JOURNAL OF COMPUTATIONAL PHYSICS 7, 385–393 (1971)

The Statistics of the Volumes Covered by Systems of Penetrating Spheres

T. W. MELNYK AND J. S. ROWLINSON

Department of Chemical Engineering and Chemical Technology, Imperial College of Science and Technology, London S. W. 7., England

Received November 10, 1970

A recently proposed model of the liquid–vapour phase transition has a configurational energy specified by the volume W covered by N freely penetrating spheres, each of volume v_0, when placed in a volume V. In this paper Monte Carlo computations are used to calculate the probability distribution of the ratio (W/V) when the spheres are placed at random. Distributions are reported for values of the density (Nv_0/V) from 1.0–5.0 in three-dimensional systems, and from 0.5–2.0 in two-dimensional systems, i.e., for disks placed on a plane. In all cases the distributions are close to normal, but there is a small asymmetry which leads to negative values of the third moment.

INTRODUCTION

Widom and Rowlinson [1] recently described a new model for the study of vapour–liquid phase transitions. In this model, the potential energy $U(\mathbf{r}_1 \cdots \mathbf{r}_N)$ is associated with a configuration of a system of N molecules confined to a volume V. About the center of each molecule, a sphere of radius σ (volume v_0) is drawn. The resulting configuration of N, in general, interpenetrating spheres covers a volume $W(\mathbf{r}_1 \cdots \mathbf{r}_N)$ and, because of the interpenetration, $W \leqslant Nv_0$. U is now defined by

$$U(\mathbf{r}_1 \cdots \mathbf{r}_N) = [W(\mathbf{r}_1 \cdots \mathbf{r}_N) - Nv_0]\,\epsilon/v_0 \qquad \text{(WR 3.1)}^1, \tag{1}$$

where $\epsilon > 0$ is an arbitrary energy parameter.

A set of dimensionless variables, ρ the density, θ the reciprocal temperature, π the pressure to temperature ratio, and ϕ the potential energy density are defined by [(WR 2.1)]:

$$\begin{aligned} \rho &= Nv_0/V, \\ \theta &= \epsilon/kT, \\ \pi &= Pv_0/kT, \\ \phi &= v_0\overline{U}/\epsilon V, \end{aligned} \tag{2}$$

[1] Equation (3.1) etc. of Ref. [1] is here denoted (WR 3.1) etc.

where k is Boltzmann's constant, T is the temperature, and P is the pressure of the system and $\overline{U} = \overline{U(\mathbf{r}_1 \cdots \mathbf{r}_N)}$ is the canonical average of $U(\mathbf{r}_1 \cdots \mathbf{r}_N)$. In these variables (1) becomes

$$\phi + \rho = \overline{W}/V. \tag{3}$$

In general, there is no analytical method for finding each $W(\mathbf{r}_1 \cdots \mathbf{r}_N)$ from a knowledge of $\mathbf{r}_1 \cdots \mathbf{r}_N$, or for finding the statistics of its distribution.

Random Distribution Approximation

If the molecules occupy positions at random (the limiting case expected at very high temperatures $\theta \simeq 0$) the mean value $\langle W \rangle$ is easily found. (Brackets $\langle\ \rangle$ indicate an average over the random configuration.) In the thermodynamic limit $N \to \infty$, $V \to \infty$, N/V constant:

$$\langle W \rangle / V = 1 - e^{-\rho}. \tag{4}$$

Through thermodynamic fluctuation theory ϕ can be expanded as a power series in θ giving

$$\phi + \rho = \sum_{n=0}^{\infty} \frac{(-)^n K_{n+1} \theta^n}{n!\, V v_0{}^n}, \tag{5}$$

where K_n is the nth cumulant of the random distribution of $W(\mathbf{r}_1 \cdots \mathbf{r}_N)$. The first cumulant $K_1 = \langle W \rangle$ is the mean of the distribution and is given by (4). Widom and Rowlinson showed that $K_2 = \langle W^2 \rangle - \langle W \rangle^2$ is given by [(WR 7.19)]

$$K_2 / V v_0 = [2b_2(\rho) - \rho]\, e^{-2\rho}, \tag{6}$$

where $b_2(x)$ is a dimensionless second cluster integral defined for dimensionality s by [(WR 6.3)]

$$v_0 b_2(x) = \frac{s}{2} a_s \int_0^{2\sigma} r^{s-1} \left[\exp\left(\frac{-x[W(r) - 2v_0]}{v_0} \right) - 1 \right] dr, \tag{7}$$

where σ is the radius of the sphere of volume v_0, a_s is the volume of an s-dimensional sphere of unit radius and $W(r)$ is the volume covered by two interpenetrating spheres each of volume v_0 when their centers are separated by distance r [(WR 6.4)]

$$W(r) = v_0 + a_{s-1} \int_0^r (\sigma^2 - \tfrac{1}{4}x^2)^{(s-1)/2}\, dx. \tag{8}$$

By expansion, in three dimensions

$$b_2(x) = \sum_{n=1}^{\infty} \sum_{m=0}^{n} \frac{3(2n)!\,(m+1)(m+2)\,x^n}{2^{m-2}(n-m)!\,(2n+m+3)!}. \tag{9}$$

In two dimensions

$$b_2(x) = -2 + (e^x + 1) \sum_{n=1,3,5}^{\infty} np^{n-1} \left[\prod_{i=1,3,5}^{n} (p^2 + i^2) \right]^{-1} + (e^x - 1) \sum_{n=2,4,6}^{\infty} np^{n-2} \left[\prod_{i=2,4,6}^{n} (p^2 + i^2) \right]^{-1}, \tag{10}$$

where $p = x/\pi$. Hence, Eq. (5) becomes [(WR 7.17)]

$$\phi = 1 - e^{-\rho} - \rho - \theta[2b_2(\rho) - \rho]\, e^{-2\rho} + O(\theta^2) \tag{11}$$

with $b_2(\rho)$ defined by (9). In order to determine the higher cumulants K_3, K_4, etc. required to extend (11) to higher powers of θ, a Monte Carlo computation was made to evaluate the volumes $W(\mathbf{r}_1 \cdots \mathbf{r}_N)$ of a large number of configurations of N randomly placed spheres in a volume V and of randomly placed circular disks on a plane.

Higher order cumulants can be expressed in terms of higher order cluster integrals by a procedure similar to that used by Widom and Rowlinson for K_2. Thus

$$K_3/Vv_0^2 = \{\rho - \tfrac{3}{2}\rho^2 - 3\rho[b_2'(\rho) - 2b_2(\rho)] - 3b_3(\rho)\}\, e^{-3\rho}, \tag{12}$$

where the prime denotes differentiation. (This expression has been obtained independently by T. J. Lie and B. Widom.) These higher cluster integrals can themselves be obtained only by numerical integration.

Other Applications

The problem of overlapping figures has been of independent statistical interest and is described in recent reviews by Moran [2].

The one-dimensional problem is encountered in the calculation of the dead-time correction to the counting rates of radioactive particles and presents few problems. The two-dimensional version arises in such problems as the coverage of a plane by impinging particles of aerosols, and was studied during the second world war by those interested in predicting bombing patterns. Garwood [3] made some experiments on the areas covered by overlapping disks on a plane. He found a

mean close to that given by Eq. (4) and obtained an empirical expression for the standard deviation:

$$K_2 = 2{\cdot}3\frac{\langle W\rangle}{V}\left[1 - \frac{\langle W\rangle}{V}\right]v_0^{1{\cdot}5}. \tag{13}$$

COMPUTATION

A system of N molecules (usually 200) is confined within a unit cube. The coordinates defining the position of each molecule are assigned by a random number generator giving rectangularly distributed random numbers between 0.0 and 1.0. In order to simulate as closely as possible the behavior of a large system, a periodic boundary condition is used, the basic cell being surrounded by replicas of itself. The value of v_0 (and hence of σ) is determined from Eq. (2) from the assigned values for ρ and N.

A Monte Carlo calculation may be performed by placing random test points inside the volume V and enquiring whether or not each test point lies inside the volume W. A count is kept of the total number of test points and of the number inside W. In the limit of a large number of test points the ratio of the number of test points inside W to the total number approaches (W/V).

Since the spheres are placed randomly in the volume V it is unnecessary here to use a random set of test points. Similar convergence to the proper distribution is expected from a fixed set of test points and a randomly moving W, and so a regular grid of $25 \times 25 \times 25$ test points uniformly filling the volume V is used (100×100 in two dimensions).

Instead of calculating the volume of a completely new configuration of N molecules each time, only one molecule is moved to a new random position in order to generate the next configuration. The operation is then repeated with each molecule in turn. It is found from sampling the set of volumes so obtained that the values cross from one side of the mean to the other on average every 25 configurations. If the configurations were totally independent a crossing would be expected on average every 2 configurations. This procedure thus slows the convergence by a factor of not more than 25, with a saving in machine time of a factor of 200, giving a net saving of a factor of 8–10 in machine time.

The apparent volume of each sphere, as assessed by the number of test points it covers, could vary by ± 10–$15\,\%$ of v_0. As well as (W/V), the value of the volume of each sphere is calculated and hence a value of Nv_0 is obtained independently of any overlap. The value of W/V is corrected for any over- or underestimation by assuming that the correction factor is proportionally the same as that required to bring the calculated value of Nv_0 to its correct value of ρV. This correction is always less than $0.3\,\%$ and typically less than $0.1\,\%$.

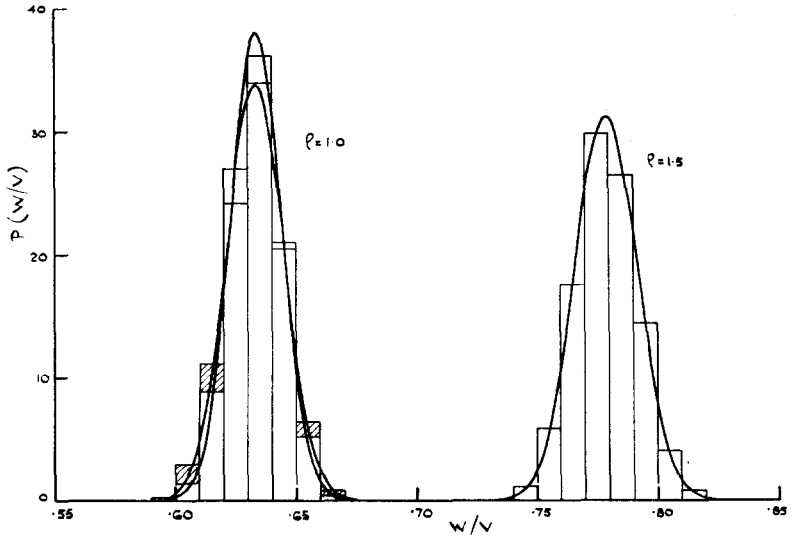

FIG. 1. Histograms and normal curves for the probability density in three dimensions at densities of $\rho = 1.0$ and 1.5.

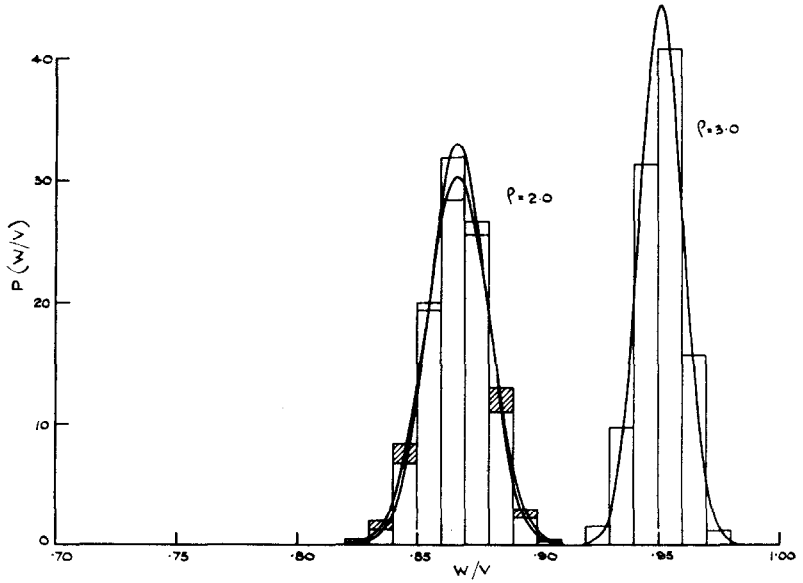

FIG. 2. Histograms and normal curves for the probability density in three dimensions at densities of $\rho = 2.0$ and 3.0.

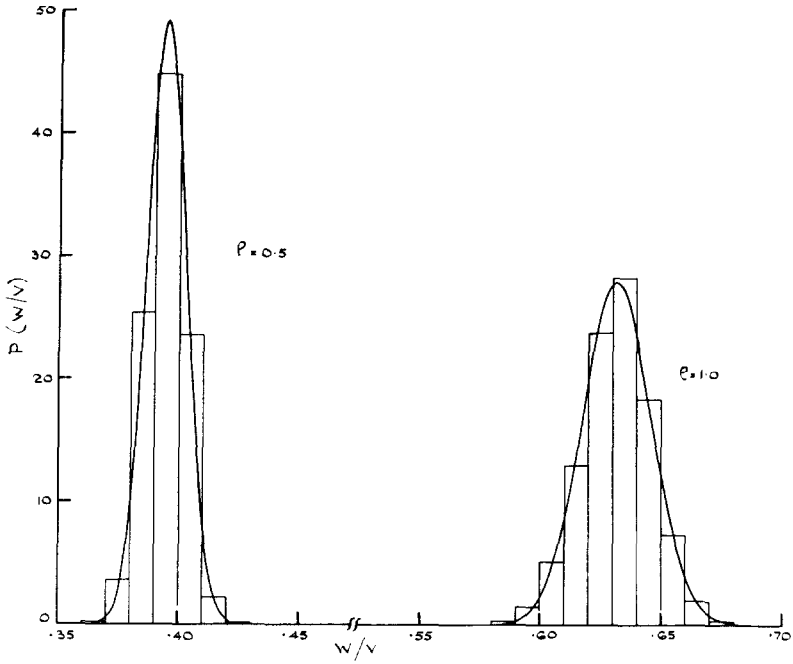

FIG. 3. Histograms and normal curves for the probability density in two dimensions at densities of $\rho = 0.5$ and 1.0.

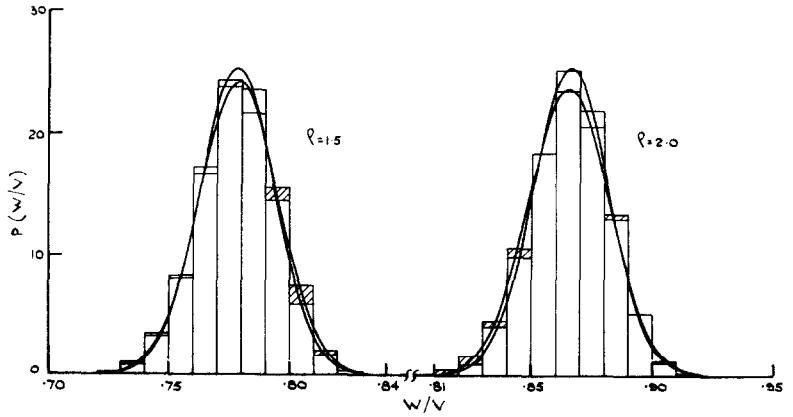

FIG. 4. Histograms and normal curves for the probability density in two dimensions at densities of 1.5 and 2.0.

At low densities a number of configurations have volume W exactly equal to Nv_0 (i.e., there is no overlap), their frequency forming a δ-function in the histogram at this value, but as ρ (as hence Nv_0) increases this δ-function rapidly becomes negligible and disappears. At high densities a similar δ-function is expected at $W/V = 1$ when the spheres completely fill the volume V. In an infinite thermodynamic system these δ-functions are absent.

RESULTS

In three dimensions two sets of 50 000 configurations were generated at densities of 1.0 and 2.0, one set of 50 000 configurations at a density of 1.5 and one set of 10 000 configurations at densities of 3.0, 4.0, and 5.0. These densities are too high for the δ-functions at $W = Nv_0$ to appear in the histogram. The δ-function at $W/V = 1$ just appears at $\rho = 5.0$, with 9 of the 10 000 configurations having this value of W/V.

The histograms are normalized to cover unit area and are shown in Fig. 1 and 2 as probability densities.

TABLE I

Three Dimensional Results

ρ	1.0	1.5	2.0	3.0	4.0	5.0
Number of Configurations	50 000[a]	50 000	50 000[a]	10 000	10 000	10 000
$\langle W\rangle/V$ [Eq. (4)]	0.6321	0.7769	0.8647	0.9502	0.9817	0.9933
$\langle W\rangle/V$ [Eq. (14)]	0.6330	0.7781	0.8660	0.9513	0.9824	0.9937
$K_2 = \mu_2\ (\times 10^3)$ [Eq. (6)]	0.1286	0.1743	0.1670	0.0938	0.0381	0.0132
$\langle W\rangle/V$	0.6334	0.7788	0.8669	0.9513	0.9824	0.9941
	0.6330		0.8667			
$K_2 = \mu_2\ (\times 10^3)$	0.110	0.164	0.146	0.0811	0.0389	0.0109
	0.139		0.173			
$K_3 = \mu_3\ (\times 10^6)$	−0.07	+0.03	−0.14	−0.22	−0.13	−0.04
	−0.21		−0.37			
$\mu_4\ (\times 10^7)$	0.35	0.79	0.67	0.21	0.05	0.05
	0.56		0.86			
$K_4 = \mu_4 - 3\mu_2^2\ (\times 10^7)$	−0.01	−0.02	+0.03	+0.01	0.00	+0.05
	−0.02		−0.04			
$\mu_5\ (\times 10^6)$	−0.08	0.00	−0.27	−0.17	−0.04	0.00
	−0.22		−0.57			
$\mu_6\ (\times 10^{10})$	0.18	0.61	0.56	0.09	0.01	0.00
	0.35		0.70			

[a] Two sets of 50 000.

TABLE II

Two Dimensional Results

ρ	0.5	1.0	1.5	2.0
Number of Configurations	50 000	50 000	50 000[a]	50 000[a]
$\langle W\rangle/V$ [Eq. (4)]	0.3935	0.6321	0.7769	0.8647
$\langle W\rangle/V$ [Eq. (14)]	0.3938	0.6330	0.7781	0.8660
$K_2 = \mu_2\,(\times 10^3)$ [Eq. (6)]	0.0584	0.1911	0.2660	0.2622
$K_2 = \mu_2(\times 10^3)$ [Eq. (13)]	0.0686	0.1891	0.2589	0.2691
$\langle W\rangle/V$	0.3945	0.6317	0.7788	0.8655
			0.7781	0.8663
$K_2 = \mu_2\,(\times 10^3)$	0.0657	0.204	0.271	0.283
			0.247	0.244
$K_3 = \mu_3\,(\times 10^6)$	−0.05	−0.41	−0.52	−1.15
			−0.40	−0.46
$\mu_4\,(\times 10^6)$	0.01	0.13	0.21	0.25
			0.19	0.18
$K_4 - \mu_4 - 3\mu_2^2\,(\times 10^6)$	0.00	$+0.00_5$	-0.00_3	$+0.00_6$
			$+0.00_3$	$+0.00_3$
$\mu_5\,(\times 10^9)$	−0.02	−0.76	−1.45	−3.42
			−0.93	−1.12
$\mu_6\,(\times 10^9)$	0.004	0.14	0.28	0.40
			0.23	0.21

[a] Two sets of 50 000.

In two dimensions, two sets of 50 000 configurations were generated at densities of 1.5 and 2.0 and one set of 50 000 configurations at densities of 0.5 and 1.0. These histograms are shown in Figs. 3 and 4. Tables I and II show the resulting means and moments of the distributions.

Equation (4) gives the correct mean $\langle W\rangle$ only for an infinite thermodynamic system. That for a finite system of N molecules is

$$\langle W\rangle = 1 - \left(1 - \frac{\rho}{N}\right)^N. \tag{14}$$

The variance is also expected to differ slightly from that given by Eq. (6) but the exact effect of the finite system with periodic boundaries on the variance is not known.

For each histogram a normal curve with the experimental mean and variance is shown in each of Figs. 1–4. In all cases the distribution is apparently represented by this normal curve. The experimental mean is very close to that given by Eq. (14) and hence the mean of the thermodynamic system is given by Eq. (4). The experi-

mental variances are not as precise as the means but do not differ greatly from those predicted by Eq. (6). For a normal distribution, all cumulants K_n, with $n \geqslant 3$ are identically zero. The experimental higher cumulants (K_n for $n > 3$) are all one or two orders of magnitude less than the corresponding moments (except perhaps for $\rho = 5.0$). Thus Eq. 5 becomes simply

$$\phi = 1 - e^{-\rho} - \rho - \theta[2b_2(\rho) - \rho]\, e^{-2\rho}. \tag{15}$$

Equation (15) is to be compared with the complete expansions of Eqs. (11) and (5). The expression to first order in θ is thus a much better approximation than (11) was initially thought to be. The higher order terms correct for the slight deviations of the random distribution of W/V from the normal distribution, and in all but one of the runs the third cumulant $K_3 = \mu_3$ is negative. These numbers are evidence of a small but probably real asymmetry in the distributions which is stronger in two dimensions than three. $K_4 = \mu_4 - 3\mu_2^2$ is always less than the experimental error.

Since higher order cumulants are almost zero, approximate expressions for the higher cluster integrals can be obtained from equations such as (12). Thus we expect that in three dimensions

$$3b_3(x) \simeq x - \tfrac{3}{2}x^2 - 3x[b_2'(x) - 2b_2(x)]. \tag{16}$$

ACKNOWLEDGMENTS

The authors are grateful to Imperial College and the University of London for a generous allotment of time on the CDC 6600 computer and to Dr. K. Singer and his colleagues for several helpful discussions. One of us (T.W.M.) acknowledges the financial support of Shell Canada Limited.

REFERENCES

1. B. WIDOM AND J. S. ROWLINSON, *J. Chem. Phys.* **52** (1970), 1670.
2. P. A. P. MORAN, *J. Appl. Prob.* **3** (1966), 453; *Adv. Appl. Prob.* **1** (1969), 73.
3. F. GARWOOD, *Biometrika* **23** (1947), 1.