PHYSICAL REVIEW A VOLUME 46, NUMBER 8 15 OCTOBER 1992

Thermodynamics, statistical thermodynamics, and computer simulation of crystals with vacancies and interstitials

William C. Swope
IBM Palo Alto Scientific Center, Palo Alto, California 94304

Hans C. Andersen
Department of Chemistry, Stanford University, Stanford, California 94305
(Received 20 May 1992)

Molecular-dynamics and Monte Carlo simulations of fluids and of crystals are typically performed using periodic boundary conditions and small numbers of particles, i.e., small compared with Avogadro's number. For fluids, if the dimensions of the simulated system are large enough compared with the correlation length of the material, the simulation can provide intensive thermodynamic properties that are very close to those for the thermodynamic limit of large system size. For crystals, however, the periodic boundary conditions and the crystal structure impose a constraint that makes it impossible, in practice, for the system to develop an equilibrium concentration of vacancies. The effect of this constraint must be taken into account when inferring thermodynamic properties from the results of computer simulations of crystals. We develop the thermodynamic and statistical thermodynamic theory of a crystal subject to such a constraint. The results include a set of stability conditions that must be satisfied and a criterion for equilibrium. We also describe a method for calculating the equilibrium concentration of vacancies in a crystal from computer simulations.

PACS number(s): 05.20.Gg, 05.70.−a, 61.70.Bv

I. INTRODUCTION

Point defects in crystalline solids, such as vacancies and interstitials, have long been recognized as having an important influence on the mechanical, elastic, and transport processes of materials. When the material is in thermodynamic equilibrium, the concentration of such defects can in principle be calculated using the methods of equilibrium statistical thermodynamics. For example, the elementary theory of the equilibrium concentration of Schottky and Frenkel defects in a crystal is discussed in standard textbooks [1,2]. There is ample reason to believe that a crystal in thermodynamic equilibrium contains defects such as vacancies and interstitials and that the concentrations of vacancies and interstitials are not necessarily equal. A modern reexamination of the theory of point defects in crystals is timely because computer simulation studies of materials have become such an important method for understanding the properties of materials and because there are many hopeful attempts to develop computer simulations into a predictive tool for real materials.

It would be especially worthwhile to be able to use computer simulations to study the melting of solids and solid-liquid phase equilibria. More generally, it would be worthwhile to be able to predict the phase diagram of a material, including crystalline phases, by computer simulations. One of the first triumphs of the molecular-dynamics method was the observation of fluid-solid phase transitions in the hard-disk and hard-sphere systems [3]. There have been many subsequent studies (too numerous to list in detail) of phase transitions involving solids. Many of them were inspired by the method of Parrinello and Rahman [4] for studying solid-solid phase transitions between phases of different crystal symmetry. The calculation of equilibrium phase diagrams requires the calculation of free energies of crystals. The method of Frenkel and Ladd [5] provides an approach to such calculations that is widely applicable. The Gibbs ensemble method of Panagiotopoulos [6] has been very successful in studying fluid-fluid equilibrium, and it might be hoped that the same or a similar method might be applied to fluid-solid or solid-solid equilibrium.

The main points of this paper are the following.

(1) The use of periodic boundary conditions in the computer simulation of a solid will impose a constraint on the simulated system. The constraint referred to here is that *in a typical computer simulation of a stable solid, the number of lattice sites is a conserved quantity,* i.e., it remains constant during the simulation and can neither fluctuate nor come to its true equilibrium value. As a result, the crystal cannot achieve its true equilibrium concentration of vacancies.

(2) If by some mechanism the system can rearrange and change its number of lattice sites, it will in general be at the cost of distortion of the shape of the unit cell or a restriction on the number of lattice sites that can be achieved or a bias favoring certain numbers over others. This distortion, restriction, or bias will influence the calculations in a spurious way, and the result will be that once again the crystal will not be able to achieve its true equilibrium concentration of vacancies.

(3) In order for a simulation method to provide a valid calculation of the equilibrium properties of a solid, it must avoid the influences of the distortion, or bias; must take the constraint into account; and must provide a way

©1992 The American Physical Society

for calculating the true equilibrium concentration of vacancies.

In Sec. II, we discuss the basis for these three assertions. In Sec. III we discuss the thermodynamics and stability of lattices that are subject to a constraint on the number of lattice sites and the criteria for equilibrium of lattices that are not subject to this constraint. In Sec. IV we discuss the corresponding statistical thermodynamics. In Sec. V we discuss a strategy for the use of computer simulations to investigate the equilibrium properties of crystals. Finally, a discussion of some of the implications is in Sec. VI.

II. CONSTRAINTS AND BIASES IN SIMULATIONS OF CRYSTALS

A typical simulation of a crystal starts with the atoms or molecules in a crystalline configuration, with the crystal axes aligned with the sides of the periodic cell so that the simulated system contains an integral number of unit cells of the crystal and hence a definite number of lattice sites. (We are not referring to simulations in which the crystal is formed by freezing of a liquid. Such solids are often highly defective and not suitable for phase equilibrium calculations.) As the simulation continues, the atoms will vibrate and change their positions, but the number of lattice sites remains the same, unless the crystal becomes unstable and melts or otherwise transforms itself. This assertion is basically an empirical statement that can be attested to by anyone who has simulated stable crystals and analyzed the structures that result.

A. Canonical and microcanonical simulations

Consider the case of simulations in which the number of particles and the size and shape of the simulation cell are held constant during a run. Suppose such a method is used to simulate a crystal with equal numbers of particles and lattice sites. As the simulation proceeds, a particle may leave its initial lattice site to form an interstitial, but it will leave behind a vacancy at its starting position. The interstitial and vacancy might both move and additional vacancy-interstitial pairs might form and recombine. However, the number of vacancies minus the number of interstitials, which we shall call the "net" number of vacancies, is fixed, in this case at the value of zero. (It is clear that in general the net number of vacancies is equal to the number of lattice sites minus the number of particles.) However, as noted in the first paragraph of this paper, there is no reason to believe that the equilibrium concentration of vacancies and interstitials is the same (i.e., that the equilibrium concentration of net vacancies is zero). Thus, a crystal simulated in this way can never be relied on to achieve true thermodynamic equilibrium.

The same conclusion is reached for simulations for which the initial number of particles is not equal to the number of lattice sites. Although such simulations can be performed with a variety of concentrations of net vacancies, the simulations themselves do not allow us to determine what the true equilibrium concentration of net vacancies actually is.

For example, let us consider a simple situation in which the following holds: (1) The equilibrium crystal structure has two atoms per unit cell, (2) the equilibrium shape of the unit cell is a cube, and (3) at thermal equilibrium there is a very small positive concentration of net vacancies, for example, one vacancy per 2000 lattice sites. Suppose the simulation cell is also a cube. If one knows the value of the equilibrium concentration of vacancies, then it is clear that one convenient way to simulate a bulk crystal would be to construct a system of 2000 particles in a cube that is ten unit cells in each dimension, remove one particle, and then use this as the starting point for the simulation. Even better would be to construct a system that is eight times larger, namely, 16 000 particles in a cube that is 20 unit cells in each dimension, then remove eight particles, and use the resulting configuration as the starting point of the simulation.

On the other hand, suppose we know the shape of the unit cell and we suspect that the equilibrium concentration of vacancies is small, but we do not know the value of the equilibrium concentration. If we simulate a system with 16 000 particles using the configuration with 8000 unit cells as a starting point, it is very likely that the system will merely vibrate around the starting configuration. In principle, with 16 000 particles there should be 16 008 lattice sites, in order to achieve the desired equilibrium concentration of vacancies, but this would require 8004 unit cells. It is hard to imagine that the system could rearrange to form exactly 8004 unit cells in the original cubic simulation cell. If the specific numbers were somewhat different and if by some rearrangement a structure with the required number of unit cells could be achieved within the strictly cubic simulation cell, then it is exceedingly likely that the shape of the unit cell would be distorted away from its equilibrium cubic shape. The properties calculated during the simulation would not then be characteristic of the undistorted crystal in the thermodynamic limit.

Similarly, if we simulate a system with a starting configuration of 8000 unit cells and 15 996 particles (i.e., four net vacancies), then it is difficult to imagine that the system could rearrange to form a structure with 16 004 lattice sites of 8002 unit cells, and even if it could do so, the unit cell would be distorted.

A real macroscopic crystal, with a fixed number of atoms, can change the number of lattice sites in the sample by, for example, having some of the atoms in the bulk leave their lattice sites, migrate to the surface, and form another layer of crystalline atoms at the surface. In this way a crystal can change its number of lattice sites without changing its number of particles. This provides a mechanism for equilibration of the bulk concentration of vacancies. Moreover, in a macroscopic system this need not lead to a distortion of the shape of the unit cell because adjustments can always be made at interfaces and boundaries and these adjustments do not affect the values of extensive thermodynamic properties.

The simplest analogous process that can take place in a periodic-boundary-condition computer simulation is for the system to contract anisotropically to create a planar void parallel to a lattice plane through the entire system, and then to have some atoms migrate to this void and form an additional plane of unit cells, leaving some vacancies in the rest of the crystal. This will create a crystal that is consistent with the periodic boundary conditions and with a different number of lattice sites than was in the initial crystal. However, this can only be accomplished at the cost of distorting the average shape of the unit cell. Moreover, only certain particular numbers of additional unit cells can be created in this way, and these numbers may not correspond to what is needed to create a true equilibrium concentration of vacancies.

There are more complicated processes that can take place under periodic boundary conditions and that can change the number of lattice sites. For example, certain types of tilting of the unit cells relative to the orientation of the simulation cell can increase or decrease the number of unit cells while still generating structures that are consistent with periodic boundary conditions. We have seen the analogous process spontaneously take place in simulations of two-dimensional crystals. But once again, if such a process were to take place, it would distort the shape of the unit cell and it might not be able to create the appropriate equilibrium number of vacancies.

The net result of these considerations is the following:

(1) If, in a simulation of a crystal with a constant number of particles in a simulation cell of constant size and shape, the system can change its number of lattice sites, there is no guarantee that the true equilibrium number of lattice sites will be achieved, even in a long simulation run. This is so because of a combination of one or more of the following.

(a) There might be no geometric way of packing the correct equilibrium number of unit cells into the simulation cell; (b) the resulting distortion of the unit cell might affect the relative probability of finding the various possible numbers of lattice sites; and (c) the rate of such structural rearrangements might be exceedingly slow, leading to averages with large statistical uncertainties.

(2) On the other hand, if the system cannot change its number of lattice sites, then it is also clear that true equilibrium cannot be achieved.

B. Grand-canonical simulations

The problems generated by the constraint are not eliminated by going to simulation methods that change the number of particles, such as the grand-canonical Monte Carlo method. A symptom of the problem can be recognized in the following considerations: In a grand-canonical Monte Carlo simulation, the volume of the system is fixed. Let us suppose that the number of lattice sites in fact does not change during the simulations. For each fixed volume, we can perform the simulations for a variety of values of μ. The resulting equilibrium systems will have a variety of different pressures. Thus, simulations at one volume can generate the pressure as a function of chemical potential over some range of chemical potentials. Now suppose the volume is changed slightly. Again let us suppose the number of lattice sites will not change. Simulations at this different volume will also generate a plot of p versus μ, but this plot will not necessarily be the same as that generated at the first volume. (To see this, note that the average volume of the unit cell will in general not be the same for the two volumes simulated. Thus, the two simulations sample quite different states.) But for a one-component system in thermodynamic equilibrium, p is a unique function of μ. The resolution of this paradox is that the system in the simulations do not have the true equilibrium concentration of vacancies consistent with the specified chemical potential, even though they can change their particle numbers. In grand-canonical simulations, changes in particle numbers and changes in numbers of net vacancies are always coupled together in such a way that they cannot both separately come to equilibrium. In fact, the number of particles plus the number of net vacancies is constrained to be a constant, equal to the number of lattice sites.

The above discussion of grand-canonical simulations was based on the assumption that the number of lattice sites remains constant during a simulation and does not change even if the volume is changed slightly. If, on the other hand, the system can change its number of lattice sites, it will be at the cost of introducing distortions of the unit cell that can influence and bias the results, as in the previous discussion of simulations that conserve the number of particles.

C. Other ensembles

Similar considerations can be used to show that constant pressure simulation methods [7] also do not resolve the difficulty, nor can true liquid-solid equilibrium be achieved in the Gibbs-ensemble method [6], as it is currently implemented for other types of phase equilibrium.

Finally, consider the Parrinello-Rahman [4] simulation method that allows the dimensions and the shape of the simulation cell to change. This is the one method that can adapt itself to relieve any distortion that could arise if the number of lattice sites were to change. However, it is not clear that the proper number of lattice sites could always be achieved in such a system subject to periodic boundary conditions.

D. Remarks

Because of the spurious effects, such as unit cell distortion, that can arise if the number of lattice sites changes, it is in fact preferable to perform simulations under conditions in which the system stays in one crystalline configuration, with the total number of lattice sites constant and the unit cell undistorted in shape. Under these conditions, the system is clearly not able to equilibrate completely and is in effect subject to a constraint on its number of lattice sites or net vacancies.

The problem generated by a constraint on the number of lattice sites is in some ways analogous to the problem that would arise if a simulation were being performed of a chemically reactive fluid on time scales so short that the chemical reaction could not take place. The concentrations of each of the reactant and product species could then be varied independently and the various systems could be simulated, whereas only certain compositions are consistent with chemical equilibrium. The correct way of determining the equilibrium concentrations is well known, however, in principle. We should perform simulations for a variety of concentrations, obtain the chemical potentials of each of the species as a function of composition, and then use the law of mass action to determine the states or compositions at which the system is in chemical equilibrium. In other words, the thermodynamics of the system subject to the constraint of no chemical reactions can be studied and then thermodynamic reasoning used to see what states would be in equilibrium if the constraint were released.

The same type of strategy can be used to attack the problem of interest here, and this is the subject of the rest of this paper. We investigate the thermodynamics and statistical thermodynamics of particles in lattices constrained to have a fixed number of lattice sites. Then we use thermodynamic reasoning to derive a law (analogous to the law of mass action) that determines the true equilibrium number of lattice sites as a function of the number of particles, the density, and the temperature. We also derive a set of stability conditions that must be satisfied if the crystal is to be stable against the development of inhomogeneities in the distribution of particles and lattice sites.

III. THERMODYNAMICS OF LATTICES

A. Unconstrained lattices

We first consider a one-component system and construct its thermodynamic description using the approach of Callen [8]. For simplicity we consider a one-component monatomic material, but the generalization to two-component and molecular materials is straightforward. We assume that the system is capable of achieving total thermodynamic equilibrium, including the adjustment of the concentration of net vacancies if it is a solid and the possibility of making a phase transition to another phase if that is required for achieving equilibrium. The state of such a system in equilibrium is determined by three extensive properties, the number of atoms N, the volume V, and the internal energy E. In particular, the entropy S is an extensive function of these extensive state variables. In the usual way, we are led to define a Helmholtz free energy $A(N,V,T)$, defined as $E-TS$, where T is the thermodynamic temperature. Its differential is

$$dA=-S\,dT-p\,dV+\mu\,dN ,$$

where the pressure P and chemical potential μ are

$$-p=\left[\frac{\partial A}{\partial V}\right]_{T,N} ,\qquad \mu=\left[\frac{\partial A}{\partial N}\right]_{V,T} . \tag{1}$$

It is convenient to define $v=V/N$, the volume per atom. Since A is an extensive property, it can be written in the form

$$A(N,V,T)=Na(v,T) ,$$

where $a(v,T)$ is the Helmholtz free energy per particle. It follows that

$$-p=\left[\frac{\partial a}{\partial v}\right]_T ,$$

$$\mu=a(v,T)-v\left[\frac{\partial a}{\partial v}\right]_T .$$

From these results, it is clear that p and μ are functions of the intensive variables v and T only. Moreover, it is straightforward to use these two equations to show that, at constant temperature,

$$d\mu=v\,dp . \tag{2}$$

B. Constrained lattices

Next we consider a crystal that is constrained to have a certain number of lattice sites M and constrained to remain a crystal. Then the thermodynamic state is specified by four extensive properties N, V, E, and M, and there is an extensive entropy $S_c(N,V,E,M)$ that is a function of these state variables. (In the discussion of a constrained system, we are led to define quantities, like entropy, that are clearly related to, but not necessarily equal to, those for the unconstrained system when the latter happens to adopt a crystalline state. Such variables for the constrained system will be denoted by a subscript c.) Similarly, we can define a Helmholtz free energy

$$A_c(N,V,T,M)=E_c(N,V,T,M)-TS_c(N,V,T,M) .$$

Its differential is

$$dA_c=-S_c dT-p_c dV+\mu_c dN+\nu_c dM , \tag{3}$$

where

$$-p_c=\left[\frac{\partial A_c}{\partial V}\right]_{T,N,M} ,\qquad \mu_c=\left[\frac{\partial A_c}{\partial N}\right]_{V,T,M} ,\qquad \nu_c=\left[\frac{\partial A_c}{\partial M}\right]_{V,T,N} . \tag{4}$$

The latter quantity ν_c plays a central role in the current theory and might be called the chemical potential of lattice sites or the chemical potential of net vacancies.

For the constrained system, it is convenient to define the following intensive variables: $\tilde{v}=V/M$ is the volume per lattice site and $\tilde{n}=N/M$ is the number of atoms per lattice site. Then, clearly,

$$v=V/N=\tilde{v}/\tilde{n} . \tag{5}$$

The extensive nature of the A_c guarantees that it can be expressed as

$$A_c(N,V,T,M)=M\tilde{a}_c(\tilde{n},\tilde{v},T) , \tag{6}$$

where $\tilde{a}_c$ is the Helmholtz free energy per lattice site. It follows that

$$\begin{aligned} -p_c &= \left[\frac{\partial \tilde{a}_c}{\partial \tilde{v}}\right]_{\tilde{N},T}, \\ \mu_c &= \left[\frac{\partial \tilde{a}_c}{\partial \tilde{n}}\right]_{\tilde{v},T}, \end{aligned} \tag{7}$$

$$\begin{aligned} \nu_c &= \tilde{a}_c - \tilde{n}\left[\frac{\partial \tilde{a}_c}{\partial \tilde{n}}\right]_{\tilde{v},T} - \tilde{v}\left[\frac{\partial \tilde{a}_c}{\partial \tilde{v}}\right]_{\tilde{n},T} \\ &= \tilde{a}_c + p_c\tilde{v} - \mu_c\tilde{n} \; . \end{aligned} \tag{8}$$

It is clear that $-p_c$, μ_c, and ν_c are functions of the intensive variables $\tilde{n}$, $\tilde{v}$, and T. A large number of Maxwell relations can be derived, relating the various derivatives of $-p_c$, μ_c, and ν_c. Moreover, from these equations it is easily shown that, at constant temperature,

$$d\nu_c = \tilde{v}\, dp_c - \tilde{n}\, d\mu_c \; . \tag{9}$$

In order to be thermodynamically stable, the constrained system must satisfy certain stability conditions, which we now derive. In the constrained system, N, V, and M are conserved quantities. There exists the possibility that the system might become inhomogeneous by, for example, forming two phases with different concentrations of particles and lattice sites in such a way as to preserve the total values of N, V, and M. Consider the following situation: A homogeneous system, characterized by the values N, V, and M, is imagined to be two smaller adjacent subsystems, each characterized by the values $N/2$, $V/2$, and $M/2$. (Throughout the discussion we keep the temperature T fixed in the system and both subsystems.) The sum of the free energies of the two subsystems is equal to the total free energy of the original system, because the free energy is extensive. Next we imagine transferring δN atoms, a volume δV, and δM lattice sites from one of the subsystems to the other. The free-energy change associated with this change in state is

$$\begin{aligned} \Delta A_c = {} & A_c(N/2+\delta N, V/2+\delta V, M/2+\delta M, T) \\ & + A_c(N/2-\delta N, V/2-\delta V, M/2-\delta M, T) \\ & - 2A_c(N/2, V/2, M/2, T) \; . \end{aligned}$$

This is a change that is not prevented from happening in the constrained system. By the second law of thermodynamics, a necessary condition for the stability of the system against development of such an inhomogeneity is that

$$\Delta A_c \geq 0$$

for all possible choices of δN, δV, and δM. In particular, we consider small values of these changes and perform a Taylor-series expansion in powers of δN, δV, and δM. ΔA_c is then expressed in terms of the second partial derivatives of A_c with regard to N, V, and M and the values of δN, δV, and δM. Since ΔA_c must be greater than or equal to zero for any choice of δN, δV, and δM, it follows that the matrix of second derivatives, namely,

$$\begin{bmatrix} \dfrac{\partial^2 A_c}{\partial N^2} & \dfrac{\partial^2 A_c}{\partial N \partial V} & \dfrac{\partial^2 A_c}{\partial N \partial M} \\ \dfrac{\partial^2 A_c}{\partial V \partial N} & \dfrac{\partial^2 A_c}{\partial V^2} & \dfrac{\partial^2 A_c}{\partial V \partial M} \\ \dfrac{\partial^2 A_c}{\partial M \partial N} & \dfrac{\partial^2 A_c}{\partial M \partial V} & \dfrac{\partial^2 A_c}{\partial M^2} \end{bmatrix}$$

must be positive semidefinite. Here it is understood that the independent variables in the partial differentiation are N, V, M, and T and that the derivatives are to be evaluated for the state in which the arguments are equal to $N/2$, $V/2$, $M/2$, and T. It follows that, in order for the homogeneous system to be stable against development of inhomogeneities, (1) the diagonal elements of this array must be nonnegative; (2) the determinant of the 2×2 array obtained by eliminating the ith row and the ith column from this array must be non-negative for each $i=1,2,3$; and (3) the determinant of this 3×3 array must be non-negative.

We note that this 3×3 matrix is symmetric, because the off-diagonal elements are mixed second partial derivatives. It can easily be verified, by evaluating the matrix elements using Eq. (6), that the third column is a linear combination of the first two columns and that the third row is a linear combination of the first two rows. Hence, the 3×3 determinant is zero, and so the third condition is never violated. All the elements of this matrix can then be expressed in terms of the 11, 12, and 22 elements. It can then easily be verified that the various 2×2 determinants are simply proportional to one another, with proportionality constants that are positive. Thus the conditions generated by the various 2×2 determinants are all equivalent. Finally, it is straightforward to show that the entire set of conditions is equivalent to just two of them, namely, that the central matrix element be non-negative and that the determinant obtained by deleting the third row and the third column be non-negative. The necessary conditions for the stability of the system can then be expressed as

$$-\left[\frac{\partial p_c}{\partial \tilde{v}}\right] \geq 0 \; , \tag{10}$$

$$-\left[\frac{\partial p_c}{\partial \tilde{v}}\right]\left[\frac{\partial \mu_c}{\partial \tilde{n}}\right] - \left[\frac{\partial p_c}{\partial \tilde{n}}\right]^2 \geq 0 \; . \tag{11}$$

Here it is to be understood that the independent variables for the partial differention are $\tilde{n}$, $\tilde{v}$, T. States that satisfy one of these conditions as an equality and the other as an inequality (or that satisfy both as equalities) are called marginally stable. The marginally stable states represent the boundary separating unstable states from stable states in the space of thermodynamic states. It should be noted that, at the boundary, the second of these conditions (or perhaps both of them but not only the first) must be satisfied as an equality.

To summarize, these non-negativity conditions are *necessary* conditions for the stability of the constrained system (i.e., the system constrained to remain a crystal and constrained to have a fixed total number of lattice sites) against the development of macroscopic inhomogeneities. By analogy with the corresponding situation in gas-liquid phase transitions, we assume that the violation of these conditions would be a signal of the existence of a phase transition in the system. In the gas-liquid system, marginal stability exists at states on the spinodal curve. In this case, however, before the spinodal is reached, the system undergoes a first-order transition and the instability is not in the range of equilibrium thermodynamic states. The marginal stability condition is also satisfied at the critical point. Thus, the present stability analysis may be of relevance for critical phenomena in solids. In particular, we propose that the Kosterlitz-Thouless critical point [9] in two-dimensional crystals represents a state at which the system is marginally stable in the sense discussed here.

C. Release of the constraint

Suppose now the constrained system is allowed to change its M and thereby achieve total equilibrium. Moreover, assume that the total equilibrium state is crystalline, rather than some other phase. Callen's formulation of the second law of thermodynamics asserts that, at fixed N, V, and E, the number of lattice sites M will adopt a value that maximizes the entropy $S_c(N,V,E,M)$. Moreover, it follows that, at fixed N, V, and T, the number of lattice sites M will adopt a value that minimizes the Helmholtz free energy $A_c(N,V,T,M)$. This latter criterion is the most useful one for the present discussion. Thus, there is an equilibrium value of M, to be denoted $M_{\text{eq}}(N,V,T)$, that minimizes A_c subject to the indicated constraints. It follows that

$$A(N,V,T)=A_c(N,V,T,M_{\text{eq}}(N,V,T)) \; . \tag{12}$$

A necessary condition for equilibrium is

$$\nu_c=0 \; , \tag{13}$$

i.e., that the chemical potential of net vacancies be zero. More precisely,

$$\nu_c(N,V,T,M_{\text{eq}}(N,V,T))=0 \; , \tag{14}$$

or, expressing ν_c as a function of the intensive state variables,

$$\nu_c(\tilde{n},\tilde{v},T)=0 \; . \tag{15}$$

It follows from these results that

$$-p(N,V,T)=-p_c(N,V,T,M_{\text{eq}}(N,V,T)) \; ,$$

$$\mu(N,V,T)=\mu_c(N,V,T,M_{\text{eq}}(N,V,T)) \; .$$

Equations (13)–(15) provide the thermodynamic solution to the problem of the relationship between the constrained and unconstrained systems. (They play the same role in the present problem as does the law of mass action, expressed in terms of chemical potentials, in the theory of chemical equilibrium. See the discussion in Sec. II D.) Let us consider a fixed temperature and discuss the states of the system in terms of intensive variables. The state of a constrained system is specified by giving the values of two intensive quantities $\tilde{v}$ and $\tilde{n}$. The state of an unconstrained system is specified by giving the value of v. Equation (15) determines the locus of states in the two-dimensional state space of the constrained system that correspond to equilibrium states of the unconstrained system. Along this locus, the ratio $\tilde{v}/\tilde{n}$ gives the value of v for the corresponding unconstrained system. Moreover, the values of $-p_c$ and μ_c along the locus give the values of $-p$ and μ for the corresponding unconstrained state. It is easily shown from the stability conditions that when proceeding along the locus of $\nu_c=0$ states, the system is crossing lines of constant p_c and constant μ_c. Equation (9) leads to the conclusion that along this locus

$$\tilde{v}\,dp_c-\tilde{n}\,d\mu_c=0 \; ,$$

which is consistent with Eq. (2) for the unconstrained system (see Eq. (5)].

The results of the stability analysis for constrained lattices are relevant for unconstrained lattices as well. Consider an unconstrained system that is in stable thermodynamic equilibrium in a homogeneous crystalline state. Its state corresponds to a point that is on the locus of $\nu_c=0$ states for the constrained system. At that point, the constrained system must satisfy the stability conditions given in Eqs. (10) and (11). (This is easily proven by contradiction. Suppose that point violated the stability conditions. Then a constrained system at that point could lower its free energy by developing an inhomogeneity. It follows that the unconstrained system could lower its free energy by developing the same inhomogeneity. Therefore, the unconstrained system is unstable, contrary to the initial assumption.)

The results of the thermodynamic analysis can be summarized in the following way: Suppose for a temperature T we can determine $\tilde{a}_c(\tilde{n},\tilde{v},T)$ for a range of values of $\tilde{v}$ and $\tilde{n}$. This range then represents the range of states accessible to a homogeneous crystal that is constrained not to change its number of lattice sites. Next we determine the region in which $\tilde{a}_c$ not only exists but satisfies the stability conditions, Eqs. (10) and (11). If there are any points at which the second of these conditions is satisfied as an equality, such points are at the boundary of this region and represent marginally stable states. This region represents the range of states in which the constrained system may be stable against the development of inhomogeneities. Within this region we determine the one-dimensional locus of states for which $\nu_c(\tilde{n},\tilde{v},T)=0$. These represent the states that in addition have the equilibrium concentration of net vacancies. Finally, along this locus we must examine the chemical potential as a function of pressure and compare this with the values for other possible phases (liquids, crystals with different crystal structures, hexatics, etc.). Those points on the locus for which there is no other phase at the same pressure with a lower chemical potential are the states of the crystal that are in total equilibrium. If this locus of total equilibrium states extends all the way to a point of marginal stability, this point is presumably a critical point.

IV. STATISTICAL THERMODYNAMICS OF LATTICES

The statistical thermodynamic description of the unconstrained lattice system is the well-known one, and the corresponding description of constrained lattices is straightforward to construct. We restrict our attention to the case of classical statistical mechanics with Boltzmann statistics.

The partition function corresponding to $A(N,V,T)$ is, of course,

$$Q(N,V,T)=(N!)^{-1}h^{-3N}\times\int_V d\mathbf{r}^N\int d\mathbf{p}^N\exp(-H_N(\mathbf{r}^N,\mathbf{p}^N)/kT)\ , \tag{16}$$

where $H_N(\mathbf{r}^N,\mathbf{p}^N)$ is the Hamiltonian for N atoms. We then have

$$A(N,V,T)=-kT\ln Q(N,V,T)\ ,$$

as well as the usual relationships for $-p$ and μ as partial derivatives of $\ln Q$.

To construct the partition function for the constrained system, we imagine that in principle each point in N-particle configuration space can be analyzed to determine the number of lattice sites that the system has in that configuration [10]. For example, this might be done by Voronoi analysis followed by analysis of the spatial arrangement and connectivity of those Voronoi polyhedra that have the shape appropriate for the perfect crystal structure. For values of N, V, and T appropriate for thermodynamic stability of solids, the configurations that primarily contribute to Q in Eq. (16) will have a number of lattice sites proportional to and approximately equal to N. There are a variety of ways that the analysis could be performed, and there is no uniquely best definition, but we assume that a reasonable one has been adopted. (The definition will have to be consistent with some mathematical and physical conditions that we will discuss below.) Then we define a function $m_N(\mathbf{r}^N,V)$ as the number of lattice sites in the N-particle configuration $\mathbf{r}^N$ in volume V. The partition function for the constrained system can be defined as

$$Q_c(N,V,T,M)=(N!)^{-1}h^{-3N}\times\int_V d\mathbf{r}^N\int d\mathbf{p}^N\delta(M,m_N(\mathbf{r}^N,V))\times\exp(-H_N(\mathbf{r}^N,\mathbf{p}^N)/kT)\ , \tag{17}$$

where $\delta(m,n)$ is the Kronecker delta function. In effect, it provides a complicated set of limits on the range of integration in configuration space. The corresponding free energy is

$$A_c(N,V,T,M)=-kT\ln Q_c(N,V,T,M)\ .$$

Comparing with Eq. (3), we conclude that

$$-p_c=-kT\left[\frac{\partial\ln Q_c}{\partial V}\right]_{N,T,M}\ , \tag{18}$$

$$\mu_c=-kT\left[\frac{\partial\ln Q_c}{\partial N}\right]_{V,T,M}\ , \tag{19}$$

$$\nu_c=-kT\left[\frac{\partial\ln Q_c}{\partial M}\right]_{N,V,T}\ . \tag{20}$$

Moreover,

$$\tilde{a}_c(\tilde{n},\tilde{v})=-(kT/M)\ln Q_c(N,V,T,M)\ . \tag{21}$$

It follows that

$$Q(N,V,T)=\sum_M Q_c(N,V,T,M)\ .$$

Under conditions of N, V, and T for which the crystal structure is thermodynamically most stable, we expect that there is one value of M whose term dominates the sum on the right. This is equivalent to the M_{eq} that minimizes A_c at fixed N, V, and T. Using standard maximum term arguments, we conclude that in the thermodynamic limit,

$$\ln Q(N,V,T)=\ln Q_c(N,V,T,M_{\text{eq}}(N,V,T))\ ,$$

which is the statistical analog of Eq. (12).

V. COMPUTER SIMULATION OF LATTICES

The partition function $Q(N,V,T)$ and its associated free-energy density $a(v,T)$ should in principle be calculated in the thermodynamic limit. In the limit of a large system, no matter what the shape of the system and no matter what type of boundary conditions are used, a crystalline material with the desired shape of the unit cell can fill essentially the entire system. It can even create surfaces and domains, should that be necessary, and particles can migrate to the surface, thereby changing the number of lattice sites and equilibrating the bulk density of net vacancies. This can all be accomplished without any distortion of the shape of the unit cell, and the presence of surfaces will not affect the values of the bulk extensive thermodynamic properties. Such undistorted crystalline configurations will dominate in the calculation of the partition function Q in the thermodynamic limit.

For simulated (and hence small) systems under periodic boundary conditions, the shape of the system and the precise number of particles can have a significant effect on the results. In particular, certain combinations of shape and particle numbers may make it impossible for the atoms to arrange in an undistorted perfect crystalline arrangement. These particular combinations will lead to simulations that are of no practical utility for learning about the thermodynamic limit. Care must be exercised in constructing simulations of small systems so that the simulation samples the types of undistorted crystalline configurations that dominate the partition function Q in the thermodynamic limit. Let us assume, for the sake of discussion, that we know (or are willing to guess) what the shape of the unit cell actually is in the thermodynamic limit [10]. Then the most reasonable thing to do is to construct a simulation volume that contains an integral number of unit cells that pack together and fill the volume in a way consistent with periodic boundary conditions. If the particles are originally positioned in a way consistent with this arrangement of unit cells, perhaps with some vacancies and/or interstitials, then a simulation at constant N, V, and T (either canonical Monte Carlo or canonical molecular dynamics) can be performed and the structure analyzed as a function of time. (Grand-canonical Monte Carlo calculations can also be performed.)

It is precisely this type of simulation (the only type that for small systems is able to explore the undistorted crystalline configurations that dominate in the thermodynamic limit) in which it is most likely that the periodic boundary conditions will tend to stabilize and preserve the lattice structure and keep the number of lattice sites constant. If a constant temperature simulation is long enough for the system to explore all of the important regions of configuration space that contribute to $Q_c(N,V,T,M)$, then it is reasonable to assume that averages of dynamical variables over the states generated in the simulation are equal to averages over an ensemble characterized by fixed N, V, T, *and* M, i.e., just the ensemble whose partition function is Q_c and whose probability density is proportional to the integrand in Q_c. Thus the quantities calculated in the simulation are quantities of the type that we have indicated with the subscript c. (If a grand-canonical simulation method is used in such a simulation, the number of lattice sites will still be fixed. The averages calculated will then correspond to an ensemble in which V, T, μ, and M are all specified. This is definitely *not* the usual grand-canonical ensemble. It is a constrained ensemble whose relationship to the grand-canonical ensemble is analogous to that between Q_c and Q.)

In particular, the pressure is normally calculated in simulations by calculating the average of the virial. The proof that the average virial is equal to the pressure is usually based on a mathematical trick in which the configuration integral is expressed in terms of scaled coordinates [11]. Above we noted that the definition of Q_c was based on a function $m_N(\mathbf{r}^N, V)$, which gives the number of lattice sites in the material when the configuration of the system is $\mathbf{r}^N$. If this function is invariant to an overall dilation of the system (as is the case when it is defined in terms of Voronoi analysis of the configuration), then the standard proof that the pressure is the ensemble average of the virial holds also for the constrained ensemble. Hence when the average virial is calculated in the normal way for a simulation in which N, V, and T are held fixed, the result is in fact equivalent to the constrained pressure p_c, whose thermodynamic meaning is given in Eqs. (4) and (7) and whose statistical meaning is given in Eq. (18), rather than to the pressure p, whose meaning is given in Eq. (1).

Thus one of the partial derivatives of $\tilde{a}_c(\tilde{n},\tilde{v})$, that with respect to $\tilde{v}$, can be obtained straightforwardly by computer simulation. The derivative with regard to $\tilde{n}$ is the chemical potential μ_c. When analyzing the $\tilde{n}$ dependence of $\tilde{a}_c$, it is preferable to deal with finite differences rather than derivatives. Applying Eq. (21), we find

$$\begin{aligned}\tilde{a}_c(N/M,V/M)-\tilde{a}_c((N-1)/M,V/M) &= -\frac{kT}{M}\ln\left[\frac{Q_c(N,V,T,M)}{Q_c(N-1,V,T,M)}\right]\\ &= -\frac{kT}{M}\ln\left[\frac{1}{N\Lambda^3}\,\frac{\int d\mathbf{r}^N\delta(M,m_N(\mathbf{r}^N,V))\exp(-U_N(\mathbf{r}^N)/kT)}{\int d\mathbf{r}^{N-1}\delta(M,m_{N-1}(\mathbf{r}^{N-1},V))\exp(-U_{N-1}(\mathbf{r}^{N-1})/kT)}\right],\end{aligned} \tag{22}$$

where $U_N(\mathbf{r}^N)$ is the potential energy of N atoms at positions $\mathbf{r}^N$ and Λ is the usual thermal de Broglie wavelength. The ratio of integrals on the last line can in principle be evaluated if the $m_N(\mathbf{r}^N, V)$ function defined above satisfies a certain simple property, namely, that

$$m_N(\mathbf{r}^N,V)=m_{N-1}(\mathbf{r}^{N-1},V)\ , \tag{23}$$

where $\mathbf{r}^{N-1}$ denotes the N positions $\mathbf{r}^N$ with *any* one of the N positions deleted. The meaning of this is that when a particle is removed from a system of N particles in volume V with M sites, keeping the other $N-1$ positions fixed, the resulting configuration must still have M sites. This is actually a quite reasonable requirement to impose on the function m [12]. When it is satisfied, then the regions of integration in the numerator and denominator, which are restricted by the $\delta(M,m)$ factors, are related in just such a way that various methods such as the Widom method [13] become applicable to the evaluation of the ratio. (For example, if a test particle is inserted into an $N-1$ particle configuration that makes a nonzero contribution to the denominator of Eq. (22), Eq. (23) guarantees that the resulting configuration is contained within the range of integration in the numerator. This makes it possible to express the ratio in Eq. (22) in terms of a certain ensemble average, namely, the exponential of the energy change for insertion of the particle averaged over a constrained ensemble of $N-1$ particles and M lattice sites in volume V at temperature T.) Moreover, the grand-canonical Monte Carlo method can be used to obtain equivalent information in a slightly different way. We have invented a simulation method that is especially suited for evaluation of such ratios [14].

Although both partial derivatives of $\tilde{a}_c$ can thus be obtained from simulations, this is not enough information to evaluate v_c, which is critically important to determine the equilibrium concentration of net vacancies in a crystal. From Eq. (8) it is clear that the absolute value of $\tilde{a}_c$ is also needed. For this, at least two methods are possible.

(1) The method of Frenkel and Ladd [5] can be used to calculate the free energy of a crystal relative to an Einstein solid. This is best done for lattices with no net vacancies.

(2) If the solid is stable down to zero temperature, the temperature-dependent energy can be evaluated by simulations at various temperatures and a thermodynamic integration used to get the free energy of the crystal relative to that of the zero-temperature crystal. The theory of harmonic lattices can then be used to evaluate the absolute free energy of the solid at low temperatures.

To map out the behavior of ν_c and hence determine the locus of equilibrium states for which $\nu_c=0$, it is necessary to have three types of information: (1) the virial pressure for a range of values of $\tilde{n}$ and $\tilde{v}$, (2) the chemical potential μ_c for the same range, and (3) the absolute free energy of the crystal for at least one value of $\tilde{n}$ and $\tilde{v}$ in that range [15]. With this information the value of ν for any state in the range studied can be calculated.

VI. DISCUSSION

When computer simulations are performed on solids in order to infer their bulk thermodynamic properties, the small size of the system studied and the periodic boundary conditions can have a subtle effect that requires a reevaluation of the relationship between the quantities calculated in the simulations and the bulk properties. The averages calculated during such a simulation are averages over a constrained ensemble, one in which the number of lattice sites is a constant. In this paper, we have presented a thermodynamic and statistical thermodynamic analysis of this ensemble and have described some of the consequences of the analysis for computer simulations of solids.

The problem discussed in this paper arose for us in simulation studies of phase transitions in a system of particles moving in two dimensions and interacting through an inverse twelfth power repulsion interaction. For this problem, the present analysis, including the necessity for determining the locus of $\nu_c=0$ states and the stability analysis, is important for constructing the equilibrium phase diagram for the material [16]. We have developed a practical method, also planned to be reported elsewhere, for calculating μ_c by the insertion and deletion of particles [14].

In closing, we make a few miscellaneous comments about computer simulation of solids, based on our experience with this two-dimensional system and on the analysis presented here: (1) In simulating solids, it is very important to have some means for analyzing in detail the structure of the resulting configurations. For example, if the system is started in a crystalline configuration, with the unit cell axes oriented with regard to the simulation cell boundaries, then the system should be monitored to see if that situation persists. If it persists, then averages calculated do correspond to averages over a constrained ensemble. If the structure does not persist, but fluctuates from one crystalline arrangement to another, then it is likely that the averages have no meaning. (2) The stability analysis above suggests that even in a computer simulation with a fixed number of lattice sites, it is possible for a system to be unstable with regard to the development of inhomogeneities and that these inhomogeneities are manifestations of the possibility of a phase transition in the system. If the simulation run is long enough, the inhomogeneities may develop. In the case of a gas-liquid transition, such inhomogeneities are typically not observed in small systems because the positive interfacial free energy suppresses the inhomogeneity. However, we have seen an example of this instability in our simulations of two-dimensional crystals. It appears in systems with 0.001 or more net vacancies per lattice site. It shows itself as an average tendency for the majority of the vacancies to cluster together. Sometimes, in addition, at lower densities, vacancy-interstitial pairs will form, and the vacancies will cluster together in a small number of clusters. As just mentioned, some method for structure analysis is invaluable for detecting this effect. When this is observed, simulations in a finite system should not be expected to give valid information about the thermodynamic limit since the clustering is evidence that under the conditions of the simulation the crystal is not thermodynamically stable.

ACKNOWLEDGMENT

This work was supported by the National Science Foundation under Grant No. CHE89-18841 and made use of computer resources provided under NSF Grant No. CHE88-21737.

---

[1] D. A. McQuarrie, *Statistical Mechanics* (Harper & Row, New York, 1976).

[2] C. Kittel, *Introduction to Solid State Physics* (Wiley, New York, 1976).

[3] B. J. Alder and T. E. Wainwright, J. Chem. Phys. **27**, 1208 (1957); Phys. Rev. **127**, 359 (1962).

[4] M. Parrinello and A. Rahman, J. Appl. Phys. **52**, 7182 (1981).

[5] D. Frenkel and A. J. C. Ladd, J. Chem. Phys. **81**, 3188 (1984).

[6] A. Z. Panagiotopoulos, Mol. Phys. **61**, 813 (1987); **63**, 527 (1988).

[7] H. C. Andersen, J. Chem. Phys. **72**, 2384 (1980).

[8] H. B. Callen, *Thermodynamics and an Introduction to Thermostatistics*, 2nd ed. (Wiley, New York, 1985).

[9] J. M. Kosterlitz and D. J. Thouless, J. Phys. C **6**, 1181 (1973).

[10] Throughout this discussion we are assuming that we know the shape and symmetry of the unit cell for the thermodynamically stable crystal structure. In principle, if there

is a question of which structure is most stable, an analysis of the type discussed here could be carried out separately for each candidate structure to determine which is the equilibrium structure under various conditions of temperature and pressure.

[11] See, for example, D. A. McQuarrie, Ref. [1], pp. 263–262.

[12] It should be noted, however, that Eq. (23) must hold only for values of N and $\mathbf{r}^N$ that contribute significantly to the partition function for a crystal in volume V. The equation is clearly incorrect if it is applied iteratively to the successive removal of enough particles to destabilize the lattice.

[13] B. Widom, J. Chem. Phys. **39**, 2808 (1963). The method was developed independently by J. L. Jackson and L. S. Klein, Phys. Fluids **7**, 228 (1964).

[14] H. C. Andersen and W. C. Swope (unpublished).

[15] In our simulations of two-dimensional crystals, we have generated the information listed here. In fact it is possible to get by with slightly less information, namely, the absolute free energy at one point, the chemical potential along one line or curve that spans the entire range of $\tilde{n}$ in the region, and the pressure in the entire region. With this it is possible to construct the line integral of $d\tilde{a}_c$ from the point at which the absolute free energy is known to any point in the region.

[16] H. C. Andersen and W. C. Swope (unpublished).