A molecular based derivation of the nucleation theorem

R. K. Bowles
Department of Chemistry and Biochemistry, University of California, Los Angeles, 405 Hilgard Avenue, Los Angeles, California 90095

R. McGraw
Department of Applied Science, Brookhaven National Laboratory, Upton, New York 11973

P. Schaaf
Institut Charles Sadron (C.N.R.S.-U.L.P), 6, rue Boussingault, 67083 Strasbourg Cedex, France, and Ecole Europe`enne de Chemie, Poly`meres et Mat´eriaux de Strasbourg, 1, rue Blaise Pascal, BP 296F, 67008 Strasbourg Cedex, France

B. Senger and J.-C. Voegel
Institut National de la Sante´ et de la Recherche Medicale, Unit´e 424, F´ed´eration de Recherches ‘‘Odontologie,’’ Univerite´ Louis Pasteur, 11 rue Humann, 67085 Strasbourg Cedex, France

H. Reiss
Department of Chemistry and Biochemistry, University of California, Los Angeles, 405 Hilgard Avenue, Los Angeles, California 90095

(Received 16 March 2000; accepted 20 June 2000)

We show that, for condensation in an almost ideal vapor, the nucleation theorem is essentially a consequence of the law of mass action. The usual form of the theorem, with the effects of the translational degrees of freedom of the cluster included, is then derived using statistical mechanics and molecular theory, but only under the assumptions that the cluster excludes a volume to the surrounding vapor and that the vapor is ideal. The form of the result obtained via molecular theory is such that it appears unlikely (but not impossible) that the theorem remains valid for cases when the vapor is nonideal. This suggests that further work is necessary before the theorem can be regarded as established. We also consider the effects of the presence of a carrier gas.

I. THE NUCLEATION THEOREM AS THE LAW OF MASS ACTION

In the last few years a number of interesting articles, both theoretical and experimental, have been published concerning the so-called nucleation theorem. These articles, along with computer simulations, seem to have confirmed this rather remarkable formula which, in its most generally accepted form for a single component system, may be expressed as

\[
\left[ \frac{\partial W(n^*)}{\partial \mu} \right]_{V,T} = -(n^* - \bar{n}),
\]

where \(W(n^*)\) is the reversible work of the formation of the critical nucleus, \(\mu\) is the chemical potential of the metastable phase, \(V\) is the volume, and \(T\) is the temperature of the metastable phase. \(n^*\) is the total number of molecules in the volume designated to be the volume covered by the “nucleus,” while \(\bar{n}\) is the number of molecules that would occupy that volume in the uniform metastable phase. The relevance of Eq. (1) to the nucleation process lies in the fact that the nucleation rate depends exponentially on \(W(n^*)\) and a preexponential factor that is thought to be insensitive to \(\mu\). It is obvious that the quantity \(n^* - \bar{n}\) will vary unless the volume is chosen large enough to include some of the metastable phase. Unless \(\bar{n}\) is negligible, it becomes arbitrary to associate a molecular number with the nucleus; only the excess number can be given any meaning. Defining \(n^*\) as the number of molecules in the nucleus is inappropriate since \(n^*\) can be made arbitrarily large by defining the volume of the nucleus to be arbitrarily large.

Kashchiev originally conjectured Eq. (1), without \(\bar{n}\), on the basis of classical nucleation theory (CNT). More recently Eq. (1) was derived using both a statistical mechanics\(^2\) and a thermodynamic argument. However, both of these arguments were remiss in not supplying a clear definition of the cluster (e.g., a thought experiment indicating how one would measure the properties of the nucleus outside of measuring the rate itself). Thus, in spite of recent laboratory and computer experiments that seem to validate Eq. (1) the validity of the nucleation theorem in general remains in some doubt.

On the other hand, if the nucleating phase is essentially ideal, in particular an ideal gas, the theorem in the form originally proposed by Kashchiev\(^1\) can be derived almost trivially from the law of mass action. This law applied to a cluster containing \(n\) molecules may be expressed as

\[
\mu = n \mu_n,
\]

where \(\mu_n\) is the chemical potential of the cluster. Of course

\(^a\)Electronic mail: bowles@chem.ucla.edu
the cluster must be specified by some definition, but this is not relevant to Eq. (2). If the gas is ideal \( \mu_n \) may be expressed as

\[
\mu_n = -kT \ln \left[ \frac{V q_n^i}{n^{-3/2} \Lambda^3} \right] + kT \ln N_n,
\]

(3)

where \( \Lambda = h/(2\pi m kT)^{1/2} \), is the thermal de Broglie wave length of a single ideal gas molecule with a mass \( m \), \( q_n^i \) is the internal partition function of the cluster, and \( N_n \) is the number of clusters of size \( n \) in the system. Substitution of Eq. (2) in Eq. (3) yields, after appropriate rearrangements,

\[
N_n = \exp \left( -\frac{1}{kT} \left[ -kT \ln q_n^i n^{3/2} - kT \ln \frac{V}{\Lambda^3} - n \mu \right] \right)
\]

(4)

where, clearly

\[
W(n) = -kT \ln q_n^i n^{3/2} - kT \ln \frac{V}{\Lambda^3} - n \mu
\]

(5)

is the work required to form a cluster of size \( n \). Multiplying the right side of Eq. (4) by the number of monomers in the system, \( N_1 \), and compensating in the exponent, and then making use of the expression for the chemical potential of an ideal gas,

\[
\mu = kT \ln \left[ \frac{\Lambda^3 N_1}{V} \right].
\]

(6)

yields

\[
N_n = N_1 \exp \left( -\frac{W_{int}(n)}{kT} \right),
\]

(7)

where

\[
W_{int}(n) = -kT \ln q_n^i n^{3/2} - (n-1) \mu
\]

(8)

is what has been called the intensive free energy\(^7\,12\) of cluster formation. Equations (5) and (8) give expressions for the work required to form a cluster of any size \( n \), so that if we take their derivatives with respect to the chemical potential we would obtain what could be described as a general nucleation theorem. For example, from Eq. (5) we would obtain

\[
\left[ \frac{\partial W(n)}{\partial \mu} \right]_{n,V,T} = -n.
\]

(9)

However, in order to relate this derivative to the rate of nucleation, and hence to experiment, the derivative must be obtained at the maximum of the barrier \( W(n=n^*) \), which of course corresponds to \( n^* \). Since \( n^* \) is itself a function of the chemical potential, we have starting with Eq. (5),

\[
\left[ \frac{\partial W(n^*)}{\partial \mu} \right]_{n^*,V,T} = \left[ \frac{\partial W(n)}{\partial \mu} \right]_{n^*,V,T} + \left[ \frac{\partial W(n)}{\partial n} \right]_{n^*,V,T} \left[ \frac{\partial n^*}{\partial \mu} \right]_{n^*,V,T},
\]

(10)

but \( \left[ \frac{\partial W(n)}{\partial n} \right]_{n^*} = 0 \), so the second term on the right-hand side vanishes to give

\[
\left[ \frac{\partial W(n^*)}{\partial \mu} \right]_{V,T} = \left[ \frac{\partial W(n)}{\partial \mu} \right]_{n^*,V,T} = -n^*,
\]

(11)

where the second equality follows from Eq. (5). Similarly, Eq. (8) leads to

\[
\left[ \frac{\partial W_{int}(n^*)}{\partial \mu} \right]_{V,T} = -(n^*-1).
\]

(12)

Equations (11) and (12) are, in essence the nucleation theorem in the form first published by Kashchiev, i.e., not involving the *excess* number \( n^* - n \). As we have derived them, these relations are only valid for a gaseous metastable system that is ideal. However, as we have indicated, it is an almost trivial expression of the well-known law of mass action which renders the nucleation theorem, in spite of its usefulness, less remarkable than originally thought. Furthermore, Eqs. (11) and (12) include the effects of translation, which, due to a logarithmic dependence of \( W(n^*) \) or \( W_{int}(n^*) \) on \( n \), slightly alters the size of the critical nuclei. This effect is not included in the derivations in Refs. 2 and 3 where the cluster definitions are somewhat loose.

Equation (2) is of course not limited to ideal systems, but for a nonideal system \( \mu_n \) would have to include an appropriate thermodynamic activity coefficient that itself would depend on \( \mu \) in some as yet undetermined manner. This would then lead to an additional term appearing in Eqs. (11) and (12). In this connection it should be indicated that the derivation in Ref. 2 is restricted to ideal systems while that in Ref. 3, as it stands, is not. However, because of the lack of sharpness in the cluster definition in both Refs. 2 and 3, and the corresponding indefiniteness in the dependence of defined quantities on the chosen thermodynamic variables, even Ref. 3 does not assure the applicability of Eq. (1) to nonideal systems.

Because of these caveats, it seems advisable to attempt a fully molecular derivation of the nucleation theorem without restricting the argument to ideal systems. Such a derivation should reveal the conditions that underlie the entry of the *excess number* into the theory, and should also include translational degrees of freedom. Furthermore, it should shed light on the applicability of Eq. (1) to nonideal systems. We begin to develop such a theory in the following section. Section II describes the derivation of the nucleation theorem from a molecular perspective, before the effects of a carrier gas are examined in Sec. III. Our discussion and concluding remarks are contained in Sec. IV.

II. DERIVATION OF THE NUCLEATION THEOREM

We begin with the partition function for the system of \( N \) particles in a volume \( V \), constrained to contain at least one \( n \)-cluster\(^12\,14\)
\[ Q_n(N,V,T) = \frac{n^3}{n!(N-n)!\Lambda^{3N}} \int_V d\mathbf{R} \]
\[ \times \int_c \cdots \int_c e^{-\beta U_n(r'_2 \cdots r'_n) + U_n(r'_2 \cdots r'_N)} d\mathbf{r}'_2 \]
\[ \cdots d\mathbf{r}'_n \]
\[ \cdots \int_V \cdots \int_V e^{-\beta U_{N-n}(r_{n+1}' \cdots r'_N)} d\mathbf{r}'_{n+1} \cdots d\mathbf{r}'_N, \]
\[ (13) \]

where the thermal de Broglie wavelength is
\[ \Lambda = \hbar/(2\pi mkT)^{1/2}, \]
\[ (14) \]
in which \( m \) is the mass of a molecule, \( \mathbf{R} \) is the coordinate of the center-of-mass, and the primed coordinates, \( \mathbf{r}' = \mathbf{r} - \mathbf{R} \), are relative to the center-of-mass of the \( n \)-cluster. The factor \( n^3 \) arises from the transformation from the laboratory coordinates, \( \mathbf{r} \) to \( \mathbf{r}' \). \( U_n \) is the potential energy of the \( n \)-cluster, \( U_{N-n} \) is the potential energy of the \( N-n \) vapor molecules, while \( U_n \) is the potential energy of interaction between the cluster and the vapor. The integral over \( c \) is carried out in accordance with some cluster criterion which should avoid problems of redundancy. \[ 15 \]

We now introduce the following decoupled probability densities within the \( n \)-cluster,
\[ \Phi^0 = \frac{e^{-\beta U_n}}{\int_V d\mathbf{r}_c \cdots \int_c e^{-\beta U_n} d\mathbf{r}'_2 \cdots d\mathbf{r}'_n}, \]
\[ (15) \]
and for the remaining \( N-n \) molecules in the vapor,
\[ \Phi^0_{N-n} = e^{-\beta U_{N-n}}. \]
\[ (16) \]

Since we are interested in the effect of the translational free energy of the cluster on the nucleation theorem it is important to clearly separate the internal and translational degrees of freedom in its partition function. With this in mind, the partition function for the decoupled cluster can be written as
\[ q_n(V,T) = \frac{1}{n!\Lambda^3} \int_V d\mathbf{R} \int_c \cdots \int_c e^{-\beta U_n} d\mathbf{r}'_2 \cdots d\mathbf{r}'_n \]
\[ = q'_n(V,T)q'_n(c,T), \]
\[ (17) \]

where
\[ q'_n(c,T) = \int_c \cdots \int_c e^{-\beta U_n} d\mathbf{r}'_2 \cdots d\mathbf{r}'_n, \]
\[ (18) \]
is the internal partition function subject to a constant set of parameters in the constraints \( c \) that define the cluster, and
\[ q'_n(V,T) = \frac{V}{\Lambda^3} \]
\[ (19) \]
is the translational partition function associated with movement of the center-of-mass of the cluster throughout the volume \( V \). \( \Lambda = \hbar/(2\pi mkT)^{1/2} \) is the de Broglie wavelength of an \( n \)-cluster. Later it will be convenient to express the free energy of the cluster as an intensive free energy \[ 17,12 \] [see Eq. (8)], since it is this intensive free energy that is most simply related to the probability of directly observing an appropriately defined cluster in a simulation cell. This can be achieved by restricting the translation of the center-of-mass of the cluster to a molecular volume \[ 12 \] \( \hat{v} = V/N \) in the system, so that Eq. (19) becomes
\[ q'_n(V,T) = \frac{N^\hat{v}}{\Lambda_n^3} = Nq'_n(\hat{v},T), \]
\[ (20) \]
where \( q'_n(\hat{v},T) \) is the translational partition function of an \( n \)-cluster in a molecular volume. Equation (17) can now be written as
\[ q_n(V,T) = Nq'_n(\hat{v},T)q'_n(c,T). \]
\[ (21) \]
The partition function for the decoupled vapor is
\[ Q(N-n,V,T) = \frac{1}{(N-n)!\Lambda^{3(N-n)}} \]
\[ \times \int_V \cdots \int_V e^{-\beta U_{N-n}(r_{n+1}' \cdots r'_N)} d\mathbf{r}'_{n+1} \cdots d\mathbf{r}'_N, \]
\[ (22) \]

Using Eqs. (14) through (17) and Eq. (22) in Eq. (13), gives
\[ Q_n(N,V,T) = Q(N-n,V,T)q_n(V,T) \]
\[ \times \int_V d\mathbf{R} \int_c \cdots \int_c e^{-\beta U_n} d\mathbf{r}'_2 \cdots d\mathbf{r}'_n \]
\[ \times \int_V \cdots \int_V d\mathbf{r}'_{n+1} \cdots d\mathbf{r}'_N \Phi^0 e^{-\beta U_n} \Phi^0_{N-n} \]
\[ = Q(N-n,V,T)q_n(V,T)(e^{-\beta U_n})^0, \]
\[ (23) \]
where
\[ (e^{-\beta U_n})^0 = \int_V \cdots \int_c e^{-\beta U_n} d\mathbf{r}'_2 \cdots d\mathbf{r}'_n \]
\[ \cdots \int_V \cdots \int_V d\mathbf{r}'_{n+1} \cdots d\mathbf{r}'_N \Phi^0 e^{-\beta U_n} \Phi^0_{N-n}, \]
\[ (24) \]
and where the average \( \langle \cdot \rangle_0 \) is taken over the decoupled system. Equation (24) contains the interactions between the cluster and the remaining \( N-n \) molecules in the vapor.

As we will see, it is the evaluation of this interaction term that, on the one hand, gives rise to the validity of the Eq. (1) when the volume of the cluster is rigorously specified and the surrounding vapor is ideal, but also raises questions as to the validity of the nucleation theorem when the vapor is no longer ideal.

The probability that there exists at least one \( n \)-cluster in the system is given by
\[ \Pi_n = \frac{Q_n(N,V,T)}{Q(N,V,T)} = \frac{Q(N-n,V,T)q_n(V,T)(e^{-\beta U_n})^0}{Q(N,V,T)} = q_n(V,T)e^{\mu/kT}(e^{-\beta U_n})^0, \]
\[ (25) \]
where \( \mu \) is the chemical potential of the vapor. Furthermore, if the \( n \)-clusters are rare then we can write
where \( p_n(i) \) is the probability that there are exactly \( i \) \( n \)-clusters in the system, \( N_n \) is the equilibrium number of such clusters, and we have made use of Eq. (25). The approximations on the right can be made because the \( n \)-clusters are rare, so that \( p_n(j) \ll p_n(1), \ j > 1 \), allowing us to ignore the higher order terms of Eq. (26). In addition we will also assume the absence of smaller clusters in the vapor, i.e., that the vapor does not contain a significant number of dimers. Now \( N_1 \approx N \), so that Eq. (7) can be expressed as

\[
N_n = N e^{-W_{\text{int}}(n)/kT},
\]  

(27)

Substitution of Eqs. (21) and (26) into Eq. (27) gives

\[
W_{\text{int}}(n)/kT = -n \mu/kT - \ln q_n^i(\hat{v}, T) - \ln q_n^i(c, T) - \ln(e^{-\beta U_0})_0,
\]

(28)

where \( F_n^i(\hat{v}, T) \) and \( F_n^i(c, T) \) are the translational and internal Helmholtz free energies of an \( n \)-cluster, respectively. Before proceeding to derive the nucleation theorem, it is worthwhile to examine the content of Eq. (28) in more detail. It has already been shown that the translational free energy contribution to \( W_{\text{int}}(n) \), through \( F_n^i(\hat{v}, T) \), can under certain circumstances be of the order of \( 20kT \) which significantly affects the rate of nucleation. However, it should also be noted that translation only has an effect to terms of the order \( \ln n \) on the barrier height, thus only slightly altering the size of the critical nucleus. The overall effect of accounting for translation in the nucleation theorem will be to marginally alter the predicted size of the critical nucleus, relative to that predicted by "stationary models" which include only the small portion of the translation of the cluster associated with the fluctuation of the center-of-mass around a fixed point in the volume. Again, Eq. (28) gives the expression for the work required to form a cluster of any size \( n \), but to relate this to the nucleation rate we must restrict ourselves to considering \( n^* \).

The final term in Eq. (28) results from the interaction between the cluster and the surrounding vapor. We can best examine the importance of this term by initially deriving the nucleation theorem, assuming there is no interaction between the \( n \)-cluster and the remaining \( N-n \) molecules, not even an exclusion volume. This can be considered as the zeroth level of approximation and is achieved by setting \( U_\sigma \) in Eq. (24) to zero for all configurations. To obtain the nucleation theorem we need to differentiate Eq. (28) with respect to the chemical potential when \( W(n) \) is at its maximum, i.e., when \( n = n^* \) the size of the critical nucleus. \( F_n^i(c, T) \) is not a function of \( \mu \) and \( \ln(e^{-\beta U_0})_0 \), however examination of Eq. (20) and use of the expression for the chemical potential of an ideal gas shows

\[
F_n^i(\hat{v}, T)/kT = -\ln q_n^i(\hat{v}, T) = -\ln N \frac{V}{N^3_n} = \mu/kT.
\]

(29)

Substitution of Eq. (29) into Eq. (28) and taking the differential with respect to \( \mu \) leads to

\[
\left[ \frac{\partial W_{\text{int}}(n^*)}{\partial \mu} \right]_{v, T} = -(n^* - 1), \tag{30}
\]

which is just Eq. (12). However, most reasonable molecular based model clusters involve the specification of a volume which contains the molecules of the cluster and excludes the surrounding vapor, otherwise it would not be possible to distinguish cluster molecules from those of the vapor. We can begin to introduce some level of interaction by including a hard sphere exclusion volume so that although cluster molecules still have no individual interaction with the vapor there is a collective exclusion of the \( N-n \) vapor molecules from a spherical volume, \( v \), with radius \( R_c \) around its center-of-mass. This, in essence, is the same approximation implicit in clusters based on the small grand ensemble and density functional theory where the molecules belonging to the cluster can be identified as those within a cell of a specified volume, although, the present model differs from these earlier models in that its volume is defined relative to its center-of-mass rather than to a fixed point in the macroscopic volume. Clusters based on the Stillinger model also include an excluded volume interaction by requiring vapor molecules to be at least a set distance away from any of the particles within the cluster, which leads to an irregular excluded volume dependent on the instantaneous configuration of the cluster.

As we will now show, the excluded volume interaction plays a significant role in making the number of particles in the critical nucleus, obtained via the nucleation theorem, an excess quantity. The interaction potential can be expressed

\[
U_\sigma = \begin{cases} \infty & r_i < R_c \\ 0 & r_i \geq R_c \end{cases} (n+1 \leq i \leq N), \tag{31}
\]

where \( r_i \) is the distance of particle \( i \) of the vapor from the center-of-mass of the cluster. Since the \( n \) molecules within the cluster do not interact with the \( N-n \) molecules of the vapor other than through the exclusion volume, the integration over the degrees of freedom of the cluster can be carried out independently of the vapor. Equation (24) then reduces to

\[
\langle e^{-\beta U_\sigma} \rangle_0 = \frac{\int_{V-v} \cdots \int_{V-v} d^N v_{n+1} \cdots d^N v_{N-n} e^{-\beta U_{N-n}}}{\int_{V-v} \cdots \int_{V-v} d^N v_{n+1} \cdots d^N v_{N-n}} = \frac{Z(N-n, V-v)}{Z(N-n, V)} = e^{\Delta U_{N-n}(N-n, V-v)}, \tag{32}
\]

where \( Z(N-n, V), Z(N-n, V-v), F(N-n, V), \) and \( F(N-n, V-v) \) are the configurational integrals and Helmholtz free energies of the \( N-n \) vapor molecules in the volume \( V \) and \( V-v \), respectively. If additional interactions between the cluster and the vapor were included, terms involving the
configurational integral of the decoupled cluster under the influence of $U_\sigma$ would also appear in Eq. (32). Since $v \ll V$ the difference in Helmholtz free energy can be expressed

$$F(N-n,V)-F(N-n,V-v) = \left[ \frac{\partial F}{\partial N} \right]_V v + \left[ \frac{\partial F}{\partial a} \right]_V a$$

$$= -Pv + \sigma a,$$

where $P$ is the pressure of the vapor, $\sigma$ is the boundary free energy density or boundary tension of the vapor with the hard wall at the boundary of the exclusion volume, and $a$ is the surface area of the exclusion volume. Using Eq. (33) in Eq. (32), and substituting the result into Eq. (28) yields

$$W_{in}(n) = -n\mu + F_n^b(\hat{v},T) + F_n^d(c,T) + P v - \sigma a.$$  

Differentiation of Eq. (34) with respect to $\mu$, with $n$ set to $n^*$, yields

$$\left[ \frac{\partial W_{in}(n^*)}{\partial \mu} \right]_{V,T} = -(n^*-\bar{n} + a n^*_s - 1),$$

where $\bar{n}$ appears due to the differentiation of the $Pv$ term of Eq. (34) with respect to $\mu$, using the Gibbs relation $\partial P/\partial \mu = N/V$, so that $Nv/V = \bar{n}$. $n^*_s$ is the excess surface density (per unit area) that arises from the boundary tension of the exclusion volume with the vapor and is obtained from the Gibbs surface of tension relation

$$d\sigma = -n^*_s d\mu,$$

at constant $T$. If we assume the vapor to be ideal then the surface tension at the hard sphere boundary will be zero so that Eq. (35) reduces to

$$\left[ \frac{\partial W_{in}(n^*)}{\partial \mu} \right]_{V,T} = -(n^*-\bar{n} - 1),$$

which is the standard form of the nucleation theorem.

Before closing this section, it should be emphasized that the result, Eq. (37), has been derived, beginning with an exact expression, based on molecular theory. However, the original expression was modified by limiting it to the case of an ideal gas and, also to the restrictions of a model in which an exclusion sphere was centered on the center-of-mass of the cluster molecules. If the limitation to an ideal gas was relaxed, the dependence of the factor $(e^{-\beta U_\sigma})_0$ on $\mu$ comes into play through the surface tension between the cluster and the vapor.

These considerations throw into some doubt, the general validity of the nucleation theorem in the form of Eq. (37), and suggest that Eq. (37) may only be valid for an ideal vapor. Indirect evidence from laboratory experiments and more direct evidence from computer experiments that support Eq. (37) has been generated by systems that were almost ideal. Thus current experimental evidence cannot really substantiate the general applicability of Eq. (37). The thermodynamic derivation of the nucleation theorem arrives at Eq. (37) but without requiring the system to be ideal, while the statistical mechanical derivation that deals with essentially the same model as the thermodynamic one is clearly limited to ideal systems. Both models fail to deal with the translational degrees of freedom of the cluster, but this should be of little consequence as far as the question of ideality is concerned.

First we must ask if there are any obvious errors in the thermodynamic treatment of Ref. 3. With the exception of a few the caveats that we have already mentioned, and will address in more detail below, the present authors can find no errors. However, it is apparent that the connection between thermodynamics and molecular behavior warrants further discussion and, in particular, we must understand the role of the molecular model implicit within the thermodynamics. The most obvious questions concern the matters of translation (already mentioned) and the omission of additional possible independent thermodynamic variables such as curvature itself (distinct from its appearance in the surface tension and other parameters). The molecular theory developed here, and in earlier work, show that translation should not have a significant effect on the size of the nucleus. However, Gibbs already noted that in order to neglect curvature as a variable, the radius of the surface of tension must be large in comparison to the molecular dimension. One must keep this in mind in the extrapolation of thermodynamic ideas to the molecular level.

The thermodynamic derivation simply starts with a radial density profile to which the Gibbsian formalism of interfacial thermodynamics is applied, but one has to ask to what does this radial density gradient really refer. It has a center, but how is the center defined and where does it lie in the physical system. Several answers are possible, but none of them are reflected in the thermodynamic treatment. For example, there might always be a molecule or the center-of-mass of the system at the center of the spherical density distribution. A third possibility would involve the conceptual use of a microscope with a spherical aperture (radius to be defined) within which the molecules of the distribution must always lie. Then the center of the aperture defines the center of the cluster. In each case it is necessary to invoke a molecular based model that is implicit in the thermodynamic description. The density profile must include some sort of time averaging in order to include all fluctuations of the density distribution, and consequently, the cluster must either live long enough to explore all allowed fluctuations or it must be observed at different times in different positions so that an appropriate ensemble average could be taken. In the last case, however, the dynamic process would involve the addition or removal of molecules from the nucleus in a particular fluctuated distribution, and a problem in the order of averaging would arise.

Having said all this, there remains one problem that, in our opinion, deserves primary consideration. This concerns the uniqueness of the distribution computed via thermodynamics. One of the early merits of the nucleation theorem was the promise that it could say something universal about the nucleus without having to describe its structure. Now in Ref. 3, the properties of the nucleus are determined by finding the appropriate extremum of the Helmholtz free energy, prescribed in Eq. (6) of that article, with respect to the vari-
ables \( n_s \), \( n_f \), and \( V_f \). This step cannot be blamed on the authors of Ref. 3 alone, since it has been adopted by many other authors\(^{8,19} \) both earlier and later. However, implicit in this extremalization, is the idea that the distribution is uniquely determined by these variables.

Away from the extremum, this is certainly not the case since constraints that amount to additional variables must be applied to establish reversible paths for the variation.\(^{20} \) At the extremum itself it is usually assumed that the constraints are irrelevant. This argument is almost certainly valid for the case of stable equilibrium, but some questions can be raised for an unstable equilibrium.

Some sense of the problem can be obtained by returning to molecular considerations and referring to the fundamental uniqueness theorem of density functional theory.\(^{21} \) According to this theorem, when the system achieves global equilibrium, the density functional integral equations provide not only a unique solution for the density distribution, but also unique expressions for various thermodynamic quantities in terms of the density distribution. But the key to the validity of this statement lies in the constraints that are known to be applied to the system. Even the usual “unconstrained” equilibrium is constrained in some sense, e.g., the volume of the system may be fixed or the pressure applied to it may be specified. The theorem could be extended to unstable equilibrium but then constraints must be applied that convert it to a stable equilibrium before a meaningful solution can be obtained.

For example, in the case of a two-phase system with an interface, the usually difficult solution for the density gradient at the interface must be obtained through the imposition of boundary conditions that enforce the density difference between the two phases, and these boundary conditions are in effect constraints that allow one to separate the “unstable” solution from the global one that might even include the unstable distribution as a rare fluctuation. In the seminal work of Talanquer and Oxtoby\(^{11} \) on a dynamical theory of nucleation via density functional theory, the unstable solution is selected with, among other things, the help of an iterative “plateau” that occurs during iterations aimed at the solution of the relevant integral equation. Also it was necessary to constrain the system to a container of fixed volume in order to convert the unstable system to a stable one. Again, the restriction to the plateau imposes a constraint (somewhat hidden) on the system in addition to that associated with the confinement to the container. In effect these constraints generate different models of the system. In each case, the unstable equilibrium is converted to a stable one, at least momentarily.

Ten Wolde\(^7 \) has also computed the radial density structure of a nucleus. However, his nucleus was modeled as a modified Stillinger cluster,\(^{32} \) and the constraints implicit in the Stillinger cluster are therefore incorporated into the structure of the derived nucleus. Ten Wolde also computed the work of formation of his nucleus and was able to show\(^7 \) that the nucleation theorem applied accurately. However, it must be emphasized that the same constraints, i.e., the same model, was used in both the calculation of the density gradient and work of formation. Thus for this model, treated consistently, the nucleation theorem could be valid, i.e., the additional constraints implicit in the cluster definition could be sufficient to assure a unique solution for the density gradient.

The same comment could be made in regard to Fig. 2 due to Oxtoby and Kaschiev in Ref. 3. Here both the density gradient and the work of formation were computed using the constraints of the method employed by Talanquer and Oxtoby,\(^{11} \) and these constraints could assure a unique solution for the unstable state.

In the present work, we have also used a cluster model. The stabilizing constraint in this case is the impermeable exclusion sphere that we attach to the center-of-mass of the molecules inside of the sphere. Here, however, using this molecular approach, we have not yet been able to demonstrate the nucleation theorem for any case beyond that of an ideal vapor. This does not mean that such a demonstration would not be possible. If it proves to be impossible then we are returned to the situation in which the validity of the theorem depends upon the model. This would reduce the value of the nucleation theorem since its most remarkable feature is supposed to be its nondependence on the structure of the cluster constituting the nucleus.

In closing this section it is useful to return to the question of how the specification of \( n_s \), \( n_f \), and \( V_f \), together with restriction to an extremal or saddle point, can guarantee a unique radial density structure for the nucleus. That density gradient is the average over a certain set of fluctuations. Without the introduction of additional information in the form of constraints, how are those fluctuations to be restricted to just those that correspond to the nonuniform density distribution. After all, the stable equilibrium characterized by \( n_s = 0 \) and values of \( n_f \) and \( V_f \) is also a solution unless it is specifically excluded through the introduction of some constraint, i.e., by the choice of some cluster model.

We emphasize that we have certainly not demonstrated that the nucleation theorem is invalid, and that it may yet prove to be valid and extremely remarkable. What we have done is to demonstrate that there is still reason for the exercise of caution before assuming that it is valid without qualification.

Finally, a positive feature of the limited molecular analysis is that it validates the nucleation theorem, under the prescribed limitations, down to the molecular level whereas the thermodynamic analysis can only be approximate at the molecular level.

### III. EFFECTS OF A CARRIER GAS

The experimental determination of nucleation rates is usually carried out in the presence of an inert carrier gas in order to maintain isothermal conditions. Without such a gas the heat of condensation would raise the temperature of the nucleating clusters. It is then assumed that the carrier gas plays no additional role, however, there have been a number of experiments which suggest that the presence of the gas may have a direct effect on the rate,\(^{23,24} \) and others that suggest that there is no effect.\(^{2,25,26} \) Using the formalism presented in the preceding section, we can begin to examine the effects that the carrier gas may have on the nucleation
rate, in a very preliminary way. Since this requires simply repeating many of the steps already described in Eqs. (13) through (37) with an augmented partition function which includes interactions between the $n$ molecules in both the cluster and the $N_g$ molecules in the carrier gas, as well as interactions between the vapor and carrier gas, we will only outline both the points where we have made assumptions and the important results.

Again we start with the partition function for a supersaturated vapor of $N$ molecules in a volume $V$, at temperature $T$, constrained to contain at least one $n$-cluster, but now there are an additional $N_g$ carrier gas molecules, so

$$Q(N, N_g, V, T) = \frac{1}{N!} \frac{1}{N_g!} A^{3N_g} \Lambda_g^{3N_g},$$

(39)

and

$$\Lambda_g = h/(2 \pi m_g k T)^{1/2},$$

(40)

is the thermal de Broglie wavelength of the carrier gas. $U_{\sigma 2}$ is the interaction potential between the carrier gas and $U_{\sigma 3}$ is the interaction potential between the vapor molecules and the carrier gas. Introducing the appropriate decoupled probability density for the carrier gas

$$\Phi_0^{N_g} = \frac{e^{-\beta U_{N_g}}}{\int_V \cdots \int_V e^{-\beta U_N} d\mathbf{r}_n' \cdots d\mathbf{r}_{N+1}'}.$$

(41)

Equation (38) can be expressed

$$Q(N, N_g, V, T) = Q(N-n, V, T) Q(N_g, V, T) Q_n(V, T) \times \langle e^{-\beta (U_0 + U_{\sigma 2} + U_{\sigma 3})} \rangle_0,$$

(42)

where

$$\langle e^{-\beta (U_0 + U_{\sigma 2} + U_{\sigma 3})} \rangle_0 = \frac{1}{N!} \int_V \cdots \int_V e^{-\beta U_0} d\mathbf{r}_n' \cdots d\mathbf{r}_{N+1}'.$$

(43)

Now if we proceed as before, following Eqs. (25) through (35), and making use of the fact that, since the vapor and carrier gas do not interact, the partition function for the entire system is $Q(N, N_g, V, T) = Q(N, V, T) Q(N_g, V, T)$ and there is no surface tension term, also assuming that the cluster simply excludes the vapor and the carrier gas from the volume $v$, we obtain the expression

$$W_{\text{int}}(n) = -n \mu + F_\mu(v, T) + F_\mu(c, T) + P v + P_g v,$$

(47)

for the work of formation of the cluster, where $P_g$ is the pressure of the carrier gas. The nucleation theorem itself remains unaffected by the presence of the carrier gas because we have considered the vapor–carrier gas mixture to be ideal. If we set $n = n^*$ and take the derivative of Eq. (47) with respect to $P_g$ we find

$$\frac{\partial (kT \ln J)}{\partial P_g} \equiv - \frac{\partial W(n^*)}{\partial P_g} = -v,$$

(48)

where $J$ is the nucleation rate. This suggests that the rate of nucleation should decrease with increasing carrier gas pres-
measure. This simply reflects the additional volume work required to create the cavity in which the cluster forms due to the presence of additional gas. It does have to be made clear that the case we have considered is the very simple one in which the nucleus only interacts with the vapor and carrier gas through an excluded volume and that the vapor–carrier gas mixture is ideal. Oxtoby and Laaksonen,\textsuperscript{27} followed by Kashchiev,\textsuperscript{28} use thermodynamics to investigate the effect of the carrier gas by considering the effect the gas has on the chemical potential of the nucleating vapor. In doing so they obtain the result that, while exclusion volume effect is present, the overall effect very much depends on mixture virial coefficients, so that the rate may in fact increase with increasing pressure. They also concluded that, in most cases, the effect on the rate would be small. The effect of the carrier gas through an excluded volume and that that the vapor–surrounding vapor is ideal. These simple assumptions all-

The nucleation theorem relates the excess number of particles in the critical nucleus to the derivative of the work of the formation of the cluster with respect to the chemical potential. The comparison between experiments and theory, and more recently, studies by computer simulation appear to repeatedly confirm this seemingly remarkable result. However in almost all cases these have been performed under conditions where the surrounding medium was approximately an ideal gas. In this work we have shown in Sec. I that the theorem is simply equivalent to the well-known law of mass action if the system is ideal. We were able to rigorously derive the nucleation theorem in its usual form using statistical mechanics, but only under the assumptions that a cluster only interacts with the surrounding vapor through an excluded volume, centered on its center-of-mass, and that the surrounding vapor is ideal. These simple assumptions allowed us to explicitly evaluate the interaction factor $\langle e^{-\beta U_{\text{int}}} \rangle_0$ which gives rise directly to the term involving $n$, rendering the number of molecules in the nucleus, obtained via the nucleation theorem, an excess quantity. This raises the question as to whether Eq. (1) will remain valid in the nonideal case because the factor $\langle e^{-\beta U_{\text{int}}} \rangle_0$ can depend on the chemical potential in some way that leads to an additional excess surface contribution. In many cases the magnitude of the interaction terms may well be smaller than present experimental error, making them difficult to detect, although computer simulation techniques, and experiments on systems known to be highly nonideal may well produce measurable effects.

In closing it is essential to note that while we have dealt exclusively with the reversible work of formation of the nucleus, either in the form of $W(n^*)$ or $W_{\text{int}}(n^*)$, in deriving the nucleation theorem, it is the nucleation rate $J$ that is measured as a function of the chemical potential of the vapor in order to obtain $n^*$. Since $J$ is assumed to be proportional to the number of critical nuclei the nucleation theorem becomes

$$\frac{\partial (kT \ln J)}{\partial \mu} = - \frac{\partial W(n^*)}{\partial \mu} + \frac{\partial (kT \ln K)}{\partial \mu},$$

where $K$ is preexponential factor. In the case of CNT the growth and decay of a cluster is assumed to occur via the addition or loss of a single molecule which leads to $K$ being proportional to $N_1$. However, if we use $W(n^*)$, i.e., use of Eq. (4) for $N_{n^*}$, then it has been suggested\textsuperscript{11,13} that an additional factor in the preexponential is required to make $K \propto N_1^2$ so that the law of mass action is still satisfied. On the other hand use of the intensive free energy in Eq. (7) does not require the additional factor.

**ACKNOWLEDGMENTS**

R.K.B. and H.R. would like to thank the National Science Foundation for support under NSF Grant No. CHE 96-18163. R.M. acknowledges support from NASA as part of its interdisciplinary research program on tropospheric aerosols.