PHYSICAL REVIEW E VOLUME 60, NUMBER 1 JULY 1999

Extended (n/v)-Stillinger cluster for use in the theory of homogeneous nucleation

P. Schaaf,[1,2,*] B. Senger,[3] J.-C. Voegel,[3] and H. Reiss[4]
[1]*Institut Charles Sadron (CNRS-ULP), 6 rue Boussingault, 67083 Strasbourg Cedex, France*
[2]*Ecole Européenne de Chimie, Polymères et Matériaux de Strasbourg, 25 rue Becquerel, 67087 Strasbourg Cedex 2, France*
[3]*Institut National de la Santé et de la Recherche Médicale, Unité 424, Fédération de Recherches ''Odontologie,''*
Université Louis Pasteur, 11 rue Humann, 67085 Strasbourg Cedex, France
[4]*Department of Chemistry and Biochemistry, University of California at Los Angeles, 405 Hilgard Avenue,*
Los Angeles, California 90095-1569
(Received 12 January 1999)

In this paper we develop a theory for an extended version of the (n/v)-Stillinger cluster that has been used in nucleation theory, where n means the number of particles constituting the cluster characterized by the volume v. The ''extended cluster'' incorporates some of the surrounding supersaturated vapor. This cluster, although requiring more extensive simulation than the original (n/v)-Stillinger cluster, is almost devoid of approximation. It maintains the non-*ad-hoc* nature of the original (n/v)-Stillinger cluster implicitly. The theory of the cluster is also applicable to clusters which avoid redundancy by some other means than the so-called ''connectivity requirement.'' Simulation of the extended cluster is now being implemented and will be used in the theory of the homogeneous nucleation rate. [S1063-651X(99)08307-5]

PACS number(s): 64.60.Qb, 82.60.Nh

I. INTRODUCTION

Homogeneous nucleation is a physical process that has been known for many years. The early theory of this process, known as the classical nucleation theory (CNT), was developed and improved in the years 1920–1940 by Volmer [1], Farkas [2], Becker and Döring [3], Zeldovich [4], and Frenkel [5]. In it, clusters of n particles participating in the nucleation process are assumed to be spherical drops of volume v having the properties, e.g., surface tension and density, of the macroscopic liquid. In addition, drops were assumed to grow or decrease in size by the gain or loss of one particle (atom or molecule) at a time. This stepwise process is described by the following set of master equations:

$$\frac{df_i}{dt}=\beta_i f_i-\gamma_{i+1}f_{i+1}, \quad i\geqslant 1 \tag{1}$$

in which f_i represents the number of drops consisting of i particles. The parameters β_i and γ_{i+1} can be evaluated by assuming that the master equation remains valid in a constrained equilibrium at the same degree of supersaturation at which the nucleation rate is to be determined, although at present an equivalent method referred to the saturated vapor is often preferred [6]. In the constrained equilibrium the number of drops f_i equals its equilibrium counterpart n_i and use is made of the principle of detailed balance to obtain a relation between β_i and γ_{i+1}. Finally, the coefficients β_i are estimated by determining the collision rates between the drops and the particles of the gas, assuming that all the particles that collide with a drop stick to it (sticking coefficient equal to unity). The determination of the nucleation rate can thus be divided into two distinct steps. (i) One first determines the equilibrium numbers of clusters, and by making use of the principle of detailed balance determines the coefficient γ_{i+1} from β_i. (ii) Once these coefficients are known, the constraint on the system is released and the master equations are solved under steady state conditions.

CNT has proven to be quite successful in the prediction of critical degrees of supersaturation but much less successful in the prediction of nucleation rates. Moreover, CNT contains approximations which at first sight seem minor but which, for small systems, can lead to considerable error. In addition, the liquid is assumed to be incompressible and the translational degrees of freedom of the drop are either ignored or treated improperly. This lack of rigor seems to be at the origin of what is known as the ''replacement free energy'' problem which gave rise to a controversy that (in our opinion) has been finally resolved [7]. Due to these and other approximations, there have been attempts, during the last two decades, to develop molecular theories for the rate of nucleation. A short review of these theories can be found in Ref. [8]. To our knowledge, most of these theories are based on the set of master equations, Eq. (1), and make use of the principle of detailed balance. However, they differ in the way in which clusters participating in the nucleation process are defined, and in how the numbers of clusters, in constrained equilibrium, are evaluated. This paper concerns this last but crucial point. As already indicated, from the number of clusters in the equilibrium distribution of clusters and from an approximate expression for the coefficients β_i, it is possible to determine the evaporation coefficients γ_i. This allows the calculation of the nucleation rate. However, until recently, not much attention has been paid to a precise definition of the clusters involved in the nucleation process. Several molecular theories of homogeneous nucleation, some based on the (n/v)-Stillinger cluster, have been developed to overcome this difficulty [8–16]. This paper describes a method, free of almost any approximation, and with the aid of computer simulation, for the evaluation of equilibrium

---

[*] Author to whom correspondence should be addressed. FAX: 33 (0)3 88 41 40 99. Electronic address: schaaf@ics.u-strasbg.fr

distribution of the (n/v)-Stillinger clusters. One of the first and most important findings of these new approaches was that, in order for a cluster definition to be usable in a molecular theory, it must allow a *nonredundant* counting of clusters. The concept of the (n/v)-Stillinger cluster arose from the requirement that in order for an assembly of n particles to constitute a cluster and thus participate in the nucleation process, the assembly must have a lifetime long enough to allow it to interact with particles in the surrounding vapor, and eventually grow. Certainly, not all configurations of n particles in a vapor satisfy this property. It is thus necessary to find one (or several) parameter(s) that allow the selection of assemblies of n particles that best imitate the requisite clusters. One definition that satisfies the nonredundancy property and which, in our opinion, comes close to satisfying the second requirement was introduced by Stillinger [17] in a different context, and is based on a connectivity criterion. This definition was used in homogeneous nucleation theory by Rao, Berne, and Kalos [18] and more recently in Ref. [8], for the case of particles interacting with spherically symmetric potentials. The requirement is the following.

Consider a vapor consisting of N particles. Then n particles among them belong to a particular n cluster if (i) there exists a continuous path between any pair of them, and (ii) no such path exists between any particle in the surrounding vapor and any particle of the cluster (but can still exist within the particles of the vapor, i.e., the vapor may contain other clusters). A continuous path is defined as one joining the centers of particles, such that the distance between two connected centers is less than a prescribed distance d_c which may be denoted as the ''connectivity distance.'' Thus an n cluster consists of the assembly of configurations of the whole system that satisfy both conditions (i) and (ii). This can be expressed mathematically in the following form. The particles labeled 1,...,n belong to an n cluster if the function

$$\alpha_1(\mathbf{r}_1,\ldots,\mathbf{r}_n)\alpha_2(\mathbf{r}_1,\ldots,\mathbf{r}_n;\mathbf{r}_{n+1},\ldots,\mathbf{r}_N)=1, \tag{2}$$

where $\mathbf{r}_i$ represents the position of the ith particle. $\alpha_1(\mathbf{r}_1,\ldots,\mathbf{r}_n)=1$ if particles 1,...,n satisfy condition (i) and is 0 otherwise. $\alpha_2(\mathbf{r}_1,\ldots,\mathbf{r}_n;\mathbf{r}_{n+1},\ldots,\mathbf{r}_N)=1$ if the $(N-n)$ remaining particles from the vapor satisfy condition (ii) and is 0 otherwise. The second condition ensures that the nonredundancy requirement is satisfied. *Any nonredundant definition of a cluster can be expressed in a form equivalent to Eq. (2).* The precise expressions for $\alpha_1(\mathbf{r}_1,\ldots,\mathbf{r}_n)$ and $\alpha_2(\mathbf{r}_1,\ldots,\mathbf{r}_n;\mathbf{r}_{n+1},\ldots,\mathbf{r}_N)$ depend, however, on the precise definition of the cluster. (Note that the cluster defined in Ref. [13] forms a subset of the Stillinger cluster.)

In a previous paper [19], we derived an expression for the number of n clusters [(n/v)-Stillinger clusters] in a vapor of N particles at equilibrium in a volume V. This derivation began with the *exact* expression of the N-particle distribution function. Some approximations were, however, necessary in obtaining the final result. Use of this result, in a theory for the rate of nucleation, demonstrated that even these apparently minor approximations led to variations of several orders of magnitude in the predicted nucleation rate [8]. More generally, the fact that any approximation can affect the nucleation rate by several orders of magnitude seems characteristic of the nucleation process and is certainly one of the reasons that previous rate theories have not been completely successful. Thus it is absolutely necessary to limit approximation as much as possible, and to analyze the effects of residual approximations in detail.

In this paper we extend the definition of the (n/v)-Stillinger cluster so that it includes part of the vapor, in fact, n' particles of that vapor. This extended cluster therefore requires the introduction of an additional parameter n', so that we will refer to it as an $(n,n'/v)$-Stillinger cluster. We derive the numbers $p(n,n'/v)$ of such clusters for a system at equilibrium. It will be seen that the derivation is almost rigorous, requiring only the introduction of very minor approximations that are much less severe than those introduced previously [8,19]. The equilibrium number of clusters $p(n)=\Sigma_{n'}\Sigma_v p(n,n'/v)$, together with the principle of detailed balance, allows the estimation of the coefficients γ_i. In addition to defining the $(n,n'/v)$ cluster we present an adaptation of the grand canonical Monte Carlo simulation algorithm that allows the evaluation of its equilibrium distribution.

Before proceeding to this derivation it is appropriate to remark that the same theory is applicable to any cluster, similar in all details except that the nonredundancy requirement could be chosen in another way. It is, however, necessary to enforce nonredundancy.

II. EXTENDED CLUSTER

As already mentioned, almost all theories of homogeneous nucleation require a knowledge of the numbers of clusters present in the vapor at equilibrium. As indicated, we have developed a rigorous method that allows the determination of these numbers and is applicable to any cluster definition as long as it satisfies the *nonredundancy* requirement [19]. Our starting point was the N-particle distribution function in the canonical ensemble. After a lengthy but straightforward calculation the following expression was obtained for the number of (n/v) clusters having a *defined* volume ranging from v to $v+dv$:

$$\begin{aligned} p(n/v)dv=\frac{dv}{\Lambda^3}\exp\Bigg\{-\beta\Bigg[&-kT\ln(\Lambda^{-3}n^{3/2}V)\\ &+Pv_c(v)-\mu n+U_0-kT\\ &\times\ln\Bigg(\frac{\Lambda^{-3(n-2)}n^{3/2}}{(n-1)!}\int_v d\mathbf{r}_2'\cdots\int_v d\mathbf{r}_{n-1}'\\ &\times\exp-\beta U(\mathbf{r}_1',\ldots,\mathbf{r}_n')\alpha_1(\mathbf{r}_1',\ldots,\mathbf{r}_n')\Bigg)\Bigg]\Bigg\}, \end{aligned} \tag{3}$$

where β is $1/kT$, k is the Boltzmann constant, T the absolute temperature, V the volume of the system (usually taken as 1 cm^3), P is the pressure of the vapor, and Λ is the de Broglie wavelength. In this expression the integrations are performed with the center of mass of the n particles fixed at the center of the spherical volume v of the cluster, and the primed coordinates are in the center of mass frame. One of the n particles is maintained in the spherical shell that contains the volume, and is denoted as the ''shell molecule.'' For these reasons the integration involves only $(n-2)$ particles. $U(\mathbf{r}'_1,\ldots,\mathbf{r}'_n)$ represents the interaction energy of the n particles. To derive this expression from the rigorous starting point of the N-particle distribution function it was necessary to spatially (not energetically) decouple the cluster from the vapor. To accomplish this it was assumed that the particles of the cluster excluded a volume $v_c(v)$, a function of the volume v of the cluster itself. In the derivation of Eq. (1) of Ref. [19] it was even assumed, as one approximation, that $v_c(v)=v$. However, to each configuration of the particles in the cluster there corresponds a possibly different exclusion volume. In Ref. [8] it was shown that changing the exclusion volume can influence the nucleation rate by several orders of magnitude. A second approximation involved the assumption that $\partial F(N-n,V-v_c)/\partial v_c=-P$ where $F(N,V)$ is the Helmoltz free energy of the system composed of N particles in a volume V. Stillinger, in his pioneering paper [17], used the grand canonical ensemble to show that this relation became increasingly exact as the vapor became more attenuated. It has been shown [20] that the pressure P appearing in this expression does not exactly equal the pressure of the vapor but depends on the shape of v_c. This effect should, however, be small. Finally, the term U_0 represents the mean interaction energy between the cluster and the gas. To estimate U_0 it was assumed that the particles belonging to the cluster were uniformly distributed over the volume v and that the density of the vapor was constant up to a distance d_c that depended on the cluster definition. Interaction of the cluster with the vapor depends on the configuration of the particles within the cluster and the magnitude of the approximation introduced by the mean field character of U_0 is not clear. Fortunately, in most cases involving nucleation in vapors this interaction is small enough to be neglected.

Unfortunately, all these approximations act exponentially on the numbers of (n/v) clusters at equilibrium. Even if they seem minor they can still cause variations in the nucleation rate of several orders of magnitude [8]. Hence, to avoid these approximations and to derive the numbers of (n/v) clusters as precisely as possible, we have developed the following modified approach. In this approach $(n,n'/v)$ clusters are defined in the following way: (i) the n particles satisfy the condition (i) of our initial definition of an (n/v) cluster; and (ii) a spherical volume v_1 concentric with the sphere defining the cluster encloses n' additional particles that belong to the surrounding vapor. These n' particles thus satisfy the condition (ii) in our initial definition of the (n/v)-Stillinger cluster. v_1 is thus larger than v. It should be emphasized that the n' particles can be found *everywhere* in v_1 and thus also in v as long as they satisfy the condition (ii) of the initial cluster definition. v_1 is surrounded by the vapor and the remaining system is thus characterized by the volume $V-v_1$ and the number of particles $N-n-n'$. V-v_1 acts as a particle and energy reservoir.

We now derive the expression for the number $p(n,n'/v)dv$ of $(n,n'/v)$ clusters characterized by a volume ranging from v to $v+dv$. We follow a route similar to that used in Ref. [19]. Start with the n-particle distribution function. Then the number of configurations of the entire system composed of N particles in the volume V such that there is the center of a first particle in $d\mathbf{r}_1$, the center of a second particle in $d\mathbf{r}_2,\ldots,$ and the center of an $(n+n')$th particle in $d\mathbf{r}_{n+n'}$ is

$$\rho^{(n+n')}(\mathbf{r}_1,\ldots,\mathbf{r}_{n+n'})d\mathbf{r}_1\cdots d\mathbf{r}_{n+n'}=\frac{N!}{(N-n-n')!}\,\frac{d\mathbf{r}_1\cdots d\mathbf{r}_{n+n'}\int_V d\mathbf{r}_{n'+1}\cdots\int_V d\mathbf{r}_N\exp[-\beta U(\mathbf{r}_1,\ldots,\mathbf{r}_N)]}{Z(N,V)}, \tag{4}$$

where $Z(N,V)$ represents the configuration integral of the N particles in V. Condition (ii) of the cluster definition is represented mathematically by the function $\alpha_2(\mathbf{r}_1,\ldots,\mathbf{r}_n;\mathbf{r}_{n+1},\ldots,\mathbf{r}_{n+n'})$, equal to unity when the particles located in configuration $\mathbf{r}_1,\ldots,\mathbf{r}_{n+n'}$ satisfy the condition (ii) and is 0 otherwise. It is assumed here that the particles $1-n$ are part of the cluster, while the remaining n' particles are part of the vapor in v_1. The mean number of $(n,n'/v)$-Stillinger clusters in $d\mathbf{r}_1 d\mathbf{r}_2\cdots d\mathbf{r}_{n+n'}$ is then

$$\begin{aligned}
&R^{(n,n'/v)}(\mathbf{r}_1,\ldots,\mathbf{r}_n;\mathbf{r}_{n+1},\ldots\mathbf{r}_{n+n'})d\mathbf{r}_1\cdots d\mathbf{r}_{n+n'}\\
&\quad=\frac{N!}{Z(N,V)}\,\frac{d\mathbf{r}_1\cdots d\mathbf{r}_{n+n'}}{(N-n-n')}\int_{V-v_1}d\mathbf{r}_{n+n'+1}\cdots\int_{V-v_1}d\mathbf{r}_N\alpha_1(\mathbf{r}_1,\ldots,\mathbf{r}_n)\\
&\qquad\times\alpha_2(\mathbf{r}_1,\ldots,\mathbf{r}_n;\mathbf{r}_{n+1},\ldots,\mathbf{r}_{n+n'})\exp[-\beta U(\mathbf{r}_1,\ldots,\mathbf{r}_N)].
\end{aligned}\tag{5}$$

Expressing the coordinates of all the $(n+n')$ particles in the center of mass frame of the n particles, Eq. (5) can be written as

$$\begin{aligned}
&R^{(n,n'/v)}(\mathbf{r}'_2,\ldots,\mathbf{r}'_n;\mathbf{r}'_{n+1},\ldots,\mathbf{r}'_{n+n'},\mathbf{R})d\mathbf{R}\,d\mathbf{r}'_2\cdots d\mathbf{r}'_{n+n'}\\
&\quad=\frac{n^3N!}{Z(N,V)}\,\frac{d\mathbf{R}\,d\mathbf{r}'_2\cdots d\mathbf{r}'_{n+n'}}{(N-n-n')!}\int_{V-v_1}d\mathbf{r}_{n+n'+\mathbf{1}}\int_{V-v_1}d\mathbf{r}_N\alpha_1(\mathbf{r}'_2,\ldots,\mathbf{r}'_n)\\
&\qquad\times\alpha_2(\mathbf{r}'_2,\ldots,\mathbf{r}'_n;\mathbf{r}'_{n+1},\ldots,\mathbf{r}'_{n+n'})\exp[-\beta U(\mathbf{R},\mathbf{r}'_1,\ldots,\mathbf{r}_N)],
\end{aligned}\tag{6}$$

where $\mathbf{R}$ is the coordinate of the center of mass and where the factor n^3 is the Jacobian of the transformation to center of mass coordinates [21]. Label with n the shell molecule which defines the volume of the cluster. The number of $(n,n'/v)$ clusters having their centers in $d\mathbf{R}$ around $\mathbf{R}$ and volumes ranging between v and $v+dv$ is thus given by

$$\begin{aligned}
&d\mathbf{R}\int_v d\mathbf{r}_2'\cdots\int_v d\mathbf{r}_{n-1}'\int_{dv} d\mathbf{r}_n'\int_{v_1} d\mathbf{r}_{n+1}'\cdots\int_{v_1} d\mathbf{r}_{n+n'}' R^{(n,n'/v)}(\mathbf{r}_2',\ldots,\mathbf{r}_n';\mathbf{r}_{n+1}',\ldots,\mathbf{r}_{n+n'}')\\
&\quad=\frac{N!n^3}{(N-n-n')!n'!(n-1)!}\,\frac{d\mathbf{R}}{Z(N,V)}\int_v d\mathbf{r}_2'\cdots\int_v d\mathbf{r}_{n-1}'\int_{dv} d\mathbf{r}_n'\int_{v_1} d\mathbf{r}_{n+1}'\cdots\int_{v_1} d\mathbf{r}_{n+n'}'\\
&\qquad\times\int_{V-v_1} d\mathbf{r}_{n+n'+1}\cdots\int_{V-v_1} d\mathbf{r}_N\,\alpha_1(\mathbf{r}_2',\ldots,\mathbf{r}_n')\,\alpha_2(\mathbf{r}_2',\ldots,\mathbf{r}_{n+n'}')\exp[-\beta U(\mathbf{r}_1,\ldots,\mathbf{r}_{n+n'},\ldots,\mathbf{r}_N)].
\end{aligned}\tag{7}$$

The $(n-1)!$ reflects the fact that, because of the presence of the shell molecule, only $(n-1)$ particles in the cluster are indistinguishable. The number of $(n,n'/v)$ clusters in V (e.g., 1 cm^3) is thus given by

$$\begin{aligned}
p(n,n'/v)dv=&\frac{N!n^3}{(N-n-n')!n'!(n-1)!}\,\frac{V\,dv}{Z(N,V)}\int_v d\mathbf{r}_2'\cdots\int_v d\mathbf{r}_{n-1}'\int_{v_1} d\mathbf{r}_{n+1}'\cdots\int_{v_1} d\mathbf{r}_{n+n'}'\\
&\times\alpha_1(\mathbf{r}_2',\ldots,\mathbf{r}_n')\,\alpha_2(\mathbf{r}_2',\ldots,\mathbf{r}_{n+n'}')\exp[-\beta U(\mathbf{r}_1,\ldots,\mathbf{r}_{n+n'})]\\
&\times\int_{V-v_1} d\mathbf{r}_{n+n'+1}\cdots\int_{V-v_1} d\mathbf{r}_N\exp[-\beta U(\mathbf{r}_{n+n'+1},\ldots,\mathbf{r}_N)]\exp[-\beta W(\mathbf{r}_1,\ldots,\mathbf{r}_{n+n'};\mathbf{r}_{n+n'+1},\ldots,\mathbf{r}_N)],
\end{aligned}\tag{8}$$

where $W(\mathbf{r}_1,\ldots,\mathbf{r}_{n+n'};\mathbf{r}_{n+n'+1},\ldots,\mathbf{r}_N)$ is the interaction energy between the $(n+n')$ particles of the $(n,n'/v)$ cluster and the remaining particles of the vapor system. However, in contrast to the case where the (n/v) cluster was in direct contact with the vapor, in the $(n,n'/v)$ cluster there are only the n' particles of the vapor, contained in the volume v_1, in direct contact with the particles from the vapor outside of the volume v_1. If the difference in the radii between the volumes v and v_1 is greater than the characteristic range of the interaction between two particles then the term $W(\mathbf{r}_1,\ldots,\mathbf{r}_{n+n'};\mathbf{r}_{n+n'+1},\ldots,\mathbf{r}_N)$ is negligible. Indeed, this quantity can be related to an interfacial tension that vanishes when the properties of the vapor are identical on either side of the interface [22]. Equation (8) can then be rewritten as

$$\begin{aligned}
p(n,n'/v)dv=&\frac{N!n^3V\,dv}{(N-n-n')!n'!(n-1)!}\,\frac{Z(N-n-n',V-v_1)}{Z(N,V)}\int_v d\mathbf{r}_2'\cdots\int_v d\mathbf{r}_{n-1}'\int_{v_1} d\mathbf{r}_{n+1}'\cdots\int_{v_1} d\mathbf{r}_{n+n'}'\\
&\times\alpha_1(\mathbf{r}_2',\ldots,\mathbf{r}_n')\,\alpha_2(\mathbf{r}_2',\ldots,\mathbf{r}_{n+n'}')\exp[-\beta U(\mathbf{r}_1,\ldots,\mathbf{r}_{n+n'})].
\end{aligned}\tag{9}$$

Since the Helmholtz free energy $F(N,V)$ and $Z(N,V)$ are related by

$$F(N,V)=-kT\ln\left(\frac{Z(N,V)}{N!\Lambda^{3N}}\right)\tag{10}$$

expression (9) can be rewritten as

$$\begin{aligned}
p(n,n'/v)=&\frac{\Lambda^{3(N-n-n')}}{\Lambda^{3N}}\exp\{-\beta[F(N-n-n',V-v_1)-F(N,V)]\}\\
&\times\frac{n^3V}{n'!(n-1)!}\int_v d\mathbf{r}_2'\cdots\int_v d\mathbf{r}_{n-1}'\int_{v_1} d\mathbf{r}_{n+1}'\cdots\int_{v_1} d\mathbf{r}_{n+n'}'\,\alpha_1(\mathbf{r}_2',\ldots,\mathbf{r}_n')\,\alpha_2(\mathbf{r}_2',\ldots,\mathbf{r}_{n+n'}')\exp[-\beta U(\mathbf{r}_1,\ldots,\mathbf{r}_{n+n'})].
\end{aligned}\tag{11}$$

But in this approach the volume v_1 is spherical and independent of internal configuration which was not the case of the excluded volume $v_c(v)$ in our earlier approach. The pressure P appearing in the relation $\partial F(N-n,V-v_1)/\partial v_1=-P$ now corresponds to the pressure of the vapor exterior to the cluster. This relation is only rigorous for an infinite volume v_1 but represents a good approximation for not too small volumes v_1 [20]. Thus

$$F(N,V)-F(N-n-n',V-v_1)=-Pv_1+(n+n')\mu. \tag{12}$$

Equation (11) then becomes

$$\begin{aligned} p(n,n'/v)dv={}&\frac{\Lambda^{-3(n+n')}n^3V\,dv}{n'!(n-1)!} \\ &\times\exp\{-\beta[Pv_1-(n+n')\mu]\} \\ &\times\int_v d\mathbf{r}_2'\cdots\int_v d\mathbf{r}_{n-1}'\int_{v_1}d\mathbf{r}_{n+1}'\cdots\int_{v_1}d\mathbf{r}_{n+n'}' \\ &\times\alpha_1(\mathbf{r}_2',\ldots,\mathbf{r}_n')\alpha_2(\mathbf{r}_2',\ldots,\mathbf{r}_{n+n'}') \\ &\times\exp[-\beta U(\mathbf{r}_1,\ldots,\mathbf{r}_{n+n'})], \end{aligned} \tag{13}$$

which can be rewritten as

$$\begin{aligned} p(n,n'/v)dv={}&\frac{dv}{\Lambda^3}\frac{V}{\Lambda^3}n^{3/2}\exp\{-\beta[Pv_1-(n+n')\mu]\} \\ &\times\exp[-\beta F(n,n'/v)], \end{aligned} \tag{14}$$

where $F(n,n'/v)$ is the internal Helmoltz free energy of the $(n,n'/v)$-Stillinger cluster, and can be written in the form

$$\begin{aligned} F(n,n'/v)={}&-kT\ln\Bigg(\frac{\Lambda^{-3(n+n'-2)}n^{3/2}}{n'!(n-1)!}\int_v d2'\cdots \\ &\times\int_v d(n-1)'\int_{v_1}d(n+1)'\cdots \\ &\times\int_{v_1}d(n+n')'\alpha_1(1,...,n) \\ &\times\alpha_2(1,...,n,\ldots,n+n') \\ &\times\exp\{-\beta[U(1,...,n)+U(n+1,...,n+n') \\ &+W(1,...,n,n+1,...,n')]\}\Bigg), \end{aligned} \tag{15}$$

where di' is shorthand for $d\mathbf{r}_i'$ and $U(1,...,n)$ stands for $U(\mathbf{r}_1,\ldots,\mathbf{r}_n)$. Moreover, $W(1,...,n,n+1,...,n+n')$ represents the interaction energy between the n particles belonging to the "dense" part of the cluster and the n' particles which are part of the gas enclosed in volume v_1. It should be noted that, unlike Pv in the case of the earlier (n/v)-Stillinger cluster, Pv, in the exponent of Eq. (14), has been *rigorously* separated from the other parts of the exponent. Finally the number of (n/v) clusters is given as

$$p(n/v)=\sum_{n'=0}^{\infty}p(n,n'/v). \tag{16}$$

III. COMPUTATIONAL PROCEDURE

The number of (n/v)-Stillinger clusters can be determined using Eqs. (14)–(16) together with Monte Carlo simulation of the free energy appearing in Eq. (15). We first concentrate on the evaluation of $F(n,n'/v)$ for a given $(n,n'/v)$ cluster. It is known that computer simulation cannot yield the free energy of a system directly but only the difference in the free energy between different states. The $(n,n'/v)$ cluster is composed of two kinds of particles. These are the n particles belonging to the cluster itself and the n' particles that are part of the vapor. In a first step we evaluate the free energy $F^{(1)}(n,n'/v)$ for an $(n,n'/v)$ cluster in which the n' particles are not subject to the condition (ii) of the cluster definition. Thus

$$\begin{aligned} F^{(1)}(n,n'/v)={}&-kT\ln\Bigg(\frac{\Lambda^{-3(n+n'-2)}n^{3/2}}{n'!(n-1)!}\int_v d2'\cdots \\ &\times\int_v d(n-1)'\int_{v_1}d(n+1)'\cdots \\ &\times\int_{v_1}d(n+n')'\alpha_1(1,...,n) \\ &\times\exp\{-\beta[U(1,...,n)+U(n+1,...,n+n') \\ &+W(1,...,n,n+1,...,n')]\}\Bigg). \end{aligned} \tag{17}$$

To determine this free energy we assume that the n' particles interact with all the other particles through the potential $\lambda u(r)$ where λ is a coupling parameter that varies from 0 to 1 and $u(r)$ is the pair potential between two particles. We thus define

$$\begin{aligned} F^{(1)}(n,n'/v,\lambda)={}&-kT\ln\Bigg(\frac{\Lambda^{-3(n+n'-2)}n^{3/2}}{n'!(n-1)!}\int_v d2'\cdots \\ &\times\int_v d(n-1)'\int_{v_1}d(n+1)'\cdots \\ &\times\int_{v_1}d(n+n')'\alpha_1(1,...,n) \\ &\times\exp\{-\beta[U(1,...,n)+\lambda U \\ &\times(n+1,...,n+n') \\ &+\lambda W(1,...,n,n+1,...,n')]\}\Bigg). \end{aligned} \tag{18}$$

For $\lambda=1$ one has $F^{(1)}(n,n'/v,\lambda=1)=F^{(1)}(n,n'/v)$, whereas for $\lambda=0$ we have

$$F^{(1)}(n,n'/v,\lambda=0)=-kT\ln\left\{\frac{\Lambda^{-3(n-2)}n^{3/2}}{(n-2)!}\int_v d2'\cdots\right.$$
$$\times\int_v d(n-1)'\alpha_1(1,...,n)$$
$$\times\exp[-\beta U(1,...,n)]\frac{\Lambda^{-3n'}}{n'!}$$
$$\left.\times\int_{v_1}d(n+1)'\cdots\int_{v_1}d(n+n')'\right\}. \tag{19}$$

The first term involves only the *n* particles satisfying condition (i) of the cluster definition and corresponds to the free energy of an (n/v) cluster that can be evaluated by the methods described in Refs. [8,19]. The second term involves the n' particles corresponding to an ideal vapor of n' particles in the volume v_1 and is therefore known. The free energy $F^{(1)}(n,n'/v,\lambda=0)$ can therefore be evaluated. $F^{(1)}(n,n'/v)$ is then given by

$$F^{(1)}(n,n'/v)=F^{(1)}(n,n'/v,\lambda=0)$$
$$+\int_0^1 d\lambda\frac{\partial F^{(1)}(n,n'/v,\lambda)}{\partial\lambda}, \tag{20}$$

which corresponds to the charging method first developed by Kirkwood [23] and often used in computer simulation [24]. Using Eq. (18) we obtain

$$F^{(1)}(n,n'/v)=F^{(1)}(n,n'/v,\lambda=0)$$
$$+\int_0^1 d\lambda\frac{1}{Z'(n,n'/v,\lambda)}\int_v d2'\cdots$$
$$\times\int_v d(n-1)'\int_{v_1}d(n+1)'\cdots$$
$$\times\int_{v_1}d(n+n')'[U(n+1,...,n+n')$$
$$+W(1,...,n,n+1,...,n')]\alpha_1(1,...,n)$$
$$\times\exp\{-\beta[U(1,...,n)+\lambda U(n+1,...,n+n')$$
$$+\lambda W(1,...,n,n+1,...,n')]\}, \tag{21}$$

where

$$Z'(n,n'/v,\lambda)=\int_v d2'\cdots\int_v d(n-1)'\int_{v_1}d(n+1)'\cdots$$
$$\times\int_{v_1}d(n+n')'\alpha_1(1,...,n)$$
$$\times\exp\{-\beta[U(1,...,n)+\lambda U(n+1,...,n+n')$$
$$+\lambda W(1,...,n,n+1,...,n')]\}. \tag{22}$$

Equation (21) can be rewritten in abbreviated notation as

$$F^{(1)}(n,n'/v)=F^{(1)}(n,n'/v,\lambda=0)+\int_0^1 d\lambda\langle U(n+1,...,n')$$
$$+W(1,...,n,n+1,...,n')\rangle_\lambda, \tag{23}$$

where the angle brackets denote canonical averaging. However, $F^{(1)}(n,n'/v)$ corresponds to a cluster in which the n' particles do not satisfy condition (ii) of the cluster definition. By performing a Monte Carlo simulation based on the Metropolis algorithm but in which the n' particles are not constrained by condition (ii), one can determine the number of configurations ν_2 that satisfy this condition within the total number ν_1 of unconstrained configurations. The Helmoltz free energy $F(n,n'/v)$ of the $(n,n'/v)$ cluster, in which *both* conditions (i) and (ii) are satisfied, is then given by [19]

$$F(n,n'/v)=F^{(1)}(n,n'/v)-kT\ln\left(\frac{\nu_2}{\nu_1}\right). \tag{24}$$

Using Eq. (14) we can then get the number of clusters for a particular set of values of *n*, n', and v. This procedure can, in principle, be repeated for any value of *n*, n', and v. It will be more efficient, however, to work in the grand ensemble. Indeed, in such an ensemble the computer effort will be largely spent on the values of n' depending on their contribution to the $(n,n'/v)$ cluster.

If we fix *n* and *v* we can perform a Monte Carlo simulation in which n' can vary, as an adapted version of the grand canonical Monte Carlo algorithm. Following Mezei [25] and Yao, Greenkorn, and Chao [26] the algorithm consists of the following steps.

(1) Initialize with *n*, n' and $\mathbf{r}_1,\ldots,\mathbf{r}_{n+n'}$ where the center of mass of the particles labeled $1-n$ coincides with the center of the sphere of volume v. The shell molecule (labeled *n*) is located on the shell of the same sphere. The remaining n' particles are in the spherical volume $v_1>v$ whose center coincides with the center of v. This constitutes the configuration (a) and it must also satisfy criteria (i) and (ii) of the cluster definition.

(2) Generate a random number between 0 and $n+n'$. If this number lies between 0 and *n* proceed to step (2.1); if it lies between *n* and $n+n'$ proceed to step (2.2).

  (2.1) Randomly select one of the $(n-1)$ particles of the cluster, but exclude the shell molecule. Move this particle randomly within a small sphere of radius δ centered on the center of the particle. Then displace all the other $(n+n'-2)$ particles over the same vector distance defined such that the center of mass of the *n* particles remains fixed. If this new configuration does not satisfy *both* conditions (i) and (ii) the system is returned to configuration (a) and a new trial involving the movement of one of the $(n-1)$-cluster particles is initiated. This procedure is repeated until an allowed move is achieved. Then the configuration is denoted as (b). We then go to step (3).

- (2.2) Randomly select one of the n' particles from the vapor and move it randomly within a small sphere of radius δ centered on its center. If this new configuration does not satisfy condition (ii) or if the particle lies outside of the volume v_1 the system is returned to configuration (a) and a new trial involving the move of one of the n' particles of the vapor is initiated. This procedure is repeated until an allowed move is achieved. In this case the system configuration is also denoted as (b). Then proceed to step (3).

(3) Choose between configurations (a) and (b) according to the usual Metropolis criterion. If (b) is not chosen, (a) is restored, renamed (b), and considered as the new configuration.

(4) Select a random number ξ between 0 and 1. If $0 \leqslant \xi < 0.5$, one tries to add a particle to the system and proceeds to step (4.1). If $0.5 \leqslant \xi < 1$ one tries to remove a particle from the system and thus proceeds to step (4.2).

- (4.1) Generate a random position for the new particle somewhere in v_1. This new particle is assumed to be part of the vapor. If it does not satisfy condition (ii) of the cluster definition, the trial is rejected and a new trial is initiated until success is achieved. With success we have a new configuration (c). If the quantity $B_{\text{add}} = (v_1/\Lambda^3)[1/(n'+1)]\exp[-\beta(U_c - U_b - \mu)] > 1$, configuration (c) is accepted and we proceed to step (5). In the expression for B_{add}, U_c (respectively, U_b) refers to the energy of the configuration (c) [respectively, (b)] and μ corresponds to the chemical potential of the particles in the system. If $B_{\text{add}} \leqslant 1$ a random number ψ is generated between 0 and 1. If $B_{\text{add}} > \Psi$, configuration (c) is accepted, otherwise the configuration (b) is retained. We then proceed to step (5).
- (4.2) Select one of the n' particles at random and remove it from the system to get a new configuration which is denoted by (c). If the quantity $B_{\text{rem}} = (n'\Lambda^3/v_1)\exp[-\beta(U_c - U_b + \mu)] > 1$, configuration (c) is accepted and we proceed to step (5). If $B_{\text{rem}} \leqslant 1$, a number ψ is generated between 0 and 1. If $B_{\text{rem}} > \psi$, the configuration (c) is accepted, otherwise configuration (b) is retained. We then proceed to step (5).

(5) Take the accepted configuration, which becomes configuration (a), to be the final configuration. It will serve as the initial configuration for the next cycle. Furthermore, the number n' of particles that it contains is stored in the memory. A new sequence is then initiated at step (2).

From this simulation one extracts the relative numbers of $(n,n'/v)$ clusters for different n', the values of n and v being fixed. But the absolute number of the $(n,n'/v)$ clusters is known for a particular set $(n,n'/v)$. This allows access to the absolute numbers of $(n,n'/v)$ clusters for all values of n'. From Eq. (16) we can thus calculate the number of (n/v) clusters.

Note that this model resembles a method used by Kusaka and Oxtoby [16], except that no arbitrary selection of a volume is required. The cluster is built on the center of mass so that translation can be accommodated properly and redundancy can be avoided in the relation of the cluster to the macrovolume of the system [27].

IV. CONCLUDING REMARKS

This paper has focused on the development of a theory for a cluster that contains part of the vapor and represents an extension of the (n/v)-Stillinger cluster that formed the centerpiece of Refs. [8,19]. Some approximations that were necessary in the theory of that cluster are unnecessary in the case of the extended cluster. For example, there is no need to deal with an exclusion volume v_c that varies with the configuration of the cluster particles. On the other hand, the original (n/v)-Stillinger cluster, although demanding more approximation, had the virtue of more clearly exhibiting the mechanism by means of which it avoids an *ad hoc* character, a feature discussed in Refs. [8,19], and demonstrated semiquantitatively in the simulations of Ref. [8]. The extended version of this cluster includes the same mechanism implicitly but therefore in a nontransparent manner.

The extended cluster deals with v_c implicitly as part of the simulation. Thus although there is this loss of transparency the extended version is able to take advantage of the power of simulation to reduce the level of approximation. Since the theory of nucleation is exponentially sensitive to any approximation this is a desirable feature. On the other hand, the size of the simulation will certainly have to be increased in the implementation of the extended cluster.

The determination of the equilibrium numbers of the extended cluster is optimized by the use of the grand canonical Monte Carlo algorithm. We are currently using this algorithm to determine the equilibrium numbers of extended clusters for comparison with the numbers derived by the previous approach. Such a comparison does require extensive simulation and we plan to report on it in future.

In closing, we remark once again that the theory of the extended cluster is applicable to other clusters similar in every detail except that the method of avoiding redundancy may differ.

ACKNOWLEDGMENTS

P.S. is indebted to the ''Institut Universitaire de France'' for financial support. This work has been partially supported by the ''National Science Foundation'' and the ''Centre National de la Recherche Scientifique'' (Action Incitative No. 5939). The authors thank Professor D. S. Corti for a careful reading of the manuscript and for his constructive remarks.

[1] O. Volmer, Z. Phys. Chem., Stoechiom. Verwandtschaftsl. **119**, 277 (1926).

[2] L. Farkas, Z. Phys. Chem., Stoechiom. Verwandtschaftsl. **125**, 236 (1927).

[3] R. Becker and W. Döring, Ann. Phys. (Leipzig) **24**, 719 (1935).

[4] J. Zeldovich, J. Exp. Theor. Phys. **12**, 525 (1942).

[5] J. Frenkel, *Kinetic Theory of Liquids* (Clarendon, Oxford, 1946), Chap. VII.

[6] J. L. Katz and M. D. Donahue, Adv. Chem. Phys. **40**, 137 (1979).

[7] H. Reiss, W. K. Kegel, and J. Katz, Phys. Rev. Lett. **78**, 4506 (1997); J. Phys. Chem. **102**, 8548 (1998).

[8] B. Senger, P. Schaaf, D. Corti, R. Bowles, J.-C. Voegel, and H. Reiss, J. Chem. Phys. **110**, 6421 (1999); B. Senger, P. Schaaf, D. Corti, R. Bowles, D. Pointu, J.-C. Voegel, and H. Reiss, *ibid.* **110**, 6438 (1999).

[9] H. M. Ellerby, C. L. Weakliem, and H. Reiss, J. Chem. Phys. **95**, 9209 (1991).

[10] H. M. Ellerby and H. Reiss, J. Chem. Phys. **97**, 5766 (1992).

[11] C. L. Weakliem and H. Reiss, J. Chem. Phys. **99**, 5374 (1993).

[12] C. L. Weakliem and H. Reiss, J. Chem. Phys. **101**, 2398 (1994).

[13] P. R. ten Wolde and D. Frenkel, J. Chem. Phys. **109**, 9901 (1998).

[14] P. R. ten Wolde, M. J. Ruiz-Montero, and D. Frenkel, J. Chem. Phys. **110**, 1591 (1999).

[15] I. Kusaka, Z.-G. Wang, and J. H. Seinfeld, J. Chem. Phys. **108**, 3416 (1998).

[16] I. Kusaka and D. Oxtoby, J. Chem. Phys. **110**, 5249 (1999).

[17] F. H. Stillinger, Jr., J. Chem. Phys. **38**, 1486 (1963).

[18] M. Rao, B. J. Berne, and M. H. Kalos, J. Chem. Phys. **68**, 1325 (1978).

[19] P. Schaaf, B. Senger, and H. Reiss, J. Phys. Chem. B **101**, 8740 (1997).

[20] D. Corti and G. Soto-Campos, J. Chem. Phys. **108**, 7959 (1998).

[21] H. Reiss, Adv. Colloid Interface Sci. **7**, 1 (1977).

[22] H. Reiss and G. A. Merry, J. Chem. Phys. **85**, 3313 (1981).

[23] J. G. Kirkwood, J. Chem. Phys. **3**, 300 (1935).

[24] D. Frenkel and B. Smit, *Understanding Molecular Simulation* (Academic, San Diego, 1996).

[25] M. Mezei, Mol. Phys. **40**, 901 (1980).

[26] J. Yao, R. A. Greenkorn, and K. C. Chao, Mol. Phys. **46**, 587 (1982).

[27] H. Reiss, J. Mol. Struct. (to be published).