

Nucleation theorems, the statistical mechanics of molecular clusters, and a revision of classical nucleation theory

I. J. Ford

Department of Physics and Astronomy, University College London, Gower Street, London WC1E 6BT, United Kingdom

(Received 21 February 1997; revised manuscript received 5 June 1997)

The nucleation theorems relate the temperature and supersaturation dependence of the rate of nucleation of droplets from a metastable vapor phase to properties of the critical molecular cluster, the size that is approximately equally likely to grow or decay. They are derived here using a combination of statistical mechanics and cluster population dynamics, using an arbitrary model cluster definition. The theorems are employed to test the validity of the classical theory of homogeneous nucleation and its “internally consistent” form. It is found that the properties of the critical cluster for these models are incorrect, and it emerges that this occurs because the classical theory employs the free energy of a fixed droplet, rather than one free to take any position in space. Thus a term representing positional, or mixing, entropy is missing from the cluster free energy. A revised model is proposed, based on the capillarity approximation but with such a term included, and it is shown that it is fully consistent with the nucleation theorems. The model increases classical rates by factors of approximately 10^4 – 10^6 . Other nucleation models should be tested for internal consistency using the same methods. Finally, the nucleation theorems are used to extract the excess internal energies of molecular clusters from experimental data for several substances. [S1063-651X(97)08910-1]

PACS number(s): 64.60.Qb

I. INTRODUCTION

The formation of clouds and fog in the atmosphere is perhaps the most familiar example of the transformation of a vapor into liquid droplets. Water vapor in air that cools below the so-called dew point becomes thermodynamically unstable with respect to the liquid phase and droplets are formed. The thermodynamics of the phase change are well understood, but the dynamics are not: no fully successful theory of the rate of droplet formation has emerged in spite of seventy years of effort.

Conceptually, the process of nucleation is simple enough. Free molecules are continually colliding with one another and occasionally becoming bound: dimers and larger *clusters* of molecules are built up in this way. Clusters can also lose molecules by occasional evaporation and so individual clusters follow a fluctuating history of growth and decay. When the vapor phase is thermodynamically stable with respect to the condensed phase, decay is more likely than growth, and large clusters tend to fall back to smaller sizes, or break apart completely into free molecules.

However, when the bulk condensed phase becomes thermodynamically stable with respect to the vapor phase, occasional growth excursions by individual clusters into larger size classes can sometimes lead to continued growth. This is because when the condensed phase is stable, growth is more likely than decay for large clusters. However, such sizes cannot be reached without the prior formation of small clusters (except by very unlikely multimolecular collisions) and these often remain thermodynamically unstable with respect to the free molecules. Decay continues to win for small clusters, even when the bulk condensed phase is the thermodynamically stable state. The phase transition is therefore brought about by infrequent fluctuations by individual clusters through the unstable size range, past the critical size, beyond

which the likelihood of growth exceeds that of decay, enabling them to grow into large droplets. This is droplet nucleation, and it increases in frequency as the vapor is made more and more metastable with respect to the bulk condensed phase, since this alters the mean growth and decay rates. For the simplest systems, the parameters that control the nucleation rate are the temperature and the supersaturation of the vapor, which is the ratio of the vapor pressure to the saturated vapor pressure.

If droplet nucleation takes place on the surface of an existing particle, which is common in the atmosphere, the process is called heterogeneous nucleation, but the more fundamental homogeneous nucleation process involves only the interactions of the vapor molecules among themselves. The latter has received more attention, both theoretically and experimentally.

The statistical mechanics of molecular clustering is perhaps the most natural theoretical framework for describing nucleation. However, progress has been hampered by uncertainties in how to represent a physical cluster in statistical mechanics. The picture of growth fluctuations of embryonic droplets that has been sketched above requires a cluster to have a certain stability, so that it can truly be held to exist as a physical entity on the time scale of molecular collisions. However, physical clusters cannot be absolutely stable, or evaporation would never take place.

Mayer clusters [1] do not represent physical clusters, for example, since there is no limitation on the separation between molecules: the definition will therefore include configurations that are unbound and ephemeral. Most definitions of physical clusters involve the confinement of the molecules to a small spatial volume in the hope that all such configurations will be bound and can serve as model physical clusters as required in the kinetic treatment of nucleation, and that no bound states are excluded. This is a difficult and perhaps impossible task, and some miscounting seems to be

inescapable. The cluster definition has an effect upon the predictions of the model, as we shall see.

The most common definition used in the literature requires each molecule of a cluster to lie within a sphere centered on the center of mass of the group [2] and in further refinements, for a molecule also to lie on the surface of the sphere [3–5]. This has been proposed since it is possible to enumerate all the configurations of molecules in a system that satisfy this cluster definition. This means that the phase space integrals can be evaluated and cluster partition functions found. It is assumed that such a construction will include all the bound states of the component molecules and will include no undesired unbound states.

Other cluster definitions exist that would appear to be more natural, such as the Stillinger cluster [6], where each molecule need only lie in a particular region centered around any other molecule already in the cluster. However, the enumeration of molecular configurations satisfying this definition is not easy analytically, nor can the definition exclude unbound transitory states.

Possibly the best cluster definition of all would require all the molecules in the cluster to have a negative total energy [7,8]. It would be the intermolecular potential that determined whether a particular molecular configuration should be viewed as a physical cluster. However, this prescription would be complicated to realize in practice, and several difficulties exist [8]. In fact the definition is not at all perfect since it would fail to account for physical clusters where individual molecular energies become positive but where circumstances, such as a collision, prevent the evaporation of such an “unbound” molecule. Furthermore, one would need to calculate kinetic energies with respect to some reference frame, and it is unclear which choice to make. Perhaps the best that can be done is to demand that the molecular potential energy should be less than the equipartitioned molecular kinetic energy.

The uncertainties in the construction of statistical mechanical models [2,9], and also the poor state of knowledge about the intermolecular forces, have made it attractive to study simpler models instead. The prime example is the classical theory, which is based on the idea that clusters can be viewed as tiny droplets with a well-defined surface area and density, characterized by the properties of bulk condensed matter [10]. A related approach is to use a phenomenological cluster model, where the difficulties in the status of clusters are set aside by fitting parameters in the model to experimental data [11–13]. Nevertheless, a well-founded theory of nucleation is more likely to emerge from statistical mechanical considerations. However, in the light of the above discussion, it would be best to try to obtain results that were not dependent on any particular choice of cluster definition.

This possibility has been pursued in the present paper. Results regarding the temperature and supersaturation dependence of the nucleation rate are derived from the statistical mechanics and population dynamics of arbitrarily defined clusters. These results are the two nucleation theorems. The first nucleation theorem was originally a conjecture based on classical nucleation theory [14], which was then placed on firmer ground using statistical mechanics [15] and then thermodynamics [16]. The second nucleation theorem was derived thermodynamically by the present author [17].

These previous derivations studied the probability of occurrence [15] or the work of formation [16,17] of the critical cluster. The nucleation rate is proportional to a Boltzmann factor containing the critical cluster work of formation, and the remainder of the expression is often called the kinetic prefactor. The kinetic prefactor from classical theory was used to complete the derivation of the theorems. In the present paper, however, the temperature and supersaturation dependence of the rate of nucleation is obtained directly, without the need to estimate the kinetic prefactor. Indeed the validity of the classical prefactor can be tested using the more general results obtained. The derivation using statistical mechanics also indicates how microscopic calculations based on a grand ensemble for a single cluster can be used, together with the theorems, to predict the sensitivity of the nucleation rate to experimental conditions.

The equilibrium statistical mechanics of a subsaturated vapor are described in Sec. II. The pressure of such a vapor is less than the saturated vapor pressure at which the vapor and liquid are in thermodynamic equilibrium. Molecular configurations are classified as collections of clusters of various sizes, and the grand partition function of the vapor is represented in terms of the partition function of a single cluster, according to some chosen cluster definition. The equilibrium system is then interpreted kinetically to obtain rate coefficients for various growth and decay processes.

The nonequilibrium situation, where the vapor is supersaturated and undergoing the nucleation of condensed phase droplets, is discussed in Sec. III. An expression for the nucleation rate is derived in terms of the properties of equilibrium systems and this expression is then used to derive the nucleation theorems in Sec. IV. These are then applied in Sec. V to test the validity of the classical theory of nucleation and an “internally consistent” derivative of it. Both models fail to satisfy the theorems, and in both cases it is because the work of formation of a fixed droplet is employed. The “positional,” or “mixing” entropy discussed recently by Reiss *et al.* [18,19] is missing, and the models are therefore inconsistent with the statistical mechanics. In the light of this, we propose a model based on the same capillarity approximation, but which includes positional entropy and is in accord with the theorems. Finally, the nucleation theorems are used in Sec. VI to extract cluster properties from experimental data, and conclusions follow in Sec. VII.

II. STATISTICAL MECHANICS OF SUBSATURATED VAPOR

A. Equilibrium populations

The statistical mechanical treatment of an imperfect vapor is traditionally developed using a canonical ensemble of systems of volume V containing N molecules that have the ability to associate into bound clusters due to mutual interactions. The law of mass action that determines the cluster size distribution in terms of cluster partition functions can be derived, but the method is not entirely satisfactory since there is no external control over the pressure of the cluster mixture, and hence the supersaturation of the vapor phase. The vapor pressure depends instead on the internal parameters V and N , which need to be chosen to obtain the desired pressure.

The development in this paper will take a slightly different route by examining a grand canonical ensemble instead. The system of volume V is considered to be in contact with a particle reservoir at a chemical potential μ so that the number of molecules N in the system fluctuates. These molecules associate into clusters as before, and by varying the chemical potential, the mean pressure of the mixture of clusters in the system can be controlled. The reservoir also acts as a heat bath at a prescribed temperature T .

If $H(N)$ is the Hamiltonian of the system when N molecules are present, and k is Boltzmann's constant, the grand partition function of the system $\Xi(\mu, T, V)$ is then proportional to the integral of $\exp[-(H(N) - N\mu)/kT]$ over the entire phase space of molecular positions and momenta, summed over all N from zero to infinity:

$$\Xi(\mu, T, V) = \sum_{N=0}^{\infty} \frac{1}{N! h^{3N}} \int \prod_{j=1}^N d^3 \mathbf{x}_j d^3 \mathbf{p}_j \times \exp\{-[H(N) - N\mu]/kT\}, \quad (1)$$

where \mathbf{x}_j and \mathbf{p}_j are the position and momentum of the j th molecule, h is Planck's constant, and $N!$ corrects for molecular indistinguishability.

The calculation of Ξ can be greatly simplified by considering the system to be occupied by populations of molecular clusters. The grand partition function of the system can then be constructed from modified canonical partition functions for single clusters of i molecules (an i -cluster). The definition of a molecular cluster will be left open, but it could, for example, require that all the molecules lie within a sphere of a certain volume centered on the center of mass of the whole set of molecules. Alternatively, it might instead be required that the separation between molecules within the cluster should not exceed a maximum distance (the Stillinger cluster). One could use any rule: a cluster could be defined as a configuration of molecules lying in a single plane. A snapshot of the molecules in the system would then be classified as a collection of clusters, with the numbers of clusters present depending on the cluster definition chosen. This is illustrated in Fig. 1. Different cluster populations emerge if different cluster definitions are used. However, as has been stressed above, a nucleation theory based on a poor cluster definition is unlikely to be very successful, and one should attempt through the definition to include all physical clusters and exclude all ephemeral states.

There is a second implication of using an inappropriate cluster definition, which is the following. The calculation of the system grand partition function in terms of cluster partition functions will be accurate only if the interactions between molecules lying in different clusters are negligible. The total energy for a particular configuration of the molecules in the system can then be separated into independent contributions from each cluster. We wish to write $H(N) \approx \sum_i n_i H_c(i)$, where n_i is the number of clusters present in a given system configuration (using the chosen cluster definition) and $H_c(i)$ is the i -cluster Hamiltonian, which depends on the positions and momenta of the molecules that make up the cluster. The cluster definition may be arbitrary, but clearly the level of approximation in the evaluation of Ξ will depend upon it. This will guide the choice of

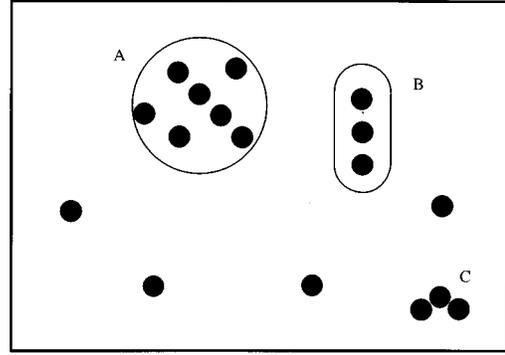


FIG. 1. A particular molecular configuration that contributes to the grand partition function of a subsaturated vapor, illustrating how different cluster definitions can affect the evaluation of cluster populations. Cluster A is defined using a spherical shell centered on the center of mass of a set of closely associated molecules. Cluster B is defined instead using a criterion that molecules should be colinear. According to this criterion, only four molecules in case A can be considered to be a cluster. Cluster C is defined by the requirement that the molecules are a fixed distance apart. Those molecules that do not satisfy a chosen cluster definition are considered to be monomers: according to definition C, therefore, none of the molecules in cases A and B is clustered.

cluster definition. For example, if intermolecular forces were isotropic, then a definition that favors spherical clusters would be a better choice than clusters based on a planar criterion. On the other hand, if the intermolecular forces were planar in character, then such a definition might not be unsuitable.

These points support the idea that a cluster should be defined as a collection of molecules that are bound energetically. This should minimize the cluster-cluster contributions to the total energy: if a nearby molecule had a strong interaction with a cluster, an energy-related definition would be likely to include the molecule within the cluster. Cluster-cluster interactions could be taken into account in the form of a virial series [20], but here we shall simply ignore any cluster-cluster interactions in the Hamiltonian.

We now introduce a modified canonical partition function \mathcal{Z}_i for a cluster containing i molecules. This is related to the integral of $\exp\{-[H_c(i) - i\mu]/kT\}$ over the phase space accessible to molecules in the cluster:

$$\mathcal{Z}_i = \frac{1}{i! h^{3i}} \int \prod_{j=1}^i d^3 \mathbf{x}_j d^3 \mathbf{p}_j \exp\{-[H_c(i) - i\mu]/kT\}. \quad (2)$$

Clearly $\mathcal{Z}_i = Q_i \exp(i\mu/kT)$, where Q_i is the canonical partition function. The cluster definition introduces constraints on the phase space available to the constituent molecules, which can act upon both molecular positions and momenta, and which is indicated in Eq. (2) as a prime on the integration symbol. The phase space of molecules that form a cluster is simply a subset of the total phase space of a collection of free molecules. The subset depends on the cluster definition, and the difficulty in actually evaluating the partition function integrals will of course depend on the definition chosen.

With a little thought, it is evident that the grand partition function of the complete system can be expressed in the following form:

$$\Xi = \sum_{\{n_i\}} \left(\prod_{i=1}^{\infty} \frac{\mathcal{Z}_i^{n_i}}{n_i!} \right), \quad (3)$$

where, as indicated, the sum is over all possible cluster size distributions $\{n_i\}$. This grand partition function then contains, as it should, contributions from all possible configurations of molecules in the system, weighted by the appropriate energies if the total Hamiltonian separates into cluster contributions. The factor of $n_i!$ corrects for overcounting due to the indistinguishability of clusters.

Many previous evaluations of a cluster partition function have proceeded from this point by defining molecular positions with respect to the cluster center of mass, and then integrating the coordinates of the cluster center of mass over the system volume. This makes the modified partition function \mathcal{Z}_i proportional to the system volume V . However, this proportionality would result from any criterion that uses the phase space configuration of the constituent molecules to determine whether they form a cluster. The fundamental point is that for any configuration of molecules that satisfies the definition, there will be others that correspond simply to spatially translated copies of the first. The summation of these contributions to the partition function introduces a proportionality to V . The center of mass definition is not exclusive, and in order to develop the statistical mechanics it is not necessary to introduce it. We note in passing that since \mathcal{Z}_i is dimensionless, it should also be inversely proportional to a quantity with the dimensions of volume. We shall return to the nature of this "scaling volume" later on.

The grand partition function Ξ in Eq. (3) is a sum of contributions over a new phase space of all sets $\{n_i\}$, i.e., all possible cluster size distributions. The molecular position and momentum coordinates are subsumed into the cluster partition function \mathcal{Z}_i . A system containing no molecules, and a system containing a single cluster filling the entire system, are among the configurations taken into account in Eq. (3). There will be many possible other arrangements, corresponding to intermediate molecular densities. There is no restriction in the grand ensemble on $\sum_i i n_i$, the number of molecules within the system.

The next step that is traditionally taken (but normally within a canonical and not a grand canonical ensemble) is to note that Ξ in Eq. (3) is dominated by a contribution from one particular size distribution $\{n_i^e\}$. In order to determine this distribution, we simply extremize the logarithm of the expression within the sum in Eq. (3). We require

$$\frac{\partial}{\partial n_i} \sum_{i=1}^{\infty} \ln \left(\frac{\mathcal{Z}_i^{n_i}}{n_i!} \right) = 0, \quad (4)$$

which leads, using Stirling's approximation, to the following expression for the most probable, or equilibrium, size distribution $\{n_i^e\}$ for the given conditions T and μ :

$$n_i^e = \mathcal{Z}_i. \quad (5)$$

Using this approximation, the grand partition function becomes

$$\Xi = \prod_{i=1}^{\infty} \frac{\mathcal{Z}_i^{n_i^e}}{n_i^{e!}} = \exp \sum_{i=1}^{\infty} n_i^e. \quad (6)$$

The pressure p_v exerted by the noninteracting clusters within the system for a size distribution $\{n_i^e\}$ is given by Dalton's law

$$p_v V = \sum_{i=1}^{\infty} n_i^e kT, \quad (7)$$

and we see that the vapor pressure is a function of μ and T , and that $\Xi = \exp(p_v V/kT)$ as required. The grand potential of the whole system is $-p_v V$. The grand canonical ensemble does not allow the vapor pressure to be fixed exactly, since fluctuations in cluster populations and therefore p_v can occur, but these are negligible for a large system, and so to a very good approximation, the vapor pressure in the system is controlled by the external parameters μ and T , as we require. Equation (5) leads to the law of mass action: $n_i^e = Q_i (n_1^e/Q_1)^i$.

Putting Eq. (5) into Eq. (7) we find

$$p_v V = kT \sum_{i=1}^{\infty} \mathcal{Z}_i. \quad (8)$$

The sum in Eq. (8) will only converge if \mathcal{Z}_i decreases suitably as a function of i . It turns out that this limits the statistical mechanical ensemble to the study of vapors at or below the saturated vapor pressure. Systems at a higher vapor pressure will not be globally stable: the true equilibrium state will be the bulk condensed phase. This constraint translates into an upper limit for the chemical potential. Since $n_1^e = Q_1 \exp(\mu/kT)$ from Eq. (5), we can introduce a reference population n_1^{es} , which is the free molecule, or monomer, population for a saturated vapor, and write $n_1^{es} = Q_1 \exp(\mu_s/kT)$. μ_s is the chemical potential that produces this reference population. We then deduce that

$$\mu = \mu_s + kT \ln S, \quad (9)$$

where we have introduced the saturation ratio $S = n_1^e/n_1^{es}$, which so far is limited to values below unity. To a good approximation, for conditions well below the critical point, $n_1^{es} kT \approx p_{vs} V$, and $S \approx p_v/p_{vs}$, where p_{vs} is the saturated vapor pressure. This confirms that while μ does not control the total pressure exactly, it does control the partial pressure of the monomers, and hence provides a very good external control over the saturation ratio.

Furthermore, the approximation that the Hamiltonian should separate into contributions from independent clusters will also fail as the system becomes denser. No attempt will be made to describe the metastable, or supersaturated vapor using statistical mechanics, since none will be needed.

The single cluster modified partition function \mathcal{Z}_i can be expressed in terms of the single cluster grand potential $\Omega(i)$:

$$\mathcal{Z}_i = \exp[-\Omega(i)/kT] = \exp\{-[F(i) - i\mu]/kT\}. \quad (10)$$

$\Omega(i)$ and $F(i)$ are the grand potential and Helmholtz free energy, respectively, of a single i -cluster in an otherwise empty system of volume V , at a temperature T and chemical potential μ . They depend on the cluster definition. From Eq. (5), the equilibrium populations for sizes i and j satisfy

$$n_i^e/n_j^e = \exp\{-[\Omega(i) - \Omega(j)]/kT\}. \quad (11)$$

B. Kinetic treatment

We now introduce a kinetic interpretation of this equilibrium situation. We consider that the equilibrium size distribution n_i^e is the stationary solution of the following set of population dynamics equations:

$$\frac{dn_i}{dt} = \sum_j n_j W_{j \rightarrow i} - n_i \sum_j W_{i \rightarrow j}, \quad (12)$$

where $W_{i \rightarrow j}$ is the coefficient that determines the mean rate at which transitions are made that convert a j -cluster into an i -cluster. Recall that since the clusters defined in the statistical mechanics are supposed to model real physical clusters, the rate coefficients in Eqs. (12) will describe such processes as molecular capture and molecular evaporation to and from embryonic droplets. The connection between rate coefficients and cluster statistical mechanics is then given by

$$\frac{W_{j \rightarrow i}}{W_{i \rightarrow j}} = \frac{n_i^e}{n_j^e} = \exp\{-[\Omega(i) - \Omega(j)]/kT\}. \quad (13)$$

The kinetic interpretation can be taken a step further by assuming that the only important transitions are those that are brought about by the addition or loss of single molecules to or from the cluster. The only nonvanishing rate coefficients are then $\beta_i = W_{i \rightarrow i+1}$ for cluster growth, and $\gamma_i = W_{i \rightarrow i-1}$ for cluster decay. They are related through

$$\frac{\beta_i}{\gamma_{i+1}} = \exp\{-[\Omega(i+1) - \Omega(i)]/kT\}. \quad (14)$$

Transitions such as dimer addition to an i -cluster to form an $(i+2)$ -cluster have been considered elsewhere [21] and in most cases they alter the nucleation rate only slightly, unless for some reason the dimer population is unusually large.

The rate of combination of single molecules (monomers) and i -clusters to form an $(i+1)$ -cluster is proportional to $n_1^e n_i^e$, so $\beta_i \propto n_1^e$. In fact, according to kinetic theory, the growth rate β_i is the molecular flux times the surface area A_i of the i -cluster, assuming all collisions stick. The usual expression is

$$\beta_i = \frac{n_1^e k T A_i}{V(2\pi m k T)^{1/2}} = S \beta_i^s, \quad (15)$$

where $\beta_i^s = n_1^{es} k T A_i / V(2\pi m k T)^{1/2}$, where m is the molecular mass. We shall use the growth coefficient β_i in Eq. (15) in the following development, which brings with it a certain temperature and supersaturation dependence. (Other growth regimes could be considered, for example, where diffusion limits the absorption rate). The decay coefficient γ_i , however, is independent of supersaturation.

Vasil'ev and Reiss [22,23] have suggested that Eq. (15) can underestimate the true absorption rate by a factor of up to 2, if the nominal surface area A_i is used. This is brought about by an attractive interaction between the cluster and a free molecule, which can alter the trajectories of approaching molecules. This enhancement will be neglected here, but would in any case only alter the nucleation rate by the same relatively small factor. A more significant implication of the effect would be to introduce a dependence of growth rate upon the carrier gas pressure. The reasoning is that inert gas molecules, which have been ignored hitherto, but which are necessarily present in nucleation experiments, might interfere with the attraction between free molecule and cluster, and change the absorption rate. Assuming that the evaporation rate is not affected by the carrier gas in an equivalent way, the nucleation rate could then be suppressed. However, this would seem to be a large effect only for high carrier gas pressures, when carrier gas molecules are often likely to be found within the attractive range of clusters, and so we will neglect it here.

The situation considered so far is limited to saturation ratios $S \leq 1$ controlled by the external chemical potential μ . At $\mu = \mu_s$ the vapor just becomes saturated: $S = 1$ and $p_v = p_{vs}$. The grand potential of the i -cluster for such conditions will be given a subscript s and written as

$$\Omega_s(i) = F(i) - i\mu_s. \quad (16)$$

The ratio of rate coefficients is then

$$\frac{\beta_i^s}{\gamma_{i+1}} = \exp\{-[\Omega_s(i+1) - \Omega_s(i)]/kT\}. \quad (17)$$

III. NUCLEATION RATE IN A SUPERSATURATED VAPOR

The statistical mechanical treatment of a subsaturated vapor in the previous section provides rate coefficients that can be used to study the population dynamics of clusters for a supersaturated vapor ($S > 1$). In this way, the process of nucleation can be modeled. We rewrite Eq. (12) in the form

$$\frac{dn_i}{dt} = \beta_{i-1} n_{i-1} - \gamma_i n_i - \beta_i n_i + \gamma_{i+1} n_{i+1} = J_{i-1} - J_i, \quad (18)$$

where $J_i = \beta_i n_i - \gamma_{i+1} n_{i+1}$ is the mean cluster current between sizes i and $i+1$: the excess of growth transitions over decay transitions between the two sizes. The populations here no longer carry the superscript e , which denoted equilibrium in a subsaturated vapor. They are now a general size distribution determined by the dynamical equations (18). The growth rate β_i is still given by Eq. (15), but with $S > 1$: we now refer to S as the supersaturation rather than the saturation ratio. This enhanced growth rate is the driving force that causes the nucleation of the new phase.

Note that we consider nucleation here to be the result of mean rates of transitions in the population dynamics of clusters, whereas the picture drawn earlier was one of individual

clusters growing and decaying stochastically, and occasionally being driven to large, stable sizes.

The population equations are solved as usual for the following boundary conditions. Firstly the monomer concentration is S times greater than that which occurs in the saturated vapor, i.e., $n_1 = S n_1^{es}$ (with $S > 1$). Secondly, the population at a size $M + 1$ is set to zero. It turns out that as long as $M + 1$ is large enough, it does not matter which particular value is chosen. The condition $n_{M+1} = 0$ prevents the system from relaxing to the global equilibrium state where the condensed phase fills the available volume. The cluster populations are therefore held somewhat artificially in a state of perpetual nucleation of the phase transition.

If we consider the steady-state solution, such that $dn_i/dt = 0$ and J_i is a constant J for all i , a process of elimination within the system of equations (18) leads to the following expression for the nucleation rate:

$$J = \frac{\beta_1 n_1}{1 + \sum_{i=2}^M \prod_{j=2}^i (\gamma_j / \beta_j)}. \quad (19)$$

The product $\prod_{j=2}^i \gamma_j / \beta_j$ can also be written as $(\beta_1 / \beta_i) \prod_{j=2}^i \gamma_j / \beta_{j-1}$ and using Eqs. (15) and (17) we then get

$$\prod_{j=2}^i \frac{\gamma_j}{\beta_j} = \frac{\beta_1}{\beta_i} \exp\{[\Omega_s(i) - \Omega_s(1)]/kT - (i-1)\ln S\}. \quad (20)$$

Substituting into the rate expression then gives

$$J = \beta_1 n_1 \left(1 + \sum_{i=2}^M \frac{\beta_1}{\beta_i} \exp\{[\Omega_s(i) - \Omega_s(1)]/kT - (i-1)\ln S\} \right)^{-1} \\ = \frac{S p_{vs} V}{kT} \exp\{[\Omega_s(1) - kT \ln S]/kT\} \left(\sum_{i=1}^M \frac{1}{\beta_i} \exp[\Omega_s(i) - ikT \ln S]/kT \right)^{-1}. \quad (21)$$

Note that the nucleation rate derived here is the number of clusters per second reaching the maximum size ($M + 1$) in the volume V . The dimensions are inverse time. Nucleation rates are more usually defined as the number of particles generated per second per unit volume, but we shall consider the total current and not its density.

The usual procedure is now to represent the sum over i as an integral between $\pm \infty$, and to expand the argument of the exponential about the point where it reaches a maximum, which defines the critical size i^* :

$$\left(\frac{\partial \Omega_s(i)}{\partial i} \right)_{i=i^*} - kT \ln S = 0. \quad (22)$$

To see that the critical size is loosely the size that is equally likely to grow or to decay, consider Eq. (14) for $i = i^*$. We have

$$\frac{\beta_{i^*}}{\gamma_{i^*+1}} \approx \exp \left[- \frac{1}{kT} \left(\frac{\partial \Omega(i)}{\partial i} \right)_{i=i^*} \right] = 1, \quad (23)$$

using Eqs. (9), (10), (16), and (22).

If we introduce the cluster work of formation

$$\epsilon(i) = \Omega_s(i) - ikT \ln S, \quad (24)$$

then to a good approximation

$$J = \frac{S p_{vs} V}{kT} \exp[\epsilon(1)/kT] Z \beta_{i^*} \exp[-\epsilon(i^*)/kT], \quad (25)$$

where Z is the so-called Zeldovich factor given by

$$Z = \left[- \frac{1}{2\pi kT} \left(\frac{\partial^2 \Omega_s(i)}{\partial i^2} \right)_{i=i^*} \right]^{1/2}. \quad (26)$$

A small contribution from the derivative of β_i with respect to i has been neglected. Note that J is proportional to β_{i^*} , so an enhancement of this coefficient according to the ideas of Vasil'ev and Reiss [22,23] would enhance the nucleation rate by the same factor.

Why is $\epsilon(i)$ the cluster work of formation? The cluster work of formation is the change in the grand potential in going from an empty system to a system containing an i -cluster, for constant external conditions of μ and T . Since the grand potential of an empty system is zero, the cluster work of formation is simply $\Omega(i)$, and Eq. (24) follows from Eqs. (9), (10), and (16).

Equation (25) takes the form that is often proposed on the following heuristic grounds. Since the critical cluster is the size that is equally likely to evaporate or to grow, the frequency of nucleation should be proportional to $\beta_{i^*} n_{i^*}$, the rate of attachment of monomers to critical clusters. The population of critical clusters is equal to $\exp[-\epsilon(i^*)/kT]$. [Strictly, this latter result is not valid for a supersaturated vapor, but is an extrapolation of Eqs. (5) and (10) for

$\mu > \mu_s$.] This accounts for the last two terms in Eq. (25). However, additional factors such as Z and $\exp[\epsilon(1)/kT]$ emerge from the more rigorous approach outlined above. Some of the factors in Eq. (25) correspond to the kinetic prefactor referred to earlier.

IV. NUCLEATION THEOREMS

A. First nucleation theorem

Now we can examine the supersaturation and temperature dependence of the nucleation rate and derive the nucleation theorems.

We start with the exact expression for J in Eq. (21) rather than the less unwieldy but more approximate version in Eq. (25):

$$J = \frac{S p_{vs} V}{kT} \exp[\Omega_s(1)/kT] \left(\sum_{i=1}^M \frac{1}{\beta_i^s} \exp[(\Omega_s(i) - ikT \ln S)/kT] \right)^{-1}, \quad (27)$$

where a factor of $1/S$ in both numerator and denominator has been canceled. Taking the derivative with respect to $\ln S$, holding T constant, we find

$$\begin{aligned} \left(\frac{\partial J}{\partial \ln S} \right)_T &= J - \frac{J}{\sum_{i=1}^M P(i)} \left(\frac{\partial}{\partial \ln S} \right)_T \left(\sum_{i=1}^M P(i) \right) \\ &= J \left(1 + \frac{\sum_{i=1}^M iP(i)}{\sum_{i=1}^M P(i)} \right), \end{aligned} \quad (28)$$

where the weighting function $P(i)$ is given by

$$P(i) = \frac{1}{\beta_i^s} \exp\{[\Omega_s(i) - ikT \ln S]/kT\}. \quad (29)$$

We therefore find that

$$\left(\frac{\partial \ln J}{\partial \ln S} \right)_T = 1 + i_m^*, \quad (30)$$

where the cluster size i_m^* is defined as the expectation value

$$i_m^* = \frac{\sum_{i=1}^M iP(i)}{\sum_{i=1}^M P(i)}. \quad (31)$$

The size i_m^* is equal to the critical size i^* under the same approximations that were used earlier to derive Eq. (25) from Eq. (21), namely, to replace β_i^s in Eq. (29) by $\beta_{i^*}^s$ and to expand the exponent to second order about i^* given by Eq. (22). Then $P(i) \approx P(i^*) \exp[-\pi Z^2(i-i^*)^2]$ with Z given by Eq. (26). The sums in Eq. (31) can then be replaced by integrals over $-\infty < i < \infty$ and so

$$i_m^* \approx \frac{\int_{-\infty}^{\infty} i \exp[-\pi Z^2(i-i^*)^2]}{\int_{-\infty}^{\infty} \exp[-\pi Z^2(i-i^*)^2]} = i^*. \quad (32)$$

The distinction between i^* and i_m^* will henceforth be dropped.

Equation (30) is the nucleation theorem [14–16]. It tells us that the supersaturation dependence of the nucleation rate is related to the critical size. It has been used previously to extract i^* from the slope of nucleation rate data plotted against supersaturation (on a log-log scale) at constant temperature. Earlier derivations concentrated on the supersaturation dependence of the critical cluster work of formation $\epsilon(i^*)$. The proof given here would appear to be more general, and the steps taken in reaching it have been carefully set out. The proof does not rely on a particular choice of cluster definition in the statistical mechanics. Some independence of cluster definition was to be expected, since the earlier derivations were made using arguments from thermodynamics. Some approximations have been made, but we believe they are tenable in most situations. A significant advantage of the present statistical mechanical–kinetic derivation is that Eq. (30) is an exact expression of the nucleation theorem, which takes into account the kinetic prefactor in the rate expression as well as the exponential term.

There is actually a hidden dependence on the cluster definition in Eq. (30). We have stressed several times that the cluster definition affects the calculation of the grand potential, and this means that each cluster definition will produce a different critical size, and therefore a different dependence of the model nucleation rate upon supersaturation. However, when we come to use Eq. (30) to analyze experimental data, we implicitly make the reasonable assumption that a ‘‘perfect’’ definition exists for physical clusters, and that the experimental data are giving us the critical size for that definition: the actual critical cluster, which is equally likely to grow or decay.

B. Second nucleation theorem

A second nucleation theorem has been derived recently [17] using the methods of small system thermodynamics [24]. It concerns the temperature dependence of the nucleation rate at constant supersaturation, and is in a sense the

conjugate to Eq. (30), which we shall now refer to as the first nucleation theorem. We now prove it within the statistical mechanical formalism. It is a rather lengthy derivation, with particular effort spent in carefully evaluating small terms that

are then neglected. The final result appears in Eq. (48), and the reader could proceed directly to that point if desired.

Taking the derivative of Eq. (27) with respect to T , we find

$$\begin{aligned} \left(\frac{\partial J}{\partial T}\right)_{\ln S} &= \left(\frac{\partial}{\partial T}\right)_{\ln S} \frac{Sp_{vs}V}{kT} \exp[\Omega_s(1)/kT] \left(\sum_{i=1}^M P(i)\right)^{-1} \\ &= J \left(\frac{p'_{vs}}{p_{vs}} - \frac{1}{T} - \frac{\Omega_s(1) - T\Omega'_s(1)}{kT^2} - \frac{\sum_{i=1}^M \{-\beta_i^{s'}/\beta_i^s - [\Omega_s(i) - T\Omega'_s(i)]/kT^2\} P(i)}{\sum_{i=1}^M P(i)} \right), \end{aligned} \quad (33)$$

where a prime indicates a partial derivative with respect to temperature. By expanding the terms in curly brackets inside the sum about i^* we can write

$$\begin{aligned} \left(\frac{\partial \ln J}{\partial T}\right)_{\ln S} &= \frac{p'_{vs}}{p_{vs}} - \frac{1}{T} - \frac{\Omega_s(1) - T\Omega'_s(1)}{kT^2} + \frac{\beta_{i^*}^{s'}}{\beta_{i^*}^s} + \frac{\Omega_s(i^*) - T\Omega'_s(i^*)}{kT^2} \\ &\quad + \frac{1}{\sum_{i=1}^M P(i)} \sum_{i=1}^M \left[\frac{1}{2kT^2} (i-i^*)^2 \left(\frac{\partial^2}{\partial i^2} [\Omega_s(i) - T\Omega'_s(i)] \right)_{i=i^*} \right] P(i). \end{aligned} \quad (34)$$

No linear term appears since with the use of previous approximations, the expectation value of $(i-i^*)$ weighted by $P(i)$ is zero.

The term involving β_i^s has been taken outside the sum since according to Eq. (15) we can write

$$\frac{1}{\beta_i^s} \frac{\partial \beta_i^s}{\partial T} = \frac{p'_{vs}}{p_{vs}} - \frac{1}{2T} + \frac{1}{A_i} \frac{\partial A_i}{\partial T}, \quad (35)$$

and if we assume that $A_i \propto (iv_i)^{2/3}$ for spherical clusters, then the final term in the last equation is simply $2v'_i/(3v_i)$ and the whole expression is independent of i .

The remaining expectation value is best dealt with by replacing the sums by integrals over i with limits $\pm\infty$, and using the approximate form for $P(i)$ used in Eq. (32). The last term in Eq. (34) can then be written as

$$\frac{1}{\sqrt{2\pi}\sigma_0 kT^2} \int_{-\infty}^{\infty} di \left[\frac{1}{2} (i-i^*)^2 \left(\frac{\partial^2}{\partial i^2} [\Omega_s(i) - T\Omega'_s(i)] \right)_{i=i^*} \right] \exp\left(-\frac{(i-i^*)^2}{2\sigma_0^2}\right) = \frac{1}{2kT^2} \sigma_0^2 \left(\frac{\partial^2}{\partial i^2} [\Omega_s(i) - T\Omega'_s(i)] \right)_{i=i^*}, \quad (36)$$

where

$$\frac{1}{\sigma_0^2} = 2\pi Z^2 = -\frac{1}{kT} \left(\frac{\partial^2 \Omega_s(i)}{\partial i^2} \right)_{i=i^*}. \quad (37)$$

The last term in Eq. (34) is therefore

$$-\frac{1}{2T} \left(\frac{\partial^2 \Omega_s(i^*)}{\partial i^2} \right)^{-1} \left(\frac{\partial^2}{\partial i^2} [\Omega_s(i^*) - T\Omega'_s(i^*)] \right) = -\frac{1}{2T} + \frac{1}{2} \frac{\partial}{\partial T} \ln \left(-\frac{\partial^2 \Omega_s(i^*)}{\partial i^2} \right). \quad (38)$$

Note the shorthand denoting the evaluation of the derivatives at the critical size. The partial derivative with respect to T in the last term is performed holding i^* constant. (The dimensionality of the argument of the logarithm can be taken care of, if wished, by the insertion of an arbitrary constant.) Equation (34) then becomes

$$\left(\frac{\partial \ln J}{\partial T}\right)_{\ln S} = \frac{2p'_{vs}}{p_{vs}} - \frac{2}{T} - \frac{\Omega_s(1) - T\Omega'_s(1)}{kT^2} + \frac{2v'_i}{3v_i} + \frac{\Omega_s(i^*) - T\Omega'_s(i^*)}{kT^2} + \frac{1}{2} \frac{\partial}{\partial T} \ln \left(-\frac{\partial^2 \Omega_s(i^*)}{\partial i^2} \right). \quad (39)$$

Now,

$$\Omega_s(i) - T \frac{\partial \Omega_s(i)}{\partial T} = F(i) - T \frac{\partial F(i)}{\partial T} - i \left(\mu_s - T \frac{\partial \mu_s}{\partial T} \right), \quad (40)$$

and since $Q_i = \exp[-F(i)/kT]$ is proportional to the integral of $\exp[-H_c(i)/kT]$ over the phase space available to the i molecules in the cluster (as determined by the cluster definition), where $H_c(i)$ is the i -cluster Hamiltonian, we can deduce that

$$F(i) - T \frac{\partial F(i)}{\partial T} = \frac{\partial \ln Q_i}{\partial [-(kT)^{-1}]} = \frac{\int \prod_{j=1}^i d^3 \mathbf{x}_j d^3 \mathbf{p}_j H_c(i) \exp[-H_c(i)/kT]}{\int \prod_{j=1}^i d^3 \mathbf{x}_j d^3 \mathbf{p}_j \exp[-H_c(i)/kT]} = \bar{E}(i), \quad (41)$$

where the integrals over the molecular positions and momenta are restricted by the cluster definition, and $\bar{E}(i)$ is the mean energy of an i -cluster in a canonical ensemble at temperature T . The bar over E emphasizes that fluctuations in the cluster internal energy are likely to be substantial for such a small system.

Now we need the derivative of μ_s with respect to T in Eq. (40). μ_s is the molecular Gibbs free energy of a bulk vapor phase at a pressure p_{vs} and we can use the Gibbs-Duhem relation

$$s_l dT - v_l dp_{vs} + d\mu_s = 0, \quad (42)$$

where s_l and v_l are the entropy and volume per molecule in the bulk liquid phase, to deduce that

$$\mu_s - T \frac{\partial \mu_s}{\partial T} = h_l - Ts_l - T \left(v_l \frac{dp_{vs}}{dT} - s_l \right) = h_l - Tv_l p'_{vs} = e_l + v_l (p_{vs} - Tp'_{vs}), \quad (43)$$

where h_l and e_l are the enthalpy and internal energy per molecule in the bulk liquid phase when in equilibrium with the vapor. Therefore Eq. (39) becomes

$$\left(\frac{\partial \ln J}{\partial T} \right)_{\ln S} = \frac{2p'_{vs}}{p_{vs}} - \frac{2}{T} - \frac{E_x(1)}{kT^2} + \frac{E_x(i^*)}{kT^2} - (i^* - 1) \frac{v_l (p_{vs} - Tp'_{vs})}{kT^2} + \frac{2v'_l}{3v_l} + \frac{1}{2} \frac{\partial}{\partial T} \ln \left(- \frac{\partial^2 \Omega_s(i^*)}{\partial i^2} \right), \quad (44)$$

where

$$E_x(i) = \bar{E}(i) - ie_l \quad (45)$$

is the excess internal energy of an i -cluster. This is the mean energy of the cluster minus the energy the i molecules would have, on average, in the bulk liquid phase at the same temperature and pressure of the vapor.

Using the Clausius-Clapeyron equation $dp_{vs}/dT = (h_v - h_l)/[(v_v - v_l)T]$ and $p_{vs}v_v = kT$, the first term on the right-hand side of Eq. (44) can be shown to be equal to $2L[1 + v_l/(v_v - v_l)]/(kT^2)$, where $L = h_v - h_l$ is the latent heat per molecule, and h_v and v_v are the enthalpy and volume per molecule in the bulk saturated vapor.

Since $\bar{E}(1)$ is the mean energy of a single vapor molecule, we can also write

$$E_x(1) = \bar{E}(1) - e_l = h_v - h_l - p_{vs}(v_v - v_l) = L - kT + p_{vs}v_l. \quad (46)$$

Equation (44) then can be written as

$$\left(\frac{\partial \ln J}{\partial T} \right)_{\ln S} = \frac{L}{kT^2} \left(1 + \frac{2v_l}{v_v - v_l} \right) - \frac{1}{T} - \frac{p_{vs}v_l}{kT^2} - (i^* - 1) \frac{v_l (p_{vs} - Tp'_{vs})}{kT^2} + \frac{E_x(i^*)}{kT^2} + \frac{2v'_l}{3v_l} + \frac{1}{2} \frac{\partial}{\partial T} \ln \left(- \frac{\partial^2 \Omega_s(i^*)}{\partial i^2} \right). \quad (47)$$

Neglecting terms of order v_l/v_v , and also the last two logarithmic terms, which in most circumstances will be small compared to the others, Eq. (47) reduces finally to the second nucleation theorem:

$$\left(\frac{\partial \ln J}{\partial T} \right)_{\ln S} = \frac{1}{kT^2} [L - kT + E_x(i^*)]. \quad (48)$$

Furthermore, it is possible to write $L - kT \approx e_v - e_l = e_x$, which defines the mean excess internal energy e_x of the vapor. The right-hand side of Eq. (48) then reduces to $[e_x + E_x(i^*)]/kT^2$.

The derivation of the theorem given here is more rigorous than the earlier treatment [17], and includes a number of small terms in Eq. (47). However, in the original version of this theorem an additional term $(L - kT)/kT^2$ appeared on

the right-hand side of Eq. (48). This arose because the thermodynamic treatment focused on the temperature dependence of the cluster work of formation: the exponential factor in the rate expression. The kinetic prefactor from classical nucleation theory was assumed, and this ultimately gave rise to the additional term. This suggests that the classical kinetic prefactor is incorrect, and this will be explored in the next section. The use of the classical prefactor to complete the derivation of the theorem meant that the excess internal energies extracted from experimental nucleation data in Ref. [17] are too small by approximately $20kT$. A reanalysis of the data is given in Sec. VI.

V. TESTS OF MODELS

A. Classical nucleation theory: The problem

The nucleation theorems can be used to test the internal consistency of the classical theory of homogeneous nucleation. This model was derived originally from thermodynamic arguments. The main assumption is that the equilibrium cluster populations are given by

$$n_{i,\text{cl}}^e = n_1^e \exp\{-[\epsilon_{\text{cl}}(i) - \epsilon_{\text{cl}}(1)]/kT\}, \quad (49)$$

where, for large i [10],

$$\epsilon_{\text{cl}}(i) - \epsilon_{\text{cl}}(1) = \sigma A_i - ikT \ln S. \quad (50)$$

$\epsilon_{\text{cl}}(i)$ is a function that plays the role of the work of formation of an i -cluster. We comment on what it represents later. The first term is the surface free energy of a spherical droplet with the bulk liquid density and bulk surface tension σ . The surface area A_i is taken to be equal to $A_0 i^{2/3}$ with $A_0 = (36\pi v_l^2)^{1/3}$. n_1^e is the monomer population, given approximately by $n_1^e = p_v V/kT$. Equation (49) is strictly valid for $S < 1$.

Using Eq. (49) we can go directly to the kinetic derivation of the nucleation rate, starting from Eq. (13) and proceeding to Eq. (25) with $\epsilon(i)$ replaced by $\epsilon_{\text{cl}}(i)$, or equivalently $\Omega_s(i)$ replaced by σA_i . Inserting the classical expression into Eqs. (22), (25), and (26) gives the classical critical size $i_{\text{cl}}^* = [2\sigma A_0 / (3kT \ln S)]^3$ and classical nucleation rate

$$J_{\text{cl}} = \left(\frac{2\sigma}{\pi m}\right)^{1/2} V \frac{S^2 p_{vs}^2 v_l}{(kT)^2} \exp\left(-\frac{\sigma A_0 i_{\text{cl}}^{*2/3}}{3kT}\right). \quad (51)$$

We now test the compatibility of the classical rate with the first nucleation theorem by calculating the derivative of $\ln J_{\text{cl}}$ with respect to $\ln S$. We find

$$\left(\frac{\partial \ln J_{\text{cl}}}{\partial \ln S}\right)_T = 2 - \frac{1}{3} \frac{\sigma A_0}{kT} \frac{(-2i_{\text{cl}}^{*2/3})}{\ln S} = 2 + i_{\text{cl}}^*. \quad (52)$$

The classical theory is therefore inconsistent with the first nucleation theorem, given in Eq. (30). It is well known that the classical theory and the law of mass action are incompatible and the failure to comply with the first nucleation theorem is another reflection of this. Both problems can be corrected by multiplying J_{cl} by a factor of $1/S$ [25], and there have been several attempts to justify this within the classical formalism.

Now we calculate the partial derivative of $\ln J_{\text{cl}}$ with respect to T :

$$\begin{aligned} \left(\frac{\partial \ln J_{\text{cl}}}{\partial T}\right)_{\ln S} &= \frac{1}{2\sigma} \frac{d\sigma}{dT} + 2 \left(\frac{L}{kT^2} - \frac{1}{T}\right) - (i_{\text{cl}}^* \ln S - 1) \frac{d \ln v_l}{dT} \\ &+ \frac{1}{kT^2} \left(\sigma - T \frac{d\sigma}{dT}\right) A_0 i_{\text{cl}}^{*2/3}. \end{aligned} \quad (53)$$

We must compare this with Eq. (47). For the classical theory

$$\frac{\partial^2 \Omega_s(i)}{\partial i^2} = \frac{\partial^2 (\sigma A_i)}{\partial i^2} = -\frac{2}{9} \sigma A_0 i^{-4/3}. \quad (54)$$

Note that the last term on the right-hand side of Eq. (47) only contributes, therefore, through the temperature dependence of the surface tension and liquid density. It is a small term, as suggested earlier, but we will retain it rather than use the approximate form of the second nucleation theorem, given in Eq. (48). However, we neglect all terms of order v_l/v_v in Eq. (47), to obtain

$$\left(\frac{\partial \ln J_{\text{cl}}}{\partial T}\right)_{\ln S} = \frac{1}{2\sigma} \frac{d\sigma}{dT} + \left(\frac{L}{kT^2} - \frac{1}{T}\right) + \frac{d \ln v_l}{dT} + \frac{E_x^{\text{cl}}(i_{\text{cl}}^*)}{kT^2}, \quad (55)$$

so the excess internal energy for the critical cluster in the model, as determined by the second nucleation theorem, is

$$E_x^{\text{cl}}(i_{\text{cl}}^*) = \left(\sigma - T \frac{d\sigma}{dT}\right) A_0 i_{\text{cl}}^{*2/3} - kT^2 \frac{d \ln v_l}{dT} i_{\text{cl}}^* \ln S + (L - kT). \quad (56)$$

Now, since $i_{\text{cl}}^* kT \ln S = (2/3) \sigma A_0 i_{\text{cl}}^{*2/3}$, the first two terms are proportional to the surface area of the critical droplet. If these were the only terms, the internal energy of the droplet would be given by a term proportional to the volume (ie_l) plus a term proportional to the surface area (the excess internal energy), and this would be consistent with the underlying capillarity approximation. However, the last term in Eq. (56) spoils this picture. Unfortunately, it is not sufficient simply to neglect it in comparison with the other terms. It is a symptom of a deeper inconsistency within classical theory that needs to be resolved.

The failure to comply with the two nucleation theorems tell us that the classical theory is incomplete. The violations arise because Eq. (50) does not represent the work of formation of a cluster correctly. The right-hand side of Eq. (50) is in fact the work of formation of a classical droplet, *which has a fixed position in the system*. We can identify $\epsilon_{\text{cl}}(i)$ with the work of formation of such a droplet and neglect $\epsilon_{\text{cl}}(1)$ in comparison. What we really need, however, is the free energy of a cluster which can appear anywhere in the system. A symptom of this problem is that $\exp[-\epsilon_{\text{cl}}(i)/kT]$ with $\epsilon_{\text{cl}}(i)$ given by the right-hand side of Eq. (50) is not proportional to V as it should be. The missing term is the positional entropy (mixing entropy [18]) arising from the translation of clusters throughout the system volume. Note this is not the same as introducing the translational kinetic energy of a cluster into the excess free energy, which we shall comment on shortly.

B. Repairing classical nucleation theory

Let us now return to Eq. (25) and see how we might be able to derive the classical rate expression taking into account positional entropy, and therefore repair classical nucleation theory. The ingredient that has to be provided by some physical model is the cluster work of formation:

$$\epsilon(i) = \Omega_s(i) - ikT \ln S = F(i) - i\mu_s - ikT \ln S. \quad (57)$$

Now, since $\mu_s = g_l$, the Gibbs free energy per molecule in the bulk liquid phase when in equilibrium with the vapor, we can write $\mu_s = f_l + p_{vs}v_l$, where f_l is the molecular Helmholtz free energy in the liquid phase, so that

$$\epsilon(i) = F_x(i) - ip_{vs}v_l - ikT \ln S \approx F_x(i) - ikT \ln S, \quad (58)$$

where $F_x(i)$ is the excess Helmholtz free energy of the i -cluster, defined by

$$F_x(i) = F(i) - if_l. \quad (59)$$

Again, this excess quantity is the free energy of the cluster minus i times the free energy per molecule in the condensed phase.

We need to calculate $\epsilon(1)$ for use in Eq. (25). We write

$$\epsilon(1) = F(1) - \mu_s - kT \ln S, \quad (60)$$

where $F(1)$ is the Helmholtz free energy of a monomer in the volume V , and μ_s is the common chemical potential of a vapor and its condensate at equilibrium at a temperature T . From elementary statistical mechanics, $F(1) = -kT \ln(V\gamma)$ and $\mu_s = -kT \ln(\gamma\rho_{vs})$, where we have approximated the monomer density in the saturated vapor by the molecular density ρ_{vs} . The factor γ is equal to $(2\pi mkT)^{3/2}/h^3$, where h is Planck's constant. Then

$$\exp[\epsilon(1)/kT] = \frac{1}{S\rho_{vs}V} = \frac{v_v}{SV}. \quad (61)$$

We shall now construct a repaired classical theory, starting by invoking the capillarity approximation, so that the excess free energy for a spherical droplet with the bulk liquid density and bulk surface tension is

$$F_x^{\text{cap}}(i) = \sigma A_i - kT \ln V/v_c. \quad (62)$$

The first term, the classical excess free energy, has been criticized on the grounds that contributions from cluster translational and rotation kinetic energy do not appear, so that the free energy does not represent all the degrees of freedom available. However, it is the *excess* free energy that is required, and it seems likely that there is no strong contribution to $F_x(i)$ from molecular kinetic energy. The molecules in the cluster probably have a similar mean kinetic energy to that which they have in the bulk liquid. This point has been debated for thirty years and has shrouded the application of statistical mechanics to nucleation in controversy. Perhaps the above argument clarifies the point: it is largely the molecular potential energy that plays a role in nucleation and not the kinetic energy terms.

A second criticism of the capillarity approximation is that small molecular clusters are most unlikely to possess the

scaled-down properties of a spherical droplet of bulk liquid. However, we shall continue to pursue this model in the spirit of trying to repair classical theory.

The second term in Eq. (62) is the positional entropy term due to the contribution to the partition function from translated copies of every cluster. It was stated in Sec. II that the cluster free energy should be proportional to the volume. The so-called scaling volume v_c appears in Eq. (62) in order to maintain the correct dimensions. It acts as a means of resolving and counting translated states in the system. This ‘‘quantum’’ of volume has been discussed extensively elsewhere [18,19]. It appears in coarse-grained statistical models where the position of a mesoscale object is a degree of freedom; droplet models are in this class, as are models of microemulsion behavior. Often the size of the scaling volume can be obtained intuitively [18]. Otherwise one needs to refer back to the statistical mechanics of the underlying system, defined in the full phase space of all the degrees of freedom.

The work of formation based on the capillarity approximation is then, to a good approximation,

$$\epsilon_{\text{cap}}(i) = \sigma A_0 i^{2/3} - kT \ln V/v_c - ikT \ln S. \quad (63)$$

This form should not be expected to apply for small i , and when it is used, it will be assumed that i is large. The critical size is found by solving $\partial \epsilon^{\text{cap}}(i^*)/\partial i = 0$, which yields the classical expression $i^* = i_{\text{cl}}^*$, if the scaling volume v_c is independent of i .

The nucleation rate is now obtained from Eqs. (25), (54), and (61):

$$J_{\text{cap}} = \left(\frac{2\sigma}{\pi m} \right)^{1/2} \frac{S p_{vs} v_l}{kT} \frac{V}{v_c} \exp\left(- \frac{\sigma A_0 i_{\text{cl}}^{*2/3}}{3kT} \right). \quad (64)$$

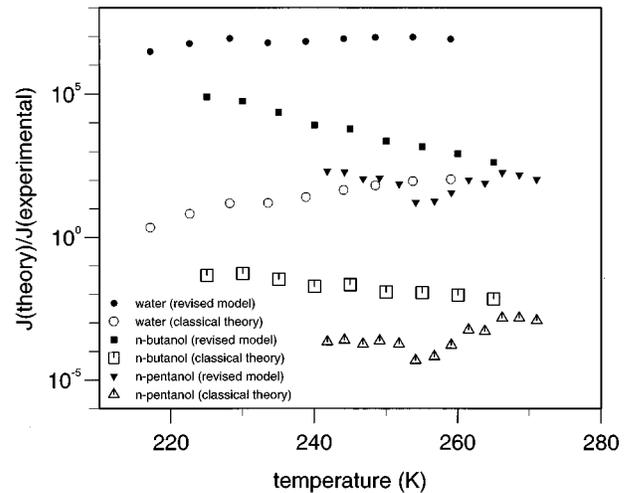


FIG. 2. The classical nucleation rate predictions divided by the experimental rates [15,29,31] are shown as open symbols for various substances and temperatures. The filled symbols show the same ratio according to a revised model based on the capillarity approximation but satisfying the nucleation theorems. The enhancement factor is $v_v/(SV_l)$.

If v_c did depend on i , the critical size of the revised model would not be the same as in the classical model. To a first approximation, however, we could ignore this influence and simply evaluate v_c in Eq. (64) at the classical critical size.

We now can see that the classical rate expression, Eq. (51), can be obtained from a more rigorous statistical mechanical–kinetic approach using the capillarity approximation, but only if we take the scaling volume in Eq. (64) to be the molecular volume in the supersaturated vapor, namely, $v_c = v_v/S = kT/(Sp_{v_s})$.

However, the fact that the classical rate does not satisfy the nucleation theorems warns us against this choice. Recent work has suggested that the scaling volume for a droplet model based on the capillarity approximation should be of the order of the molecular volume in the condensed phase, not in the supersaturated vapor [18]. Indeed Reiss *et al.* [19] have recently examined the scaling volume for a droplet model based on the capillarity approximation and proposed that for sizes relevant to nucleation, it takes the approximate form $v_c = v_l i^{1/2}$.

We can examine the properties of the scaling volume v_c by requiring that Eq. (64) should satisfy the nucleation theorems. Repeating the steps of Eqs. (52)–(56) we find that if v_c is independent of both S and T , then both nucleation theorems are satisfied, with the excess cluster internal energy being given by Eq. (56) without the unwanted final term. As forecast, this extra term was a symptom of a major problem, which is now seen to be either the total neglect of positional entropy in the classical cluster free energy, or equivalently the use of an incorrect scaling volume. We conclude that the classical rate should be corrected by a factor of $v_v/(Sv_c)$, where v_c is a temperature- and supersaturation-independent volume.

Several other nucleation models (not necessarily based on the capillarity approximation) have suggested that the classical rate should be corrected by a factor $v_v/(Sv_l)$ [2,26,27], together with additional numerical factors and powers of i^* . We see that this correction factor has a suitable form, apart from the minor temperature dependence of v_l , to ensure compliance with the nucleation theorems. The scaling volume is then indeed proportional to the molecular volume in the liquid.

We examine the effect of the rate enhancement factor $v_v/(Sv_l)$ in Fig. 2 for various substances over typical experi-

mental temperature ranges. The open symbols denote the ratio of J_{cl} to experimentally measured nucleation rates. The filled symbols denote the same ratio multiplied by the appropriate factor of $v_v/(Sv_l)$, which takes values between 10^4 and 10^6 . Unfortunately, the revised model is no more successful than the classical theory in collapsing all the data onto a single line, or even in accounting for the experimental temperature dependence. In several cases the predicted rate has an improved temperature dependence, but this is not universal, as demonstrated by the data for *n*-butanol. Nevertheless, the revised theory is better founded than classical theory, and presumably it is the capillarity approximation that leads to the poor agreement.

C. Internally consistent classical theory

The nucleation theorems allow us also to study the so-called internally consistent classical theory (ICCT) [28] for which the nucleation rate is

$$J_{\text{ICCT}} = \left(\frac{2\sigma}{\pi m} \right)^{1/2} V \frac{Sp_{v_s}^2 v_l}{(kT)^2} \exp(\sigma A_0/kT) \exp\left(-\frac{\sigma A_0 i_{cl}^{*2/3}}{3kT} \right). \quad (65)$$

We can derive this expression in two ways. The model was originally developed by employing Eq. (49) with the terms in the exponent given by

$$\epsilon_{\text{ICCT}}(i) - \epsilon_{\text{ICCT}}(1) = (\sigma A_0 i^{2/3} - ikT \ln S) - (\sigma A_0 - kT \ln S). \quad (66)$$

This is motivated by a desire for an expression that gives the correct result at $i=1$ in Eq. (49): this is the internal consistency that the model is designed to achieve. The critical size in the model is the same as for classical theory, i_{cl}^* , and the supersaturation dependence of Eq. (65) satisfies the first nucleation theorem. However, when we use the second nucleation theorem to extract the excess internal energy, the same problem we encountered in the classical theory appears. We find that

$$\left(\frac{\partial \ln J_{\text{ICCT}}}{\partial T} \right)_{\ln S} = \frac{1}{2\sigma} \frac{d\sigma}{dT} + 2 \left(\frac{L}{kT^2} - \frac{1}{T} \right) - [(i_{cl}^* - i_{cl}^{*1/3}) \ln S - 1] \frac{d \ln v_l}{dT} + \frac{1}{kT^2} \left(\sigma - T \frac{d\sigma}{dT} \right) A_0 (i_{cl}^{*2/3} - 1), \quad (67)$$

so the excess internal energy for the ICCT is

$$\begin{aligned} E_x^{\text{ICCT}}(i_{cl}^*) &= \left(\sigma - T \frac{d\sigma}{dT} \right) A_0 (i_{cl}^{*2/3} - 1) - kT^2 \frac{d \ln v_l}{dT} (i_{cl}^* - i_{cl}^{*1/3}) \ln S + (L - kT) \\ &= \left(1 - \frac{T}{\sigma} \frac{d\sigma}{dT} - \frac{2T}{3} \frac{d \ln v_l}{dT} \right) \sigma A_0 (i_{cl}^{*2/3} - 1) + (L - kT). \end{aligned} \quad (68)$$

The inconsistency this time is that the excess internal energy might be expected to be proportional to $(i_{cl}^{*2/3} - 1)$, bearing in mind the form of Eq. (66), but the last term in Eq. (69) once again spoils matters.

We now derive the ICCT rate expression by a second route, taking into account positional entropy, and determine the scaling volume, which is implicit in the model. We use the following cluster work of formation

$$\epsilon_{ICCT}^{rev}(i) = \sigma A_0 (i^{2/3} - 1) - (i - 1)kT \ln S - kT \ln V/v_c. \quad (70)$$

Equation (25) now leads, together with Eqs. (54) and (61), to the following nucleation rate:

$$J = \left(\frac{2\sigma}{\pi m} \right)^{1/2} V \frac{p_{vs} v_l}{kT} \frac{V}{v_c} \exp(\sigma A_0/kT) \exp\left(- \frac{\sigma A_0 i_{cl}^{*2/3}}{3kT} \right). \quad (71)$$

The ICCT expression in Eq. (65) is consistent with the choice of scaling volume $v_c = v_v/S$ in Eq. (71), as before. This is not surprising, since the ICCT result is simply the classical theory expression multiplied by a particular factor chosen to correct a certain apparent inconsistency within the model. In fact, it is debatable whether the ICCT corrects a true inconsistency, since the problem only arises if the capillarity approximation used in Eq. (50) is forced to apply for $i = 1$, which is clearly unphysical. Equation (49) is actually consistent by construction.

What we have shown is that the rate expressions for classical theory and the ICCT can be derived from a cluster model with a scaling volume equal to the volume per molecule in the supersaturated vapor. However, there is every indication that the scaling volume for a model based on the capillarity approximation should be virtually independent of S and T , and perhaps of the order of the volume per molecule in the liquid state. This emerges since the models fail to comply with the nucleation theorems.

VI. DATA ANALYSIS

Finally, the nucleation theorems given in Eqs. (30) and (48) will be used to extract the excess internal energy of critical clusters from experimental nucleation data. This procedure was used in Ref. [17] but using the earlier version of the second theorem, in which an additional $(L - kT)/kT^2$ term appeared on the right-hand side. As we have seen in this paper, this was due to the use of a classical kinetic prefactor in the derivation. The excess internal energies reported in [17] are therefore too small by approximately $20kT$.

Data for water [15] and n -butanol [29] have therefore been reanalyzed and the resulting plots of $E_x(i^*)$ against critical size i^* are given in Fig. 3. Data for n -heptane and n -octane, and n -nonane and n -decane [30] are also analyzed and the excess internal energies are shown in Figs. 4 and 5, respectively. Data for n -pentanol [31] and dibutylphthalate (DBP) [32] have also been examined, and the excess internal energies are shown in Fig. 6. The curves in the figures show values of $E_{xc}(i) = (\sigma - Td\sigma/dT)A_i$ based on the capillarity approximation.

Errors in the data points are small along the size axis, but are possibly more significant along the energy axis, as is

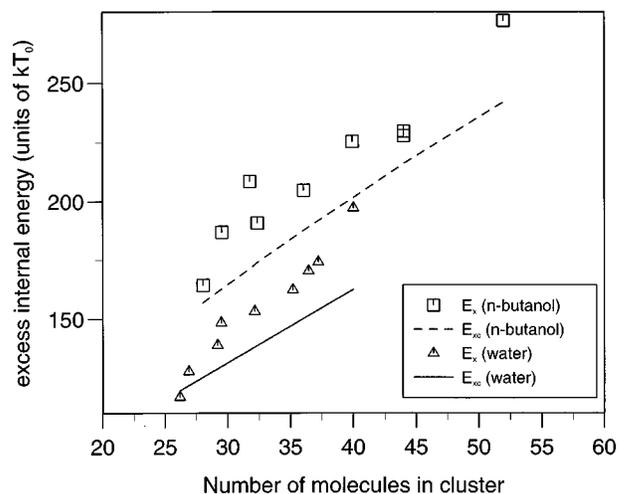


FIG. 3. Dimensionless excess internal energy E_x/kT_0 for clusters of i molecules of the substances water and n -butanol, extracted from experimental data [15,29] using the nucleation theorems. The excess internal energies according to the capillarity approximation E_{xc} are also shown as curves. The reference temperature is $T_0 = 273.15$ K.

evident in the scatter of values in some cases. The deviations between data points and the capillarity model predictions are in some cases not too large, but more often they are substantial. This is simply a reflection of the failure of the capillarity approximation to describe small clusters accurately. Studies of molecular cluster structure, using reasonable intermolecular potentials and a suitable cluster definition, are planned in order to try to account for the observed excess internal energies. Such studies are likely to be more straightforward than studies of the free energy.

VII. CONCLUSIONS

The first part of this paper gave a derivation of the two nucleation theorems using a combination of statistical mechanics and a kinetic approach. The results confirm previous

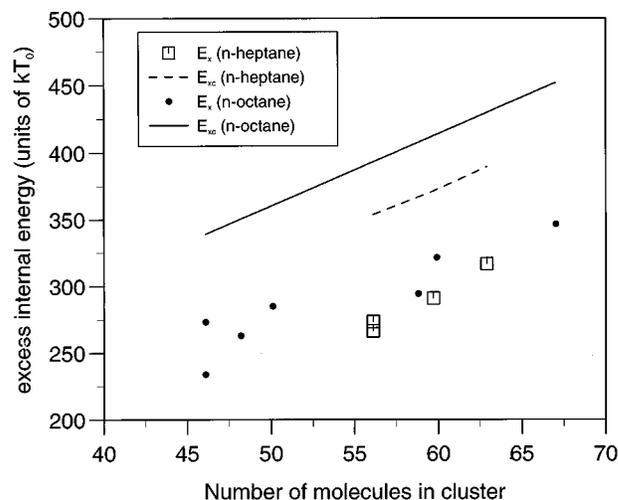


FIG. 4. Same as Fig. 3, but for the substances n -heptane and n -octane [30].

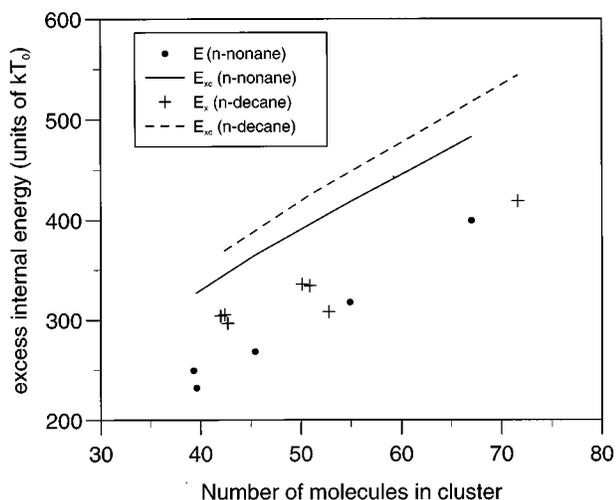


FIG. 5. Same as Fig. 3, but for the substances *n*-nonane and *n*-decane [30].

derivations from continuum and small system thermodynamics. These theorems relate the temperature and supersaturation dependence of the nucleation rate to properties of the critical molecular cluster, which is the size that has the maximum cluster work of formation, and which is loosely the size equally likely to grow or decay.

The derivation described here yields the temperature and supersaturation dependence of the nucleation rate directly. The approximations that have been made in the derivation are not unusual and the results seem to be soundly based. In particular, the results are valid for any choice of cluster definition used in the statistical mechanics. On the other hand, the suitability of the definition will determine the degree to which critical cluster properties calculated from models will correspond to the properties of real critical clusters. This choice is crucial to the success of the theory. A good definition must include all molecular configurations in phase space that correspond to the relatively long-lived states that take part in cluster population dynamics, and exclude all ephemeral states. Physical intuition, and ultimately comparison

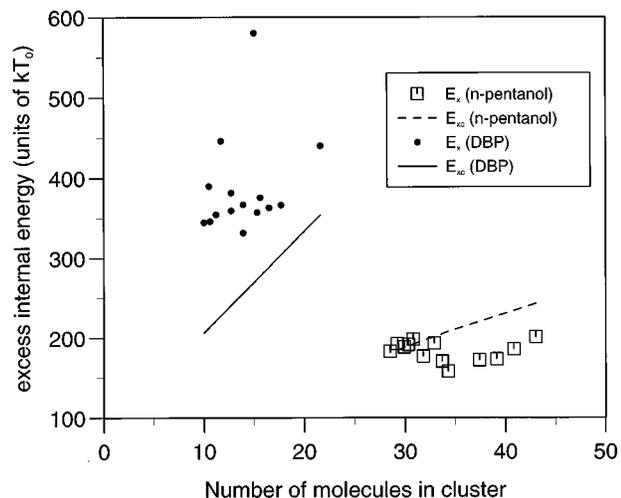


FIG. 6. Same as Fig. 3, but for the substances *n*-pentanol [31] and dibutylphthalate (DBP) [32].

with experimental data, should guide the choice of cluster definition used in calculations.

The theorems were derived previously by studying only the properties of the critical cluster work of formation, and a classical prefactor was assumed. The version of the second nucleation theorem derived here differs from the earlier version and this implies that the classical kinetic prefactor is incorrect.

To reinforce this point, it has been shown that the classical nucleation rate, and its so-called internally consistent revision, fail to comply with the nucleation theorems. The excess internal energy derived from the model nucleation rate expression does not have the expected form. The reason for this failure is the following. An approximate calculation of the grand partition function of a vapor within statistical mechanics requires the free energy of a single cluster that is able to occupy any position within the system. It turns out, though, that classical theory is based on the properties of a spherical droplet with its center fixed at one position. The excess cluster free energy, within the capillarity approximation, is proportional to the droplet surface: it is independent of the system volume V . However, the enumeration of all positions of a free cluster within the volume must introduce a $\ln V$ term into the free energy. On dimensional grounds V appears together with a scaling volume v_c , which is in effect the fundamental cell volume in the positional phase space. The existence of this positional, or mixing entropy term, is implicit in the present derivation of the nucleation theorems.

It is possible to derive the classical rate expression from a cluster work of formation that includes positional entropy by choosing the scaling volume to be the molecular volume in the supersaturated vapor v_v/S . The same conclusion can be reached for the so-called internally consistent classical theory. However, compliance with the theorems requires the scaling volume to be independent of supersaturation and temperature, at least to a first approximation, and so this justification of the classical formula is not possible. From other work [18,19] it would appear that the molecular volume in the condensate v_l is a natural scaling volume for models based on the capillarity approximation.

The population of i -clusters in subsaturated vapors is given by

$$n_i^e = n_1^e \exp\{-[\epsilon(i) - \epsilon(1)]/kT\} \quad (72)$$

with $\epsilon(i) = F(i) - i(\mu_s + kT \ln S)$ for all i . This expression is mathematically consistent at $i=1$ and in agreement with the law of mass action since $S \propto n_1^e$ and so $n_i^e \propto (n_1^e)^i$. The free energy $F(i)$ contains the term $-kT \ln(V/v_c)$ for $i \neq 1$, with v_c independent of S and T . $\epsilon(1)$ is simply $-kT \ln(SV/v_v)$. These logarithmic terms reflect the freedom for droplets to nucleate anywhere in the system. If v_c is given by v_l , as seems likely, then the classical nucleation rate is corrected by a factor $v_v/(Sv_l)$. This factor is strongly temperature dependent and of order $10^4 - 10^6$ numerically. It has appeared several times before in theories of nucleation [2,26,27].

The revised rate may be more acceptable on formal grounds, but it still relies on the capillarity approximation, namely, that the cluster density and surface properties are those of the bulk condensate. For this reason, it should not be

expected to be an accurate model. However, the revised theory rests on a much firmer basis in statistical mechanics than the original classical theory. Disagreement with experiment is likely to be due to the physical assumptions made rather than any mathematical inconsistencies. More realistic models might be used within the same framework to make

progress towards successful nucleation rate predictions.

Finally, the nucleation theorems have been used to plot cluster excess internal energies against cluster size for a number of substances, using experimental nucleation rate data. This information will provide useful points of comparison with microscopic models of clusters.

-
- [1] J. E. Mayer and M. G. Mayer, *Statistical Mechanics* (Wiley, New York, 1940).
- [2] H. Reiss, J. L. Katz, and E. R. Cohen, *J. Chem. Phys.* **48**, 5553 (1968).
- [3] H. Reiss, A. Tabazadeh, and J. Talbot, *J. Chem. Phys.* **92**, 1266 (1990).
- [4] H. M. Ellerby and H. Reiss, *J. Chem. Phys.* **97**, 5766 (1992).
- [5] C. L. Weakliem and H. Reiss, *J. Chem. Phys.* **101**, 2398 (1994).
- [6] F. H. Stillinger, *J. Chem. Phys.* **38**, 1486 (1963).
- [7] C. F. Clement, in *Nucleation and Atmospheric Aerosols*, edited by N. Fukuta and P. E. Wagner (A. Deepak Publishing, Hampton, VA, 1992), p. 327.
- [8] A. A. Lushnikov and M. Kulmala (unpublished).
- [9] J. Lothe and G. M. Pound, *J. Chem. Phys.* **36**, 2080 (1962).
- [10] F. F. Abraham, *Homogeneous Nucleation Theory* (Academic, New York, 1974).
- [11] A. Dillmann and G. E. A. Meier, *J. Chem. Phys.* **94**, 3872 (1991).
- [12] A. Laaksonen, I. J. Ford, and M. Kulmala, *Phys. Rev. E* **49**, 5517 (1994).
- [13] V. I. Kalikmanov and M. E. H. van Dongen, *J. Chem. Phys.* **103**, 4250 (1995).
- [14] D. Kashchiev, *J. Chem. Phys.* **76**, 5098 (1982).
- [15] Y. Viisanen, R. Strey, and H. Reiss, *J. Chem. Phys.* **99**, 4680 (1993).
- [16] D. W. Oxtoby and D. Kashchiev, *J. Chem. Phys.* **100**, 7665 (1994).
- [17] I. J. Ford, *J. Chem. Phys.* **105**, 8324 (1996).
- [18] H. Reiss and W. K. Kegel, *J. Phys. Chem.* **100**, 10 428 (1996).
- [19] H. Reiss, W. K. Kegel, and J. L. Katz, *Phys. Rev. Lett.* **78**, 4506 (1997).
- [20] I. J. Ford, *J. Aerosol. Sci.* **23**, 447 (1992).
- [21] C. F. Clement and I. J. Ford, in *Proceedings of the 14th International Conference on Nucleation and Atmospheric Aerosols*, edited by M. Kulmala and P. E. Wagner (Pergamon, Oxford, 1996), p. 121.
- [22] O. V. Vasil'ev and H. Reiss, *J. Chem. Phys.* **105**, 2946 (1996).
- [23] O. V. Vasil'ev and H. Reiss, *Phys. Rev. E* **54**, 3950 (1996).
- [24] T. L. Hill, *Thermodynamics of Small Systems* (Dover, New York, 1994). (Originally published by W.A. Benjamin Inc., New York, in the *Frontiers in Chemistry* series, 1963, 1964).
- [25] W. G. Courtney, *J. Chem. Phys.* **35**, 2249 (1961).
- [26] J. C. Barrett, *J. Chem. Phys.* (to be published).
- [27] I. J. Ford, J. C. Barrett, and M. Lazaridis, *J. Aerosol. Sci.* **24**, 581 (1993).
- [28] S. L. Girschick and C.-P. Chu, *J. Chem. Phys.* **93**, 1273 (1990).
- [29] Y. Viisanen and R. Strey, *J. Chem. Phys.* **101**, 7835 (1994).
- [30] M. M. Rudek, J. A. Fisk, V. M. Chakarov, and J. L. Katz, *J. Chem. Phys.* **90**, 1856 (1989).
- [31] J. Hruby, Y. Viisanen, and R. Strey, *J. Chem. Phys.* **104**, 5181 (1996).
- [32] K. Hämeri and M. Kulmala, *J. Chem. Phys.* **105**, 7696 (1996).