A typical coordinate rotation computing sequence is shown in Table II. (Note that in this computation, the vector is rotated through the negative X axis.)

Since the change of magnitude will be exactly known beforehand, it may be compensated for exactly either by scaling or by a magnitude-correction multiplication. It may then be said that, except for the practical consideration of limiting the number of digits and the number of steps to some finite value, both algorithms produce an exact solution. In a practical computer, no approximations are necessary except round-off.

In applying this computing technique to practical problems, the complete solution may be programmed by considering the computer as the digital equivalent of an analog resolver.

If the analog-resolver solution flow for the particular problem is known, then the number of operations and the information-flow diagram can be obtained simply by substituting for each resolver, on a time-shared basis, a CORDIC operation.

Monte Carlo Calculations in Statistical Mechanics
W. W. WOOD† AND J. D. JACOBSON†

I. STATISTICAL MECHANICAL INTRODUCTION

According to classical statistical mechanics, the thermodynamic properties of a system of N molecules at temperature T and volume V are obtainable from the Gibbs configurational phase integral,

\[ Z_N(T, V) = \int_N \cdots N \int_V e^{-U/kT} d\mathbf{r}_1 \cdots d\mathbf{r}_N, \]

where \( k \) is Boltzmann's constant; \( U \) is the potential energy of the system of \( N \) molecules, and is a function of the position vectors \( \mathbf{r}_\alpha, \alpha = 1, 2, \cdots, N, \) of the molecules. For suitably simple molecules one usually assumes \( U \) to be expressible as a sum of spherically symmetric pair interactions \( u(r) \):

\[ U(r_1, \cdots r_N) = \frac{1}{2} \sum_{\alpha=1}^{N} \sum_{\beta=1}^{N} u(r_{\alpha\beta}), \]

where \( r_{\alpha\beta} = ||r_\beta - r_\alpha|| \), and the prime indicates omission of terms for which \( \alpha = \beta \).

Most of the thermodynamic functions are expressible in terms of "ensemble averages" of some related function of the configurational coordinates. In these averages the factor \( e^{-U/kT} \) appears as a weighting factor, so that \( Z_N \) given by (1) is the associated normalizing factor. For example, the pressure \( p \) is given by the average of the "virial" \( V_B \) of the total intermolecular force:

\[ pV/NkT = 1 - (1/3NkT)(V_B), \]

where

\[ (V_B) = (1/Z_N) \int_N \cdots N \int_V \left( 1/2 \sum_{\beta} \sum_{\alpha} r_{\alpha\beta} u(r_{\alpha\beta}) / dr_{\alpha\beta} \right) \times e^{-U/kT} d\mathbf{r}_1 \cdots d\mathbf{r}_N \]

The "radial distribution function" \( g(r) \) is of considerable importance in the study of fluids. Let us first define \( n(r) \), the "cumulative radial distribution function," giving the average number of molecules lying within the distance \( r \) from any representative molecule:

\[ n(r) = (1/N) \left( \sum_{\alpha} \sum_{\beta} A(r_{\alpha\beta}, r) \right), \]

where \( A(x, r) \) is the step function

\[ A(x, r) = \begin{cases} 1 & 0 \leq x < r \\ 0 & r \leq x \end{cases} \]

Then \( g(r) \), which is the number density of molecules at the distance \( r \) from any representative molecule, relative to the over-all macroscopic density \( N/V \), is given by
In terms of \( g(r) \) (3) can be written
\[
\rho V / N k T = 1 - \frac{2 \pi}{3} \frac{N}{V} \int_0^\infty r u'(r) g(r) r^2 dr.
\]

Although more involved potential functions \( u(r) \) are required to predict the thermodynamic properties of real molecular systems [1], there is much theoretical interest in the simple hard sphere (sometimes called billiard ball, or elastic sphere) model for which
\[
\begin{align*}
u(r) = \infty & \quad r < \sigma, \\
0 & \quad r \geq \sigma;
\end{align*}
\]

\( \sigma \) is the molecular diameter. This form of \( u(r) \) also leads to considerable computational simplification in the Monte Carlo procedure, and we shall refer most of our discussion to this case. Eq. (8) then becomes
\[
\rho V / N k T = 1 + \frac{2 \pi}{3} \frac{N}{V} \sigma^3 g(\sigma),
\]
\[
= 1 + \frac{\sigma^3}{6} g(\sigma);
\]

the pressure depends only on the value of the radial distribution function at the surface of the molecule.

II. Theoretical Basis of the Monte Carlo Method

For a macroscopic system, \( V \sim 1 \) cm\(^3\), \( N \) is of the order of \( 10^{23} \), which is a considerable obstacle to either analytical or numerical treatment. Just as in the various theories in one way or another radically reduce the dimensionality of the problem, a similar reduction is necessary before undertaking numerical work. Here this is done by applying the periodic boundary condition to a basic system of perhaps a few hundred molecules. The basic volume \( V \) is taken to have a shape such that it fills space compactly under successive translation along three axes in space. There are \( N \) molecules in each replica \( V \), in the same relative positions, and in evaluating \( U \) from (2) the sum over \( \alpha \) is limited to the \( N \) molecules in a particular reference \( V \), but the sum over \( \beta \) is allowed to extend over the adjacent replicas out to some appropriately determined cutoff point. Particularly at high densities, near close-packed, it is important that \( N \) be chosen so that a regular lattice of \( N \) points fixed in \( V \) generates under the above translations the corresponding regular lattice throughout space. Most of our work has used a cubical \( V \) for which the natural choice of lattice is cubic-close-packed (face-centered cubic); then the possible choices of \( N \) are \( 4n^3 \), \( n = 1, 2, \ldots \), or \( N = 4, 32, 108, 256, 500, \ldots \). \( N = 32 \) is the first reasonably realistic choice, and most of our calculations have been done for this value. The periodic boundary condition is illustrated in Fig. 1 for a two-dimensional square \( V \) containing four molecules.

For hard spheres \( \rho V / N k T \) can readily be shown to be a function only of the reduced volume parameter \( \tau \), defined by
\[
\tau = \frac{N}{V \sigma^2 / 2}.
\]

(The parameter so defined is the ratio of the volume \( V \) actually occupied by the system to the volume occupied at regular close packing.) For finite values of \( N \), with the periodic boundary condition, some additional dependence upon \( N \), for fixed \( \tau \), may be expected, and one object of the investigation should be to estimate this dependence and perhaps extrapolate to \( N = \infty \), since it is this limiting value which will correspond to the desired thermodynamic value.

By use of the periodic boundary condition the problem has been approximated by one of a few hundred degrees of freedom, which is in the region where analytical methods fail but where Monte Carlo methods on electronic calculators may be expected to be feasible. However, the most obvious Monte Carlo procedure, the evaluation of \( Z_N \) as given by (1) by means of independent sampling of points in \( 3N \) dimensional configuration space, proves not to be feasible at interesting values of \( V / N \) [2]. The method to be described here is due to Metropolis, et al. [3], and involves instead a Markov chain whose states are points in \( 3N \) dimensional configuration space. A Markov chain may be regarded as the next level of stochastic complexity beyond the simple series of independent events such as a series of throws of a die, or the usual Monte Carlo integration procedure just mentioned. In contrast to the latter, in a Markov chain the probability of obtaining a given result at the \( t \)th step depends explicitly on the result on the \( t-1 \)th step (and therefore implicitly on all preceding steps). The simplest example is the well-known random walk, and the present method is in fact a random walk in \( 3N \) dimensional configuration space. The Markov chain method usually does not result in an evaluation of \( Z_N \) (1), but instead permits the direct evaluation of con-
figurational averages such as appear in (4) and (5). The Markov chain is so constructed that the average of any function of configuration over all states in the chain tends toward the corresponding ensemble average [e.g., (4)] as the chain is developed indefinitely.

The discussion of Markov chains in which the admissible states form a discrete, finite set is somewhat simpler than that required for the case of a continuum of states, and we shall therefore take advantage of the fact that the use of digital calculation itself imposes a “quantization” of configuration space determined by the number of binary bits (in our calculations, 17) used to specify the molecular coordinates. Let us then represent possible configuration states by the letters \( j, k \), and let \( p(j, k) \) be the probability, if at time \( t \) the configuration state is \( j \), that at time \( t + 1 \) the state will be \( k \). To describe the development of the chain we use the usual “time” terminology; it should be emphasized that this is not a time directly related to molecular times, but may conveniently be thought of as corresponding to calculating machine time. The \( p(j, k) \) are independent of \( t \), and for the desired averaging property must satisfy the following conditions [1]:

\[
\sum_k p(j, k) = 1, \tag{12}
\]

and

\[
p(j, k)e^{-U(j)/kT} = p(k, j)e^{-U(k)/kT}. \tag{12a}
\]

In addition, there is an ergodic condition on the matrix \( p(j, k) \) all states included (with nonzero weight) in the integral in (4) must belong to a single ergodic class [5]. That is, for every pair of such states \( j, k \) there must exist at least one finite \( n \) such that \( p^{(n)}(j, k) \neq 0; p^{(n)}(j, k) \) is the probability of passing from state \( j \) at time \( t \) to state \( k \) at time \( t + n \):

\[
p^{(1)}(j, k) = p(j, k); p^{(n)}(j, k) = \sum_l p^{(n-1)}(j, l)p(l, k). \tag{13}
\]

Among the \( p(j, k) \) matrices satisfying (12) and (12a) we use the following:

\[
p(j, k) = A(j, k) \quad \text{if} \quad U(k) \leq U(j); \quad A(j, k) = A(j, k) \exp \left[ -\left( U(k) - U(j) \right)/kT \right], \quad \text{if} \quad k \neq j; \tag{14}
\]

\[
p(j, j) = 1 - \sum_k p(j, k), \tag{15}
\]

where

\[
A(j, k) = \frac{1}{N} \delta_{jk} e^{\alpha}, \quad \text{if} \quad r(\alpha; k) = r(\alpha; j), \quad \alpha = 1, 2 \cdots \beta - 1, \beta + 1 \cdots N;
\]

\[
\text{and} \quad r(\beta, k) \subset C_j \{ r(\beta; j) \}, \quad \beta = 1, 2 \cdots N; \quad \tag{16}
\]

\[
0, \quad \text{otherwise.}
\]

In (16) \( C_j \{ r(\beta; j) \} \) denotes a cube of half edge \( \delta \), oriented parallel to the coordinate planes, with center at \( r(\beta; j) \); the latter notation means the position vector of molecule \( \beta \) in configuration \( j \). The length \( \delta \) is a parameter of the Monte Carlo process whose value is found to influence the rate of convergence of the process. It is ordinarily small (see Section IV) compared to the edge of the cube \( V \), and in (16) it is expressed in units of the smallest increment in the molecular coordinates; in our calculations, therefore, in units of \( 2^{-17} \) times the edge of the cube \( V \).

The motivation underlying this specification of \( p(j, k) \) is as follows. The factor \( A(j, k) \) corresponds to a uniform choice among the limited number of neighboring states contained in \( C_j \); this limitation is related to the small mean free path in molecular systems at interesting densities. The unsymmetrical form of the Boltzmann weighting factor (14) is chosen in preference to the alternative symmetrical form

\[
p(j, k) = A(j, k) \frac{e^{-U(j)/kT}}{e^{-U(j)/kT} + e^{-U(k)/kT}} \tag{17}
\]

mostly for computational convenience, since it leads to fewer calculations of the exponential function (in the case of more complicated models than hard spheres), but also because the unsymmetrical form leads to a more rapid motion in configuration space.

For hard spheres (14) reduces to

\[
p(j, k) = A(j, k), \quad \text{if} \quad U(k) = 0,
\]

\[
= 0, \quad \text{if} \quad U(k) = \infty, \tag{18}
\]

assuming \( U(j) = 0 \), the only case with which we will be concerned since the initial state of the chain has \( U = 0 \).

It can be easily shown that if \( u(r) \) has only point infinities (say, at \( r = 0 \)), then \( p(j, k) \) defined above also satisfies the ergodic condition. However, the hard sphere \( u(r) \) is infinite over a finite interval in \( r \), with the result that \( U \) in \( 3N \) dimensional space is infinite over regions of considerable extent, from which the state point representing our system is excluded. The topological question of whether the accessible region in which \( U \) is zero is then connected or not, depending on the values of \( N/V \) and \( \sigma \), is one to which the answer is unknown, and as a result the equivalence of the Markov and “ensemble” averages is somewhat doubtful. Under such circumstances the \( p(j, k) \) defined above may separate the configuration states into separate ergodic classes. Within each class (i.e., depending on the choice of initial configuration) the Markov average will converge to an average equivalent to the ensemble average taken over a restricted region of configuration space. However, it should also be mentioned that the \( p(j, k) \) defined by (14) and (15) may actually connect spatially disconnected regions of configuration space, if the thickness of the dividing barrier is not too large compared to the parameter \( \delta \). (This is the only circumstance in which the value of \( \delta \) will affect the convergent average of the Markov chain.)
Such a compartmentalization of configuration space also raises the well-known quasi-ergodic problem in statistical mechanics concerning the equivalence of "ensemble" averages to the time-average behavior of an actual physical system of $N$ molecules moving according to the Newtonian equations of motion for the given potential function. In this connection we call attention to the work of Alder and Wainwright [6], in which these equations of motion for the same systems of hard spheres with periodic boundary conditions have been integrated. Thus it is possible to compare the results of this kinetic theory calculation with the present statistical calculation. In a later part of the paper some comparisons of this sort will be made.

As mentioned already, the topological question of the connectivity of configuration space has not been solved. It is quite evident, for example, that at sufficiently high densities, near close packed, it will not be possible to interchange the positions of two molecules without crossing a region in which $U$ is infinite. Configuration space is then evidently compartmentalized into something like $N!$ disconnected regions. However, for systems of $N$ identical molecules (i.e., a one-component system) this compartmentalization is of no consequence since all ensemble averages have identical values in all such compartments (i.e., for hard spheres, the volumes of the compartments are all the same). It is the question of whether, associated with this trivial partitioning of the configuration space, there is a further, nontrivial partitioning into regions of different properties, which is uncertain. In particular, for example, if such a nontrivial compartmentalization does occur, how does it depend on the choice of $N$, at constant $N/V$ and $\sigma$? Is it an artifact of the periodic boundary condition? We will have occasion to return to these questions when we describe below the results obtained with the method.

III. CALCULATIONAL PROCEDURES

It is convenient to take as unit of distance the edge of the cubical volume $V$, so that in these units $V=1$. For a given choice of $N$, specification of $\tau$ then determines the value of $\sigma$, by (11). We first outline briefly, then discuss in greater detail, the steps by which the overall, associated with this trivial partitioning of the system, this compartmentalization is of no consequence since all ensemble averages have identical values in all such compartments (i.e., for hard spheres, the volumes of the compartments are all the same). It is the question of whether, associated with this trivial partitioning of the configuration space, there is a further, nontrivial partitioning into regions of different properties, which is uncertain. In particular, for example, if such a nontrivial compartmentalization does occur, how does it depend on the choice of $N$, at constant $N/V$ and $\sigma$? Is it an artifact of the periodic boundary condition? We will have occasion to return to these questions when we describe below the results obtained with the method.

1) Assume that at "time" $t$ the system is in the state $j(t)$, corresponding to subscript $j$ in (14)–(16).
2) A random choice of one of the $N$ molecules is made, corresponding to subscript $j$ in (16). We call this molecule $\beta(t)$.
3) Each coordinate of molecule $\beta(t)$ is given a tentative random displacement, uniform on the interval $(-\delta, \delta)$, corresponding to $r(\beta, k)$ in (16). Call the configuration in which molecule $\beta(t)$ is in this displaced position, and the other $N-1$ molecules $\alpha$ have position $r(\alpha, j(t))$ configuration $j'(t)$.
4) Configuration $j'(t)$ is tested for overlaps; i.e., one or more pairs of molecules whose distance between centers is less than the molecular diameter $\sigma$.
5) a) If no overlap is found in Step 4, then the next configuration in the chain is $j'(t)$; i.e., $j(t+1)=j'(t)$.
   b) If configuration $j'(t)$ contains an overlap, the configuration at time $t+1$ is identical with that at time $t$: $j(t+1)=j(t)$.
6) Certain procedures concerned with the averaging process are performed; the procedure then repeats beginning at Step 1, except for occasional interruptions for checking and census procedures. We now describe each of these steps in greater detail.

A. Specification of a Configuration

Here we describe the information carried in the calculator memory for purposes of specifying a particular configuration. Fundamentally all that is required are the values of the $3N$ coordinates $r(\alpha)$, which are stored in the $r$-Table. For economy of calculating time, however, it is desirable to carry along additional redundant information, since the calculation requires a rather large amount of machine time. For example, in connection with the calculation of the "cumulative radial distribution function" $n(r)$ (5), it is desirable to have available for each configuration $j(t)$ a tally, called the C-Table, of the intermolecular distances (squared) $r_{ab}^2$ into the $N_B$ intervals $C(\nu) = [\sigma^2 + \nu \cdot 2^{-1} \cdot \sigma^2 + (\nu + 1) \cdot 2^{-1}]$, $\nu = 0, 1, \ldots, N_B - 1$. Here $s$ is chosen to be an integer in order to expedite the tallying process. When we are interested only in estimating the pressure, using (10), so that only $g(\sigma)$ is required, the range of the C-Table can be made quite small, with a consequent economy of machine time, as will appear below.

The search for overlaps in Step 4 can be considerably expedited by a device which avoids the necessity of examining all $N-1$ values of $r_{ab}^2$. The basis for the device is the fact that the displacement parameter $\delta$ is in all interesting cases relatively small compared to unity (edge of $V$), so that any actual overlap $r_{ab}$ which occurs must be with a molecule $\alpha$ which was fairly close to $\beta$ in configuration $j(t)$. We proceed as follows. At some point $t_0$ in the chain (in particular, at its beginning $t=0$) we establish a set of tables called the $\mu(\alpha)$-Tables, one for each $\alpha=1, 2, \ldots, N$. Each $\mu(\alpha)$-Table consists of an indefinite number of entries $\mu(\alpha, \gamma)$, $\gamma=1, 2, \ldots$, and gives all molecule numbers $\mu(\alpha, \gamma)$ for which $r_{ab}^2 < d_{ab}^2$ in configuration $j(t)$. The parameter $d_{ab}$ must satisfy the relation

$$d_{ab}^2 > d_{ab}^2 = \sigma^2 + 2^{-1}N_B$$

but is otherwise arbitrary, and can be adjusted for optimum calculation rate. Also at $t_0$ the positions of all $N$

4 Note that since configurations $j'(t)$ and $j(t)$ differ only in the coordinates of molecule $\beta$, only these $N-1$ distances have changed. Thus if $j(t)$ is free of overlaps (which will be the case unless an error has occurred at some previous value of $t$), any overlap in $j'(t)$ will occur among this set of $N-1$ interactions.
molecules are recorded in the \( r_\alpha \)-Table; we shall call the configuration specified by this \( r_\alpha \)-Table “configuration \( j_\alpha \).” A simple application of the triangle inequality shows that as long as all molecules remain within a distance

\[
d = (d_4 - d_3)/2
\]

of their \( r_\alpha \)-Table positions, only those molecules in the \( \mu(\alpha) \)-Table can approach closer than \( d_3 \) to molecule \( \alpha \). Thus as long as this “diffusion condition” is satisfied we need test in Step 4 only the molecules contained in the \( \mu(\beta(t)) \)-Table.\(^3\)

The recent availability of considerably larger core memories on the Los Alamos 704 calculators has made it possible to consider carrying other redundant information in order to increase further the calculation rate. Some of these possibilities are mentioned later in the discussion.

The starting configuration most often used is a regular face-centered lattice arrangement.

B. Choosing Molecule \( \beta(t) \)

This is done by the usual technique of taking the integer part of the product \( \xi \times \eta \), where \( \xi \) is a nominally randomly, uniformly distributed variable on the interval \([0, 1]\), obtained from a sequence of pseudo-random numbers generated by the middle square process.\(^7\) We use a sequence of 70 bit numbers (generated by double precision techniques) which has not been statistically analyzed except to verify that there are no repetitions among the first \( 10^6 \) 70 bit numbers. Faster techniques\(^8\) can be used to generate pseudo-random numbers, and are currently being incorporated in the codes. The generation of all the random numbers required in the complete calculation uses at most about 7 per cent of the calculation time, however, so that no large improvement can be made by this change alone.

C. Displacing Molecule \( \beta(t) \)

The procedure is similar to that in Step 2, the displaced position being given by

\[
r[\beta(t); j'(t)] = r[\beta(t); j(t)] + \delta \xi
\]

where the three components of \( \xi \) are independent random numbers uniform on the interval \((-1, 1)\). We may mention that as a consequence of the periodic boundary condition, if molecule \( \beta(t) \) due to the displacement given by (21) leaves the volume \( V \) through one face, it re-enters \( V \) through the opposite face; thus the number of molecules \( N \) contained in \( V \) remains constant (see Fig. 1).

Steps 2 and 3 together require about 19 per cent of the average calculating time per configuration, with the present code.

D. Testing \( j'(t) \) for Overlaps, and Tallying its Interactions into the \( C \)-Table

The distances\(^4\) \( || r[\beta(t); j'(t)] - r[\alpha; j_\alpha] || \) are computed for \( \gamma = 1, 2, \ldots \); if for any \( \gamma \) an overlap is found, control passes immediately to Step E2. Those distances less than \( d_3 \) are tallied into an \( E \)-Table whose structure is identical with that of the \( C \)-Table already described. If none of the entries in \( \mu(\beta(t)) \) produce an overlap, the diffusion test is made to see whether \( r[\beta(t); j'(t)] \) is within the distance \( d \) of the \( r_\alpha \)-Table position stored for molecule \( \beta \). If it is, control passes immediately to Step E1. If the diffusion condition is violated, control enters a routine called “refresh” which has the following functions:

1) All \( N-1 \) interactions \( || r[\beta(t); j'(t)] - r[\alpha; j_\alpha] || \) for \( \alpha \neq \beta(t) \) are calculated, and the values of \( \alpha \) for which these distances are less than \( d_4 \) are entered in a temporary \( W_\alpha \)-Table, which may eventually replace the current \( \mu(\beta(t)) \)-Table.

2) As the procedure described in 1) is carried out, the distance \( || r[\beta(t); j'(t)] - r[\alpha; j(t)] || \) is also calculated for every value of \( \alpha \) which is placed in the \( W_\alpha \)-Table but which is not found in \( \mu(\beta(t)) \). If any of these interactions in configuration \( j'(t) \) is found to be an overlap, control passes immediately to Step E1. Those which lie in the interval \( (\sigma, d_4) \) are tallied into the \( E \)-Table already mentioned.

3) Those values of \( \alpha \) which are entered in the \( W_\alpha \)-Table, but which are not contained in \( \mu(\beta(t)) \) are tabulated, and if no overlap is found in \( j'(t) \), the entry \( \beta(t) \) is inserted in \( \mu(\alpha) \) for these values of \( \alpha \).

4) Those values of \( \alpha \) which are \( \omega \) entered in the \( W_\alpha \)-Table, and which appear in \( \mu(\beta(t)) \), are tabulated, and if no overlap is found in \( j'(t) \), the entry \( \beta(t) \) is removed from \( \mu(\alpha) \) for these values of \( \alpha \).

5) If no overlap is found in \( j'(t) \), \( r[\beta(t); j'(t)] \) replaces \( r[\beta(t); j_\alpha] \), and control passes to Step E1.

The testing of \( j'(t) \) for overlaps, exclusive of the time spent in the “refresh” routine, is the most time-consuming part of the calculation, using about 46 per cent of the total time; the “refresh” routine uses about 16 per cent of the time. These figures are for values of the parameters controlling the calculation rate which seem to be optimum (see below).

E1) No Overlap in \( j'(t) \)

In this case, the next configuration of the chain, \( j(t+1) \), is configuration \( j'(t) \). The necessary modificatio-

\(^3\) Actually, one might set up two \( \mu(\alpha) \)-tables, \( \mu(\alpha) \) giving interactions which are potential overlaps with \( \alpha \) (entries being determined by \( d_4 = a+2d \) (22)); \( \mu(\alpha) \) defined as in the text, giving potential \( C \)-Table interactions. This might be worthwhile when \( d_4 \) is relatively large compared to \( \sigma \) (for example, if \( g(\sigma) \) is to be calculated over a range of several molecular diameters). In the present investigations we have concentrated on determining the pressure \( P \); as we have seen since only \( g(\sigma) \) must then be determined, \( d_4 \) may be quite close to \( \sigma \) (we use \( d_4 = 1.05 \sigma \)), and not much advantage would be gained from the additional complication.

\(^4\) When we speak of computing a distance between two molecules we always mean the computation of the square of the distance; no square root procedures are required.
tion of the \( r_\alpha \)-Table and of the \( \mu \)-Tables has been described under Step 4. The only further complication appears in the modification of the \( C \)-Table. Evidently what is required is to remove the interactions \([r(\beta(t); j(t)) - r(\mu(\beta(t), \gamma); j(t))]\) which have been tallied into the \( C \)-Table at earlier \( t \), and replace them by the interactions tallied into the \( E \)-Table during Step 4. The present code requires the recomputation of these \( j(t) \) interactions at this point; current improvements will add to the redundant information already carried a set of \( \phi(\alpha) \)-Tables recording for each \( \alpha-\mu(\alpha, \gamma) \) interaction in \( j(t) \) the corresponding interval in the \( C \)-Table. In any case, the procedure consists of a tally of these \( j(t) \) interactions into the \( E \)-Table, but with a negative unit increment; at the conclusion of this procedure the \( E \)-Table then contains the desired changes in the \( C \)-Table, so that the next procedure is \( C + E \rightarrow C \). The modification of the \( r \)-Table is simply \([r(\beta(t); j'(t)) \rightarrow r(\beta(t); j(t))]\). If the “refresh” procedure for modifying the \( \mu \)-Tables was required in Step 4, we now replace \( \mu(\beta(t)) \) by the \( W_\gamma \)-Table described there. This delay is necessary since \( \mu(\beta(t)) \) must be used for the computation of \( j(t) \) interactions.

We have so far not established a good criterion of the rapidity of the convergence process, other than simply a visual inspection of the trend of the results. In order to obtain some quantitative criterion for the choice of the parameter \( \delta \), we calculate the sum of the squares of the distances in \( 3N \) dimensional space between successive configuration points:

\[
\phi(t) = \sum_{i'=1}^{t} \|r(\beta(i'); j(i')) - r(\beta(i'), j(i' - 1))\|^2. \tag{22}
\]

At present about 7 per cent of the calculation time is spent in Step E1.

E2) Overlap in \( j'(t) \)

In this case \( j(t+1) = j(t) \), no changes in the tables are required, and there is no contribution to \( \phi(t) \).

F. Census and Checking Procedures

For convenience in monitoring the progress of the calculation, averages of the \( C \)-Table are taken over successive intervals of fixed length in time \( t \), as well as over the entire length of the chain. For the former purpose an \( A \)-Table is carried which is set to zero at the beginning of each time interval, and at successive time steps is incremented by the \( C \)-Table. For the over-all average an \( S \)-Table is carried which is augmented by the \( A \)-Table at the end of each time interval. The cumulative distribution function \( n(\sigma) \), (5), is then estimated by

\[
n(\sigma^2 + \nu \cdot 2^{-\nu}) = \frac{1}{Nt} \sum_{\nu = 0}^{\infty} S(\nu; t). \tag{23}
\]

Numerical differentiation (performed later by hand) then gives \( g(\sigma) \), and the pressure \( p \).

Numerous checks are built into the problem, particularly in the “refresh” routine in Step 4, and at the census intervals just described, where the current \( C \)-Table is checked by calculating all \( N(N-1)/2 \) interactions. These procedures have proven to be very worthwhile in giving notice of the occurrence of machine errors.

About 12 per cent of the calculation time is involved in these procedures.

IV. Results and Discussion

The Monte Carlo method was previously applied to the system of hard spheres by Rosenbluth and Rosenbluth [9]. However, the results obtained by Alder and Wainwright [6] at an early point in their kinetic investigation suggested that the Rosenbluth results might be partly in error, so that we undertook a concurrent re-investigation of the question. The new results shown in Fig. 2 are indeed in rather good agreement with those of Alder and Wainwright.

The new Monte Carlo results have been described briefly elsewhere [9] from the standpoint of their statistical mechanical interest. Here we shall consider them briefly from the standpoint of computational interest, concentrating on aspects of the behavior which are in some sense exceptional.

Plotted as the reduced pressure \( p\nu_0/kT \) vs the reduced volume \( \nu \), where \( \nu_0 \) is the close-packed volume per molecule \( \nu_0 = \sigma^3/\sqrt{2} \), the results fall on two distinct curves, as shown in Fig. 2. This behavior may well be related to the presence of a first-order phase transition; we leave aside this point in the present discussion. In the region \( \nu < 1.6 \) the figure shows double-valued pressures. In the region \( \nu = 1.52-160 \) these arise because of the behavior of a representative single chain shown in Fig. 3. These chains show in addition to the usual statistical fluctuation, a tendency to oscillate between two rather well-defined classes of states, with interclass transitions being so rare (about four hours of 704 time units were required to generate the chain pictured in Fig. 3) as to make the over-all averages very poorly determined within a single chain, and differing widely in independent chains generated for the same values of \( N \) and \( \nu \). If the two classes are averaged separately, however, the inter- and intrachain agreement is quite good; it is these separate averages which are plotted in Fig. 2, and which give rise to the double-valuedness. It is noteworthy that Alder and Wainwright observe the same rare interclass transitions.

In the course of investigating the effect of varying the parameter \( \delta \), we generated 14 different Markov chains for the 32-molecule system at \( \nu = 1.55 \), the chain shown in Fig. 3 being one of these. These chains comprised a total of 12,800,000 configurations, the longest chain in the set containing 3,500,000 configurations and the shortest, 560,000 configurations. A total of 14 well-defined interclass transitions were observed; four chains produced no transitions, eight produced one transition,
and two chains showed three transitions. Grouped all together, the results suggest that a transition occurs on the average of about once in 900,000 configurations. These 14 chains gave values of \( p v_0/kT \) (averaged over the entire length of each chain, not the separate high- and low-class averages mentioned previously) ranging from 5.45 to 6.92, with the average over all 14 chains being 6.23. This value of the average is obtained for two different weightings: 1) each configuration of every chain is given equal weight; or 2) each chain is weighted with its final value of \( \phi(t) \) (22). The calculated standard deviation of this average is 0.11. We do not wish to stress the significance of this average, though it is interesting that it lies midway between the two curves of Fig. 2.

We interpret this transition behavior to be due to a compartmentalization of phase space (see the discussion of this point in Section II) into two regions which are narrowly connected in the region \( 1.52 < \tau < 1.60 \). At larger values of \( \tau \) the distinction between the two regions disappears rather quickly, and the over-all averages are easily determined. At smaller values of \( \tau \) neither we nor Alder and Wainwright have observed any interclass transitions, and it is not known whether the two regions of configuration space corresponding to the two branches of Fig. 2 are in fact connected or not for \( \tau < 1.52 \); the lower curve is obtained when the initial configuration is taken as the regular lattice arrangement; the upper curve was obtained from initial configurations generated by a step-wise compression of the system (i.e., gradually increasing \( \sigma \) as \( t \) increases), starting from a randomly chosen configuration from the upper class in the region \( \tau = 1.55-1.60 \). The behavior does not change as \( N \) is varied, and is therefore probably not an artifact of the periodic boundary condition.

For optimum calculation time it has been found best to choose \( d_4/\sigma \approx 1.3 \) for \( N = 32, \approx 1.45 \) for \( N = 256 \), both for \( \tau \approx 1.5 \). The choice of this parameter does not affect the sequence of configurations generated in the chain, but only the rate (in terms of machine time) at which they are generated. The parameter \( \delta \), on the other hand, affects both the calculation rate and the motion through configuration space. The dependence of the optimization parameter \( [\phi(t)/t]^{1/2} \) on the choice of \( \delta \) is shown in Fig. 4, for the 32-molecule system at \( \tau = 1.55 \). Here \( \theta \) is the average actual machine (IBM 704) time required for the generation of one configuration in the chain, and \( [\phi(t)/t]^{1/2} \) (taken for large \( t \)) is the root mean square displacement per configuration. Thus the maximum in Fig. 4 corresponds to a maximum movement through configuration space per unit of machine time, on the basis of the root-mean-square displacement criterion; it occurs at a value of \( \theta \approx 20 \) milliseconds/configuration. Also shown in Fig. 4 is the behavior of the "reject ratio" \( R \), which is the fraction of recurring configurations due to overlaps in \( j'(t) \). It is rather striking that the apparent optimum calculation rate occurs at a rather large value of \( R \approx 0.90 \). So large an optimum value of \( R \) was something of a surprise to us, but is at least partially supported by our qualitative impressions of the convergence of the radial distribution function, and probably also by the frequency of interclass transitions, which appears to be greatest at similarly large values of \( R \).

It is of some interest to inquire how the rate at which the 704 calculator moves the state point through configuration space compares with that at which an actual molecular system executes its thermal movement, even though the two processes are of radically different character. If we take the temperature to be \( 25^\circ \text{C} \), the molecular diameter \( \sigma = 3.5 \times 10^{-8} \text{ cm} \) and the molecular mass as 40 atomic weight units (values which are approximately correct for the argon molecule), we find that a
A possible improvement in method lies in the question of how fine a subdivision of configuration space is required in order for a digital calculation to give results practically equivalent to the continuum of states implied by the analytical integral in (1). As already mentioned, we currently use \((2^{17})^3 = 2^{51}\) points in \(3N\) dimensional space. If the number of bits required to represent the molecular coordinates can be reduced from the present 17 to perhaps 13–14 or less, then in the computation of the intermolecular distances we can replace arithmetic multiplication by a faster table-look-up process. With this and other similar modifications we hope to gain a considerable factor, but certainly less than 10, of improvement in speed.

The “Stretch” calculator currently under contract with IBM, for which coding is now in progress, will increase the calculation rate by a factor of 20–50, and we hope that the combination of improved programming and faster calculator will make possible a determination of the over-all average in the region near \(\tau = 1.55\). In another direction in computer design, it is probably worthwhile to mention that this problem falls in the class of those for which parallel computation by several arithmetic units with common memory and control could result in a considerable increase in calculating rate.

We shall close with a brief mention of some applications in which the results have been interesting without being confusing. For the hard sphere system at \(\tau > 1.6\), where the chains are well convergent, the results are believed to be essentially exact and have been of considerable utility in calibrating various theoretical approximations leading to analytical treatments. At sufficiently low densities the method gives results in agreement with the “virial expansion” whose first five coefficients are known [9]. On the other hand, at sufficiently high densities \(\tau < 1.3\) the lower of the two pressures calculated from the two separate classes agrees increasingly well with the “free-volume” approximation. This suggests that the latter may be asymptotically correct at high densities.

In the case of “Lennard-Jones molecules,” the method yields results [1] which applied to argon agree very well with experimental observations at pressures below about 4000 atmospheres, over a region where no analytical approximation has given agreement. At higher pressures there is disagreement, with a suggestion that the experimental observations may be in error. If the disagreement is real it has interesting consequences in the study of intermolecular forces. Just as with hard spheres, there is agreement with the free-volume approximation at very high densities.

The method has also been tested successfully on the so-called “lattice gas” with nearest neighbor interactions, by comparison with known analytical results [4].

We hope in the future to apply the method to systems of molecules of different kinds (mixtures) and to non-spherical molecules.
A PARTICULAR EXAMPLE ILLUSTRATING THE ANALYSIS TECHNIQUE

THE techniques which are described in this paper apply to the digital mechanization of any system represented as an ordinary linear differential equation with constant coefficients—inhomogeneous or homogeneous.

As an example, consider the sinusoidal loop as generated by a DDA excited by a sampled external forcing function, \( f(t) \), as shown in Fig. 1. It is

\[
\dot{x} + \omega_0^2 x = f(t)
\]

which can also be written as

\[
\dot{x}(t) - \dot{x}(0) + \omega_0^2 \int_0^t x(r) \, dr = \int_0^t f(r) \, dr.
\]

However, reference to Fig. 1 reveals that the DDA lashup does not mechanize (2) but does mechanize

\[
\int_0^t \dot{x}(r) \, dr - \int_0^t x(0) \, dr + \omega_0^2 \int_0^t \int_0^t x_d(r) \, dr \, dr = \int_0^t \int_0^t f(r) \, dr \, dr
\]

where \( \int \) represents the integration operation used in the computer.

Let \( \int \) represent the rectangular rule for integration, and designate the Laplace transform of \( \int \) by \( \dot{g}(s) \),

\[
\dot{g}(s) = \frac{1}{s^2 + \omega_0^2 s h} \quad \text{where} \quad h = \frac{1}{1 - e^{-sh}}
\]

Taking the transform of (3), we obtain

\[
x_D(s) = \frac{\dot{x}(0)}{s^2 + \omega_0^2 s h} + \frac{s x(0)}{s^2 + \omega_0^2 s h} + \frac{\ddot{x}}{s^2 + \omega_0^2 s h}
\]

where \( x_D(s) \) is the Laplace transform of the digital solution, \( x_D(t) \). Consider then the inverse Laplace transforms:

\[
M_D(l) = L^{-1} \left[ \frac{1}{s^2 + \omega_0^2 s h} \right] = L^{-1} \left[ \frac{1}{s^2 + \omega_0^2 s h} \right]
\]

and

\[
N_D(l) = L^{-1} \left[ \frac{s}{s^2 + \omega_0^2 s h} \right] = L^{-1} \left[ \frac{s}{s^2 + \omega_0^2 s h} \right]
\]