrije energie wordt indirect uitgerekend met de methode van stapsgewijze bemonstering ('multistage sampling') van Valleau e.a.

Om te beginnen is de methode toegepast op een 'harde bollen' vloeistof. De vrije energie blijkt zeer nauwkeurig verkregen te worden. De standaarddeviatie is slechts 0.2 % met een beperkt aantal stappen (ongeveer 7.10^6).
Vervolgens is de methode toegepast op starre bolvormige moleculen, waarbij elk molecuul een binding kan aangaan met een of meer buren (maximaal 4 bindingen per molecuul). De aldus gevormde bindingen zijn gericht in de 4 hoofdrichtingen van een tetraeder. Lange afstand interactie is in dit model niet aanwezig.

Van een systeem met deze modelmoleculen zijn de toestandsvergelijking, de energie, de vrije energie, de entropie en de distributiefuncties berekend (inclusief een aantal orientatie distributiefuncties).

Ook in dit geval is de vrije energie tamelijk nauwkeurig verkregen. De standaarddeviatie (voor de excess-vrije energie) bedraagt ongeveer 0.5 %.

Verder is een analytische methode ontwikkeld om de berekende eigenschappen te transformeren in de eigenschappen van een overeenkomstig systeem maar dan met polaire moleculen. (polair, maar zonder lange afstand interactie).

De eigenschappen van het laatste systeem kunnen vergeleken worden

met de physische eigenschappen van water. Het watermolecuul is eveneens polair en heeft voorkeursrichtingen voor het aangaan van waterstofbindingen, die maar weinig afwijken van de tetraederrichtingen. Evenwel is er in water wel lange afstand interactie en er is een van der Waals attractie tussen de moleculen. Bovendien zijn de watermoleculen niet star.
Het is daarom duidelijk dat de eigenschappen van physisch water en van het model niet gelijk zijn, maar in sommige opzichten is er overeenstemming. De entropie blijkt van de goede grootte te zijn en verder is er, net als in water, ook in het model een anomalie in de thermische uitzetting bij constante druk en in de isotherme compressibiliteit. Deze verschijnselen treden echter op bij een heel andere druk.

Dus het is gerechtvaardigd om het verschijnsel van de anomalie in de thermische uitzetting van water toe te schrijven aan de voorkeur van de moleculen voor tetraedrische coordinatie.

Bij de energie, de vrije energie, de soortelijke warmte en ook in het drukgebied waar anomalie in de uitzetting optreedt, is er een aanzienlijk verschil tussen ons model en echt water. Wanneer men echter een sferische attractie invoert van de grootte zoals bij water voorkomt, dan blijkt dat verschil grotendeels te verdwijnen.

In de loop van een computer simulatie kunnen veel details van het modelsysteem nader in beschouwing genomen worden. Deze gelegenheid is gebruikt om de klassieke roostertheorie te testen. Deze theorie gaat uit van een bepaalde structuur ('vibrationally-averaged structure') van de vloeistof met een zeker 'vrij volume' voor elk molecuul. De gemiddelde grootte van het vrije volume is bij de computer simulatie uit te rekenen en er blijkt een grote discrepantie tussen de zo berekende vrije energie en de werkelijke vrije energie te zijn. Het betreft een factor van niet minder dan 10^{5N} als er N moleculen in het systeem aanwezig zijn.

Het is dus niet mogelijk om te spreken van 'de' structuur van een vloeistof. Men moet veel meer denken aan een enorm aantal gelijkwaardige structuren.

Het is zinloos om te speculeren over geometrische rangschikkingen die meer dan vier of zes moleculen omvatten (zoals in de literatuur pleegt te worden gedaan).
Deze conclusie sluit aan bij wat verkregen wordt uit physische metingen. De roentgendiffractie bijvoorbeeld verschaft informatie over korte afstand ordening, maar zegt niets over eventuele rangschikkingen van grotere afmetingen.

LEVENSBERICHT

Willem Bol werd in 1924 op Charlois (Rotterdam) geboren en was van het begin af aan een dromer.
Het gevolg was een duidelijke spanning, die in zijn latere leven een belangrijke rol is gaan spelen. Immers, wie opgroeit in Rotterdam ontmoet weinig waardering voor bespiegelingen, maar leert dat het in de eerste plaats op daden aan komt. De wens om de twee manieren om het leven te leven, bespiegelend en bedrijvend, met elkaar in overeenstemming te kunnen brengen is altijd gebleven.

Zo was het een aangename ontdekking, te vernemen, dat men scheikundig ingenieur kan worden. Want scheikunde houdt de belofte in zich van het dieper inzicht, terwijl het ingenieur zijn constructieve creativiteit veronderstelt. De studie voor scheikundig ingenieur werd in 1950 voltooid. De keuze van een werkkring was niet moeilijk. Philips in Eindhoven was een industrie die belangrijk bijdroeg aan het herstel van de welvaart, dat juist moeizaam op gang was gekomen en anderzijds een industrie die een grote faam had op het gebied van fundamenteel onderzoek.

Na zestien jaar bij Philips, zes jaar op het Natuurkundig Laboratorium en tien jaar als produktontwikkelaar bij de fabrikage van kondensatoren, werd duidelijk dat een verdere carrière in de industrie steeds verder van de theoretische verdieping zou moeten leiden.

Daarom werd in 1966 de stap naar de Technische Hogeschool Eindhoven gedaan. Daar deed het vloeistofonderzoek een krachtig beroep op zijn creativiteit in het theoretische vlak, terwijl het onderwijs en met name het begeleiden van praktijkstages van studenten zijn industriële interesses tot hun recht deed komen.

Het in dit proefschrift behandelde onderwerp leefde al zeer lang in zijn gedachten. Op ongeveer achtjarige leeftijd was de jeugdige Willem Bol tot de konklusie gekomen dat water geen kontinuum kon zijn, maar uit deeltjes zou moeten bestaan.

De meester op school had verteld van het bestaan van microscopen, instrumenten waarmee men in een druppel water allerlei wezentjes kon zien. Wezentjes, die te klein waren om met het blote oog waar te nemen.
Terwijl je met het blote oog toch al zulke kleine dingen kunt zien.

Daarvan uitgaande is de volgende stap naar een submicroscopische wereld licht gezet.
Staande aan de oever van de Nieuwe Maas, op het Charloisse Hoofd, is het idee geboren van de vloeistof die in laatste instantie uit deeltjes zou bestaan. Deeltjes, die te klein zijn om met het microscoop te kunnen zien.

Met die theorie kon een belangrijke moeilijkheid worden opgelost: het water stroomt namelijk moeiteloos om een hindernis, een ducdalf of een steigerpaal heen. Daarbij splitst het zich en even later herenigt het zich weer zonder ook maar een spoor van een scheurtje of anderszins verandering van eigenschappen te vertonen.

Zoals de lezer gemakkelijk kan verifiëren kan men zich niet indenken, hoe een echt kontinuum dat zou kunnen doen. Met een deeltjesvloeistof daarentegen is dat alles juist vanzelfsprekend.

De deeltjes waren rond, grauwgroen van kleur, glibberig en gleden beweeglijk langs elkaar heen.

Het is curieus dat in diezelfde tijd aan de andere kant van de Noordzee Bernal en Fowler hun baanbrekende gedachten over hetzelfde onderwerp aan het formuleren waren *).
Zij het met meer kennis van zaken.

Het moet wel zo zijn dat hun gedachten over het water meegevoerd zijn door de wind en dat flarden ervan het Charloisse Hoofd bereikt hebben.

---

*) J.D. Bernal & R.H. Fowler, J.Chem.Phys. 1, 515 (1933).

D A N K W O O R D

Graag wil ik dank brengen aan allen die bijgedragen hebben aan het tot stand komen van dit proefschrift.

In het bijzonder aan mijn vriend Maarten Donkersloot. De vele diskussies, die wij gevoerd hebben zijn van wezenlijk belang geweest.

STELLINGEN

1. Bij röntgendiffractie van vloeistoffen moet ten zeerste worden aangeraden om, waar mogelijk, de gebruikelijke schaling op basis van de verstrooiing bij grote hoeken aan te vullen met een schaling op basis van de nulhoek-verstrooiing.

2. De methode van Sarkisov, Dashevsky en Malenkov voor het berekenen van de vrije energie uit Monte Carlo simulaties is ondeugdelijk.

   G.N. Sarkisov, V.G. Dashevsky & G.G. Malenkov, Mol.Phys. 27, 1249 (1974).

3. In het licht van de resultaten die de laatste tien jaar bereikt zijn met de computersimulatie van water is de controverse tussen Walrafen en anderen (zoals Schiffer) over het "two-state model" versus het "continuum model" als opgelost te beschouwen zonder dat het ongelijk van een van beide partijen kan worden vastgesteld.

   G.E. Walrafen, J.Chem.Phys. 50, 567 (1969).
   J. Schiffer, J.Chem.Phys. 50, 566 (1969).

4. Gezien de toekomstige ontwikkelingsmogelijkheden van hout als energiebron is het nodig dat in statistieken en andere publicaties die van belang zijn voor het energiebeleid, het brandhout opgenomen wordt als gelijkwaardig met olie, aardgas en steenkool.

   Vergelijk bijvoorbeeld "OECD Energy Balances 1973-75", Organisation for Economic Co-operation and Development, Paris 1977,
   met:
   Yearbook of Forest Products 1977, FAO. Rome 1979, pag 87, 88.

5. Voor het vergemakkelijken van het transport van biomassa, bestemd voor energieproductie, moet overwogen worden het materiaal te brengen in de vorm van houtskoolbriketten, gestapeld op pellets.

   J.T. Wassink, "Houtskool", Kon. Inst. v.d. Tropen, November 1974.

6. De stelling van Hartog:

   > "..... dat sociale lasten nu eenmaal kosten zijn die teweeggebracht worden door het inzetten van de produktiefactor arbeid. Elke andere heffingswijze dan koppeling aan het loon versluiert de juiste kostenverhoudingen, trekt dus de calculatie scheef en benadeelt daardoor de welvaart. Sociale lasten behoren krachtens deze gedachtengang op het loon te drukken en nergens anders op."

   is in het algemeen voor discussie vatbaar.
   Maar voor zover het de premie voor weduwen- en wezenvoorzieningen betreft, is deze stelling een onaanvaardbare overdrijving van het risico van productieve arbeid en voor zover het de premie voor werkloosheidsvoorziening betreft, is hij apert onjuist.

   F. Hartog, NRC. Handelsblad, 31 Jan 1979.

7. De mogelijkheid om algemene kosten (sociale lasten of belasting) te verschuiven van de productiefactor arbeid naar de productiefactor kapitaal (productieve investeringen) kan een bruikbaar beleidsinstrument zijn om de werkgelegenheid te beheersen. Men verlaagt op die wijze de kosten die voor arbeid in rekening gebracht moeten worden zonder direct macro-economisch effect.

8. De vraag van Pen "Is macro-economie een wetenschap?" zou met meer recht positief kunnen worden beantwoord als met behulp van computersimulatie de relatie tussen micro-economie en macro-economie zou worden bestudeerd zoals in dit proefschrift de relatie tussen moleculaire en macroscopische eigenschappen van water werden bestudeerd.

   J. Pen en L.J. v. Gemerden, "Macro-Economie", Aula 612 Het Spectrum, Utrecht 1977

9. De instelling van het "jubeljaar", zoals in de bijbel beschreven (Leviticus 25:23-34.), houdt een relativering van de privé eigendom in die nadere bestudering verdient.

10. Het verdient aanbeveling de benaming - en daarmee de doelstelling - van gemeentelijke plantsoenendiensten te veranderen in bijvoorbeeld "Dienst voor het Beheer van de Stedelijke Flora en Fauna".

11. Spelen in de loterij impliceert kapitalistische aspiraties. Wie derhalve er een gewoonte van maakt in de loterij te spelen kan niet staand houden dat hij wars is van het kapitalisme.

12. Omdat het effect van het aan huis bezorgen van goederen even gunstig is als het effect van openbaar vervoer van personen, dient deze vorm van dienstverlening in gelijken mate als het openbaar vervoer bevorderd te worden.

13. De Nederlandse vertaling van het engelse "quartz" is kwarts en niet quartz, zoals de handelaars in horloges suggereren.

Eindhoven, 18 september 1979 W. Bol