"Equation of State Calculations by Fast Computing Machines"

Nicholas Metropolis, Arianna W. Rosenbluth, and Marshall N. Rosenbluth

Los Alamos Scientific Laboratory, Los Alamos, New Mexico

and

Augusta H. Teller* and Edward Teller†

Department of Physics, University of Chicago, Chicago, Illinois

* Now at Los Alamos Scientific Laboratory, Los Alamos, New Mexico

† Now at the Radiation Laboratory of the University of California, Livermore, California

---

Abstract

A general method, suitable for fast computing machines, for investigating such properties as equations of state for substances consisting of interacting individual molecules is described. The method consists of a modified Monte Carlo integration over configuration space. Results for the two-dimensional rigid sphere system have been obtained on the Los Alamos Maniac and are presented here. These results are compared to the free volume equation of state and to a four term virial coefficient expansion.

I. Introduction

The purpose of this paper is to describe a general method, suitable for fast electronic computing machines, of calculating the properties of any substance which may be considered as composed of interacting individual molecules. Classical statistics is assumed, only two-body forces are considered, and the potential field of a molecule is assumed spherically symmetric. These are the usual assumptions made in theories of liquids. Subject to the above assumptions, the method is not restricted to any range of temperature or density. This paper will also present results of a preliminary two-dimensional calculation for the rigid sphere system. Work on the two-dimensional case with a Lennard-Jones potential is in progress and will be reported in a later paper. Also, the problem in three dimensions will be investigated.

II. The General Method for an Arbitrary Potential between the Particles

In order to reduce the problem to a feasible size for numerical work we can, of course, consider only a finite number of particles. This number N may be as high as several hundred. Our system consists of a square* containing N particles. In order to minimize the surface

---

*We will use the two-dimensional nomenclature here since it is easier to visualize. The extension to three dimensions is obvious.

---

effects we suppose the complete substance to be periodic, consisting of many such squares, each square containing N particles in the same configuration. Thus we define $\dot{d}_{AB}$, the minimum distance between particles A and B, as the shortest distance between A and any of the particles B, of which there is one in each of the squares which comprise the complete substance. If we have a potential which falls off rapidly

with distance, there will be at most one of the distances AB which can make a substantial contribution; hence we need consider only the minimum distance d_{AB}.

Our method in this respect is similar to the cell method except that our cells contain several hundred particles instead of one. One would think that such a sample would be quite adequate for describing any one-phase system. We do find, however, that in two-phase systems the surface between the phases makes quite a perturbation. Also, statistical fluctuations may be sizable.

If we know the positions of the N particles in the square, we can easily calculate, for example, the potential energy of the system.

$$E = 1/2 \sum_{\substack{i=1 \\ i \neq j}}^{N} \sum_{j=1}^{N} V(d_{ij}) \qquad (1)$$

(Here V is the potential between molecules, and d_{ij} is the minimum distance between particles i and j as defined above.)

In order to calculate the properties of our system we use the canonical ensemble. So, to calculate the equilibrium value of any quantity of interest F,

$$\bar{F} = \left[\int F \exp(-E/kT)\, d^{2N}p\, d^{2N}q\right] \Big/ \left[\int \exp(-E/kT)\, d^{2N}p\, d^{2N}q\right] \qquad (2)$$

where $(d^{2n}p\ d^{2n}q)$ is a volume element in the 4N-dimensional phase space. Moreover, since forces between particles are velocity independent, the momentum integrals may be separated off and we need perform only the integration over the 2N-dimensional configuration space. It is evidently impractical to carry out a several hundred dimensional integral by the usual numerical methods, so we resort to the

Monte Carlo method.* The Monte Carlo method for many dimensional

---

* This method has been proposed independently by J.E. Mayer and by S. Ulam. Mayer suggested the method as a tool to deal with the problem of the liquid state, while Ulam proposed it as a procedure of general usefulness. B. Alder, J. Kirkwood, S. Frankel, and V. Lewinson discussed an application very similar to ours.

---

integrals consists simply of integrating over a random sampling of points instead of over a regular array of points.

Thus the most naive method of carrying out the integration would be to put each of the N particles at a random position in the square (this defines a random point in the 2N-dimensional configuration space), then calculate the energy of the system according to equation (1), and give this configuration a weight exp.(-E/kT). This method, however, is not practical for close-packed configurations, since with high probability we choose a configuration where exp (-E/kT) is very small, hence a configuration of very low weight. So the method we employ is actually a modified Monte Carlo sc heme, where, instead of choosing configurations randomly, then weighting them with exp (-E/kT), we choose configurations with a probability exp (-E/kT), and weight them evenly.

This we do as follows: We place the N particles in any configuration, for example, in a regular lattice. Then we move each of the particles in succession according to the following prescription:

$$X \longrightarrow X + \alpha \xi_1 \qquad (3)$$
$$Y \longrightarrow Y + \alpha \xi_2$$

where α is the maximum allowed displacement, which for the sake of this argument is arbitrary, and ξ_1 and ξ_2 are random numbers*

*It might be mentioned that the random numbers that we used were generated by the middle square process. That is, if ξ_n is an m-digit random number, then a new random number, ξ_{n+1}, is given as the middle m digits of the complete 2m-digit square of ξ_n.

between (-1) and 1. Then after we move a particle it is equally likely to be anywhere within a square of side 2α centered about its original position. (In accord with the periodicity assumption, if the indicated move would put the particle outside the square, this only means that it reenters the square from the opposite side.)

We then calculate the change in energy of the system, ΔE, which is caused by the move. If $\Delta E < 0$, i.e., if the move would bring the system to a state of lower energy, we allow the move and put the particle in its new position. If $\Delta E > 0$, we allow the move with a probability exp $(-\Delta E/kT)$; i.e., we take a random number ξ_3 between 0 and 1, and if $\xi_3 < \exp(-\Delta E/kT)$, we move the particle to its new position. If $\xi_3 > \exp(-\Delta E/kT)$, we return it to its old position. Then, whether the move has been allowed or not, i.e., whether we are in a different configuration or in the original configuration, we consider that we are in a new configuration for the purpose of taking our averages. So

$$\overline{F} = (1/M) \sum_{j=1}^{M} F_j \qquad (4)$$

where F_j is the value of the property F of the system after the j^{th} move is carried out according to the complete prescription above. Having attempted to move a particle we proceed similarly with the next one.

We now prove that the method outlined above does choose configurations with a probability exp (-E/kT). Since a particle is allowed to move to any point within a square of side 2α with a finite probability, it is clear that a large enough number of moves will enable it to reach any point in the complete square.* Since

*In practice it is, of course, not necessary to make enough moves to allow a particle to diffuse evenly throughout the system since configuration space is symmetric with respect to interchange of particles.

this is true of all particles we may reach any point in configuration space. Hence, the method is ergodic.

Next consider a very large ensemble of systems. Suppose for simplicity that there are only a finite number of states** of the

**A state here means a given point in configuration space.

system, and that ν_r is the number of systems of the ensemble in state r. What we must prove is that after many moves the ensemble tends to a distribution

$$\nu_i \propto \exp(-E_i/kT)$$

Now let us make a move in all the systems of our ensemble. Let the a priori probability that the move will carry a system in state r to state s be P_{rs}. (By the a priori probability we mean the probability before discriminating on exp (-ΔE/kT).) First it is clear that $P_{rs} = P_{sr}$, since according to the way our game is played a particle is equally likely to be moved anywhere within a square of side 2α centered about its original position. Thus, if states r and s differ from each other only by the position of the particle moved, and if these positions are within each other's squares, the transition probabilities

are equal; otherwise they are zero. Assume $E_r > E_s$. Then the number of systems moving from state r to state s will be simply $\nu_r P_{rs}$, since all moves to a state of lower energy are allowed. The number moving from s to r will be $\nu_s P_{sr} \exp(-(E_r - E_s)/kT)$, since here we must weight by the exponential factor. Thus the net number of systems moving from s to r is

$$P_{ij}\left(\nu_j \exp(-(E_i - E_j)/kT) - \nu_i\right) \tag{5}$$

So we see that between any two states r and s, if

$$(\nu_i/\nu_j) > [\exp(-E_i/kT)/\exp(-E_j/kT)] \tag{6}$$

on the average more systems move from state r to state s. We have seen already that the method is ergodic; i.e., that any state can be reached from any other, albeit in several moves. These two facts mean that our ensemble must approach the canonical distribution. It is incidentally clear from the above derivation that after a forbidden move we must count again the initial configuration. Not to do this would correspond in the above case to removing from the ensemble those systems which tried to move from s to r and were forbidden. This would unjustifiably reduce the number in state s relative to r.

The above argument does not, of course, specify how rapidly the canonical distribution is approached. It may be mentioned in this connection that the maximum displacement α must be chosen with some care; if too large, most moves will be forbidden, and if too small, the configuration will not change enough. In either case it will then take longer to come to equilibrium.

For the rigid sphere case, the game of chance on exp (-ΔE/kT) is , of course, not necessary since ΔE is either zero or infinity. The

particles are moved, one at a time, according to equation (3). If a sphere, after such a move, happens to overlap another sphere, we return it to its original position.

III. Specialization to Rigid Spheres in Two Dimensions

A. The Equation of State

The virial theorem of Clausius can be used to give an equation of state in terms of $\bar{n}$, the average density of other particles at the surface of a particle. Let $\sum_i^{(tot.)}$ and $\sum_i^{(int.)}$ represent the total and the internal force, respectively, acting on particle i, at a position $\vec{r}_i$. Then the virial theorem can be written

$$\overline{\sum_i \vec{X}_i^{(tot)} \cdot \vec{r}_i} = 2PA + \overline{\sum_i \vec{X}_i^{(int)} \cdot \vec{r}_i} = 2E_{KIN} \qquad (7)$$

Here P is the pressure, A the area, and E_{kin} the total kinetic energy,

$$E_{KIN} = N m \overline{v^2}/2$$

of the system of N particles.

Consider the collisions of the spheres for convenience as represented by those of a particle of radius d_o, twice the radius of the actual spheres, surrounded by $\bar{n}$ point particles per unit area. Those surrounding particles in an area of $2\pi d_o V \cos\phi \Delta t$, travelling with velocity V at an angle ϕ with the radius vector, collide with the central particle provided $|\phi| < \pi/2$. (See Fig. 1) Assuming elastic recoil, they each exert an average force during the time Δt on the central particle of

$$2mv\cos\phi/\Delta t$$

One can see that all ϕ's are equally probable, since for any velocity independent potential between particles the velocity distribution will

just be Maxwellian, hence isotropic. The total force acting on the central particle, averaged over ϕ, over time, and over velocity, is

$$\overline{F_i} = m\overline{v^2}\,\pi d_o\,\bar{n} \tag{8}$$

The sum $\sum_i \overrightarrow{X}_i^{(int)} \cdot \vec{r}_i$ is $-1/2 \sum_i \left\{ \sum_j r_{ij} F_{ij} \right\}$, with F_{ij} the magnitude of the force between two particles and r_{ij} the distance between them. We see that $r_{ij} = d_o$ and $\Sigma_j F_{ij}$ is given by equation (8), so we have

$$\overline{\sum_i \overrightarrow{X}_i^{(int)} \cdot \vec{r}_i} = -(N m \overline{v^2}/2)\,\pi d_o^2 \bar{n} \tag{9}$$

Substitution of (9) into (7) and replacement of $\frac{N}{2} m \overline{v^2}$ by E_{kin} gives finally

$$PA = E_{kin}(1 + \pi d_o^2 \bar{n}/2) \equiv NkT(1 + \pi d_o^2 \bar{n}/2) \tag{10}$$

This equation shows that a determination of the one quantity $\bar{n}$, according to equation (4), as a function of A, the area, is sufficient to determine the equation of state for the rigid spheres.

B. The Actual Calculation of $\bar{n}$

We set up the calculation on a system composed of N = 224 particles (i=0,1···223) placed inside a square of unit side and unit area. The particles were arranged initially in a trigonal lattice of fourteen particles per row by sixteen particles per column, alternate rows being displaced relative to each other as shown in Fig. 2. This arrangement gives each particle six nearest neighbors at approximately equal distances of d = 1/14 from it.

Instead of performing the calculation for various areas A and for a fixed distance d_o, we shall solve the equivalent problem of

leaving $A = 1$ fixed and changing d_o. We denote by A_o the area the particles occupy in closely packed arrangement (see Fig. 3). For numerical convenience we defined an auxiliary parameter ν, which we varied from zero to seven, and in terms of which the ratio (A/A_o) and the forbidden distance d_o are defined as follows:

$$d_o = d(1 - 2^{\nu-8}); \quad d = (1/14) \tag{11}a$$

$$(A/A_o) = 1/(3^{1/2} d_o^2 N/2) = 1/0.98974329(1-2^{\nu-8})^2 \tag{11}b$$

The unit cell is a parallelogram with interior angle 60°, side d_o, and altitude $3^{1/2}d_o/2$ in the close packed system.

Every configuration reached by proceeding according to the method of the preceding section was analyzed in terms of a radial distribution function $N(r^2)$. We chose a $K>1$ for each ν and divided the area between πd_o^2 and $K^2\pi d_o^2$ into sixty-four zones of equal area ΔA^2:

$$\Delta A^2 = (K^2-1)\pi d_o^2/64$$

We then had the machine calculate for each configuration the number of pairs of particles N_m $(m = 1,2\cdots64)$ separated by distances r which satisfy

$$(m-1)\Delta A^2 + \pi d_o^2 < \pi r^2 \leqslant m\Delta A^2 + \pi d_o^2 \tag{12}$$

The N_m were averaged over successive configurations according to equation (4), and after every sixteen cycles (a cycle consists of moving every particle once) were extrapolated back to $r^2 = d_o^2$ to obtain $N_{1/2}$. This $N_{1/2}$ differs from $\bar{n}$ in equation (10) by a constant factor depending on N and K.

The quantity K was chosen for each ν to give reasonable statistics for the N_m. It would, of course, have been possible by choosing fairly large K's, with perhaps a larger number of zones, to obtain $N(r^2)$ at large distances. The oscillatory behaviour of $N(r^2)$ at large distances is of some interest. However, the time per cycle goes up fairly rapidly with K, and with the number of zones in the distance analysis. For this reason only the behaviour of $N(r^2)$ in the neighborhood of d_o^2 was investigated.

The maximum displacement α of equation (3) was set to $(d-d_o)$. About half the moves in a cycle were forbidden by this choice, and the initial approach to equilibrium from the regular lattice was fairly rapid.

IV. Numerical Results for Rigid Spheres in Two Dimensions

We first ran for something less than sixteen cycles in order to get rid of the effects of the initial regular configuration on the averages. Then about forty-eight to sixty-four cycles were run at

$$\nu = 2,\ 4,\ 5,\ 5.5,\ 6,\ 6.25,\ 6.5,\ \text{and}\ 7$$

Also, a smaller amount of data was obtained at $\nu = 0$, 1, and 3. The time per cycle on the Los Alamos Maniac is approximately three minutes, and a given point on the pressure curve was obtained in four to five hours of running. Fig. 4 shows $(PA/NkT)-1$ versus $(A/A_o)-1$ on a log - log scale from our results (curve A), compared to the free volume equation of Wood[1] (curve B), and to the curve

---

[1] William W. Wood, J. Chem. Phys., 20, 1334 (1952).

---

given by the first four virial coefficients (curve C). The last two virial coefficients were obtained by straightforward Monte Carlo integration on the Maniac (see Section 5). It is seen that the agreement between curves (A) and (B) at small areas, and between curves (A) and (C) at large areas, is good. Deviation from the free volume theory begins with a fairly sudden break at $\nu = 6$ (A/A_o 1.8).

A sample plot of the radial distribution function for $\nu = 5$ is given in Fig. 5. The various types of points represent values after sixteen, thirty-two, and forty-eight cycles. For $\nu = 5$, least square fits with a straight line to the first sixteen N_m values were made, giving extrapolated values of $N_{1/2}^{(1)} = 6367$, $N_{1/2}^{(2)} = 6160$, and $N_{1/2}^{(3)} = 6377$. The average of these three was used in constructing PA/NkT. In general, least square fits of the first sixteen to twenty N_m's by means of a parabola, or, where it seemed suitable, a straight line, were made. The errors indicated in Fig. 4 are the root mean square deviations for the three or four $N_{1/2}$ values. Our average error seemed to be about 3 per cent.

Table I gives the results of our calculations in numerical form. The columns are ν, A/A_o, (PA/NkT)-1, and, for comparison purposes, (PA/NkT - 1) for the free volume theory and for the first four coefficients in the virial coefficient expansion, in that order, and finally PA_o/NkT from our results.

V. The Virial Coefficient Expansion

One can show[2] that

---

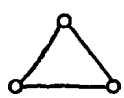
[2] J. E. Mayer and M. G. Mayer, Statistical Mechanics, pp. 277-291, John Wiley and Sons, Inc., New York, 1940.

$$(PA/NkT)-1 = C_1(A_0/A)+C_2(A_0/A)^2+C_3(A_0/A)^3+C_4(A_0/A)^4+O(A_0/A)^5,$$

$$C_1=\pi/3^{1/2}\ ,\ C_2=4\pi^2A_{3,3}/9\ ,\ C_3=\pi^3(6A_{4,5}-3A_{4,4}-A_{4,6})/3^{3/2}\ ,\ C_4=(8\pi^3/135)\cdot[\tag{13}$$

$$[12A_{5,5}-60A'_{5,6}-10A''_{5,6}+30A'_{5,7}+60A''_{5,7}+10A'''_{5,7}-30A'_{5,8}-15A''_{5,8}+10A_{5,9}-A_{5,10}]$$

The coefficients $A_{i,k}$ are cluster integrals over configuration space of i particles, with k bonds between them. In our problem a bond is established if the two particles overlap. The cluster integral is the volume of configuration space for which the appropriate bonds are established. If k bonds can be distributed over the i particles in two or more different ways without destroying the irreducibility of the integrals, the separate cases are distinguished by primes. For example, A_{33} is given schematically by the diagram

$A_{3,3}$: 

and mathematically as follows: if we define $f(r_{ij})$ by

$$f(r_{ij}) = 1 \quad \text{if} \quad r_{ij} < d,$$

$$f(r_{ij}) = 0 \quad \text{if} \quad r_{ij} > d, \text{ then}$$

$$A_{3,3} = \frac{1}{\pi^2 d^4}\int\cdots\int dx_1\, dx_2\, dx_3\, dy_1\, dy_2\, dy_3 (f_{12}f_{23}f_{31})$$

The schematics for the remaining integrals are indicated in Fig. 6.

The coefficients $A_{3,3}$; $A_{4,4}$; and $A_{4,5}$ were calculated algebraically, the remainder numerically by Monte Carlo integration. That is, for $A_{5,5}$ for example, particle 1 was placed at the origin, and particles 2, 3, 4, and 5 were put down at random, subject to $f_{12} = f_{23} = f_{34} = f_{15} = 1$. The number of trials for which $f_{45} = 1$, divided by the total number of trials, is just $A_{5,5}$.

The data on $A_{4,6}$ is quite reliable. We obtained

$$A_{4,6}/A_{4,4} = 0.752(\pm 0.002)$$

However, due to the relatively large positive and negative terms in C_4 of equation (13), the coefficient C_4, being a small difference, is less accurate. We obtained

$$C_4 = 8\pi^3(0.585)/135 \qquad (\pm \sim 5 \text{ per cent})$$

Our final formula is

$$(PA/NkT) - 1 = 1.813799(A_o/A) + 2.57269(A_o/A)^2 + \\ + 3.179(A_o/A)^3 + 3.38(A_o/A)^4 + 0(A_o/A)^5 \qquad (14)$$

This formula is plotted in curve (C) of Fig. 4 and tabulated for some values of (A/A_o) in column 5 of Table I. It is seen in Fig. 4 that the curve agrees very well with our calculated equation of state for $(A/A_o) > 2.5$. In this region both the possible error in our last virial coefficient and the contribution of succeeding terms in the expansion are quite small (less than our probable statistical error) so that the virial expansion should be accurate.

VI. Conclusion

The results are in agreement with the free volume approximation for $A/A_o < 1.8$ and with a four term virial expansion for $A/A_o > 2.5$. No indication of a phase transition appears.

Table I

Results of this calculation for $(PA/NkT) - 1 = X_1$ compared to the free volume theory (X_2) and the four-term virial expansion (X_3). Also (PA_o/NkT) from our calculations.

ν	(A/A_o)	X_1	X_2	X_3	(PA_o/NkT)
2	1.04269	49.17	47.35	9.77	48.11
4	1.14957	13.95	13.85	7.55	13.01
5	1.31966	6.43	6.72	5.35	5.63
5.5	1.4909	4.41	4.53	4.02	3.63
6	1.7962	2.929	2.939	2.680	2.187
6.25	2.04616	2.186	2.323	2.065	1.557
6.5	2.41751	1.486	1.802	1.514	1.028
7	4.04145	0.6766	0.990	0.667	0.4149

Captions for the Figures

Fig. 1 Collisions of rigid spheres

Fig. 2 Initial trigonal lattice

Fig. 3 The close packed arrangement for determining A_o

Fig. 4 A plot of $(PA/NkT) - 1$ versus $(A/A_o) - 1$. Curve A (solid line) gives the results of this paper. Curves B and C (dashed and dot-dashed lines) give the results of the free volume theory and of the first four virial coefficients, respectively.

Fig. 5 The radial distribution function, N_m, for $\nu = 5$, $(A/A_o) = 1.31966$, $K = 1.5$. The average of the extrapolated values of $N_{1/2}$ in $\bar{N}_{1/2} = 6301$. The resultant value of $(PA/NkT) - 1$ is $64\bar{N}_{1/2}/N^2(K^2-1)$ or 6.43. Values after 16 cycles: • ; after 32: ✕ ; and after 48: ⊙

Fig. 6 Schematic diagrams for the various area integrals

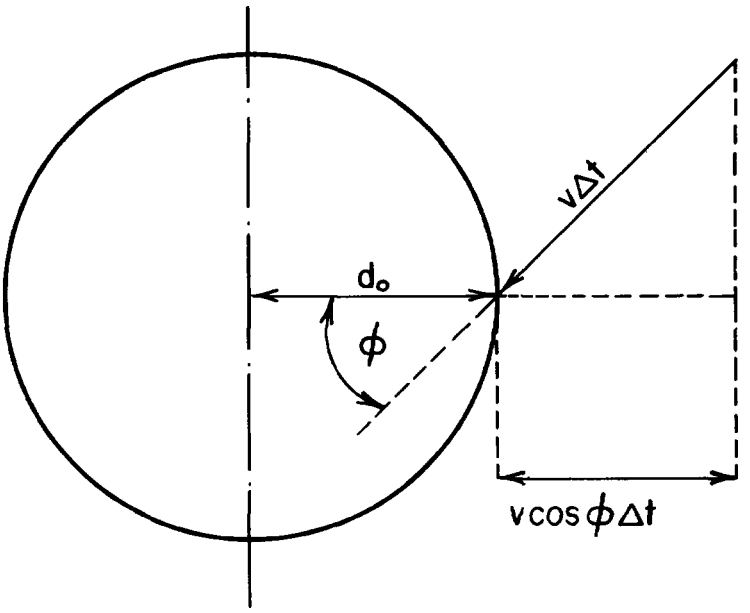


Fig. 1

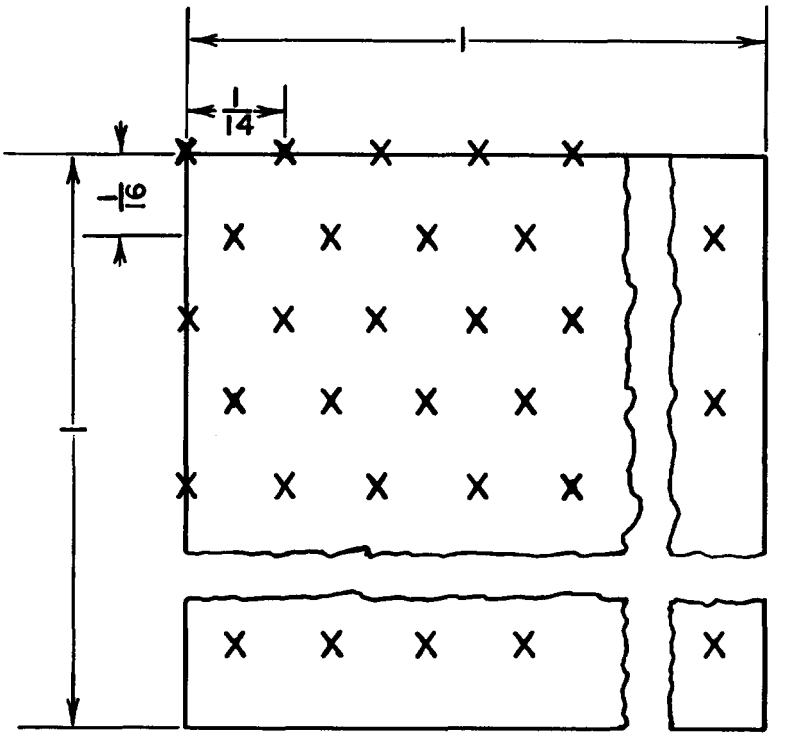


Fig. 2.

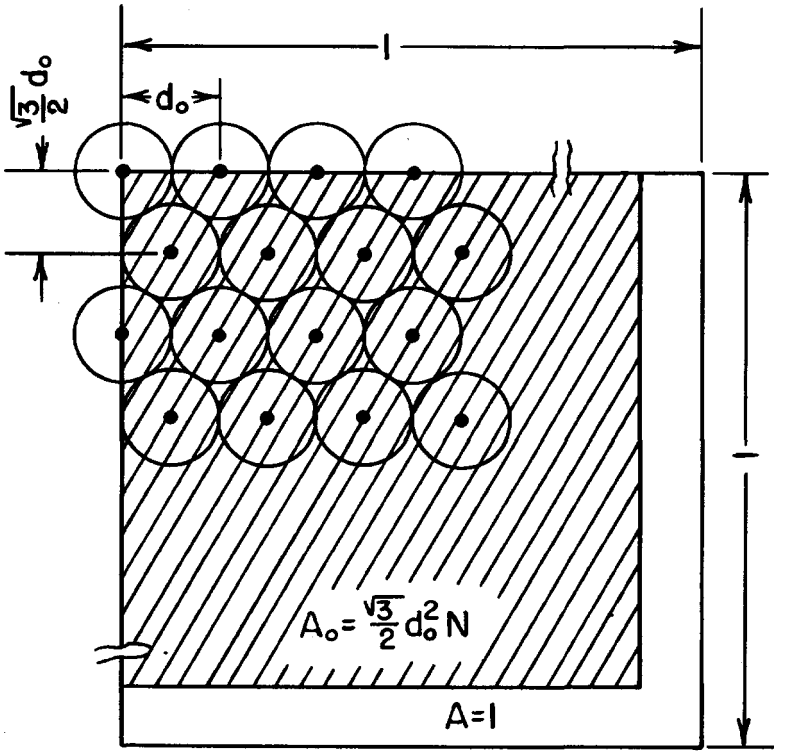


Fig 3

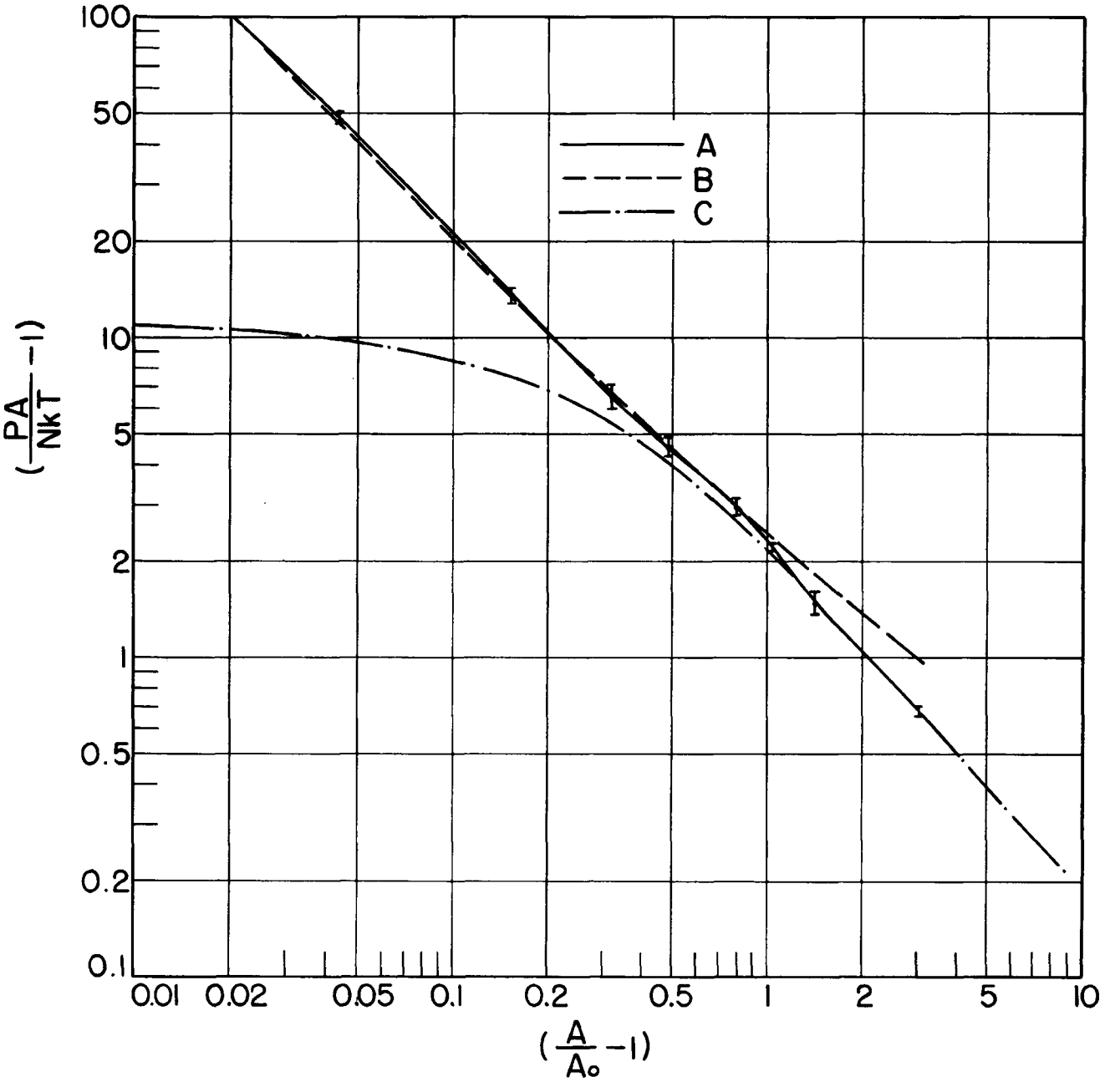
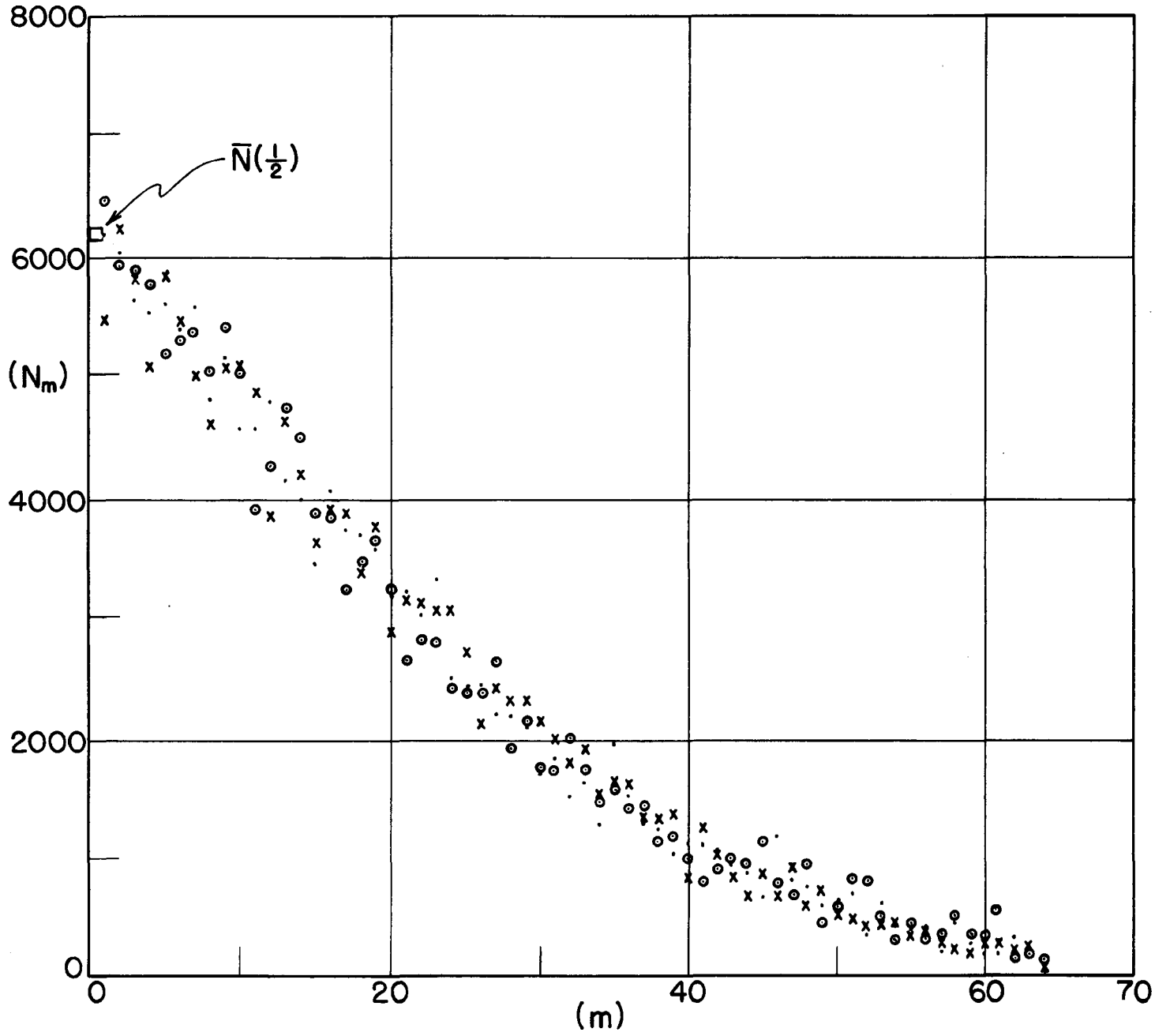

100
50
20
10
5
2
1
0.5
0.2
0.1
0.01
0.02
0.05
0.1
0.2
0.5
1
2
5
10
A
B
C
$(\frac{PA}{NkT}-1)$
$(\frac{A}{A_o}-1)$

Fig. 2

Fig. 5



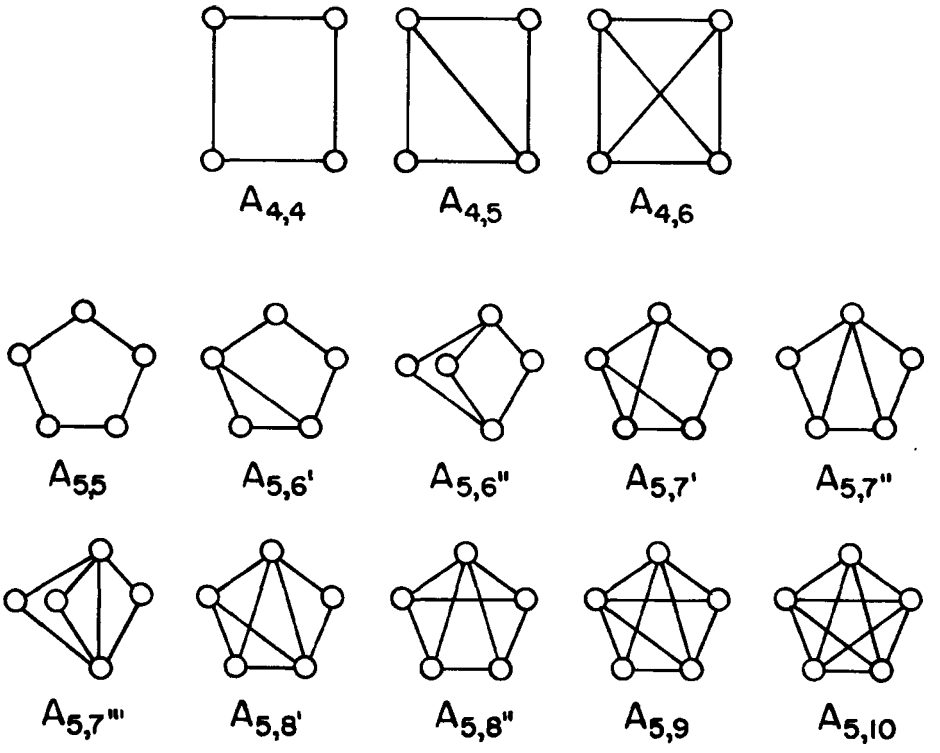


Fig. 6