INTERNATIONAL ATOMIC ENERGY AGENCY
UNITED NATIONS EDUCATIONAL, SCIENTIFIC AND CULTURAL ORGANIZATION
INTERNATIONAL CENTRE FOR THEORETICAL PHYSICS
I.C.T.P., P.O. BOX 586, 34100 TRIESTE, ITALY, CABLE: CENTRATOM TRIESTE

SMR.550 - 39

SPRING COLLEGE IN MATERIALS SCIENCE ON

"NUCLEATION, GROWTH AND SEGREGATION IN MATERIALS SCIENCE AND ENGINEERING"

(6 May - 7 June 1991)

SELECTED PAPERS ON

SURFACE DIFFUSION AND GRAIN BOUNDARY SIMULATIONS

A.F. VOTER
Los Alamos National Laboratory
Los Alamos, New Mexico 87545
U.S.A.

These are preliminary lecture notes, intended only for distribution to participants.

Strada Costiera, 11 Tel. 22401 Telefax 224163 / 224559 Telex 460392 ADRIATICO GUEST HOUSE Via Grigiano, 9 Tel. 224241 Telefax 224531

Dynamical corrections to transition state theory for multistate systems: Surface self-diffusion in the rare-event regime

Arthur F. Voter and Jimmie D. Doll
Chemistry Division, Los Alamos National Laboratory, Los Alamos, New Mexico 87545

(Received 10 August 1984; accepted 17 September 1984)

We derive an expression for the classical rate constant between any two states of a multistate system. The rate is given as the transition state theory rate of escape from the originating state, multiplied by a dynamical correction factor in the form of a time-correlation function which is evaluated using molecular dynamics techniques. This method is desiged to treat cases in which reactive state-change events are so infrequent (e.g., at low temperature) that direct molecular dynamics calculations are unfeasible. In this regime where dynamical recrossings occur much more quickly than the average time between reactive state changes, the concept of a rate between two nonadjacent states becomes meaningful. We apply the method to the surface diffusion of Rh on Rh(100) at the temperatures employed in field ion microscope experiments.

I. INTRODUCTION

The direct simulation of chemical processes via classical molecular dynamics (MD) techniques is now a widely used and powerful approach,[1] due largely to the high speed of modern computers. There remains, however, a class of systems for which direct molecular dynamics simulation is unfeasible. These systems are characterized by the "rare-event" nature of their dynamical evolution from reactant to product states. Because there is a "bottleneck" in phase space (e.g., a high activation barrier) through which the system must pass to change states, direct integration of the equations of motion may require many years of computer time before a single reactive state-change event is observed. Examples of these rare-event processes include thermal desorption from solid surfaces, surface diffusion at low temperatures, and any chemical reaction with an activation barrier which is high relative to the temperature.

For such systems an elegant alternative approach has been developed,[2-9] which yields exact dynamical quantities without the computational effort growing proportional to the rareness of the event of interest (as in direct MD). The key to this approach is the factoring of the rate constant into two parts: (1) an equilibrium factor, defined as the flux through a dividing surface separating the two states, and (2) a dynamical correction factor, which accounts for the fact that this flux contains spurious crossings which do not correspond to true reactive state-change events. The rare-event nature of the process is included in the first factor, which is simply the transition state theory (TST) rate constant. Because this is an equilibrium property of the system, and does not require knowledge of the dynamics, a variety of methods are available for computing it efficiently.[10,11] The second factor requires explicit use of molecular dynamics, but is easily evaluated because the trajectories need only be followed for relatively short times.

For many problems the dynamical correction factor is close to unity (especially if the TST dividing surface is chosen well), and thus the TST approximation is often very good. Indeed, TST has proven to be, and remains, a most valuable tool for computing reaction rates.[12] However, it is important to have the capability of computing dynamically exact rates for two reasons: (1) the quality of TST can be tested for a certain type of system by comparing to the exact result for a representative case, and (2) for many systems, the TST approximation is too poor to be useful.

Though the concept of a dynamical correction factor has been around for many years,[13] the first use of MD for its exact evaluation was due to Keck.[2] More recently, this type of approach has been applied to describe diffusion in solids,[6] gas phase reactions,[3,4,14,15] intramolecular rearrangements[8,16] and reactions in solution,[9] and thermal desorption from solids.[17-19] All work to date (with two exceptions[16,19] discussed below) has been on two-state systems. In this paper we present a generalization of the dynamical-correction formalism to the many-state case. This should be useful for a variety of problems, such as diffusion on or in a solid, polymer chain dynamics, etc. As an example, we apply the method to calculate the surface self-diffusion rate for a Rh atom on a Rh(100) surface at 300 K, a task which would require $>10^8$ years of computer time[20] using direct MD.

Our derivation of the many-state result is identical in spirit to the two-state formalism presented by Chandler,[8] and Skinner and Wolynes.[9] From the properties of the fluctuation-fluctuation autocorrelation function for a system at equilibrium, they extracted an expression which gives a precise prescription for evaluating the dynamical correction factor using classical trajectories. Montgomery, Holmgren, and Chandler[16] have examined the appropriate time correlation functions to estimate the importance of multiple-state transitions in alkane chains, and

 0021-9606/85/010080-13$02.10

Adams and Doll[19] have derived a matrix formulation for the many-state case, but in neither work did the authors actually extract elementary rate constants. We show here that the elementary rate constants can be extracted and, as in the two-state case, the formalism yields a simple prescription for computing the necessary dynamical correction factors. Moreover, an interesting concept which emerges naturally from this formalism is that of a rate constant between two states which are not adjacent in configuration space.

This paper is organized as follows: We review the two-state derivation in Sec. II, and extend it to the many-state case in Sec. III. Section IV contains a description of how molecular dynamics techniques are applied to compute the dynamical correction factors, and Sec. V contains the application to Rh on Rh(100) surface diffusion.

II. THEORY: TWO-STATE SYSTEM

We first derive the method for making dynamical corrections to transition state theory (TST) for a two-state classical system. This is a simplified version of Chandler's original derivation.[8] As shown in Fig. 1, we are considering a one-dimensional system with two stable configurations, A and B, separated by an energy barrier at $x = q$. We assume that for a canonical ensemble of such systems, a first-order rate law applies, so that

$$\dot{N}_A = -k_{A\to B}N_A + k_{B\to A}N_B$$
$$\dot{N}_B = k_{A\to B}N_A - k_{B\to A}N_B, \tag{2.1}$$

where N_A and N_B are the populations of state A and state B, respectively, $k_{A\to B}$ and $k_{B\to A}$ are the elementary rate constants, and the dot indicates a time derivative. We can define fluctuations of N_A and N_B away from their equilibrium values by

$$\delta N_A = N_A - \bar{N}_A$$
$$\delta N_B = N_B - \bar{N}_B, \tag{2.2}$$

where conservation of particles requires

$$\delta N_A + \delta N_B = 0. \tag{2.3}$$

Combining Eqs. (2.1), (2.2), and (2.3) leads to a rate law for the fluctuations,

$$\begin{aligned}\delta\dot{N}_A(t) &= k_{B\to A}\delta N_B - k_{A\to B}\delta N_A \\ &= -k_{B\to A}\delta N_A - k_{A\to B}\delta N_A \\ &= -k_{\rm eff}\delta N_A(t),\end{aligned} \tag{2.4}$$

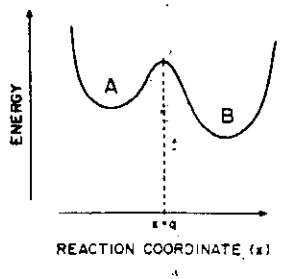


FIG. 1. Energy diagram for one-dimensional two-state system. The boundary between the two states is at $x = q$.

where we have defined an effective rate constant as

$$k_{\rm eff} = k_{A\to B} + k_{B\to A}. \tag{2.5}$$

Equation (2.4) gives the response of the system to an artificial displacement from equilibrium. From the fluctuation-dissipation theorem,[21] we expect that when the system is at equilibrium, it responds in this same way to spontaneously occurring fluctuations. Thus, we deduce the behavior of the fluctuation–fluctuation autocorrelation function to be

$$\langle\delta N_A(0)\delta\dot{N}_A(t)\rangle = -k_{\rm eff}\langle\delta N_A(0)\delta N_A(t)\rangle. \tag{2.6}$$

The angular brackets indicate the usual canonical-ensemble average

$$\langle Y\rangle = \frac{\int dp\int dx\, Ye^{-\beta H}}{\int dp\int dx\, e^{-\beta H}},$$

where p is the conjugate momentum, H is the Hamiltonian for the system, and $\beta = 1/k_BT$ (k_B = the Boltzmann constant, T = temperature). Noting that

$$\langle\delta N_A(0)\delta\dot{N}_A(t)\rangle = -\langle\delta\dot{N}_A(0)\delta N_A(t)\rangle,$$

we may rewrite Eq. (2.6) as

$$k_{\rm eff} = \frac{\langle\delta\dot{N}_A(0)\delta N_A(t)\rangle}{\langle\delta N_A(0)\delta N_A(t)\rangle}. \tag{2.7}$$

Equation (2.7) gives an expression for the dynamically correct rate constant in terms of equilibrium properties, though it is not yet in a particularly useful form.

Noting that the denominator in Eq. (2.7) is a slowly varying function of time compared to the numerator, we make the following approximation, which we discuss later:

$$\langle\delta N_A(0)\delta N_A(t)\rangle \approx \langle\delta N_A(0)\delta N_A(0)\rangle. \tag{2.8}$$

With this assumption, and making use of Eq. (2.3), Eq. (2.7) becomes

$$k_{\rm eff}(t) = \frac{-\langle\delta\dot{N}_A(0)\delta N_B(t)\rangle}{\langle\delta N_A(0)\delta N_A(0)\rangle}, \tag{2.9}$$

where we have explicitly indicated the time dependence of $k_{\rm eff}$. Without loss of generality, we can make the simplifying requirement that

$$N_A + N_B = 1,$$

so that

$$N_A(t) = \theta[q - x(t)],$$
$$N_B(t) = \theta[x(t) - q],$$

where θ is the standard step function. The fluctuations in N_A and N_B are thus given by

$$\delta N_A(t) = \theta[q - x(t)] - \chi_A$$
$$\delta N_B(t) = \theta[x(t) - q] - \chi_B, \tag{2.10}$$

where χ_A and χ_B are the equilibrium mole fractions of components A and B, respectively (note that $\chi_A = \bar{N}_A$, $\chi_B = \bar{N}_B$). The autocorrelation function in the denominator of Eq. (2.9) becomes

$$\begin{aligned}&\langle\delta N_A(0)\delta N_A(0)\rangle \\ &\quad= \langle\{\theta[q - x(0)] - \chi_A\}\{\theta[q - x(0)] - \chi_A\}\rangle \\ &\quad= \langle\theta[q - x(0)]\theta[q - x(0)]\rangle - \chi_A^2 \\ &\quad= \chi_A - \chi_A^2.\end{aligned} \tag{2.11}$$

From detailed balance, we know that

$$\frac{k_{A\to B}}{k_{B\to A}} = \frac{\chi_B}{\chi_A}$$

so that Eq. (2.5) becomes

$$k_{\rm eff} = k_{A\to B}\left(1 + \frac{\chi_A}{\chi_B}\right),$$

and from Eqs. (2.9) and (2.11) we have

$$\begin{aligned}k_{A\to B}(t) &= \frac{-\langle\delta\dot{N}_A(0)\delta N_B(t)\rangle}{[1 + (\chi_A/\chi_B)](\chi_A - \chi_A^2)} \\ &= \frac{-\langle\delta\dot{N}_A(0)\delta N_B(t)\rangle}{\chi_A}.\end{aligned} \tag{2.12}$$

To simplify the numerator of Eq. (2.12), we need the time derivative of δN_A; from Eq. (2.10) we have

$$\delta\dot{N}_A(0) = -\dot{x}(0)\delta[q - x(0)],$$

where $\delta(\cdots)$ indicates the Dirac delta function. Thus, we obtain

$$k_{A\to B}(t) = \frac{\langle\dot{x}(0)\delta[x(0) - q]\theta[x(t) - q]\rangle}{\chi_A}. \tag{2.13}$$

The numerator in Eq. (2.13) can be interpreted as the particle flux through the $x = q$ dividing surface, modified by the step function $\theta[x(t) - q]$. The particles crossing the dividing surface in the $+x$ direction (i.e., moving from state A to state B) at time zero that reside in state B at time t contribute positively to the ensemble average, while particles residing in state B at time t that were making a $-x$ crossing at time zero contribute negatively to the average.

If every particle that crosses the dividing surface in the $+x$ direction remains in state B (at least until a time greater than t), and every particle crossing in the $-x$ direction remains in state A, then the ensemble average in Eq. (2.13) is simply the one-directional flux through the dividing surface. This is the transition state theory (TST) assumption, that every dividing surface crossing corresponds to a reactive state change. We can modify Eq. (2.13) to reflect the TST approximation by setting t to some small value ϵ, chosen such that no recrossings can occur between $t = 0$ and $t = \epsilon$. This gives

$$\begin{aligned}k_{A\to B}^{\rm TST} &= \frac{\langle\dot{x}(0)\delta[x(0) - q]\theta[x(\epsilon) - q]\rangle}{\chi_A} \\ &= \frac{\langle\dot{x}(0)\delta[x(0) - q]\theta[\dot{x}(0)]\rangle}{\chi_A}.\end{aligned} \tag{2.14}$$

It is easy to show that Eq. (2.14) is equivalent to more familiar expression for the TST rate. Noting fir that, in a canonical ensemble, the forward and revers flux through $x = q$ have equal magnitudes, we can writ

$$k_{A\to B}^{\rm TST} = \frac{\frac{1}{2}\langle|\dot{x}(0)|\delta[x(0) - q]\rangle}{\chi_A}. \tag{2.15}$$

We can express χ_A as

$$\begin{aligned}\chi_A &= \langle\theta[q - x(0)]\rangle \\ &= \frac{\int dp\int dx\,\theta[q - x(0)]e^{-\beta H(x,p)}}{\int dp\int dx\, e^{-\beta H(x,p)}} \\ &= \frac{\int dp\int_A dx\, e^{-\beta H(x,p)}}{\int dp\int_{A+B} dx\, e^{-\beta H(x,p)}},\end{aligned}$$

where the subscripts A and A + B indicate that th configuration-space integrals are over state A and al space, respectively. Equation (2.15) thus becomes

$$\begin{aligned}k_{A\to B}^{\rm TST} &= \left[\left[\frac{1}{2}\frac{\int dp\int_{A+B} dx|\dot{x}(0)|\delta[x(0) - q]e^{-\beta H}}{\int dp\int_{A\to B} dx\, e^{-\beta H}}\right]\right. \\ &\quad\left.\times\left[\frac{\int dp\int_{A+B} dx\, e^{-\beta H}}{\int dp\int_A dx\, e^{-\beta H}}\right]\right] \\ &= \frac{1}{2}\langle|\dot{x}(0)|\delta[x(0) - q]\rangle_A,\end{aligned} \tag{2.16}$$

where the subscript A indicates that the configuration space part of the ensemble average is integrated only over state A (plus the infinitesimal region necessary to include the TST dividing surface). Equation (2.16) is a standard expression for $k_{A\to B}^{\rm TST}$,[8,11,22] interpretable as the flux of particles through the TST dividing surface bordering state A, divided by two since half the flux is going the wrong direction. Note that the expression is independent of the nature of state B[23]; $k_{A\to B}^{\rm TST}$ is simply the *rate of escape* from state A across the boundary surface.

The dynamical correction factor we seek is obtained by taking the ratio of $k_{A\to B}$ [Eq. (2.13)] to $k_{A\to B}^{\rm TST}$ [Eq. (2.14)], which yields

$$f_d(t) = \frac{k_{A\to B}}{k_{A\to B}^{\rm TST}} = \frac{\langle\dot{x}(0)\delta[x(0) - q]\theta[x(t) - q]\rangle}{\langle\dot{x}(0)\delta[x(0) - q]\theta[x(\epsilon) - q]\rangle}. \tag{2.17}$$

Implementation of Eq. (2.17) requires an understanding of the time dependence of the correlation function in the numerator. As mentioned above, the TST assumption that every particle crossing the dividing surface will come to rest without recrossing is invalid for many systems. After a crossing, there is a chance that the particle will recross the surface one or more times before ultimately

thermalizing in state A or state B. (We are assuming that the system is coupled to other degrees of freedom, which are constantly exchanging energy with the coordinate considered explicitly.) We define τ_{corr} as the time scale on which these correlated dynamical recrossing events occur; i.e., a particle crossing $x = q$ at time zero will have thermalized by time τ_{corr}. We also define a time scale associated with the reaction, $\tau_{rxn} \approx (k_{eff})^{-1}$, which corresponds to the average time between *reactive* crossings of the dividing surface. At very short times ($t \lesssim \epsilon$), the dynamical correction factor given by Eq. (2.17) is unity, since the numerator and denominator are equivalent. As t is increased to τ_{corr}, $f_d(t)$ decays to some value less than one, because trajectories which recross contribute negatively, or fail to contribute, to the numerator. If τ_{rxn} is much greater than τ_{corr}, then $f_d(t)$ will reach a plateau value at $t = \tau_{corr}$, as shown in Fig. 2. If $f_d(t)$ is examined on a very long time scale (i.e., units of τ_{rxn}), it will be seen to decay exponentially, and it ultimately tends to zero, because the direction of travel of a particle making a crossing at time zero is uncorrelated with its position at $t \gg \tau_{rxn}$. Thus, only if this plateau region exists, and if Eq. (2.17) is evaluated during that time, can meaningful results be obtained. The dynamically correct rate constant will be given by

$$k_{A\to B} = k_{A\to B}^{TST} f_d(t_p), \tag{2.18}$$

where t_p satisfies

$$\tau_{corr} < t_p \ll \tau_{rxn}. \tag{2.19}$$

We can now examine the approximation made in Eq. (2.8). Because the correlation function $\langle \delta N_A(0)\delta N_A(t)\rangle$ decays on a time scale of τ_{rxn}, and because we restrict ourselves to times much smaller than τ_{rxn}, the zero time assumption is perfectly reasonable. Note that Eq. (2.7), which does not include this assumption, gives an expression for the dynamically correct rate constant which is valid for *all* times greater than τ_{corr}. As the numerator in Eq. (2.7) decays towards zero for $t \gg \tau_{rxn}$, the denominator also decays, leading to a constant ratio. Thus, if we used Eq. (2.7) to compute k_{eff}, we would not need to require a separation of time scales between τ_{corr} and τ_{rxn}; we could simply pick any $t > \tau_{corr}$ to evaluate the right-hand side of Eq. (2.7). However, evaluating the denominator of Eq. (2.7) requires more computational work than the numerator, and can become prohibitively expensive when the reaction is slow. For this reason Eq. (2.17) becomes the method of choice when the requirements of Eq. (2.19) can be met.

The concept of time-scale separation merits further discussion. We can view this approach to calculating $k_{A\to B}$ in the following way: We imagine watching the motion of a single particle which has resided in state A for a time greater than τ_{corr}. When that particle finally makes a crossing of the dividing surface, we close our eyes for the next τ_{corr} time units, and then note the state of the particle upon reopening our eyes. If the particle is in state B, we say that a reactive crossing has occurred, but if the particle is in state A, we pretend that nothing happened—the particle never left state A. We now continue to monitor the motion of this particle, and follow the same prescription each time a crossing occurs (in either direction). After observing a large number of reactive hops, we can compute $k_{A\to B}$ as the average of the inverse of the time between a reactive hop into state A and the subsequent reactive exit from state A. Because $\tau_{corr} \ll \tau_{rxn}$, we make only a negligible error using this approach—there is a negligible probability that an important event (i.e., an extra reactive crossing) occurs during the infinitesimal fraction of time that our eyes are closed.

Now consider how Eq. (2.17) actually performs this counting. Figure 3 shows an idealized plot of the time evolution of a system executing the four possible types of surface crossings: A → B, A → A, B → A, and B → B, each consisting of multiple recrossings. Only the A → B event [Fig. 3(a)] should contribute to the calculation of $k_{A\to B}$. If we examine a canonical ensemble of particles at a given instant of time, we find that each of the five crossings in Fig. 3(a) are present. The time-correlation function $\langle \dot{x}(0)\delta[x(0) - q]\theta[x(t_p) - q]\rangle$ instructs us to allow each of these five systems to evolve for a time τ_{corr}, and then examine the final state. Because each system in our example will be in state B, so that $\theta[x(t_p) - q] = 1$, they will contribute positively or negatively to the average, depending on the sign of $\dot{x}(0)$. Crossings 1, 3, and 5 contribute positively, while 2 and 4 contribute negatively. Thus, the net contribution will be one crossing, as will be the case for any number of recrossings. [Note that only the *phase* of $\dot{x}(0)$ is important, not the magnitude;

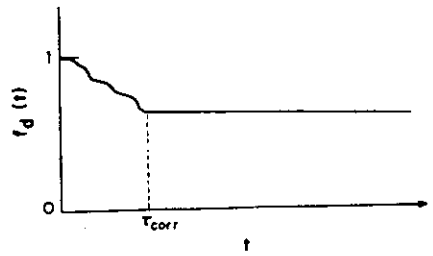


FIG. 2. Typical behavior of the two-state dynamical correction factor [Eq. (2.17)] for $\tau_{corr} \ll \tau_{rxn}$.

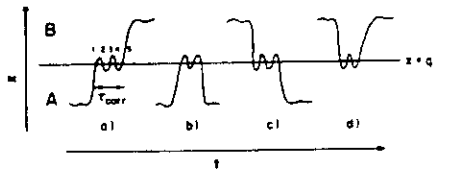


FIG. 3. Idealized diagram of the four possible multiple-crossing events for the two-state system. The numbers 1 to 5 in (a) label the five crossings that comprise the single A → B reactive crossing event. The A → A, B → A, and B → B crossing events shown in (b), (c), and (d) make zero contribution to the numerator in Eq. (2.17), and hence do not contribute to the dynamically correct $k_{A\to B}$.

...e magnitude appears in the ensemble average because we are sampling from a velocity-weighted Boltzmann distribution, corresponding to a snapshot in time of the particles crossing the dividing surface. This will be demonstrated in Sec. IV.] The A → A and B → A events shown in Figs. 3(b) and 3(c) contribute nothing to the average because $\theta[x(t) - q] = 0$. The B → B event shown in Fig. 3(d) makes no net contribution to the average because there are an equal number of $\dot{x}(0) > 0$ and $\dot{x}(0) < 0$ crossings.

Computation of the dynamically exact classical rate constant is accomplished in two stages. First, the TST rate constant is evaluated via Eqs. (2.15) or (2.16). Then the dynamical correction factor f_d is computed from Eq. (2.17), using the same dividing surface as in the TST calculation. The ensemble averages in Eq. (2.17) are evaluated by following the behavior of a swarm of classical trajectories which start from the TST dividing surface. This procedure is described in detail in Sec. IV. It is worth noting that the rate constant computed in this manner is independent of the location of the dividing surface. While choosing the surface to be at the free-energy maximum between the two states will yield the most accurate $k_{A\to B}^{TST}$ (and f_d closest to unity), other positions will also work. This is helpful in many-dimensional systems, because the position which will maximize f_d is often not easily defined.

III. THEORY: MULTISTATE SYSTEM

In the last section we showed that the dynamical correction factor for a two-state system can be expressed as a ratio of equilibrium correlation functions, which may be evaluated using classical trajectory techniques. In this section we extend the analysis to a system with an arbitrary number of states, and with the same assumption that $\tau_{corr} \ll \tau_{rxn}$, we show that an analogous expression exists for the rate of reaction between any two states of the system.

We begin with the matrix formulation of the rate law used by Adams and Doll.[19] The system is assumed to be composed of n states, with no restriction on the connectivity of the states. An example is the potential for an atom diffusing on a solid surface. In this case, there are an infinite number of states, each corresponding to one binding site, with each state bordering between two and four other states (depending on the crystal face). Because the total number of particles N is held fixed, it is sufficient to specify the populations of $n - 1$ states. We thus define the $(n - 1)$-dimensional column vector of time-dependent populations $\mathbf{N}(t)$ and fluctuations away from equilibrium,

$$\delta\mathbf{N}(t) = \mathbf{N}(t) - \bar{\mathbf{N}}, \tag{3.1}$$

where $\bar{\mathbf{N}}$ is the column vector of equilibrium populations. We again assume that the populations, and hence the fluctuations, obey a first-order rate law, which we may write as

$$\delta\dot{\mathbf{N}}(t) = -\mathbf{k}\delta\mathbf{N}(t), \tag{3.2}$$

where $\mathbf{k}$ is a $(n - 1) \times (n - 1)$ matrix of rate constants. It is important to note that the elements of $\mathbf{k}$ are *not* the elementary rate constants (indeed, our task will be to find the relation between the two). As in the two-state case, we take the behavior of the fluctuation correlation functions to be governed by the same rate law, leading to

$$\langle \delta\dot{\mathbf{N}}(t)\delta\mathbf{N}^T(0)\rangle = -\mathbf{k}\langle \delta\mathbf{N}(t)\delta\mathbf{N}^T(0)\rangle,$$

which was obtained by right multiplying Eq. (3.2) by the row vector $\delta\mathbf{N}^T(0)$ [the transpose of $\delta\mathbf{N}(0)$], and taking an ensemble average. [Note that $\langle \delta\mathbf{N}(t)\delta\mathbf{N}^T(0)\rangle$ represents a matrix of correlation functions.] Taking advantage of

$$\langle \delta\dot{\mathbf{N}}(t)\delta\mathbf{N}^T(0)\rangle = [\langle \delta\mathbf{N}(0)\delta\dot{\mathbf{N}}^T(t)\rangle]^T$$

and

$$\langle \delta\mathbf{N}(0)\delta\dot{\mathbf{N}}^T(t)\rangle = -\langle \dot{\mathbf{N}}(0)\delta\mathbf{N}^T(t)\rangle,$$

we obtain an expression for the elements of the transpose of $\mathbf{k}$,

$$\mathbf{k}^T = [\langle \delta\mathbf{N}(0)\delta\mathbf{N}^T(t)\rangle]^{-1}\langle \delta\dot{\mathbf{N}}(0)\delta\mathbf{N}^T(t)\rangle. \tag{3.3}$$

We now deviate from Adams and Doll.

We wish to express the *elementary* rate constants $k_{i\to j}^{el}$ in terms of time correlation functions, as we did for the two-state case. We can write the rate law as

$$\begin{aligned}\delta\dot{N}_i(t) &= \sum_j^n (-\delta N_i k_{i\to j}^{el} + \delta N_j k_{j\to i}^{el}) \\ &= -\sum_j^n \delta N_i f_{i,j} + \sum_j^n \delta N_j f_{j,i},\end{aligned} \tag{3.4}$$

where δN_i is the ith element of $\delta\mathbf{N}(t)$, and we have introduced the $n \times n$ matrix of elementary rate constants $\mathbf{f}$ defined by

$$(\mathbf{f})_{i,j} = f_{i,j} = k_{i\to j}^{el}, \tag{3.5}$$

to avoid confusing $k_{i\to j}^{el}$ with an element of $\mathbf{k}$. Now consider the following expression, which gives $\mathbf{k}$ in terms of elementary rates:

$$(\mathbf{k})_{i,j} = -f_{ji} + f_{ni} + \delta_{ij}\sum_k^n f_{ik}; \quad i,j < n \tag{3.6}$$

(here δ_{ij} is the Kronecker delta function). We will prove Eq. (3.6) by inserting it into Eq. (3.2) and comparing with Eq. (3.4). Rewriting Eq. (3.2) and inserting Eq. (3.6) gives

$$\begin{aligned}\delta\dot{N}_i(t) &= -\sum_k^{n-1} (\mathbf{k})_{ik}\delta N_k(t) \\ &= -\sum_k^{n-1}\left(-f_{ki} + f_{ni} + \delta_{ik}\sum_j^n f_{ij}\right)\delta N_k(t) \\ &= \sum_k^{n-1} f_{ki}\delta N_k(t) - f_{ni}\sum_k^{n-1}\delta N_k(t) - \sum_k^n f_{ik}\delta N_i(t).\end{aligned}$$

Extending the range of the first two sums from $n - 1$ to n (the extra terms cancel), and noting that

$$\sum_k^n \delta N_k(t) = 0 \tag{3.7}$$

we obtain

$$\delta\dot{N}_i(t) = \sum_k^n f_{ik}\delta N_k(t) - 0 - \sum_k^n f_{ik}\delta N_i(t),$$

which is the same as Eq. (3.4), proving Eq. (3.6). Now consider another expression:

$$k_{i\to j}^{\mathrm{el}} = f_{ij} = \sum_k^{n-1} \chi_k(\mathbf{k})_{jk} - (\mathbf{k})_{ji}; \quad i,j < n, \tag{3.8}$$

where χ_k is the equilibrium mole fraction in state k. We will prove this by inserting Eq. (3.6) into the right-hand side of Eq. (3.8), which gives

$$\sum_k^{n-1} \chi_k\left(-f_{kj} + f_{nj} + \delta_{jk}\sum_l^n f_{jl}\right) + f_{ij} - f_{nj} - \delta_{ji}\sum_k^n f_{jk}.$$

Because we are not interested in $k_{i\to i}^{\mathrm{el}}$, we may discard the last term ($\delta_{ji}\sum_k^n f_{jk}$). We may also extend the range of the first sum from $n-1$ to n, because the extra terms cancel, as long as $j < n$. (We are currently restricting i and j to be less than n, but our ultimate expression for $k_{i\to j}^{\mathrm{el}}$ will be valid for all $i \neq j$.) This leads to

$$-\sum_k^n \chi_k f_{kj} + f_{nj}\sum_k^n \chi_k + \sum_l^n f_{jl}\chi_j + f_{ij} - f_{nj},$$

and using

$$\sum_k^n \chi_k = 1$$

and

$$\chi_i f_{ij} = \chi_j f_{ji},$$

we obtain

$$\sum_k^n (-\chi_k f_{kj} + \chi_j f_{jk}) + f_{nj}\left(\sum_k^n \chi_k\right) + f_{ij} - f_{nj}$$
$$= 0 + f_{nj} + f_{ij} - f_{nj}$$
$$= f_{ij},$$

which proves Eq. (3.8). Substituting Eq. (3.3) into Eq. (3.8) gives the elementary rate constants in terms of correlation functions, but first we wish to simplify the "denominator" matrix in Eq. (3.3).

As in the two-state case, we will approximate the fluctuation-fluctuation correlation function by its time-zero value

$$\langle\delta\mathbf{N}(0)\delta\mathbf{N}^T(t)\rangle \approx \langle\delta\mathbf{N}(0)\delta\mathbf{N}^T(0)\rangle. \tag{3.9}$$

Taking $N = 1$ (with no loss of generality), so that $\bar{N} = \chi_n$, it is easy to show (see the Appendix) that

$$[\langle\delta\mathbf{N}(0)\delta\mathbf{N}^T(0)\rangle]_{ij} = \delta_{ij}\chi_i - \chi_i\chi_j, \tag{3.10}$$

and that the desired inverse matrix is given by

$$\{[\langle\delta\mathbf{N}(0)\delta\mathbf{N}^T(0)\rangle]^{-1}\}_{ij} = \frac{1}{\chi_n} + \frac{\delta_{ij}}{\chi_i}. \tag{3.11}$$

Using Eqs. (3.3), (3.8), and (3.11), we can express the elementary rate constants in terms of the correlation-function matrix elements

$$C_{ij} = \langle\delta\dot{N}_i(0)\delta N_j(t)\rangle.$$

From Eqs. (3.3) and (3.11) we have

$$(\mathbf{k}^T)_{ij} = \sum_k^{n-1}\left(\frac{1}{\chi_n} + \frac{\delta_{jk}}{\chi_j}\right)C_{ki},$$

so that

$$(\mathbf{k})_{ij} = \sum_k^{n-1}\left(\frac{1}{\chi_n} + \frac{\delta_{jk}}{\chi_j}\right)C_{kj},$$

and inserting into Eq. (3.8) gives

$$k_{i\to j}^{\mathrm{el}} = \sum_l^{n-1}\chi_l\left[\sum_k^{n-1}\left(\frac{1}{\chi_n} + \frac{\delta_{lk}}{\chi_l}\right)C_{kj}\right] - \sum_k^{n-1}\left(\frac{1}{\chi_n} + \frac{\delta_{jk}}{\chi_j}\right)C_{ki}$$
$$= \left(\sum_l^{n-1}\chi_l\right)\left(\sum_k^{n-1}\frac{C_{kj}}{\chi_n}\right) + \sum_l^{n-1}\frac{\chi_l}{\chi_l}C_{lj} - \sum_k^{n-1}\frac{C_{kj}}{\chi_n} - \frac{C_{ij}}{\chi_i}$$
$$= \left(\frac{1-\chi_n}{\chi_n}\right)\sum_k^{n-1}C_{kj} + \sum_k^{n-1}C_{kj} - \frac{1}{\chi_n}\sum_k^{n-1}C_{kj} - \frac{C_{ij}}{\chi_i}$$
$$= -C_{ij}/\chi_i,$$

and substituting for C_{ij} yields

$$k_{i\to j}^{\mathrm{el}} = \frac{-\langle\delta\dot{N}_i(0)\delta N_j(t)\rangle}{\chi_i}. \tag{3.12}$$

Equation (3.12) is the central result of this paper. It states that, when $\tau_{\mathrm{corr}} \ll \tau_{\mathrm{rxn}}$, the elementary rate constant between any two states, connected or not, is given by a simple equilibrium correlation function. Note that the form of Eq. (3.12) is identical to Eq. (2.12) for the two-state system. Note also that there is no dependence on state n in Eq. (3.12); we are free to choose any state other than i or j as the nth state, and thus we may abandon the $i, j < n$ restriction. As in the two-state case, we will tinker with Eq. (3.12) to obtain a computationally convenient form.

Because systems of interest will rarely be one dimensional, we extend our scope to a many-dimensional system defined by the coordinate $\mathbf{R}$. The time-dependent population of state i is given by

$$\theta_i(t) = \theta\{F_i[\mathbf{R}(t)]\},$$

where $F_i(\mathbf{R})$ is a continuous, differentiable function with the property that

$$F_i(\mathbf{R}) \begin{cases} >0, & \text{if } \mathbf{R} \text{ is in state } i, \\ =0, & \text{if } \mathbf{R} \text{ is on the boundary to state } i, \\ <0, & \text{if } \mathbf{R} \text{ is outside state } i. \end{cases}$$

The population fluctuation is thus

$$\delta N_i(t) = \theta_i(t) - \chi_i \tag{3.13}$$

the time derivative becomes

$$\delta\dot{N}_i(t) = \delta\{F_i[\mathbf{R}(t)]\}\nabla F_i\cdot\mathbf{v}$$
$$= -\delta_i(t)v_i(t), \tag{3.14}$$

where v_i is the velocity normal to the dividing surface (defined as positive when the system is *exiting* state i),

$$v_i = \frac{-\nabla F_i\cdot\mathbf{v}}{|\nabla F_i|},$$

and we have defined

$$\delta_i(t) = \delta\{F_i[\mathbf{R}(t)]\}|\nabla F_i[\mathbf{R}(t)]|. \tag{3.15}$$

Equation (3.12) becomes

$$k_{i\to j}^{\mathrm{el}} = \frac{\langle v_i(0)\delta_i(0)\theta_j(t)\rangle}{\chi_i}, \tag{3.16}$$

and writing the TST rate of *escape* from state i as

$$k_{i\to}^{\mathrm{TST}} = \frac{\frac{1}{2}\langle|v_i(0)|\delta_i(0)\rangle}{\chi_i} \tag{3.17}$$

[see Eq. (2.15)], we obtain

$$k_{i\to j}^{\mathrm{el}} = \frac{\langle v_i(0)\delta_i(0)\theta_j(t)\rangle}{\frac{1}{2}\langle|v_i(0)|\delta_i(0)\rangle}\,k_{i\to}^{\mathrm{TST}}. \tag{3.18}$$

As in the two-state case, we have expressed the true rate constant as the transition state theory rate modified by a ratio of ensemble averages. The expression is only valid if the numerator can be evaluated at a time t such that

$$\tau_{\mathrm{corr}} < t \ll \tau_{\mathrm{rxn}}. \tag{3.19}$$

Before describing how to evaluate Eq. (3.18) via molecular dynamics, we conclude this section with a few important points.

As indicated by Eq. (3.17), $k_{i\to}^{\mathrm{TST}}$ depends only on the nature of state i, and the shape of its boundary surface. Thus, in computing $k_{i\to j}^{\mathrm{el}}$, the same TST rate constant is employed for any final state j. All dependence on state j is contained in the time-correlation function $\langle v_i(0)\delta_i(0)\theta_j(t)\rangle$, which accounts for the ultimate fate of particles which are at the boundary to state i at time zero.

Note that state i and state j *need not be connected*, $k_{i\to j}^{\mathrm{el}}$ may be computed from Eq. (3.18) for any two states of the system. While this may seem contrary to the conventional concept of a rate constant, we feel that $k_{i\to j}^{\mathrm{el}}$ is in fact well defined for nonadjacent i and j. As in the two-state case, we appeal to the notion of "closing our eyes" during the time τ_{corr} following an initial crossing event. If a particle initially in state i exits and reappears in some distant state j when we reopen our eyes, then it is reasonable to say that it has made a direct flight from state i to state j. The possibility that we are missing some important event is negligibly small, since our eyes are closed only $\tau_{\mathrm{corr}}/\tau_{\mathrm{rxn}}$ of the time. This concept can also be stated another way. Though states i and j are not connected in configuration space, we can think of them as being connected by a sort of "tunnel" in phase space, which passes through the states connecting i and j. When a particle exits state i with the right set of phase-space coordinates for entering this tunnel, it is automatically guided to state j. Even though it passes through the configuration space of some state k, this tunnel region of phase space is inaccessible to state k, and can only be entered from state i (or state j, by time-reversal symmetry—assuming that all coordinates are explicitly included in the system). A particle can pass from i to j through this tunnel in a time less than τ_{corr}, while actually hopping from i to k, and then k to j, would require $\sim 2\,\tau_{\mathrm{rxn}}$.

As in the two-state case, it is easy to visualize how the ensemble average in Eq. (3.18) counts only the true reactive events. Figure 4 shows a schematic trajectory which, in time τ_{corr}, passes from state i through states k, l, and m, and into state j. Each of the crossing points, labeled 1 to 7, will appear in the canonical ensemble, and after evolving for time τ_{corr}, will reside in state j. Because of the form of the time-correlation function $\langle v_i(0)\delta_i(0)\theta_j(t)\rangle$, this set of trajectories can only contribute to $k_{i\to j}^{\mathrm{el}}$, $k_{k\to j}^{\mathrm{el}}$, $k_{l\to j}^{\mathrm{el}}$, and $k_{m\to j}^{\mathrm{el}}$. However, the net contributions to $k_{k\to j}^{\mathrm{el}}$, $k_{l\to j}^{\mathrm{el}}$, and $k_{m\to j}^{\mathrm{el}}$ will be zero because for each of these three initial states, there are an equal number of initially entering trajectories [$v_i(0) < 0$] and initially exiting trajectories [$v_i(0) > 0$]. (For example, trajectories 4 and 6 contribute negatively to $k_{m\to j}^{\mathrm{el}}$, while 5 and 7 contribute positively.) Thus, as we would hope, the only nonzero contribution is to $k_{i\to j}^{\mathrm{el}}$.

Figure 5 shows the expected time dependence of the dynamical correction factor, defined as

$$f_d(i\to j) = k_{i\to j}^{\mathrm{el}}/k_{i\to}^{\mathrm{TST}}. \tag{3.20}$$

If states i and j are adjacent, the behavior is similar to the two-state case; $f_d(i\to j)$ decays from unity and reaches a plateau value by $t = \tau_{\mathrm{corr}}$. If states i and j are not adjacent, $f_d(i\to j)$ is initially zero (because a finite amount of time is required for any trajectory initially exiting state i to reach state j), and then *rises* towards a plateau value at τ_{corr}. For both cases, the approach to the plateau value is not necessarily a monotonic function, since trajectories passing through state j on their way to another final state will cause a temporary increase (or decrease) in $f_d(i\to j)$

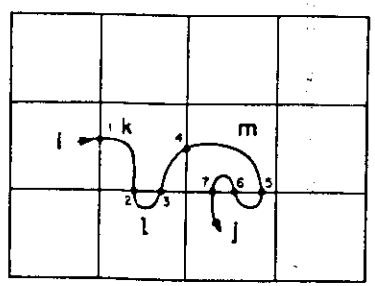


FIG. 4. Schematic example of a trajectory that passes from state i to state j in time τ_{corr}, via states k, l, and m. Each box represents a different state of the system. Each of the seven surface crossings shown will be represented in the canonical ensemble. Using Eq. (3.18) to evaluate the dynamically correct rate constants shows that this trajectory contributes only to $k_{i\to j}$, while at the TST level of approximation this trajectory contributes to $k_{i\to}^{\mathrm{TST}}$, $k_{k\to}^{\mathrm{TST}}$, $k_{l\to}^{\mathrm{TST}}$, $k_{i\to m}^{\mathrm{TST}}$, $k_{k\to m}^{\mathrm{TST}}$, $k_{m\to j}^{\mathrm{TST}}$, and $k_{l\to m}^{\mathrm{TST}}$.

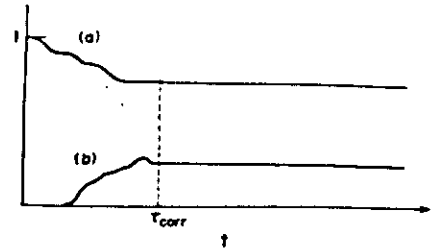


FIG. 5. Typical behavior of the many-state dynamical correction factor [Eq. (3.20)] for $\tau_{corr} \ll \tau_{rxn}$. In case (a), state j is adjacent to state i, and the behavior is very similar to the two-state case. In case (b), states i and j are not adjacent, so that the correction factor is initially zero, and rises to a plateau value.

(this is also true for the two-state case). As in the two-state case, if the time-zero approximation [Eq. (3.9)] had not been imposed, the computed value for $f_{ij}(i \to j)$ would remain constant for all $t > \tau_{corr}$, and the requirement that $\tau_{corr} \ll \tau_{rxn}$ would not be necessary (though this approach would require as much work as a direct molecular dynamics calculation). While in the two-state case this time-zero approximation caused a slow exponential decay in the plateau region, in the many-state case the plateau behavior is more complicated (though still slow), since a number of terms arising from the matrix multiplication shown in Eq. (3.3) are missing.

IV. EVALUATION OF THE RATE CONSTANTS VIA MOLECULAR DYNAMICS

The first stage in calculating $k_{i\to j}^{el}$ consists of computing the TST rate $k_{i\to}^{TST}$. This can be accomplished via Eq. (3.17), or from an ensemble average over only the phase space belonging to state i,

$$k_{i\to}^{TST} = \frac{1}{2}\langle |v_i(0)|\delta_i(0)\rangle_i. \tag{4.1}$$

We have shown elsewhere[24] that this can usually be further simplified to give an average over only the *configuration space* of state i,

$$k_{i\to}^{TST} = \frac{1}{2}\left(\frac{2k_BT}{\pi m}\right)^{1/2}\langle\delta_i(0)\rangle_i, \tag{4.2}$$

where $\delta_i(0)$ is given by Eq. (3.15), and m is the particle mass, assuming that $F_i(\mathbf{R})$ (i.e., the dividing surface) involves the coordinates of one particle only. If $F_i(\mathbf{R})$ is a function of the coordinates of more than one particle, the separation implied by Eq. (4.2) is still possible provided that the effective mass[25] of the coordinate perpendicular to the dividing surface is constant over the whole dividing surface. For this case, the mass in Eq. (4.2) is replaced by the effective mass of this perpendicular coordinate [$S(\mathbf{R})$], and the gradient operator in Eq. (3.15) operates with respect to S rather than $\mathbf{R}$, so that

$$\delta_i(t) = \delta(F_i)\left|\frac{\partial F_i}{\partial S}\right|. \tag{4.3}$$

Equation (4.2) may be evaluated using Metropolis Monte Carlo techniques,[26–28] once a form for the potential energy function has been chosen. The Monte Carlo approach has the advantage that the exact TST rate is obtained, with statistical error bars that can be made arbitrarily small. Alternatively, some form of harmonic approximation may be employed; this tends to work very well at low temperature where the system resides predominately in the harmonic region of the potential.

The second stage, which is the focus of this section, is the evaluation of the dynamical correction factors. (Either f_{ij} or k^{TST} may be evaluated first; the only requirement is that the *same* TST dividing surface be used for each calculation.) If we define the weighting function for a Maxwellian-flux distribution at the boundary to state i by

$$w_i(\mathbf{R}, \mathbf{p}) = |v_i(0)|\delta_i(0)e^{-\beta H(\mathbf{R},\mathbf{p})} \tag{4.4}$$

(i.e., w_i corresponds to a snapshot of the particles crossing the dividing surface), Eq. (3.18) becomes

$$k_{i\to j}^{el} = \frac{2\iint \frac{v_i(0)}{|v_i(0)|}\theta_j(t)\,v_i\,d\mathbf{R}\,d\mathbf{p}\Big/\iint e^{-\beta H}\,d\mathbf{R}\,d\mathbf{p}}{\iint w_i\,d\mathbf{R}\,d\mathbf{p}\Big/\iint e^{-\beta H}\,d\mathbf{R}\,d\mathbf{p}}k_{i\to}^{TST}$$
$$= 2\left\langle\frac{v_i(0)}{|v_i(0)|}\theta_j(t)\right\rangle_{w_i}k_{i\to}^{TST}, \tag{4.5}$$

where the subscript w_i indicates that the ensemble average is over the distribution given by Eq. (4.4). Equation (4.5) gives a simple prescription for evaluating the dynamical correction factors. From a configuration-space Metropolis walk restricted to remain on (or very near) the TST dividing surface for state i, we select a number of configurations. For each of these configurations we assign a velocity to the perpendicular coordinate S, chosen randomly from a Maxwellian-flux distribution [$P(v) \propto |v_i|\exp(-\beta m v_i^2/2)$], and assign a velocity chosen from a Maxwellian distribution [$P(v) \propto \exp(-\beta m v^2/2)$] to each other coordinate. Each of these N configurations becomes a starting point for a classical trajectory, and we assign a phase to each one according to

$$\text{phase } (I) = \frac{v_i(0)}{|v_i(0)|},$$

so that the phase of a trajectory is positive if it is initially exiting state i, and negative if it is initially entering state i. (Note that these N initial conditions could instead be generated using Bennett's method[6] of selecting snapshots

TABLE I. Monte Carlo TST rate constants (total rate for escape in any direction) and diffusion constants for Rh on Rh(100). The rates have been extrapolated to zero slab width and the error estimates are two standard deviations.

T (K)	k^{TST} (s^{-1})	D^{TST} (cm^2 s^{-1})
200	3.32×10^{-13}	$(6.38 \pm 0.48) \times 10^{-29}$
300	1.39×10^{-4}	$(2.68 \pm 0.22) \times 10^{-20}$
500	1.30×10^{3}	$(2.49 \pm 0.19) \times 10^{-11}$
1000	2.10×10^{9}	$(4.03 \pm 0.37) \times 10^{-4}$

om the path of a trajectory that is artificially reflected back and forth across the TST dividing surface.) The N trajectories are propagated until $t = \tau_{corr}$, at which time surface-crossing events should have ceased, leaving each trajectory thermalized in some final state. For every state j ($\neq i$) in the system, the elementary rate constant $k_{i\to j}^{el}$ can be computed from

$$k_{i\to j}^{el} = k_{i\to}^{TST}\left[\frac{2}{N}\sum_{I=1}^{N}\text{phase } (I)\theta_j(I)\right], \tag{4.6}$$

where $\theta_j(I) = 1$ if trajectory I resides in state j at time τ_{corr}, and is zero otherwise.

Two properties of Eq. (4.6) are evident. First, $k_{i\to j}^{el}$ can only be nonzero for those states j in which trajectories terminate—as we would expect. Second, it appears possible to obtain a negative value for $k_{i\to j}^{el}$, which would be meaningless. It is easy to prove, however, that for a perfect sampling of trajectories (i.e., if N is made large enough), only nonnegative rates will be computed. This is because in the true canonical ensemble, for every trajectory with negative phase, there is an associated trajectory with positive phase that will terminate in the same state. This can be verified by inspection of Fig. 4; e.g., trajectory 2 is initially entering state I (and thus contributes negatively to $k_{i\to j}^{el}$), but to terminate in state j, it has to exit state I, thus generating the starting point for trajectory 3, which has positive phase and cancels trajectory 2.

V. RHODIUM SURFACE DIFFUSION

Surface diffusion, the migration of an atom or molecule (or collection of molecules) on a solid surface, is an important part of a variety of physical processes, including crystal growth, defect formation, epitaxial layer growth, and heterogeneous catalysis. Understanding the factors that influence the rate and mechanism of this migration is thus of central importance, and has received considerable attention in recent years.[29]

Use of the field ion microscope[30] (FIM), which is capable of observing a single adatom on a clean crystal face, has yielded high-quality surface diffusion constants for a variety of metal-on-metal systems.[31–38] These diffusion constants generally exhibit Arrhenius behavior, with activation energies and preexponentials that vary widely with the choice of metal and crystal face. These results thus allow a test of our qualitative understanding of the microscopic features of the adatom dynamics, and provide quantitative results against which we can test our theoretical methods.

A number of theoretical studies of single-adatom diffusion have appeared, applying either direct MD,[39–45] or some form of TST[24,32,37,46–49] to the problem. All the calculations employed a finite cluster of atoms, bound together by Lennard-Jones or Morse pairwise potentials, to simulate the solid surface. These simulations have proved quite helpful in understanding qualitative features of the diffusion dynamics. For example, Halichioglu *et al.*[45,46] were able to demonstrate that the cross-channel diffusion of Ir on Pt(110) proceeds by a channel-wall knockout mechanism. However, in making quantitative comparisons to experimental diffusion constants, none of these studies has been able to reduce the possible types of error to a single source.

Assuming classical mechanics is valid for describing the dynamics of a particle as heavy as a transition metal atom, the accuracy of a MD simulation should only be limited by the quality of the potential energy function employed. However, at the temperature used in FIM studies (i.e., 350 K and below), the diffusive hops between surface binding sites represent a rare-event process. Thus, MD simulations have only been feasible at much higher temperatures, so that no direct comparisons could be made between theory and experiment. TST diffusion constants have been calculated in the FIM temperature range, but there was no way to tell whether the approximate dynamics or the approximate potential energy function was responsible for the observed disagreement.

The multistate dynamical corrections method derived in Secs. III and IV allows us to directly address this problem, by computing dynamically exact surface diffusion constants at the same temperature as the FIM experiments. Any discrepancy between theory and experiment can then be attributed to the approximate potential function.

The calculations presented here all employ the Lennard-Jones 6-12 template model described previously.[24,42,43] Briefly, this model consists of one or more layers of movable atoms affixed to a rigid template of atoms that has the geometry appropriate for the desired crystal face. The potential energy at each geometry is computed from a Lennard-Jones 6-12 pairwise potential with spherical cutoff

$$V = \sum_{i>j} v'_{ij}, \tag{5.1}$$

where

$$v'_{ij} = \begin{cases} v_{ij}(r_{ij}) - v_{ij}(r_c) & (0 < r_{ij} \leq r_c) \\ 0 & (r_{ij} > r_c) \end{cases} \tag{5.2}$$

and

$$v_{ij}(r_{ij}) = 4\epsilon_{ij}\left[\left(\frac{\sigma_{ij}}{r_{ij}}\right)^{12} - \left(\frac{\sigma_{ij}}{r_{ij}}\right)^{6}\right]. \tag{5.3}$$

Here r_{ij} is the distance between atoms i and j, ϵ_{ij} and σ_{ij} are the Lennard-Jones well depth and distance parameters for that pair of atoms, and r_c is a cutoff distance. This form for V eliminates the need to calculate interactions between atoms separated by more than r_c, while maintaining a potential that is continuous (though the derivatives are discontinuous at $r_{ij} = r_c$). For the Rh on Rh(100) case considered here, all the interactions are identical ($\sigma_{ij} = \sigma$, $\epsilon_{ij} = \epsilon$), and the parameters were chosen to match previous MD[40–43] and TST[24,49] studies ($\epsilon/k_B = 7830$ K, $\sigma = 2.47$ Å, $r_c = 2.2\sigma$). These values for ϵ and σ were obtained from bulk thermodynamic data.[42] There are 32 atoms per layer in the cluster, with one layer allowed to move and three layers frozen; the total cluster including the adatom thus consists of 129 atoms.

The lattice spacing at all temperatures was chosen to give a nearest-neighbor distance of $2^{1/6}\sigma$ (the minimum in v_{ij}), and periodic boundary conditions were employed.

The TST dividing surface for one binding site is shown in Fig. 6. It is defined by connecting the four atoms around a binding site with planes that extend infinitely in the $+z$ and $-z$ directions. Rather than allowing these planes to move with these four atoms (as in our previous TST studies[24,49]), we simply fixed them to the equilibrium positions of the atoms. This gives a TST surface that depends on the adatom coordinates only, eliminating the need for an effective-mass correction and simplifying the calculation of the trajectory initial velocities. To compute the TST rate constant (Table I), each plane was replaced by a slab with finite thickness b. A Monte Carlo procedure was used to determine the ensemble average in Eq. (4.2), using uniform weighting over the thickness of the slab as a prelimit form for the Dirac delta function. By making independent runs with different values for b, the TST rate was extrapolated to that for a zero-width slab, as shown in Fig. 7. We found that using $b \lesssim 0.1$ bohr was sufficient to produce agreement with the $b = 0$ limit. At these low temperatures, accurate evaluation of the Monte Carlo averages requires the use of importance sampling techniques, due to the rare-event nature of the process. We recently presented one such method,[49] but in the present work we had better success with an alternative approach when the slab width was made very small. This new approach exploits the fact that the desired ensemble average can be written as a ratio of partition functions for two states (A and B),

$$\langle \delta_i(0) \rangle_i = \lim_{b \to 0} \left(\frac{1}{b} \frac{Q_B}{Q_A} \right). \tag{5.4}$$

where state A is the whole binding site bounded by the TST dividing surface (the center of the slab), and state B is the hypothetical state generated by restricting the adatom to be inside the slab. The problem thus becomes one of evaluating the free-energy difference between states A and B, and the method is a variation on an existing

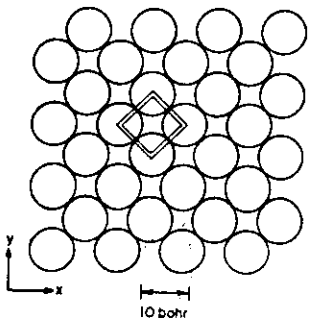


FIG. 6. The LJ template model used for modeling the Rh(100) surface. Only the top layer is shown. For one binding site the TST boundary surface is shown, represented by four finite-width slabs (see the text), which extend infinitely in the $+z$ and $-z$ directions.

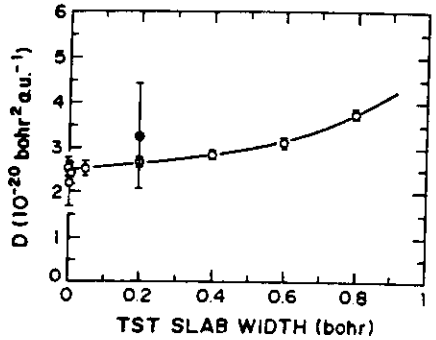


FIG. 7. Extrapolation of the TST diffusion constant at 300 K to zero slab width (see Fig. 6). The error bars are two standard deviations. The shaded point was obtained using TST slabs that moved along with the top-layer atoms (see Ref. 49) and is seen to be in good agreement with the fixed-slab results. The TST rates are shown in Table I.

technique[6] for this type of problem. This method is described in detail elsewhere.[50]

The TST diffusion constant is obtained from

$$D^{\mathrm{TST}} = \frac{l^2}{2d} k_{i\to}^{\mathrm{TST}}, \tag{5.5}$$

where d is the dimensionality of the space ($d = 2$ in this case; $d = 1$ for channeled surfaces) and l is the distance between adjacent binding sites; i.e., the hop length. In addition to the TST assumption that adatoms crossing the dividing surface do not recross, Eq. (5.5) assumes that successive hops are directionally uncorrelated, so that an adatom has equal probability of exiting to each of the (in this case, four) adjacent sites, regardless of the direction from which it entered. Of course, both of these assumptions are approximations, which can be eliminated using dynamical corrections.

Using the procedure described in Sec. IV, classical trajectories were started from the fixed TST surface, allowing calculation of $k_{i\to j}^{\mathrm{TST}}$ from Eq. (4.6) for all possible binding sites (j) to which the adatom can "hop." The integration was performed using a fourth order Runge–Kutta–Gill algorithm, with a time step of 100 atomic units (1 a.u. = 2.418×10^{-17} s) and an energy rescaling every 11 steps. The highest temperature we report here is 1000 K; above this temperature the time scales (τ_{corr} and τ_{rxn}) begin to be inseparable, so that the elementary rate constants cannot be cleanly extracted. Because of the symmetry of the lattice and the TST surface, it is sufficient to start all trajectories in the same direction from only one of the four TST planes shown in Fig. 6; the fate of the set of trajectories in the true canonical ensemble (which enter and exit state i from all four planes) can be determined from the unique set of trajectories using symmetry considerations.

Figure 8 shows the behavior of 200 trajectories run at 1000 K. At short times, all the trajectories reside in the state that they were entering at $t = 0$ (state B). After ~20 000 a.u., dynamical recrossing events begin to occur. While the majority of trajectories are trapped in state B,

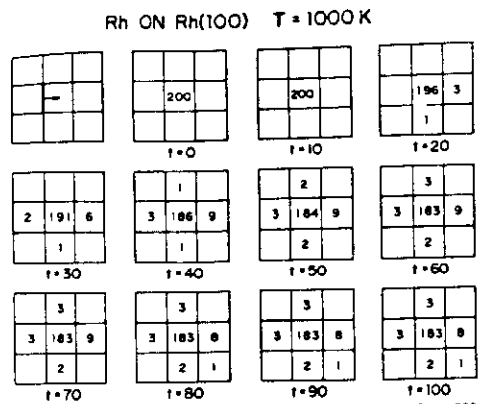


FIG. 8. Time dependence of the state populations resulting from 200 trajectories initiated at the TST plane. All trajectories were initiated from the same TST plane, in the same direction, as indicated by the arrow in the first subfigure. The time units are 10^3 a.u. = 2.42×10^{-14} s. The coordinate system for this representation of the fcc(100) lattice is rotated by 45° compared to Fig. 6.

17 of the trajectories exit in various directions before thermalizing in nearby states. These correlated dynamical events have terminated by 80 000 a.u. (continuing 100 of these trajectories until 200 000 a.u. caused only one additional crossing), so that $\tau_{\mathrm{corr}} = 80\,000$ a.u. Taking $\tau_{\mathrm{rxn}} \approx (k^{\mathrm{TST}})^{-1}$ leads to

$$\frac{\tau_{\mathrm{rxn}}}{\tau_{\mathrm{corr}}} = \frac{2.1 \times 10^8}{80\,000} = 2.6 \times 10^3,$$

indicating that the condition of Eq. (3.19) is satisfied. Applying symmetry to generate all possible starting conditions (trajectories entering and exiting from all four sides) and applying Eq. (4.6) leads to the dynamical correction factors shown in Fig. 9. The time dependence is as discussed in Sec. IV, with the correction factors for nonadjacent states initially zero and then rising to a plateau value. The sum of all the f_d values is 0.90,

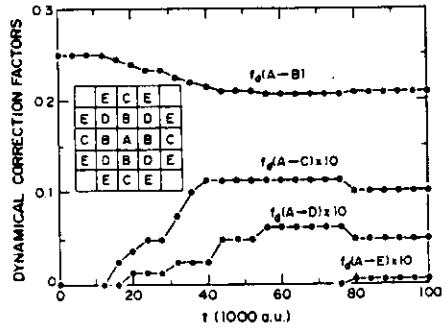


FIG. 9. Dynamical correction factors [Eq. (3.20)] resulting from the 200 trajectories shown in Fig. 8. Symmetry operations were used to construct the additional trajectories present in the total "ensemble," in which trajectories enter and exit from all four directions. (The distribution was also made isotropic by symmetry averaging.)

indicating that the rate at which adatoms escape from binding site is 10% slower than that predicted by TST.

To compute the dynamically corrected diffusi constant, we replace Eq. (5.5) with a sum over all possi binding sites to which the adatom can directly "hop,"

$$D = \frac{1}{2d} \sum_{j \neq i} k_{i\to}^{\mathrm{TST}} l_{ij}^2, \tag{5}$$

where l_{ij} is the distance between states i and j. Note t in the limit where the TST assumptions are valid, (5.6) reduces to Eq. (5.5). Taking the ratio of Eq. (5.6 Eq. (5.5) leads to the dynamical correction factor for diffusion constant,

$$\frac{D}{D^{\mathrm{TST}}} = \sum_{j \neq i} f_d(i \to j) \left(\frac{l_{ij}}{l} \right)^2. \tag{}$$

These are shown for various temperatures, along with $f_d(i \to j)$ values, in Table II. While the average hop is decreased by the dynamical corrections, the diffus constant is *increased*, because the average squared len of a hop is increased. Of course, dynamical correcti in general will not necessarily increase D. For exam if the TST surface were poorly chosen, so that $k_{i\to}^{\mathrm{TST}}$, hence D^{TST}, were artificially high, then the dynam corrections would act to lower D to the classically cor value. Another example is an effect we have observe Rh on Rh(111) diffusion; at certain temperatures thei an enhanced probability that the adatom that is ente a binding site will bounce off the far wall of this bin site and re-exit in the direction from which it came. the Rh(100) surface no such effect was observed.

Inspection of Table II also shows that the dynam effects decrease as the temperature is lowered. Thi because at low temperatures the particle typically has enough energy to pass over the barrier and is ea trapped in the adjacent state, while at higher temperat there is an increased probability that the particle have energy in excess of that required to cross the bar and may thus make another crossing before this en is dissipated. This temperature dependence has b observed previously at higher temperatures.[39-44]

An Arrhenius plot of the diffusion constants is sh in Fig. 10, along with the experimental FIM result, Ayrault and Ehrlich.[35] Fitting to the Arrhenius form

$$D = D_0 e^{-E_A/k_B T} \tag{}$$

leads to $E_A = 23.82 \pm 0.05$ kcal/mol and $D_0 = \pm 0.6) \times 10^{-3}$ cm^2 s^{-1}, compared to the FIM value $E_A = 20.2 \pm 1.7$ kcal/mol and $D_0 = 1 \times 10^{-3}$ cm^2 The theoretical value for D at 300 K is smaller than experimental values by a factor of 60. As discussed ab this discrepancy can, for the first time, be attribute the approximate potential function. This LJ poten which was fit to Rh bulk thermodynamic data, is necessarily suited to describing atoms on the surface two reasons, we make no attempt here to modify potential to achieve better agreement with experin First, the purpose of the present study is to demons the feasibility of obtaining exact dynamical results i rare-event regime, not to match experiment. Second

TABLE II. Dynamical correction factors for Rh on Rh(100). The state designations (i and j) used in $f_d(i \to j)$ are defined in Fig. 9. N_{traj} is the number of trajectories used to determine the dynamical corrections; N_{corr} is the number of these that come to rest in a state other than B. Note that $f_d(A \to B) = 0.25$ corresponds to TST.

T (K)	N_{traj}	N_{corr}	$f_d(A \to B)$	$f_d(A \to C)$	$f_d(A \to D)$	$f_d(A \to E)$	D/D^{TST}
200	100	0	0.25	0.0	0.0	0.0	1.00
300	100	0	0.25	0.0	0.0	0.0	1.00
500	150	1	0.2467	0.0017	0.0	0.0	1.013
1000	200	17	0.2088	0.0100	0.0050	0.000 62	1.060

of a number of arbitrary modifications could be made to the potential to increase the agreement with experiment, but we feel this would have limited meaning. It is possible, however, that calculations like those presented here could be used in conjunction with a variety of other data to generate more accurate potential functions.

Probably the most important result of this Rh on Rh(100) study is that the dynamical corrections are essentially negligible in the temperature range of the FIM experiments (~300 to 330 K). (Even at 1000 K, D/D^{TST} is only 1.06.) While it would be premature to assume this is true of all surface dynamics, it is consistent with other results we have obtained for single metal atoms on metal surfaces. Thus, it appears that future studies of this type can benefit from the significant computational savings of using TST, assuming that a good TST dividing surface can be constructed. For those systems in which there is doubt about the accuracy of TST, the method presented here may be used to obtain the exact dynamics many orders of magnitude more quickly than using full MD simulations.

VI. CONCLUSIONS

We have extended the two-state dynamical corrections formalism to treat the general many-state case. For processes characterized by rare-event dynamics, the method allows the computation of (classical) dynamically exact rate constants between any two states of a system, a task which is unfeasible using direct molecular dynamics. The key to the method is embodied in Eq. (3.12), which expresses the arbitrary elementary rate constant ($k^{el}_{i\to j}$) as an equilibrium time-correlation function. Equation (3.18) shows that $k^{el}_{i\to j}$ is simply the TST rate of escape from state i multiplied by a dynamical correction factor, and Eq. (4.6) indicates how this correction factor can be evaluated using molecular dynamics techniques. Dynamical properties of the system may be obtained by following the evolution of a pseudodynamical simulation, in which the system executes a biased random walk from state to state, with weighting governed by the precalculated elementary rate constants. By combining this approach at low temperatures with direct molecular dynamics at high temperatures, dynamical properties may be computed at any temperature. (The middle range of temperatures, where reactive transitions are slightly too frequent to be classified as rare events, then becomes the most computationally intensive.)

An interesting result that emerges from this work is the definition of a rate constant between states which are not adjacent in configuration space. The concept of a "direct" transition between nonadjacent states is intuitively meaningful, and in the rare-event regime, Eq. (3.12) gives a precise definition for this rate.

For the Rh on Rh(100) system, we have calculated the first dynamically exact single-atom diffusion constants in the temperature range of the FIM experiments. The discrepancy between theory and experiment (which is surprisingly small, considering the simplicity of the LJ potential) can now be attributed to the approximate potential function. At 1000 K and below, the dynamical corrections were found to be negligible, indicating that TST is a very good approximation for this type of system. Moreover, we have shown recently[49] that simple harmonic TST, which requires energies at only a few geometries, is a good approximation to exact TST at these low temperatures.[51] Thus, it should be feasible to calculate accurate diffusion constants using *ab initio* electronic structure methods, once relative energies can be obtained with better than 1 kcal accuracy.

Finally, we note that although the Rh diffusion example presented here exhibited only minor dynamical corrections, the method will work equally well if the corrections are large. The method can be applied to any multistate system, as long as molecular dynamics can be run on the potential energy function which is employed. Examples of systems which would benefit from this treatment include the diffusion of atom clusters on a metal surface, the diffusion of vacancies or impurities through a solid, and the dynamical evolution of a polymer chain.

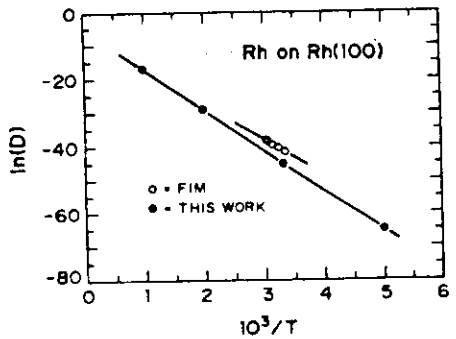


FIG. 10. Surface self-diffusion constants for Rh on Rh(100), plotted in Arrhenius form. The units are D (cm^2/s) and T (K). The error bars (two standard deviations) are not shown since they are smaller than the plotted points. The experimental results are from Ref. 35.

ACKNOWLEDGMENTS

We wish to thank David Chandler for helpful discussions regarding this work, and J. Winterkamp of Los Alamos National Laboratory's Group ESS-5 for computational support.

APPENDIX

We wish to show that the inverse of the time-zero fluctuation-fluctuation autocorrelation function is indeed given by Eq. (3.11). We take $N = 1$, so that Eq. (3.13) holds, and note that

$$\langle \theta_i^2(0) \rangle = \langle \theta_i(0) \rangle = \chi_i \tag{A1}$$

and

$$\langle \theta_i(0)\theta_j(0) \rangle = \delta_{i,j}\chi_i. \tag{A2}$$

For an arbitrary element of the matrix we have

$$\begin{aligned}\langle \delta N_i(0)\delta N_j(0) \rangle &= \langle [\theta_i(0) - \chi_i][\theta_j(0) - \chi_j] \rangle \\ &= \langle \theta_i(0)\theta_j(0) \rangle - \chi_i\langle \theta_j(0) \rangle - \chi_j\langle \theta_i(0) \rangle + \chi_i\chi_j \\ &= \delta_{i,j}\chi_i - \chi_i\chi_j,\end{aligned} \tag{A3}$$

as asserted by Eq. (3.10). If the right-hand side of Eq. (3.11) is the inverse of the matrix defined by Eq. (A3), it will satisfy the condition for the inverse of a matrix A:

$$\sum_k (A^{-1})_{ik}(A)_{kj} = \delta_{i,j}. \tag{A4}$$

Inserting Eqs. (A3) and (3.11) into the left-hand side of Eq. (A4) yields

$$\begin{aligned}\sum_k^{n-1} (\delta_{ik}\chi_i - \chi_i\chi_k)\left(\frac{1}{\chi_n} + \delta_{kj}\frac{1}{\chi_k}\right) &= \frac{\chi_i}{\chi_n} + \delta_{ij}\frac{\chi_j}{\chi_j} - \sum_k^{n-1}\frac{\chi_i\chi_k}{\chi_n} - \frac{\chi_i\chi_j}{\chi_j} \\ &= \delta_{ij} + \frac{\chi_i}{\chi_n}\left(1 - \sum_k^{n-1}\chi_k\right) - \chi_i \\ &= \delta_{ij},\end{aligned} \tag{A5}$$

proving Eq. (3.11). The last step in Eq. (A5) made use of

$$\sum_k^n \chi_k = 1.$$

[1] For a recent review, see J. Kushick and B. J. Berne, in *Modern Theoretical Chemistry*, edited by B. J. Berne (Plenum, New York, 1977), Vol. 6.
[2] J. C. Keck, Discuss. Faraday Soc. 33, 173 (1962); Adv. Chem. Phys. 13, 85 (1967).
[3] B. Woznick, AVCO-Everett Research Laboratory Report No. RR223, Everett, Massachusetts, 1965.
[4] R. L. Jaffe and J. B. Anderson, J. Chem. Phys. 54, 2224 (1971); R. L. Jaffe, J. M. Henry, and J. B. Anderson, *ibid* 59, 1128 (1973); J. Am. Chem. Soc. 98, 1140 (1976), and references therein.
[5] J. B. Anderson, J. Chem. Phys. 58, 4684 (1973).
[6] C. H. Bennett, in *Diffusion in Solids: Recent Developments*, edited by J. J. Burton and A. S. Nowick (Academic, New York, 1975), p. 73.
[7] C. H. Bennett, in *Algorithms for Chemical Computation*, edited by R. E. Christofferson (American Chemical Society, Washington, D.C., 1977).
[8] D. Chandler, J. Chem. Phys. 68, 2959 (1978); J. A. Montgomery, Jr., D. Chandler, and B. J. Berne, *ibid* 70, 4056 (1979).
[9] J. L. Skinner and P. G. Wolynes, J. Chem. Phys. 69, 2143 (1978); 72, 4913 (1980).
[10] See, for example, D. G. Truhlar and B. C. Garrett, Acc. Chem. Res. 13, 440 (1980), and references therein.
[11] J. E. Adams and J. D. Doll, J. Chem. Phys. 74, 1467 (1981).
[12] For a recent review, see P. Pechukas, Annu. Rev. Phys. Chem. 32, 159 (1981).
[13] A. Marcelin, Ann. Phys. (Paris) 3, 120, 185 (1915).
[14] R. L. Jaffe, Chem. Phys. 40, 185 (1979).
[15] R. N. Porter, D. L. Thompson, L. M. Raff, and J. M. White, J. Chem. Phys. 62, 2429 (1975).
[16] J. A. Montgomery, Jr., S. L. Holmgren, and D. Chandler, J. Chem. Phys. 73, 3688 (1980).
[17] E. K. Grimmelmann, J. C. Tully, and E. Helfand, J. Chem. Phys. 74, 5300 (1981); J. C. Tully, Surf. Sci. 111, 461 (1981).
[18] J. E. Adams and J. D. Doll, Surf. Sci. 103, 472 (1981); J. Chem. Phys. 80, 1681 (1984).
[19] J. E. Adams and J. D. Doll, Surf. Sci 111, 492 (1981).
[20] This estimate assumes the use of the fastest currently available computer.
[21] B. J. Berne and G. D. Harp, Adv. Chem. Phys. 17, 63 (1970).
[22] J. D. Doll, J. Chem. Phys. 74, 1074 (1981).
[23] Of course, Eq. (2.16) does depend on state B in the sense that it depends on the location of the boundary line ($x = q$) between state A and state B.
[24] A. F. Voter and J. D. Doll, J. Chem. Phys. 80, 5832 (1984).
[25] Harold S. Johnston, *Gas Phase Reaction Rate Theory* (Ronald, New York, 1966), p. 48.
[26] N. Metropolis, A. Rosenbluth, M. N. Rosenbluth, A. Teller, and E. Teller, J. Chem. Phys. 21, 1087 (1953).
[27] J. P. Valleau and S. G. Whittington, in *Modern Theoretical Chemistry*, edited by B. J. Berne (Plenum, New York, 1977), Vol. 5.
[28] For applications of Monte Carlo to calculating TST rates, see Refs. 11, 24, and 29.
[29] For a recent review, see G. Ehrlich and K. Stolt, Annu. Rev. Phys. Chem. 31, 603 (1980).
[30] E. W. Muller and T. T. Tsong, *Field Ion Microscopy: Principles and Applications* (Elsevier, New York, 1969).
[31] G. Ehrlich and F. G. Hudda, J. Chem. Phys. 44, 1039 (1966).
[32] G. Ehrlich and C. F. Kirk, J. Chem. Phys. 48, 1465 (1968).
[33] D. W. Bassett and M. J. Parsley, J. Phys. D 3, 707 (1970).
[34] W. R. Graham and G. Ehrlich, Phys. Rev. Lett. 31, 1407 (1973); Thin Solid Films 25, 85 (1975).
[35] G. Ayrault and G. Ehrlich, J. Chem. Phys. 60, 281 (1974).
[36] P. Cowan and T. T. Tsong, Phys. Lett. A 53, 383 (1975); 22nd International Field Emission Symposium, Atlanta, Georgia, 1975.
[37] D. W. Basset and P. R. Webber, Surf. Sci. 70, 520 (1978).
[38] P. G. Flahive and W. R. Graham, Surf. Sci. 91, 463 (1980).
[39] M. R. Mruzik and G. M. Pound, J. Phys. F 11, 1403 (1981).
[40] H. K. McDowell and J. D. Doll, Surf. Sci. 121, L537 (1982).
[41] J. D. Doll and H. K. McDowell, Surf. Sci. 123, 99 (1982).
[42] J. D. Doll and H. K. McDowell, J. Chem. Phys. 77, 479 (1982).
[43] H. K. McDowell and J. D. Doll, J. Chem. Phys. 78, 3219 (1983).
[44] J. C. Tully, G. H. Gilmer, and M. Shugard, J. Chem. Phys. 71, 1630 (1979).
[45] S. H. Garofalini and T. Halicioglu, Surf. Sci. 104, 199 (1981); 112, L775 (1981); also see H. P. Bonzel, *ibid.* 121, L531; S. H. Garofalini and T. Halicioglu, *ibid* 121, L535 (1982).
[46] T. Halicioglu and G. M. Pound, Thin Solid Films 57, 241 (1979).
[47] P. Wynblatt and N. A. Gjostein, Surf. Sci. 22, 125 (1970).
[48] J. R. Banavar, M. H. Cohen, and R. Gomer, Surf. Sci. 107, 113 (1981).
[49] A. F. Voter and J. D. Doll, J. Chem. Phys. 80, 5814 (1984).
[50] A. F. Voter and J. D. Doll, J. Chem. Phys. (in press).
[51] The agreement between simple TST and exact TST should be even better than that shown in Ref. 49 if all force constants are evaluated at the minimum and the saddle point. This can be accomplished with current *ab initio* techniques.

Computing Classically Exact Diffusion Constants Using Short-Time Trajectories

Arthur F. Voter

Theoretical Division (T-12, MS J369), Los Alamos National Laboratory, Los Alamos, New Mexico 87545

(Received 17 August 1988)

The classical diffusion constant of a point defect in an infinite lattice of binding sites is shown to be expressible as transition-state-theory rates multiplied by dynamical-correction factors computed from short-time classical trajectories initiated at the site boundaries. The expression, which results from time differentiating the lattice-discretized mean square displacement, is valid at any temperature for which the site lattice is well defined. It thus avoids both the time-scale limitations of direct molecular dynamics and the rare-event requirements of standard dynamical-corrections methods.

PACS numbers: 66.30.Lw, 66.30.Jt, 68.35.Fx

The utility of molecular-dynamics (MD) simulations in computing surface and bulk diffusion constants is well established.[1,2] When the time scale for diffusive jumps is less than $\sim 10^{-10}$ s, i.e., when the temperature is high relative to the diffusion barrier, direct MD simulation can be employed, in which the system is evolved classically and the diffusion constant (D) is computed from the mean square displacement or the velocity autocorrelation function of the diffusing particle. As the temperature is lowered, the average time between diffusive jumps gets longer, making direct MD impractical, and finally, impossible.

For temperatures low enough that diffusion is truly a rare-event process, an elegant alternative to direct MD can be employed, in which the exact dynamical rate constant between reactant and product (in this case, the binding sites occupied before and after a diffusive jump) is expressed as the transition-state-theory[3] (TST) rate multiplied by a dynamical-correction factor.[2,4-6] The TST rate constant, which is the outgoing flux through the boundary surface dividing the reactant and product, is an equilibrium property of the system and is thus easily evaluated. The dynamical-correction factor accounts for TST-violating dynamical events (i.e., correlated recrossings of the boundary) that occur during the relatively short time that the system is activated. This dynamical-correction factor can be evaluated from classical trajectories initiated at the TST boundary, as first demonstrated by Keck[4] for gas-phase systems. The first application of this approach to the condensed phase was Bennett's[2] investigation of bulk diffusion. Alternatively, the correction factor can be obtained from a detailed examination of the phase-space properties of the TST saddle surface.[7-9]

Because direct MD is practical only at high temperatures and the dynamical-corrections formalism is rigorous only at low temperatures,[2,5] computing D at intermediate temperatures remains a problem. For example, for adatom self-diffusion on the Lennard-Jones fcc (111) surface, direct MD methods can be applied down to reduced temperatures of $T^* \cong 0.11$,[10,11] while the rare-event assumption breaks down above $T^* \cong 0.04$.[12] In this Letter, a new method is presented that bridges this gap. By taking the time derivative of the mean square displacement over a site-discretized coordinate system, a new expression for D is derived that is valid at any temperature for which the lattice binding sites is well defined. As in standard dynamical-corrections methods, the expression is evaluated using TST and short-time trajectories initiated at the TST boundaries; however, no rare-event assumptions are required. It is shown elsewhere[13] that this method can also be derived starting from the many-state dynamical-corrections formalism,[6] by expressing the squared hop length as a rate-constant-weighted average over all states to which the diffusing defect can jump. Despite the fact that the individual rate constants become meaningless at high temperature, the diffusion constant is valid.

In the case of the fcc-(111)-surface diffusion mentioned above, the dynamical-correction factor, D/D^{TST}, exhibits an interesting dip below unity at low temperature.[12,14] Use of the method presented here allowed a determination of the position and depth of the minimum, which occurs at $T^* = 0.04$ (the edge of the rare-event regime), and the cause has been identified.[12] This previously unobserved dip is due to trajectories in which the adatom undergoes a well-aimed, direction-reversing collision with the substrate atom on the far side of the triangular binding site, causing a recrossing of the TST boundary. At lower temperatures, the adatom is less likely to recross because it has less kinetic energy, while at higher temperatures the effect is masked by forward multiple jumps which increase D/D^{TST}.

In the following derivation, the system is assumed to consist of a particle diffusion in an infinite lattice of binding sites (e.g., a crystal impurity atom in the lattice of all possible interstitial sites). A discretized coordinate for the diffusing particle is defined by

$$\mathbf{R}^d(t) = \sum_i \theta_i(t)\mathbf{R}_i , \tag{1}$$

where

$$\theta_i(t) = \theta[F_i[\mathbf{R}(t)]] . \tag{2}$$

Here $\theta\{\cdots\}$ is the standard step function, $\mathbf{R}(t)$ is the configuration-space coordinate of the particle at time t, $\mathbf{R}_i$ is the average position for the particle in state i, and $F_i[\mathbf{R}]$ is a continuous, differentiable function with the property that

$$F_i[\mathbf{R}] \begin{cases} > 0, & \text{if } \mathbf{R} \text{ is in state } i, \\ = 0, & \text{if } \mathbf{R} \text{ is on the boundary to state } i, \\ < 0, & \text{if } \mathbf{R} \text{ is outside of state } i. \end{cases} \tag{3}$$

Thus, $\mathbf{R}^d(t)$ takes on the value of the average position of the state in which the particle resides at time t. Note that this discretized coordinate is only meaningful if the binding sites are defined such that they tile all of configuration space. The time derivative of this discretized coordinate is zero except when the particle is at the boundary between two states. We obtain

$$\frac{d\mathbf{R}^d(t)}{dt} = -\sum_i \delta_i(t) v_i(t) \mathbf{R}_i , \tag{4}$$

where $v_i(t)$ is the velocity component normal to the boundary at $\mathbf{R}(t)$,

$$v_i = -\frac{\nabla F_i \cdot \mathbf{v}}{|\nabla F_i|} \tag{5}$$

(v_i is positive if the particle is *exiting* state i), and $\delta_i(t)$ is nonzero only at the boundary to state i,

$$\delta_i(t) = \delta[F_i[\mathbf{R}(t)]]\,|\nabla F_i[\mathbf{R}(t)]| . \tag{6}$$

($\delta\{\cdots\}$ is the Dirac delta function.) This convention matches that used previously.[6]

The diffusion constant is given by

$$D = \frac{1}{2a} \lim_{t\to\infty} \left\{ \frac{d}{dt} \langle |\mathbf{R}(t) - \mathbf{R}(0)|^2 \rangle \right\} , \tag{7}$$

which, for diffusion over the lattice binding sites, can be written as

$$D = -\frac{1}{2a} \frac{d}{dt_0} \langle |\mathbf{R}^d(t) - \mathbf{R}^d(t_0)|^2 \rangle , \tag{8}$$

where a is the dimensionality of the space and $t - t_0$ is assumed to be longer than the memory time of the diffusion particle. The brackets $\langle \cdots \rangle$ indicate the usual canonical ensemble average,

$$\langle P \rangle = \frac{\int P e^{-\beta H} d\mathbf{R}\, d\mathbf{p}}{\int e^{-\beta H} d\mathbf{R}\, d\mathbf{p}} , \tag{9}$$

where P is some property, $\mathbf{p}$ is the particle momentum, $H = H(\mathbf{R},\mathbf{p})$ is the classical Hamiltonian, and $\beta = 1/k_B T$, with T the temperature and k_B the Boltzmann constant. An average over a weighting factor (w) other than $e^{-\beta H}$ will be indicated by a right-hand subscript. An average restricted to the configuration space of state i will be written $\langle P \rangle_i$, where

$$w_i = \theta_i(0) e^{-\beta H} . \tag{10}$$

Carrying out the time derivative in Eq. (8) over the discretized coordinates (setting $t_0 = 0$) leads to

$$\begin{aligned} D &= \frac{1}{2a} \left\langle 2 \left[\frac{d\mathbf{R}^d(t)}{dt} \right]_{t=0} \cdot [\mathbf{R}^d(t) - \mathbf{R}^d(0)] \right\rangle \\ &= \frac{1}{a} \sum_i \chi_i \left\langle \left[\frac{d\mathbf{R}^d(t)}{dt} \right]_{t=0} \cdot [\mathbf{R}^d(t) - \mathbf{R}^d(0)] \right\rangle_i \\ &= \frac{1}{a} \sum_i \chi_i \left\langle -\sum_j \delta_j(0) v_j(0) \mathbf{R}_j \cdot [\mathbf{R}^d(t) - \overline{\mathbf{R}}_{ij}] \right\rangle_i . \end{aligned} \tag{11}$$

Here χ_i is the mole fraction for state i,

$$\chi_i = \langle \theta_i(t) \rangle , \tag{12}$$

and

$$\mathbf{R}^d(0) = \overline{\mathbf{R}}_{ij} = \tfrac{1}{2}(\mathbf{R}_i + \mathbf{R}_j) \tag{13}$$

as a result of the definition of the standard step function, $\theta(0) = \frac{1}{2}$. Because the ensemble average is restricted to state i, the sum over j will collapse to those states having a common border with state i. Defining

$$\theta_j^{\pm}(0) = \theta_j(0+) + \theta_j(0-) , \tag{14}$$

which picks out those trajectories that are either entering or exiting state j at $t = 0$, the terms in j become

$$-\delta_j(0) v_j(0) \mathbf{R}_j = \begin{cases} -\theta_j^{\pm}(0)\delta_i(0) v_i(0) \mathbf{R}_i, & \text{if } j = i, \\ +\theta_j^{\pm}(0)\delta_i(0) v_i(0) \mathbf{R}_j, & \text{if } j \text{ borders } i, \\ 0, & \text{otherwise,} \end{cases} \tag{15}$$

where we have made use of $v_i(0) = -v_j(0)$ for a trajectory at the i-j border. Equation (11) thus becomes

$$D = \frac{1}{a} \sum_i \chi_i \sum_j \langle \theta_j^{\pm}(0)\delta_i(0) v_i(0) [\mathbf{R}_j - \mathbf{R}_i] \cdot [\mathbf{R}^d(t) - \overline{\mathbf{R}}_{ij}] \rangle_i . \tag{16}$$

Equation (16) expresses the diffusion constant in terms of trajectories that are at the TST boundaries at $t = 0$, as shown in Fig. 1. It can be put into a form that makes the computational solution more apparent. The TST rate of escape from state i is given by[15,16]

$$k_i^{\mathrm{TST}} = \langle |v_i(0)|\,\delta_i(0) \rangle_i . \tag{17}$$

fining the weighting function for a Maxwellian flux of particles through the TST border to state i,

$$w_{MF_i} = \theta_i(0)\delta_i(0)|v_i(0)|e^{-\beta H}, \tag{18}$$

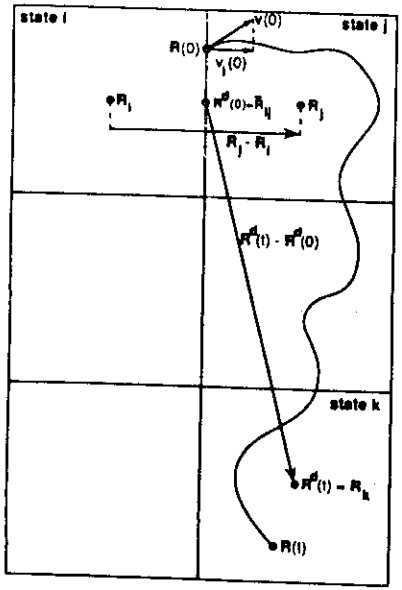
ST can be written as

$$k_{i\rightarrow}^{TST} = \frac{\int w_{MF_i} d\mathbf{R}\, d\mathbf{p}}{\int w_i\, d\mathbf{R}\, d\mathbf{p}}, \tag{19}$$

hat Eq. (16) becomes

$$D = \frac{1}{a}\sum_i \chi_i k_{i\rightarrow}^{TST} \sum_j \left\langle \theta_j^{\pm}(0) \frac{v_i(0)}{|v_i(0)|} [\mathbf{R}_j - \mathbf{R}_i]\cdot[\mathbf{R}^d(t) - \bar{\mathbf{R}}_{ij}] \right\rangle_{MF_i}. \tag{20}$$

Equation (20) is the central result of this paper. It es a prescription for computing D using TST along h trajectories initiated at the TST boundaries. These jectories are integrated long enough that the memory e of the diffusing particle is exceeded. The method is id at any temperature for which the lattice of binding s is well defined. [In spite of the appearance of the T rate constant in Eq. (20), the expression is valid n at temperatures so high that the diffusive jumps do correspond to a first-order rate process. This can be ified by inspection of Eq. (16), in which no rate con- nts appear.] It thus provides a unified computational approach, bridging the gap between the low-temperature, rare-event regime, where existing dynamical-corrections methods are valid, and the high-temperature regime where direct MD can be employed.



IG. 1. Schematic illustration of the factors in Eqs. (16) (20). The curved line represents a trajectory that is pass- etween states i and j at time $t=0$, and resides in state k at t.

Evaluation of Eq. (20) at time $t=0+$ leads to the TST approximation to D, which assumes that every border crossing corresponds to a diffusive jump to the adjacent binding site and all crossings are uncorrelated. Integrating the trajectories out to longer times corrects for these assumptions, accounting for both multiple recrossings and correlated multiple jumps. The quantity $D(t)/D(0+)$ thus represents the dynamical-correction factor for diffusion, D/D^{TST}. That $D(t=0+)$ corresponds to D^{TST} is most easily demonstrated by considering a system with only one unique binding site (thus collapsing the sum over i—see below) and one unique type of jump, such as the square lattice in Fig. 1. For this case, evaluation of Eq. (20) at $t=0+$ yields the well-known expression

$$D(t=0+) = D^{TST} = \frac{1}{2a} k^{TST} l^2, \tag{21}$$

where $l = |\mathbf{R}_j - \mathbf{R}_i|$ (i and j adjacent) is the length of a single jump. The computational procedure is now discussed.

By symmetry, the sum over i needs to include only the representative set of states that can be replicated to generate the complete infinite system. The integration over all space, e.g., in the definition of χ_i in Eq. (12), should be similarly restricted to this subspace of representative states. Once the TST boundaries have been specified, the evaluation of Eq. (20) involves computing $k_{i\rightarrow}^{TST}$ for each unique state, e.g., using a configuration-space Monte Carlo method,[16] and then integrating MD trajectories that are initiated at each unique TST boundary. The trajectory initial conditions can be obtained by sampling from a configuration-space Metropolis walk[17] restricted to the TST boundary. For the diffusing particle, the initial momentum along the direction normal to the TST boundary is chosen randomly from a Maxwellian flux distribution, as dictated by the weighting function in Eq. (18). The moments for the other coordinates in the system, including the two perpendicular coordinates of the diffusing particle, are chosen from a standard Maxwellian distribution. Alternatively, the "TST-flux" trajectory intitial conditions can be generated using Bennett's method of sampling from a trajectory that is artificially reflected back and forth across the TST boundary,[2] or using the approach of Gillan, Harding, and Tarento[18] in which a trajectory is constrained to lie exactly on the TST surface. These trajectories are then integrated long enough that the memory of the initial conditions are lost, after which time Eq. (20) maintains a plateau value with statistical fluctuations, as shown in Refs. 1 and 13. The classically exact diffusion constant, with associated error bars, is obtained by averaging over these fluctuations in $D(t)$.

The author thanks J. M. Cohen and B. L. Holian for critically reviewing this manuscript.

[1]J. D. Doll and A. F. Voter, Annu. Rev. Phys. Chem. 38, 413 (1987).

[2]C. H. Bennett, in *Diffusion in Solids: Recent Developments*, edited by A. S. Nowick and J. J. Burton (Academic, New York, 1975), p. 73; in *Algorithms for Chemical Computation*, edited by R. E. Christofferson (American Chemical Society, Washington, DC, 1977), p. 63.

[3]S. Glasstone, K. J. Laidler, and H. Eyring, *The Theory of Rate Processes* (McGraw-Hill, New York, 1941).

[4]J. C. Keck, Discuss. Faraday Soc. 33, 173 (1962); Adv. Chem. Phys. 13, 85 (1967).

[5]D. Chandler, J. Chem. Phys. 68, 2959 (1978); J. A. Montgomery, Jr., D. Chandler, and B. J. Berne, *ibid.* 70, 4056 (1979).

[6]A. F. Voter and J. D. Doll, J. Chem. Phys. 82, 80 (1985).

[7]M. Toller, G. Jacucci, G. DeLorenzi, and C. P. Flynn, Phys. Rev. B 32, 2082 (1985).

[8]M. Marchese, G. Jacucci, and C. P. Flynn, Phys. Rev. B 36, 9469 (1987).

[9]G. Jacucci, G. DeLorenzi, M. Marchese, C. P. Flynn, and M. Toller, Phys. Rev. B 36, 3086 (1987).

[10]M. R. Mruzik and G. M. Pound, J. Phys. F 11, 1403 (1981).

[11]J. D. Doll and H. K. McDowell, J. Chem. Phys. 77, 479 (1982).

[12]J. M. Cohen and A. F. Voter (to be published).

[13]A. F. Voter, J. D. Doll, and J. M. Cohen, J. Chem. Phys. 90, 2045 (1989).

[14]A. F. Voter and J. M. Cohen, J. Vac. Sci. Technol. A 7, 1859 (1989).

[15]Equation (17) differs from expressions given in Ref. 6 and by A. F. Voter [J. Chem. Phys. 82, 1890 (1985)] by a factor of 2 because, in the present work, $\theta_i(0)$ in the weighting function w_i [Eq. (10)] cuts $\delta_i(0)$ in half.

[16]Voter, Ref. 15.

[17]N. Metropolis, A. Rosenbluth, M. N. Rosenbluth, A. Teller, and E. Teller, J. Chem. Phys. 21, 1087 (1953).

[18]M. J. Gillan, J. H. Harding, and R.-J. Tarento, J. Phys. C 20, 2331 (1987).

Reprinted from SPIE Vol. 821—Modeling of Optical Thin Films

Simulation of the Layer-Growth Dynamics in Silver Films: Dynamics of Adatom and Vacancy Clusters on Ag(100)

Arthur F. Voter

Theoretical Division, MS J569, Los Alamos National Laboratory
Los Alamos, New Mexico 87545

ABSTRACT

Simulation results are presented for some dynamical processes occurring in the growth of (100) layers of silver. The overlayer dynamics are evolved using a recently developed method that, in the regime where surface diffusion consists of discrete hops, yields classically exact dynamics for an arbitrary interatomic potential. The time-scale limitations of direct molecular dynamics simulations are thus overcome. The Ag/Ag(100) system is modeled using a sophisticated form of interaction potential, similar to the embedded atom method, in which the energy is given by a sum of pairwise interactions plus a term for each atom that depends on the local atomic density. This type of potential includes the many-body terms necessary to describe a variety of atomic environments, such as the perfect fcc metal, free surfaces, vacancies, interstitials, and even the diatomic molecule, but with the computational scaling of a simple pair potential. The present study focuses on some of the dynamics in a single layer of silver: the diffusion and dissociation of clusters of adatoms and vacancies. Some interesting features are observed, including a nonmonotonic decrease in diffusion constant with increasing cluster size, and a roughly constant mean square distance a cluster migrates before dissociation (ejection of a monomer).

1. INTRODUCTION

The growth of crystals and thin films, because of their widespread technological importance,[1,2] has received considerable theoretical attention,[3] often in the form of atomistic computer simulations.[3-8] To make predictions about film growth, these simulations must model large systems over long time scales. To facilitate this, simplifying assumptions are made about the microscopic details of the growth process. Because some macroscopic properties of a system are relatively insensitive to the details of the underlying microscopic force laws and dynamics, a model system that is properly chosen and properly applied can yield quite useful results. For example, the simple solid on solid model[3] correctly predicts the roughening transition and other features of a growing crystal, even though the crystal packing is nonphysical, surface diffusion is omitted, and the interatomic forces are reduced to a single parameter. In more sophisticated simulations, commonly employed simplifications include restricting the film to two dimensions and the use of hard spheres or hard spheres with additive interactions.

As an alternative to the model system approach, one can attempt to simulate thin film growth while maintaining a connection to the true microscopic force laws. In other words, for an arbitrary interatomic potential (which can, in principle, be chosen to be fairly realistic for the system of interest) one follows the growth process using classical dynamics.

The most direct way to accomplish this is by integrating a classical trajectory involving a few hundred or thousand atoms representing the substrate plus thin film, while periodically introducing new atoms to the system from the gas phase.[9] Since the system's evolution is governed by the proper force laws, the resulting features (columnar growth, formation of voids, etc.) are physically correct for the chosen interatomic potential. However, this direct molecular dynamics (MD) approach is currently limited to simulation of a few thousand atoms for approximately a nanosecond or less, due to the present speed of computers. With these restrictions, the most meaningful results are obtained by simulating close-packed surfaces (e.g., fcc(111), which have lower activation barriers for diffusion, at relatively high temperatures.

We have recently developed a different approach to this type of problem.[10] With certain restrictions, the method provides classically exact surface dynamics over long time scales. The method is designed to treat the overlayer dynamics for a submonolayer coverage of adsorbates on a perfect crystal face. As such, it is well suited for modeling Frank-van der Merwe (layer-by-layer) film growth. Ideally, one would like to treat film growth in the general case, with adsorption, desorption, three-dimensional clusters, voids, etc. In principle, this method can be extended to handle such processes (as discussed below), but it is currently in its infant stages. At the present time, this approach offers access to dynamical properties that cannot be calculated by direct MD, and, as such, the two methods are complimentary. Even though neither approach is capable of a full-scale simulation of general film growth, the results can be used to improve the parameterization of more macroscopic models, or to test assumptions in these models.

The study presented here is meant to be a demonstration of the feasibility of this new method for calculating dynamical properties important to film growth using a realistic interatomic potential. In particular, we examine the diffusion and dissociation of two-dimensional clusters of silver atoms on the Ag(100) surface. The silver interactions are provided by a state-of-the-art embedded-atom type of potential, described in Section 2. To the extent that this potential is realistic, the calculated dynamical properties should be appropriate for a (100) layer of a silver film that is thick enough to be unperturbed by the underlying substrate.

2. INTERATOMIC POTENTIAL FOR SILVER

Because the results of a molecular dynamics or Monte Carlo simulations are only as accurate as the underlying interatomic potential, there is a motivation to develop high-quality potentials. Daw and Baskes[11] have recently presented a new form of interatomic potential, known as the embedded atom method, in which the total energy of a homonuclear system of atoms is written as

$$E = \frac{1}{2}\sum_{i\neq j}\phi(r_{ij}) + \sum_{i} F[\rho_i] \quad , \tag{2.1}$$

where

$$\rho_i = \sum_{j(\neq i)} \rho(r_{ij}) \quad . \tag{2.2}$$

Here, r_{ij} is the scalar distance between atoms i and j, $\phi(r)$ is the pairwise interaction term, $\rho(r)$ is a function that is roughly the electron density at a distance r from an atom, and $F[\rho]$ is the "embedding function." This type of potential is discussed in detail elsewhere[11-13] and has been quite successful in providing a reasonably quantitative description of metallic systems in a variety of simulations.[14] We briefly discuss here some of the key features of this form of potential, and how we have chosen to parameterize $\phi(r)$, $\rho(r)$ and $F(\rho)$ for Ag.

If Eq. (2.1) is truncated after the first sum, a simple pair potential results. Such a potential has the virtue of computational simplicity and ease of parameterization, but suffers from some defects. For example, the vacancy formation energy (ΔE_{vac}) is equal to the cohesive energy (E_{coh}), and the elastic constants c_{12} and c_{44} are also equal, while neither of these equalities holds for real metal crystals. In a Lennard-Jones pairwise description of silver, the melting point is ~2.8 times too high.[15] The simple pair potential can be improved by introducing a volume dependence, i.e., a pair potential is derived that is appropriate for a particular density, or volume, of the system.[16] The motivation for this arises from viewing the metal crystal as ions immersed in a background electron gas.[17-19] The energy will depend strongly on the density of the electron gas and, hence, the volume of the system. While this modification to the pair potential corrects the three problems mentioned above,

it introduces the restriction that major defects such as a free surface cannot be treated, since the system cannot be characterized by a constant volume.

The embedded atom method allows a solution to this problem. In Eq. (2.1), the pair potential is augmented by a sum of embedding functions which, in effect, provide a local volume contribution to the energy for each atom. This is because ρ_i acts as a sensor for the crowding of atoms about atom i. Hence, with this form of potential, systems with a wide variation in density can be treated. The embedding function, F, provides a many-body contribution to the energy (assuming F is not purely linear), but the computational work involved in evaluating the energy and derivatives from Eq. (2.1) scales the same as a simple pair potential. However, there is no angular dependence in Eq. (2.1), so it is inappropriate for systems with strongly directional bonding.

We now briefly describe the parameterization used in the present study for fcc silver. The details of both the fitting procedure and the parameterization philosophy can be found elsewhere.[21,22] The pair potential is taken to be a Morse potential,

$$\phi(r) = D_M\{1 - \exp[-\alpha_M(r-R_M)]\}^2 , \qquad (2.3)$$

with three adjustable parameters, R_M, D_M, and α_M. The density function is defined as

$$\rho(r) = r^6[e^{-\beta r} + 512e^{-2\beta r}] , \qquad (2.4)$$

which is designed to roughly mimic the electron density of a transition metal atom,[23] with adjustable parameter B. Both $\phi(r)$ and $\rho(r)$ are modified to go smoothly to zero at a cutoff distance, r_{cut}, which is also an adjustable parameter. Following Foiles et al.,[14,24] for each choice of these five parameters, $F(\rho)$ is defined by specifying that the total energy [from Eq. (2.1)] of an fcc crystal follow the form

$$E(a^*) = -E_{coh}(1 + a^*)e^{-a^*} \qquad (2.5)$$

where a^* is a reduced distance variable given by

$$a^* = (a/a_o - 1) / (E_{coh}/9B\Omega)^{1/2} . \qquad (2.6)$$

Here a is the lattice constant, a_o is the equilibrium lattice constant, B is the bulk modulus, E_{coh} is the cohesive energy of the fcc crystal, and Ω is the atomic volume. The use of Eq. (2.5) to define $F(\rho)$ is motivated by the finding of Rose et al.[25] that results of density functional calculations on solids show a behavior similar to Eq. (2.5). This approach ensures reasonable behavior over a wide range of atomic densities, and leads to a potential that gives exact agreement with the experimental lattice constant, cohesive energy and bulk modulus.

The five parameters (R_M, D_M, α_M, β and r_{cut}) are optimized using a simplex search procedure,[26] by minimizing the root-mean-square deviation (χ_{rms}) between the calculated and experimental values for the three cubic elastic constants (c_{11},c_{12}, and c_{44}), the vacancy formation energy (ΔE_{vac}), and the bond length (R_e) and bond strength (D_e) of the diatomic molecule Ag_2. The bcc and hcp crystal structures are also required to be higher in energy than fcc. The resulting fit, shown in Table I, has χ_{rms} = 0.15%, and the parameter values are D_M = 0.672 eV, R_M = 2.570 Å, α_M = 1.826 Å^{-1}, β = 3.906 Å^{-1}, and r_{cut} = 5.542 Å.

TABLE I. - Experimental properties used in fit for Ag potential, along with the calculated values from the best fit. The values for a_o, E_{coh}, and B are fit exactly, due to the way $F(\rho)$ is obtained.

	experiment	calculated
a_o (Å)	4.09[a]	
E_{coh} (eV)	2.85[b]	
B (10^{12} erg cm^{-3})	1.04[c]	
c_{11} (10^{12}erg cm^{-3})	1.24[c]	1.24
c_{12} (10^{12}erg cm^{-3})	0.934[c]	0.935
c_{44} (10^{12}erg cm^{-3})	0.461[c]	0.461
ΔE_{vac} (eV)	1.1[d]	1.10
D_e (eV)	1.66[e]	1.66
R_e (Å)	(2.5)[f]	2.504

[a]Reference 27
[b]Reference 28
[c]Reference 29
[d]Reference 30
[e]Reference 31
[f]Reference 32

3. OVERLAYER DYNAMICS METHOD

We describe here the method used to follow the dynamics of an evolving pattern of adsorbed atoms. This recently developed method has been presented in detail elsewhere.[10] The key feature of this approach is that, for a certain class of systems, it can provide classically exact, course-grained dynamics for the evolution of an overlayer pattern, starting from an arbitrary interatomic potential. The method is restricted to systems in which the adsorbed species bind in registry with the surface, and to temperatures below the melting point of the surface (since a melted surface would not have well-defined binding sites). In the present work, we further impose the condition that the temperature be low enough that transition state theory is valid, as discussed below. Since the focus here is on Ag adatoms on the Ag(100) surface, the method will be described as it applies to that type of system.

3.1. Single-atom diffusion

Consider a single adatom on an fcc(100) face. For silver the stable binding site is in the fourfold hollow, and diffusive jumps to adjacent binding sites take place via the twofold saddle point which separate these fourfold sites, as shown in Fig. 1. Given a particular interatomic potential, which defines the adatom-substrate and substrate-substrate interactions, the classical thermal rate constant for escape from this binding site can be rigorously computed using transition state theory[33,34] (TST) combined with dynamical corrections. In the TST approximation, the rate constant for transfer of a system between two states is given by the thermal flux through a dividing surface that separates the states. (For the fcc(100) surface, a good choice for the TST dividing surface is a square whose corners are defined by the four atoms holding the adatom. When the adatom coordinates, projected onto the plane parallel to the surface, coincide with this square, the system is at the TST surface.) The TST rate constant is thus an equilibrium property of the system and can be rigorously computed using Metropolis Monte Carlo methods, as described elsewhere.[35,36]

Even if the TST surface is positioned such that the transition flux is minimized, TST does not give the true dynamical rate constant, because each crossing of the TST dividing surface does not necessarily correspond to a reactive state-change event. For example, a system passing from state A to state B may jiggle around in the region of the dividing surface, recrossing it many times before thermalizing in state B. Moreover, for a system consisting of surface binding sites, an energized adatom may make multiple hops, thermalizing in a binding site many sites away from where it started to hop. The effects of these "correlated dynamical events" can be accounted for in a rigorous fashion (regardless of the position of the TST surface) using a many-state dynamical corrections method,[37] which is an extension of a two-state formalism presented by Chandler.[38] This formalism is strictly applicable when the time duration of the correlated dynamical events (τ_{corr}) is much shorter than the average time between reactive transitions (τ_{rxn}),

$$\tau_{corr} << \tau_{rxn} . \qquad (3.1)$$

The method is implemented computationally by initiating a statistical sampling of classical trajectories at the TST dividing surface and integrating them until the correlated events cease. The result is that the true rate constant between state i and any other state (j) of

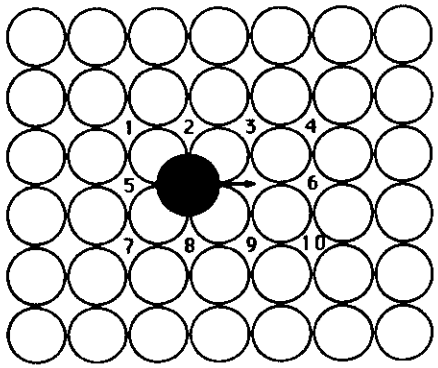


Fig. 1. Top layer of the substrate pad used to compute the rate constants for Ag/Ag(100). The ten binding sites making up the local environment are shown. All the 43 atoms shown (plus between zero and ten environment adatoms) are allowed to relax in each of the 2^{10} harmonic TST computations.

the system can be written as

$$k_{i\to j} = k_{i\to}^{TST}\, f_d(i\to j)\ , \qquad (3.2)$$

where $k_{i\to}^{TST}$ is the TST rate of escape through the dividing surface enclosing state i, and $f_d(i\to j)$ is the dynamical correction factor obtained from an analysis of the half-trajectories that originated at the boundary to state i. The total (diffusive) escape rate from state i is given by

$$k_{i\to} = \sum_{j(\neq i)} k_{i\to}^{TST}\, f_d(i\to j)\ . \qquad (3.3)$$

The surface diffusion process provides a good example of the motivation for expressing the rate constant using TST with dynamical corrections. Using Eq. (3.2), the diffusion constant for a single adatom on a two-dimensional periodic surface becomes

$$D = \frac{1}{4}\, k_{i\to}^{TST} \sum_{j(\neq i)} f_d(i\to j)\ \ell_{ij}^2\ , \qquad (3.4)$$

where ℓ_{ij} is the distance between binding sites i and j. In the temperature regime where Eq. (3.1) holds, it is relatively straightforward to compute D from Eq. (3.4). Alternatively, D can be computed from one long classical trajectory, from the time derivative of the mean squared displacement of the adatom position. However, this is completely unfeasible at low temperatures, due to the long time between diffusive hops of the adatom. In contrast, the trajectories in the dynamical corrections method are only integrated for a short time, since they *start* at the TST dividing surface, i.e., in the middle of a hop, and are only run until the hop is completed.

In previous studies on Lennard-Jones systems,[37] we have found that at low to moderate temperatures the dynamical corrections are usually negligible, affecting the diffusion constant by only a few percent (compared to the TST approximation). Moreover, we have shown[10] that reasonably accurate TST rates can be calculated using a harmonic approximation, given by

$$k^{HTST} = \nu_o \exp[-(E_{saddle}-E_{min})/k_BT] \qquad (3.5)$$

where ν_o is a harmonic frequency, k_B is the Boltzmann constant, T is the temperature, and E_{min} and E_{saddle} are the energies at the adatom minimum and the saddle point between two binding sites, respectively. These energies are found by performing a Newton-Raphson search for stationary points in the hyperspace defined by the Cartesian coordinates of all moving atoms. The frequency factor ν_o is computed from the second derivatives of the energy at the minimum and the saddle point.[10]

For computational convenience, we make use of this harmonic TST approximation in the present work.

3.2. Multiple, interacting adatoms

Now consider a distribution of interacting, adsorbed adatoms. The escape rate constant for any adatom in this overlayer will depend on its environment of nearly adatoms. Using the methods described above, this rate can be calculated for a hop in a particular direction. If these hopping rate constants are known for all atoms in the system at all times, then by using the appropriate thermodynamic statistics, the overlayer pattern can be dynamically evolved. If the rate constants are classically exact, then the overlayer dynamics will also be classically exact, in a course-grained sense; one can only ask about dynamical behavior for time scales much longer than the time between adatom hops. The spirit of the overlayer dynamics method employed here is to precalculate all possible rate constants that might be needed and store them in a catalog that is referenced during the dynamical simulation. More specifically, we define a ten-atom environment surrounding an adatom and the empty binding site in which it will land upon hopping to the right, as shown in Fig. 1. Since each of the ten sites can be empty or occupied, there are 2^{10}=1024 rate constants to be calculated (some of which will be equivalent by symmetry). Computation of this rate catalog is carried out in an automated fashion using the harmonic TST approximation described above. Once the catalog has been generated, the dynamics of an arbitrary overlayer pattern of N atoms is evolved as follows:

(i) Use the catalog to look up the appropriate value of the directional escape rate, k_{esc}, for each of the $4N_{atom}$ possible hops (each atom has four possible hop directions). For hops that are blocked by an adatom in the adjacent binding site, k_{esc}=0.

(ii) Increment the clock by

$$\Delta t_{hop} = \left[\sum_{i=1}^{4N_{atom}} k_{esc}(i)\right]^{-1} \qquad (3.6)$$

which is the time, on average, before *some* atom in the overlayer makes a hop.

(iii) Randomly select one of the $4N_{atom}$ possible hops, weighting the probability of selection of each hop by k_{esc}.

(iv) Go to (i).

As stated above, for the chosen interatomic potential, this procedure leads to classically exact, course-grained dynamics for the overlayer, including the motion and response of the substrate, within the assumption that the rate catalog is exact. The two approximations in the cataloged rates are of a controlled nature: they can be tested, improved and, in principle, eliminated. The first of these approximations lies in the method used to compute the rate constants; harmonic TST is employed, and the dynamical corrections are omitted. This approximation is good at low temperatures. The second approximation arises from truncating the environment at ten atoms, since the true rates will clearly depend on more distant adatoms. The motivation for this lies in the fact that 2^m rates must be calculated for an m-site environment. The effect of more distant adatoms can be tested, as discussed in the next section. In the present study, the approximation having the greatest effect on comparison to experiment is probably the silver interatomic potential described in Section II, even though it is at the present state of the art.

All dynamical properties that can be calculated with standard lattice-gas methods, such as tracer and chemical diffusion constants, autocorrelation functions of adsorbate density fluctuations and island nucleation rates, can also be computed using the present method. The method can be used with any interatomic potential, as long as the adsorbate binds in registry and Eq. (3.1) is satisfied. This approach, for the first time, provides a connection between stochastic lattice-gas techniqu and the interatomic potential.

To conclude this section, we briefly discu extensions of the method necessary to treat the mo general case of thin film growth. None of the extensions are employed in the present work.

Treating desorption with this method is straigh forward. The rate constant catalog is augmented include rate constants for desorption of an adatom, that during the simulation, each adatom has possibility of desorbing or diffusing that depends on environment. The TST desorption rate constant for given environment is computed from the flux throug a plane, parallel to the surface, that caps the top of t binding site. Dynamical corrections, if desired, can computed by initiating trajectories from this plane a following them for a time τ_{corr}.

Treating adsorption, as is necessary to model fil growth, is also straightforward. Dynamic corrections are computed in the same way as f desorption, and the *same* set of trajectories can even b used if the impinging atoms are assumed to be thermal equilibrium with the growing film (i.e., t dynamical correction factor for desorption is the sam as the thermal sticking coeffieient if the TST surface far enough from the surface). During the simulatio new atoms are introduced to the system at a rat dictated by the temperature and pressure of the vap phase and the dynamical correction factors give th relative probabilities of sticking, sticking in a near binding site, or completely bouncing off the surfac Note that if a nonthermal distribution of incomin atoms is assumed (e.g., for molecular beam epitaxy then the dynamical correction factors must b computed for that distribution, and will be differen than the desorption correction factors.

The simulation of general film growth requires tha the surface be allowed to have terraces, vacancies, etc rather than being defect free as assumed above. I principle, the rate catalog can be extended to includ jumps up or down steps, jumps in the presence of surface or bulk vacancy, and so forth. In practice, thi increases the number of environment sites (m substantially, so that calculating 2^m rate constant

becomes unfeasible. However, by judiciously choosing the environment patterns that are important to include for a particular case or type of process, it may be possible to perform this type of simulation.

4. Ag/Ag(100) CLUSTER DYNAMICS

In this section we discuss the simulation of the diffusion of two-dimensional clusters of adatoms and vacancies on Ag(100). There have been very few previous investigations of cluster motion (for a review, see Ref. 39), and until recently,[10] little was known about the motion of clusters larger than a few atoms. Understanding their dynamical behavior is useful, as the coalescence, diffusion and dissociation of clusters certainly play a role in layer-by-layer film or crystal growth, and in mass transport diffusion.

4.1 Calculational details

The Ag substrate used in the TST rate calculations was a four-layer block of atoms with 42 atoms per layer as shown in Fig. 1. Periodic boundary conditions were employed in the two directions parallel to the surface. For each harmonic TST calculation, the primary adatom, along with between zero and ten environment adatoms, were placed on top of this block, and these adatoms plus all of the first layer adatoms were allowed to move in the Newton-Raphson search. The layer size and block depth were made large enough to eliminate any boundary effects on the moving atoms. Generation of the rate catalog was automated as described previously.[10] For each of the 1024 environment patterns, a preexponential factor (ν) and an activation barrier ($E_A = E_{saddle} - E_{min}$) were stored, so that, using Eq. (3.5), the catalog could be employed to drive dynamics at any temperature (though the rates become more approximate at higher temperatures). Unless otherwise noted, calculations reported here were performed at T= 500K.

The effect of allowing relaxation of more than one substrate layer was tested by computing TST activation barriers for a single adatom. The barrier heights, expressed in kcal/mol, were found to be 11.68, 11.27, 11.18, 11.16 and 11.16, respectively, for relaxation of zero through six layers. Thus, as found previously,[35] accurate TST barrier heights for the fcc(100) surface can be obtained with only one layer relaxed.

The effect of the truncation of the catalog at ten binding sites was examined by generating a rate catalog for the 13-site environment shown in Fig. 2. This catalog includes six representative sites from the second shell (A-F), along with the seven sites from the first shell (1,2,3,5,7,8,9) thought to most strongly interact with the outer six. For each of the 4096 environment patterns, E_{min} and E_{saddle} were computed, along with E_∞, the energy with the primary adatom removed from the system. The energy required to desorb the primary adatom from the minimum is

$$E_1 = E_\infty - E_{min} \quad , \tag{4.1}$$

the energy to desorb the adatom from the saddle point is

$$E_2 = E_\infty - E_{saddle} \quad , \tag{4.2}$$

and the activation energy can be written as

$$E_A = E_1 - E_2 \quad . \tag{4.3}$$

The analysis of these energies is shown in Table 2. For each of the 13 sites, the effect (ΔE) of removing the adatom from that site was determined. For example, removing the adatom from site A can raise E_A by as much as 0.76 kcal/mol, or lower it by 0.07 kcal/mol. These values are determined by scanning the 2048 pairs of patterns (patterns in a pair are the same except for the occupation of site A), and finding the largest deviations in E_A. The magnitude of ΔE gives an indication of the effect a particular site has on the catalog energies, while the difference between ΔE_{max} and ΔE_{min} is a measure of the nonadditivity of the interactions. Note that the nonaddivity effects on E_A can be greater than 5 kcal/mol (see site 3). Of the sites omitted from the present work (A-F), site F appears to be most important, with ΔE_A ranging from -0.98 kcal/mol to 0.18 kcal/mol.

From an examination of ΔE_1 and ΔE_2, it is clear that different sites are important for the accuracy of E_1 than for E_2. For example, the sites having more than 0.5 kcal/mol effect on E_1 are 1,2,3,5,7,8,9,A, and E, whereas the sites having more than 0.5 kcal/mol effect on E_2 are 1,2,3,5,7,8,9,E, and F. Thus, using the 11-atom environment shown in Fig. 3a to compute a catalog of E_{min} values and the 14-atom environment in Fig. 3b to compute a catalog of E_{saddle} values would lead to a combined catalog converged such that any site added to the environment would affect E_A by less than 0.5 kcal/mol. (The union of these two environments has 15 sites, which would require 2×2^{15} calculations, rather than $2^{11} + 2^{14}$ calculations.)

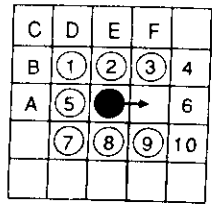
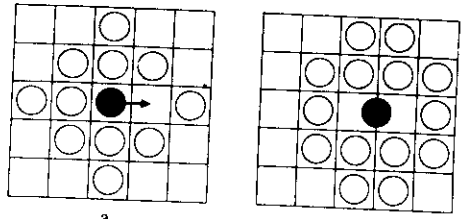


Fig. 2. The 13 sites used in computing a rate catalog to test the effect of environment size (see Table 2). To simplify the comparison with the 10-site environment used in the present work, the numbering of the first ten sites is maintained, even though only the circled sites were included in the catalog (4,6, and 10 were omitted). Sites A-F are outside the environment of the 10-site catalog used in the present dynamical calculations.

Table 2. Analysis of the 13 sites in Fig. 2, according to how much the desorption energy and activation energy of the primary adatom changes when the site atom is removed. ΔE is the change in desorption energy (kcal) upon site removal; ΔE_1 refers to desorption of the primary adatom from the minimum; ΔE_2 refers to the desorption of the primary adatom from the saddle point; ΔE_A refers to the activation energy of the primary adatom; min = lowest ΔE found in list of all configurations; max = highest ΔE found in list of all configurations.

	ΔE_1		ΔE_2		ΔE_A	
site	min	max	min	max	min	max
1	-0.85	-0.02	-0.32	0.87	-1.56	-0.07
2	-7.24	-5.17	-7.39	-2.05	-3.99	0.55
3	-0.88	-0.31	-7.98	-2.49	1.74	7.38
5	-7.25	-5.40	-1.79	-0.18	-6.82	-4.37
7	-0.91	-0.07	-0.32	0.72	-1.44	-0.16
8	-7.31	-5.76	-7.45	-2.40	-3.95	0.63
9	-0.95	-0.42	-7.85	-2.86	2.06	7.30
A	-0.08	0.76	-0.14	0.27	-0.07	0.76
B	-0.05	0.32	-0.14	0.09	-0.07	0.30
C	-0.06	0.09	-0.11	0.03	-0.05	0.12
D	-0.05	0.38	-0.16	0.39	-0.26	0.19
E	-0.09	0.76	-0.15	1.22	-0.56	0.50
F	-0.03	0.33	-0.10	1.17	-0.98	0.18



a b

Fig. 3. Environments deduced from Table 2 for computation of a catalog good to better than 0.5 kcal/mol (see text). a) 11-site environment for computing E_{min}; b) 14-site environment for computing E_{saddle}.

To improve the accuracy of the vacancy cluster simulations presented here, a second 10-site catalog was computed, in which all the binding sites outside of the 10-atom environment were occupied with Ag adatoms. Comparing this catalog to the first 10-site catalog shows a maximum deviation in E_A of 2.96 kcal/mol, a value that is in line with the findings above (989/1024 cases differed by less than 2 kcal/mol, and 807/1024 cases differed by less than 1 kcal/mol). The indices in this catalog were then transformed using a left-right reflection with occupation inversion,

$$\begin{aligned} 1 &\rightarrow \bar{4} \\ 2 &\rightarrow \bar{3} \\ 3 &\rightarrow \bar{2} \\ 4 &\rightarrow \bar{1} \\ 5 &\rightarrow \bar{6} \\ 6 &\rightarrow \bar{5} \\ 7 &\rightarrow \bar{10} \end{aligned}$$

$$8 \rightarrow \overline{9}$$
$$9 \rightarrow \overline{8}$$
$$10 \rightarrow \overline{7}$$

to make the rates appropriate for a vacancy hop to the right (i.e., an adatom hop to the left). Using this special catalog, vacancy dynamics were evolved in exactly the same way as adatom dynamics.

The cluster diffusion constants were computed using an independent dynamical simulation for each cluster size. Simulations were performed on a square grid using periodic boundary conditions, with the grid size chosen to prevent interaction between the cluster and its periodic image. The following describes the procedure for one simulation.

The initial cluster configuration of n atoms was generated by placing atoms at random on the grid, rejecting any placement that broke the connectivity of the cluster. Two atoms are considered "connected" if they are first or second nearest neighbors (i.e., an adatom can be directly connected to up to 8 other atoms). This cluster configuration was then evolved in time, using the procedure described in Section 3, with the restriction that if a chosen hop broke the cluster connectivity, that hop was rejected (and the clock was not incremented). This nonphysical restriction was imposed so that a diffusion constant could be computed for a well-defined cluster. An alternative approach would be to compute a diffusion constant from the motion of clusters that have not yet dissociated. These two approaches become equivalent at very low temperature[39,40] (probably lower than 500K). After a suitable warmup period, the diffusion constant was computed from the time derivative of the mean-squared displacement of the cluster center of mass,

$$D_n = \lim_{t\to\infty} \frac{1}{4}\frac{d}{dt} < \Delta R^2_{cm}(t) > \qquad (4.4)$$

by examining equal-time snapshots taken from a long trajectory. Typical trajectories were run for millions of steps, corresponding to tens of milliseconds. The mean squared displacement of a ten-atom cluster is shown in Fig. 4.

4.2 Results and discussion

To gain a qualitative understanding of the overlayer dynamics, it is instructive to examine the rate constants as a function of environment pattern. Figure 5 shows some representative activation barriers from the adatom rate catalog and from the vacancy rate catalog. The barrier to diffusion for a single adatom is 11.3 kcal/mol. In principle, this activation energy can be measured experimentally using field ion microscopy,[41] but to our knowledge, no Ag/Ag(100) measurements have been performed. The diffusion barrier for a vacancy is 10.7 kcal, so single vacancies are somewhat more mobile than single adatoms. When the energy of formation from a kink site is taken into account, the predicted mass transport diffusion barrier is 26.1 kcal for an adatom and 22.6 kcal for a vacancy. Thus, this potential predicts that vacancy monomer diffusion dominates adatom monomer diffusion in mass transport.

The barrier for an adatom to jump away from a nearest-neighbor atom is 17.8 kcal/mol, 6.5 kcal greater than free migration, indicating a strong tendency for the silver adatoms to cluster, as expected. In contrast, the barrier for separating two nearest-neighbor vacancies is only 1 kcal/mol higher than free migration of a vacancy. In fact, the vacancy dimer is more likely to dissociate (Fig. 5b) than to make a

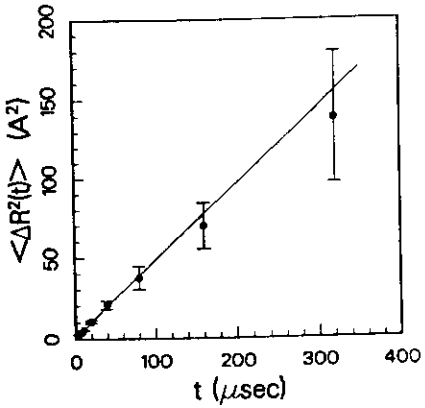


Fig. 4. Mean squared displacement of a 10-atom Ag cluster at T=500K.

diffusive step (Fig. 5c). For both the adatoms and the vacancies, the barrier for breaking out of the corner (Fig. 5d) or side (Fig. 5e) of a block is substantially larger than free monomer migration.

Figure 5f shows that an adatom can move easily along the edge of a perfect block, with a barrier of only 5.9 kcal/mol. The barrier for liberating an edge runner from a kink site on a cluster edge is 11.6 kcal/mol (Fig. 5g). At low temperature, where clusters tend to form tightly packed blocks, edge running can dominate the dynamics. This makes cluster diffusion simulation difficult, since edge running alone will not allow a cluster to diffuse. For vacancy clusters, this edge running has a barrier of 11.0 kcal, which will not dominate the dynamics. The vacancy edge-running that *does* occur is that of an edge-vacancy in the vacancy cluster (i.e., an adatom), but this adatom gets trapped at the corners, and thus does not so greatly dominate the dynamics.

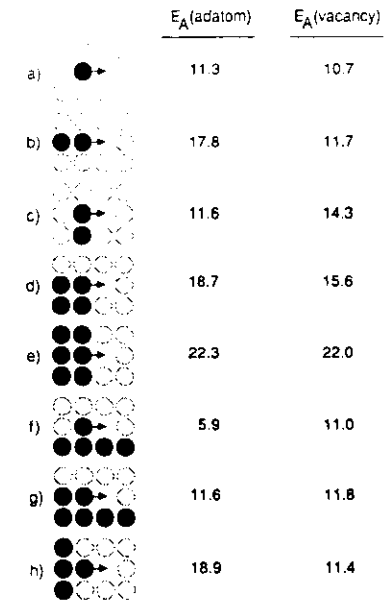


Fig. 5. Some representative activation barriers (kcal/mol) from the Ag/Ag(100) rate catalogs. Note that by particle-hole symmetry, the barriers for the adatom and the vacancy should be the same for (e) and also for (f). They differ due to the way the catalogs were generated.

Vacancy clusters also differ from adatom clusters in their greater propensity toward dissociation, as shown in Fig. 5h. Cluster dissociation rates at T=500K, estimated from the number of dissociating steps that were rejected during the run (this is not a rigorous dissociation rate), were $\sim 10^3$ times faster for the vacancy clusters when compared to adatom clusters of the same size.

Figure 6 shows the diffusion constants for clusters up to n=100 at T=500K. As expected, monomers (n=1) diffuse the fastest, but D_n does not decrease monotonically with increasing cluster size. This is because clusters that can form stable blocks (e.g., n=4,6,8) diffuse more slowly due to the long time required to break out of a perfect block structure (see Figs. 5d and 5e). Thus, for example, an n=4 cluster should have an activation barrier of 18.7 kcal/mol. In contrast, an n=5 cluster can diffuse with a maximum barrier of 12.8 kcal/mol, by following the sequence shown in Fig. 7. These values are in agreement with diffusional activation energies for n=4 and n=5 clusters obtained from Arrhenius plots. Both the adatom clusters and the vacancy clusters show this stable-block effect.

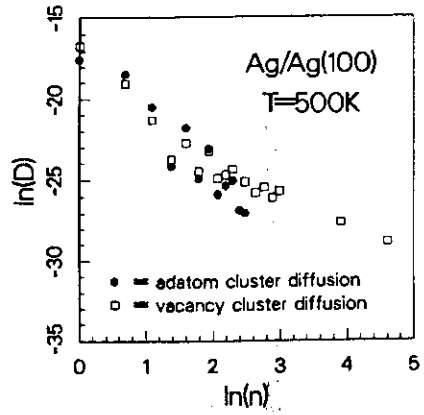
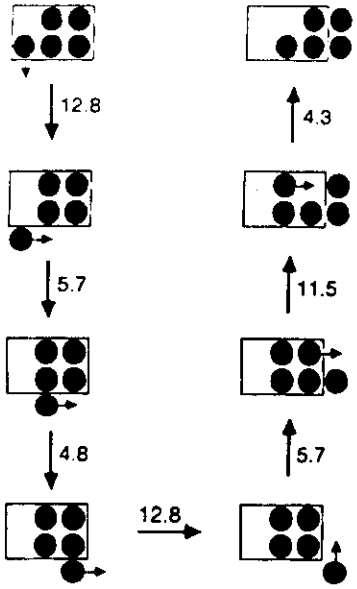


Fig. 6. Cluster diffusion constants (cm^2/sec) for Ag or Ag vacancies on Ag(100).

As in the previous study of Lennard-Jones usters,[10] the dominant diffusion mechanism for large g clusters is found to be edge motion. To actually ove the cluster, an adatom must climb onto a fresh lge from a kink site, surmounting a barrier of 19.2 al/mol, as shown in Fig. 8. This is in excellent reement with the activation energy (19±2 kcal/mol) tained from an Arrhenius plot of the diffusion of a 0-atom Ag cluster between T=600K and T=800K. gure 9 shows successive snapshots of an n=100 atom cluster and an n=100 vacancy cluster. The erall diffusion rates are seen to be roughly mparable. The 25-configuration sequence (5 msec) presents 1/40 of the total trajectory used to compute 100 for the vacancy cluster. In the adatom cluster, nproductive edge running results in three times as ny steps per unit time than for the vacancy cluster, king computation of a diffusion constant more ficult



. 7. Diffusion pathway for n=5 adatom cluster, in ich the cluster is replicated one binding site to the ht after 7 moves. The numbers are activation riers in kcal/mol. The maximum barrier is 12.8 l/mol.

As mentioned above, the cluster lifetime (before dissociation) can be estimated by dividing the total trajectory time by the number of cluster-dissociating steps that were rejected during the run. This is not a rigorous calculation, but is a good approximation if many more steps are accepted than rejected (not always the case in the present work). Knowing this dissociation lifetime (τ_n) and the diffusion constant, the root mean square distance that a cluster will diffuse before dissociation can be estimated by

$$d_n = [4D_n\tau_n]^{1/2} \quad . \qquad (4.5)$$

It is interesting to note that for adatom clusters d_n lies between 1Å and 10 Å for a large range of cluster sizes (up to n=40) and temperature (d_n decreases slowly as n increases and as T increases). For vacancy clusters the same effect is observed, but with d_n about ten times smaller. (Actual values for d_n are not given here due to

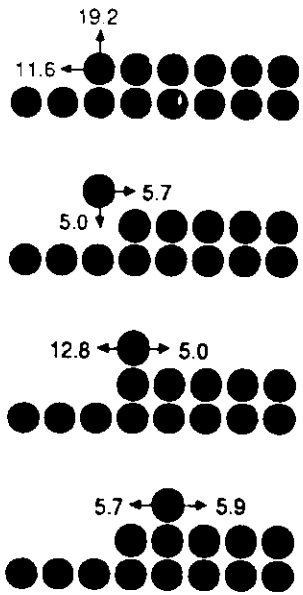


Fig. 8. Mechanism for an adatom to climb onto a fresh edge from a kink site. The maximum barrier in this process is 19.2 kcal/mol, in agreement with the Arrhenius activation energy for diffusion of a 100-atom Ag cluster.

the uncertainty in τ_n.) Under the right thermodynamic conditions, cluster diffusion may contribute significantly to mass transport, but these small values for d_n imply that the transport will *not* be via clusters moving long distances intact. Rather, a picture emerges in which slowly moving clusters are in equilibrium with rapidly moving monomers. A cluster captures any adatom that comes too close, but then ejects a monomer after it has diffused just a few angstroms (or a few tenths of an angstrom in the case of vacancy clusters). Thus, a cluster that has migrated a long distance (e.g. hundreds of angstroms), probably consists mostly of different atoms than when it started.

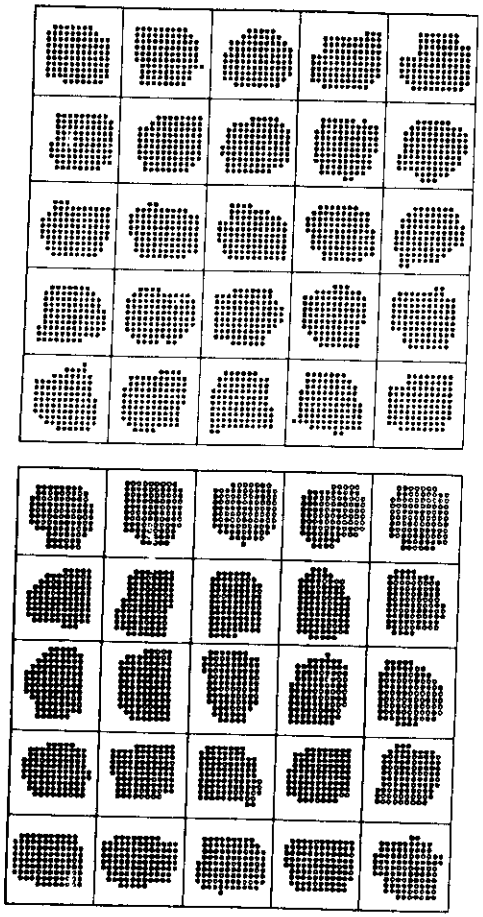


Fig. 9. Equal time shapshots of an n=100 adatom cluster (top) and an n=100 vacancy cluster (bottom) on Ag(100) at T=500K. The time between snapshots is 0.2 msec for both cases. Between successive snapshots the adatom cluster made ~300,000 steps and the vacancy cluster made ~95,000 steps.

Another interesting feature is that a passing monomer can act as a sort of catalyst to enhance the diffusion of a small block cluster. For example, below room temperature, an n=4 cluster is virtually immobile compared to an n=5 cluster. Thus, when the n=4 cluster absorbs a monomer, it suddenly starts to actively diffuse. When it later ejects a monomer, it becomes immobile again.

5. CONCLUSIONS

The overlayer dynamics method, in conjunction with an embedded atom-style potential, has been applied to the classical dynamics of clusters of Ag adatoms and vacancies on the Ag(100) surface. The adatom cluster diffusion is qualitatively similar to the diffusion observed for Lennard-Jones clusters on fcc(100).[10] Monomers diffuse the fastest, with E_A=11.3 kcal/mol. Larger clusters diffuse more slowly, though the decrease is not monotonic, because some small clusters can form stable blocks with high activation barriers to diffusion. The rate determining step for diffusion of large clusters is the 19.2 kcal/mole barrier for an adatom to climb onto a fresh edge from a kink site.

Vacancy clusters were also examined, and show the same features as adatom clusters, though they dissociate more rapidly (by ejection of a monomer), and nonproductive edge-running is less prevalent. A vacancy monomer diffuses faster (E_A = 10.7 kcal/mol) than an adatom on the Ag(100) surface, and is expected to contribute to mass transport with an activation energy of 22.6 kcal/mol, whereas the mass transport barrier for adatom diffusion is 26.1 kcal/mol.

From an estimate of the cluster dissociation lifetime, it appears that clusters at T=500K migrate between 1Å and 10 Å before dissociation (clusters

much larger than 100 atoms survive for a shorter distance), and vacancies survive for about 0.1 Å to 1 Å.

The treatment of some types of thin film growth with this method should be possible, by incorporating adsorption, desorption and multi-layer adatom jumps into the catalog.

REFERENCES

1. R.W. Vook, Int. Met. Rev. 27, 209 (1982).
2. J.A. Venables and G.D.T. Spiller, in Surface Mobilities on Solid Materials, edited by V.T. Binh, Vol. 86, (Plenum, New York, 1983), p. 341.
3. J.D. Weeks and G.H. Gilmer, Adv. Chem. Phys. 40, 157 (1979).
4. G.H.J. Gilmer and J.Q. Broughton, J. Vac. Sci. Tech. B 1, 298 (1983).
5. F.F. Abraham and G.M. White J. Appl. Phys. 41, 1841 (1970).
6. B. Liao and H.A. Macleod, Proc. SPIE-Int. Soc. Opt. Eng. 540, 150 (1985).
7. K.H. Müller, Appl. Phys.A, 40, 209 (1986).
8. J. Salik, J. Appl Phys. 57, 5017 (1985).
9. M. Schneider, A. Rahman, and I.K. Schuller, Phys. Rev. 34, 1802 (1986).
10. A.F. Voter, Phys. Rev. B. 34, 6819 (1986).
11. M.S. Daw and M.I. Baskes, Phys. Rev. B 29, 6443 (1984).
12. M.W. Finnis and J.E. Sinclair, Phil. Mag. A 50, 45 (1984).
13. M. Manninen, Phys. Rev. B 34, 8486 (1986).
14. S.M. Foiles, M.I. Baskes, and M.S. Daw, Phys. Rev. B 33, 7983 (1986), and references therein.
15. Using the Lennard-Jones parameters from T. Halichioglu and G.M. Pound, Phys. Status Solidi A 30, 619 (1975), and along with the melting temperature of a Lennard-Jones crystal, which is 0.7ε.
16. W.A. Harrison, Pseudopotentials in the Theory of Metals (Benjamin, New York, 1966).
17. J.K. Norskov and N.D. Lang, Phys. Rev. B 21, 2131 (1980).
18. M.J. Stott and E. Zaremba, Phys. Rev. B 22, 1564 (1980).
19. M.J. Puska, R.M. Nieminen and M. Manninen, Phys. Rev. B 24, 3037 (1981).
20. J.K. Norskov, Phys. Rev. B 26, 2875 (1982).
21. A.F. Voter, to be published.
22. A.F. Voter and S.P. Chen, Proceedings of the Materials Research Society, Boston, Winter 1986, Symposium I (MRS Volume 82).
23. Equation (4) is the electron density for a 4s Slater orbital, appropriate for a first-row transition metal. This was found to work well for the six fcc metals fit in Ref. 21.
24. S.M. Foiles, Phys. Rev. B 32, 7685 (1985).
25. J.H. Rose, J.R. Smith, F. Guinea and J. Ferrante, Phys. Rev. B 29, 2963 (1984).
26. J.A. Nelder and R. Mead, Comp. J. 7, 308 (1965).
27. C. Kittel, "Introduction to Solid State Physics," 5th edition (Wiley, New York, 1976), p. 31.
28. Metal Reference Book, 5th ed., edited by C.J. Smith (Butterworths, London 19876), p. 186.
29. G. Simmons and H. Wang, Single Crystal Elastic Constants and Calculated Aggregate Properties: A Handbook (MIT Press, Cambridge, 1971).
30. R.W. Balluffi, J. Nucl. Mater. 69 & 70 240 (1978).
31. K.P. Huber and G. Herzberg, Constants of Diatomic Molecules (Van Nostrand Reinhold, New York, 1979), p. 8.
32. R_e for Ag_2 was estimated by scaling R_e for Cu_2 (Ref. 31, p. 198) using the fcc lattice constants for Cu and Ag from Ref. 23.
33. S. Glasstone, K.J. Laidler and H. Eyring, The Theory of Rate Processes (McGraw Hill, New York 1941).
34. J.D. Doll, J. Chem. Phys. 73, 2760 (1980); 74, 1074 (1981).
35. A.F. Voter and J.D. Doll, J. Chem. Phys. 80, 5832 (1984).
36. A.F. Voter, J. Chem. Phys. 82, 1890 (1985).
37. A.F. Voter and J.D. Doll, J. Chem. Phys. 82, 80 (1985).
38. D. Chandler, J. Chem. Phys. 68, 2959 (1978).
39. J.D. Doll and A.F. Voter, Ann. Rev. Phys. Chem. 38, 413 (1987).
40. J.C. Tully, G.H. Gilmer and M. Shugard, J. Chem. Phys. 71, 1630 (1979).
41. J.A. Panitz, J. Phys. E 15, 1281 (1982).

Computer simulation on surfaces and [001] symmetric tilt grain boundaries in Ni, Al, and Ni_3Al

S.P. Chen, D.J. Srolovitz,* and A.F. Voter
Los Alamos National Laboratory, Los Alamos, New Mexico 87545

(Received 21 April 1988, accepted 29 September 1988)

We have used "local volume" (embedded atom) type potentials to study the surfaces and grain boundaries of Ni, Al, and Ni_3Al. The simulations show that with appropriately fit potentials, the surface and grain boundary structure can be realistically calculated. The surface rippling and relaxation show good agreement with experiments. The energies of most surfaces and grain boundaries also agree with existing data. The structural unit model for grain boundaries in Ni_3Al shows the same generic units as in pure metals, but with large variations due to distortions and multiplicity. The utility of the structural unit model is thus more limited for alloys. The grain boundary energies were found to be the highest for Al-rich Ni_3Al grain boundaries, and depend significantly on the local composition of the grain boundary. The cusps in the grain boundary energy as a function of misorientation angle are different for different grain boundary stoichiometries. The Ni_3Al grain boundaries have approximately the same grain boundary energy and cohesive energy as that of Ni.

I. INTRODUCTION

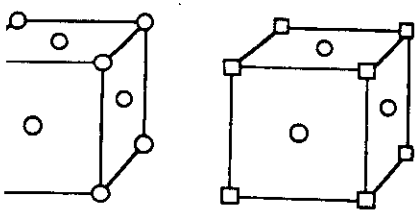
The important role that grain boundaries play in the physical properties of polycrystalline materials (diffusion and electrical transport, creep resistance, fracture stress, etc.) is well established. Consequently, grain boundaries in pure metals have received substantial experimental and theoretical attention[1-3] over the last few decades. The main aim of these studies has been to develop an understanding of the structure, energetics, and phenomena of grain boundaries. Such knowledge is paramount if the physical properties of polycrystalline materials are to be understood and controlled.

Several models of the structure and/or energetics of grain boundaries have been proposed. Perhaps the earliest and most successful are dislocation models,[4] which have been shown to correctly predict grain boundary structures and energies. These models, however, are valid only in the low (tilt or twist) angle limit where the spacing between dislocations is large compared to the dislocation core size. The coincident site lattice (CSL)[5-8] and O-lattice[9] models have attracted a great deal of attention because of their simplicity and the possibility of relating grain boundary structure to physical properties through geometric parameters such as the reciprocal coincident site density, Σ. While these models can be applied to arbitrarily large angle grain boundaries, recent computer simulations[2,3] have pointed out that they are of limited utility since they do not take into account atomistic relaxation at the grain boundary itself.

These same simulations have led to new grain boundary models which are known as structural unit[10] or polyhedral models.[11] The structural unit model[10] maintains that for a given misorientation, the grain boundary structure consists of a combination of two basic structural units. Within a well-defined range of misorientations, the grain boundaries are composed of the same two structural units. This model has been justified[12] on the basis of geometrical necessity and continuity conditions. In addition to predicting grain boundary structure, the structural unit model has been successfully employed[13] to calculate grain boundary energies as a function of misorientation. Comparison of the atomistic structures determined by computer simulation with experimental (electron microscopy and x-ray scattering) determinations has shown good agreement in some cases[14] and poor agreement in others.[15] In cases where the agreement has been poor, further analysis of the simulation results has shown[15] that a number of possible structures exist which are very close in energy. In those cases, averaging over the multiplicity of the nearly degenerate structures results in good agreement between simulation[15] and experiment.[16]

In this paper, we employ recently developed "local volume" potentials[17] in a series of computer simulations[18-21] designed to elucidate the structure and energies of grain boundaries in metals. In particular, we concentrate on pure Ni, pure Al, and on the ordered alloy Ni_3Al. These metals were chosen for a variety of reasons. First, there has been a great deal of interest in grain boundaries in Ni_3Al due to its brittleness in the polycrystalline form.[22] Second, Ni, Al, and Ni_3Al share the same underlying crystal structure (Ni_3Al has the $L1_2$ structure, which is topologically equivalent to the fcc structure of Ni and Al, as indicated in Fig. 1). Finally, the surfaces of Ni, Al, and their ordered alloys have been the subject of numerous experimental studies (see references cited in Ref. 18) and hence provide a quantitative indication of the validity of the description of the atomic interactions employed.

* Permanent address: Department of Materials Science and Engineering, University of Michigan, Ann Arbor, Michigan 48109.



Ni (fcc) Ni_3Al ($L1_2$)

○ = Ni □ = Al

Face-centered cubic (fcc) crystal structure of Ni (a) and $L1_2$ of Ni_3Al (b). The Ni and Al atoms are represented by circles ures, respectively, in this figure and throughout this paper. Note Al has two sets of (100) planes: one is 100% Ni atoms and the 50% Ni and 50% Al atoms.

e begin by describing the simulation technique and cal volume" (embedded atom-like[23]) potentials em-

Quantitative comparison[18-20] of the simulated atomic res of surfaces in Ni, Al, and Ni_3Al with experi- data indicates the reliability of the potentials and idity of the relaxation technique. In Ni, Al, and the displacements of the atomic planes away from in boundary show a decaying oscillatory relaxa- perimposed on a net expansion which is very similar seen in the vicinity of free surfaces. The net lat- pansion or excess volume associated with the grain ry is found to vary relatively smoothly with misori- n and to be very similar in Ni and Al. In Ni_3Al, r, the variation is less regular and there is a depen- on the local grain boundary stoichiometry. While ng a reasonable description of the grain boundary es in pure metals, the structural unit model appears f limited utility in describing the ordered alloy. This to the complexity arising from the different chemi- stituents and the large distortions of the structural The grain boundary energies are found to scale lin- ith the excess grain boundary volume in both the tals and the ordered alloy. The positions and depths nima (cusps) in the grain boundary energy vs mis- ion plots for Ni_3Al are shown to depend on the lo- n boundary stoichiometry. The Al-rich boundaries have the highest energies and generally exhibit the hesion.

OMIC INTERACTION POTENTIALS FOR Ni, D Ni_3Al

grain boundary simulations presented here em- ocal-volume form of potential closely related to the ed atom method of Daw and Baskes.[23,17] We use "local volume" to indicate that there is a term in the energy expression for each atom that depends on the local atomic density (or "volume") about that atom. While the details of the form and fit of the potential are presented elsewhere,[17,24] we present a summary here.

For a pure metal the energy of an n-atom system is

$$E = \sum_i^n E_i , \tag{1}$$

where the energy of atom i is given by

$$E_i = \frac{1}{2} \sum_{j \neq i} \Phi(r_{ij}) + F(\rho_i) . \tag{2}$$

Here r_{ij} is the scalar distance between atoms i and j, Φ is a pairwise interaction potential, and ρ_i is the density at atomic site i due to all its neighbors,

$$\rho_i = \sum_{j \neq i} \rho(r_{ij}) \tag{3}$$

The embedding energy, $F(\rho_i)$, can be interpreted as the energy arising from embedding atom i in an electron gas of density ρ_i. To first order, this embedding energy contains all, and only, the quantum mechanical interaction between the embedded atom and the host, leaving the pairwise potential to account for the classical electrostatic interaction.[25] To mimic roughly the shape of the classical electrostatic interaction between two frozen, neutral atoms, while allowing fitting flexibility, $\Phi(r)$ is taken to be a Morse potential,

$$\Phi(r) = D_M\{1 - \exp[-\alpha_M(r - R_M)]\}^2 - D_M . \tag{4}$$

The three parameters, D_M, R_M, and α_M, respectively define the depth, distance to the minimum, and a measure of the curvature at the minimum of the Morse potential. The density function, $\rho(r)$ is taken as

$$\rho(r) = r^4[e^{-\beta r} + 512e^{-2\beta r}] , \tag{5}$$

where β is an adjustable parameter. This is the density (ignoring normalization) of a hydrogenic 4s orbital, with the second term added to ensure that $\rho(r)$ decreases monotonically with r over the whole range of possible interaction distances. The 4s shape is chosen to be appropriate for a first row transition metal, and is found to work well for aluminum, too.[24]

Following Foiles *et al.*,[26] $F(\rho)$ is specified by requiring that the energy of the fcc crystal behaves properly as the lattice constant is varied. Rose *et al.*[27] have shown that the cohesive energy of most metals can be scaled to a simple universal function, which is approximately

$$E_u(a^*) = -E_0(1 + a^*) \exp(-a^*) \tag{6}$$

where a^* is a reduced distance variable and E_0 is the depth of the function at the minimum ($a^* = 0$). The appropriate scaling is obtained by taking E_0 as the equilibrium cohesive energy of the solid (E_{coh}), and defining a^* by

$$a^* = (a/a_0 - 1)/(E_{coh}/9B\Omega)^{1/2} , \tag{7}$$

where a is the lattice constant, a_0 is the equilibrium lattice constant, B is the bulk modulus, and Ω is the equilibrium atomic volume. Thus, knowing E_{coh}, a_0, and B, the embedding function is defined by requiring that the energy of the perfect crystal [Eq. (6)] matches the energy from Eq. (2) for all values of a^*. By fitting $F(\rho)$ in this way, the potential should behave properly over a large range of densities.

To be suitable for use in atomistic simulations, the interatomic potential and its first derivatives with respect to nuclear coordinates should be a continuous function of r. This is accomplished by forcing $\Phi(r)$, $\Phi'(r)$, $\rho(r)$, and $\rho'(r)$ to go smoothly to zero at $r = r_{cut}$ (see Ref. 24), where r_{cut} is optimized in the fitting procedure and the primes indicate differentiation with respect to the argument. $F(\rho)$ is also modified to go smoothly to zero when the expanded crystal has a nearest neighbor distance equal to r_{cut}.

Given the above functional forms for $\Phi(\rho)$ and $\rho(r)$, the five parameters, R_M, D_M, α_M, β, and r_{cut}, are determined by minimizing the deviation between the calculated and experimental values for the three cubic elastic constants (C_{11}, C_{12}, and C_{44}), the vacancy formation energy ($\Delta E'_{1_f}$), and the bond length (R_e) and bond energy (D_e) of the diatomic molecule, and by requiring that the hcp and bcc crystal structures be less stable than the fcc. Note that because of the way $F(\rho)$ is determined, and because it is recomputed for each new choice of the five parameters defining $F(\rho)$ and $\rho(r)$, the potential always gives a perfect fit to the experimental values of E_{coh}, a_0, and B. Table I shows the experimental data employed in the fits for Ni and Al, along with the calculated values and the rms deviation.

For alloy systems, the energy expression for atom i becomes

$$E_i = \frac{1}{2} \sum_{j \neq i} \Phi_{t_i t_j}(r_{ij}) + F_{t_i}(\rho_i) , \tag{8}$$

The density at atom i is now

$$\rho_i = \sum_{j \neq i} \rho_{t_j}(r_{ij}) \tag{9}$$

where the subscripts t_i and t_j indicate the elemental types of atoms i and j, respectively. To generate potentials for the Ni–Al binary alloy system, the first step is to fit the pure-element functions $\Phi_{NiNi}(r)$, $\Phi_{AlAl}(r)$, $\rho_{Ni}(r)$, $\rho_{Al}(r)$, $F_{Ni}(\rho)$, and $F_{Al}(\rho)$, as described above. The cross potential, $\Phi_{NiAl}(r)$, is then fit using thermodynamic data on Ni–Al alloys. To obtain a good fit, one can also take advantage of two additional transformations (three parameters: S_{Al}, g_{Ni}, g_{Al}) under which this form of potential is invariant.[17] Assuming a Morse potential for $\Phi_{NiAl}(r)$ leads to a total of seven parameters, which are optimized by fitting to the Ni_3Al lattice constant, cohesive energy, elastic constants, ordering energy, vacancy formation energy, (111), and (100) antiphase boundary (APB) energies, the superlattice intrinsic stacking fault (SISF) energy, and the lattice constant and cohesive energy of the B2 phase NiAl (CsCl structure).[28-40] The "best fit" is shown in Table II, and the optimized parameters are given in Table III. The inclusion of the Ni_3Al APB and SISF energies in the fit is quite important since their omission leads to potentials which yield negative APB or SISF energies, indicating that the $L1_2$

TABLE II. Data used to fit the Ni_3Al cross potential. Superscripts are the experimental references (a = Ref. 36, b = Ref. 37, c = Ref. 38, d = Ref. 39, e = Ref. 40, f = Ref. 28).

Ni_3Al properties	Expt	Calc
a_0 (Å)	3.567[a]	3.573
E_{coh} (eV)	4.57[b]	4.59
C_{11} (10^{12} erg/cm^3)	2.30[c]	2.46
C_{12} (10^{12} erg/cm^3)	1.50[c]	1.37
C_{44} (10^{12} erg/cm^3)	1.31[c]	1.23
$\Delta E'_{1_f}$ (eV)	1.6 ± 0.2[d]	1.64 (Ni), 1.87 (Al)
SISF (111) (mJ/m^2)	10 ± 5[e]	13
APB (100) (mJ/m^2)	140 ± 14[e]	83
APB (111) (mJ/m^2)	180 ± 30[e]	142
B2 NiAl properties		
a_0 (Å)	2.88[f]	2.87
E_{coh} (eV)	4.51[b]	4.38

TABLE I. Data used in fitting the pure metal potentials. Calculated values of a_0, E_{coh}, and B match experiment exactly due to the way $F(\rho)$ is determined. Superscripts are the experimental references (a = Ref. 28, b = Ref. 29, c = Ref. 30, d = Ref. 31, e = Ref. 32, f = Ref. 33, g = Ref. 34, h = Ref. 35).

	Ni		Al	
Property	Experimental	Calculated	Experimental	Calculated
a_0 (Å)	3.52[a]	...	4.05[a]	...
E_{coh} (eV)	4.45[b]	...	3.36[c]	...
B (10^{12} erg/cm^3)	1.81[d]	...	0.79[d]	...
C_{11} (10^{12} erg/cm^3)	2.47[d]	2.44	1.14[d]	1.07
C_{12} (10^{12} erg/cm^3)	1.47[d]	1.49	0.619[d]	0.652
C_{44} (10^{12} erg/cm^3)	1.25[d]	1.26	0.316[d]	0.322
$\Delta E'_{1_f}$	1.60[e]	1.60	0.75[f]	0.73
D_e (eV)	1.95[g]	1.94	1.60[g]	1.54
R_e (Å)	2.2[g]	2.23	2.47[h]	2.45
rms deviation (%)	0.75		3.85	

TABLE III. Potential parameters optimized from fits to experimental data. The parameters S_{Al}, g_{Ni}, and g_N are defined and discussed in Ref. 17.

	Ni	Al	Ni–Al	
D_M (eV)	1.5335	3.7760	3.0322	S_{Al} = 0.61723
R_M (Å)	2.2053	2.1176	2.0896	g_{Ni} = 6.5145 eV Å^3
α_M (Å^{-1})	1.7728	1.4859	1.6277	g_N = −0.2205 eV Å^3
r_{cM} (Å)	4.7895	5.5550	5.4639	
β (Å^{-1})	3.6408	3.3232		

structure is unstable. The data on B2 NiAl properties are included to broaden the range of stoichiometries over which the potential is valid.

III. SIMULATION PROCEDURE

The unrelaxed, starting configurations of the grain boundaries were constructed by overlapping two infinite perfect crystals and performing the appropriate rotation (θ) about the [001] axis to achieve the desired misorientation. The grain boundary plane was chosen such that the grain boundary was symmetric. In other words, crystal 1 was rotated by $\theta/2$ and crystal 2 by $-\theta/2$, where θ is measured from the (011) plane in the unrotated frame. Each of the two original infinite crystals was then truncated such that only atoms from crystal 1 were on one side of the grain boundary and only atoms from crystal 2 were on the other.

In pure metals, there is only one unique way of performing the initial overlap of the two infinite crystals. However, in an ordered binary alloy, one must choose the elemental decoration of each lattice site. Since the $L1_2$ structure (Fig. 1) consists of four interpenetrating simple cubic lattice (three occupied by Ni and one by Al in Ni_3Al), a choice must be made as to which type of atom occupies each sublattice. For an $L1_2$ structure, this ambiguity permits four distinct atomic arrangements in each of the two crystals meeting at the grain boundary and sixteen possible different grain boundaries with the same indices (rotation axis and rotation angle). For the case of [001] symmetric tilt boundaries, many of these possible grain boundaries are equivalent, resulting in three unique compositions for each set of indices. This is because every plane containing the [001] axis contains either 100% Ni or 50% Ni and 50% Al. The three unique compositions can thus be classified by specifying the percentage of Ni (50 or 100) in the first layer of each of the two grains, i.e., (100/100), (100/50), and (50/50). Note that the (100/50) boundaries have the same stoichiometry as bulk Ni_3Al, while the (100/100) boundaries are Ni-rich and the (50/50) boundaries are Al-rich. The starting unrelaxed structures of the Σ5 (210) grain boundary in Ni, Al, Ni_3Al (100/100), Ni_3Al (100/50), and Ni_3Al (50/50), constructed by the method described above, are shown in Fig. 2.

The calculational cell used in the present study is indicated in Fig. 3. The cell extended at least 18 lattice parameters (a_0) on either side of the grain boundary (i.e., in the $\pm z$ directions). In the direction perpendicular to the grain boundary, the cells were terminated by free surfaces that were chosen to have the same composition as the grain boundary. In this way, the structure and energy of the grain boundary and surface were determined simultaneously. Since there is periodicity in the atomic structure of the grain boundary, periodic boundary conditions were employed in the x and y directions.

The structure of the grain boundary was relaxed to the minimum energy configuration using an energy gradient method. The relaxation was considered complete when the force on each atom (computed using the potentials described in Sec. II) was less than 10^{-4} eV/Å. Since the calculational cell includes a large number of atoms, a net shift of the top grain with respect to the bottom grain requires the collective motion of many atoms and hence is extremely inefficient if atoms are moved individually. To surmount this difficulty, the total force on the top grain due to the bottom grain was calculated and the relative position of the two grains was shifted accordingly.[41] By alternating individual atom relaxations with rigid shift relaxations, these two modes were allowed to occur in tandem.

Although the top and bottom grains are free to translate with respect to each other, the translational state of the relaxed structure may not correspond to the lowest energy configuration, because multiple energy minima may exist. In fact, Wang, Sutton, and Vitek[42] have demonstrated that for many boundaries, multiple energy minima exist with very similar energies and very different atomic structures. To find the global minimum among the many local minima, simulations for each set of grain boundary parameters were performed starting with different shifts of the top grain with respect to the bottom grain. To accomplish this in an efficient way, grain translation energy maps were constructed. This was done by freezing the atom positions in the top and bottom grain, fixing an x-y translational state of the top grain with respect to the bottom grain, and performing a relaxation of the z-shift (i.e., separation between grains in the z direction) of the top grain with the bottom grain held fixed. Performing a grid of these calculations in the x and y directions yields a crude energy map, as shown in Fig. 4 for (210) and (310) grain boundaries in Al. While each of these contour maps shows multiple minima and maxima, certain of these are equivalent by symmetry. For example, the points labeled 1 and 2 in Fig. 4(b) are equivalent [the slight differences in appearance are due to the finite grid size (0.1 a_0) employed in generating the contours]. In searching for the absolute minimum energy configuration in the full relaxation, at least 10 simulations were performed starting with different x-y shifts. The full relaxation generally yields x-y shifts very close to the minima predicted by the energy map. The final x-y shift obtained following full relaxation depends on the choice of the initial x-y shift, indicating a finite basin of attraction associated with each possible minimum.

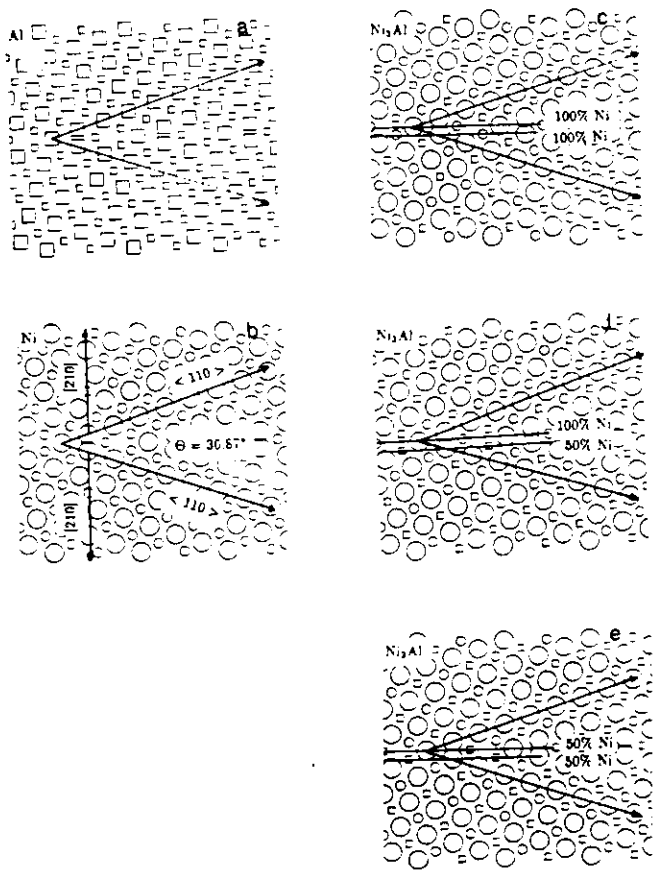
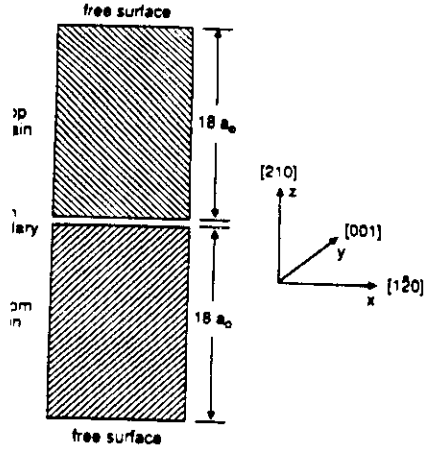


FIG. 2. The unrelaxed Σ5 (210)/[001] symmetric tilt boundary structures of Al (a), Ni (b), Ni_3Al 100/100 (c), Ni_3Al 100/50 (d), and Ni_3Al 50/50 (e). The Ni_3Al case has three unique GB stoichiometries due to the two types of planes existing in Ni_3Al, as shown in Fig. 1.

IV. FREE SURFACE RESULTS: A TEST OF INTERATOMIC POTENTIALS

While the experimental analysis of grain boundaries is complicated by the fact that these interfaces are buried, free surfaces may be analyzed by more direct means. Therefore, it is not surprising that significantly more progress has been made in understanding the structure of free surfaces in metals.[43] As a consequence of the recent flurry of activity in this area, a large body of experimental data exists for surfaces,[43–49] some of which has been confirmed in multiple studies using different techniques. As such, theories, computer simulations, and experiments on the atomic structure of surfaces may be more easily compared than for grain boundaries. To gauge the accuracy of the interatomic potentials, we compare our simulation results with experimental data on the structure of metallic surfaces. Detailed comparison can be found in Refs. 18 and 20.

Our results on the surface structures are reported in terms of the percentage changes in interlayer spacing perpendicular to the free surface. This change is simply the strain, indicated by the symbol $\Delta d_{n,n+1}$, where the $n, n+1$ indicates the change in the separation between the n^{th} and $(n+1)^{th}$ atomic planes. The nature of the agreement between the simulations and experiments is indicated in Table IV. While our results are in good agreement with the LEED results[44,45] on Al(110), The agreement with the LEED[46] and medium energy ion scattering[47] data on Ni(110) is off by approximately a factor of two (corresponding to a discrepancy of −0.002 nm). Nonetheless, the main features of the experimental data are reasonably well repro-



Schematic drawings of the simulational blocks of atoms. The ...eriodic in the x and y directions, but has two free surfaces in the ...ion so that surface information and grain boundary information are ...d at the same time. The cells are 36 a_0 in the z direction to avoid ...ference between the grain boundary relaxation and the surface re... All the atoms are allowed to relax and the top and bottom grains to shift with respect to each other.

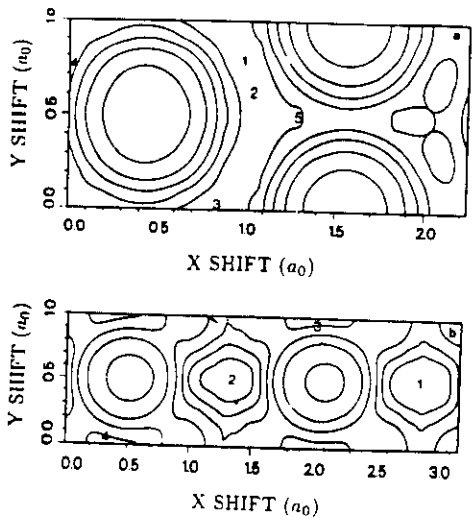


FIG. 4. The constant grain boundary energy contour plot of x and y-shift for (210) (a) and (310) (b) GB in the Ni_3Al. Some distortions are introduced by the finite size of the grid. In (a) the points 1, 2, 3, and 5 converge to the same minimum while 4 converges to another minimum after full relaxation. In (b) 1 and 2 converge to one minimum while 3 and 4 converge to another minimum.

The presence of an oscillation in the sign of the ...ion and the rapid decay of the oscillation amplitude ...her general in that they occur in both Al and Ni for ...he surfaces studied, as shown in Fig. 5. ...r ordered alloys, surfaces with the same index are ...cessarily identical. For example, layers parallel to ...0) and (100) surfaces of Ni_3Al alternate between ...and 50% Ni. For the (100) surface of Ni_3Al, we find ...: 50% Ni surface is more stable, in agreement with ...electron spectroscopy.[48] Recent LEED data show ...Ni) ≈ −2.8% compared with −2.73% from our ...tion. The first layer was found to ripple (Al out, Ni ...0.002 ± 0.003 nm in the LEED study[48] and by ...nm in the simulation. $Ni_3Al(111)$ has only one ter... ...n (75% Ni). LEED results[48] on this surface show $\Delta d_{1,2}$(Ni) = −0.48%, $\Delta d_{1,2}$(Al) = +2.4%, and a Ni–Al surface rippling of 0.006 nm (Al out, Ni in). For comparison, the present simulation study yields $\Delta d_{1,2}$(Ni) = −0.33%, $\Delta d_{1,2}$(Al) = +3.18%, and a Ni–Al surface rippling of 0.007 nm. Based on these comparisons and similar comparisons on other surfaces (in Ni, Al, Ni_3Al, and NiAl),[18,20] we conclude that the potentials employed provide a good qualitative description of surface relaxation and quantitative accuracy better than 0.01 nm.

While the above results show that the potentials employed in the present study provide reasonable predictions of atomic positions, another requirement of a suitable potential is its ability to accurately predict defect energies

IV. Comparison of the percentage changes in the interlayer spacings (Δd) near the relaxed Al and Ni (110) surfaces.

Al (110)			Ni (110)		
Present work	LEED[46]	LEED[47]	Present work	LEED[46]	MEIS[47]
−10.35	−8.6 ± 0.8	−8.5 ± 1.0	−4.87	−8.7 ± 0.5	−9.0 ± 1.0
+ 3.14	+5.0 ± 1.1	+5.5 ± 1.1	+0.57	+3.0 ± 0.6	+3.5 ± 1.5
− 2.75	−1.6 ± 1.2	+2.2 ± 1.3	−0.86	−0.5 ± 0.7	
+ 1.41	+1 ± 1.3	+1.6 ± 1.6	+0.34		
− 0.61			−0.15		

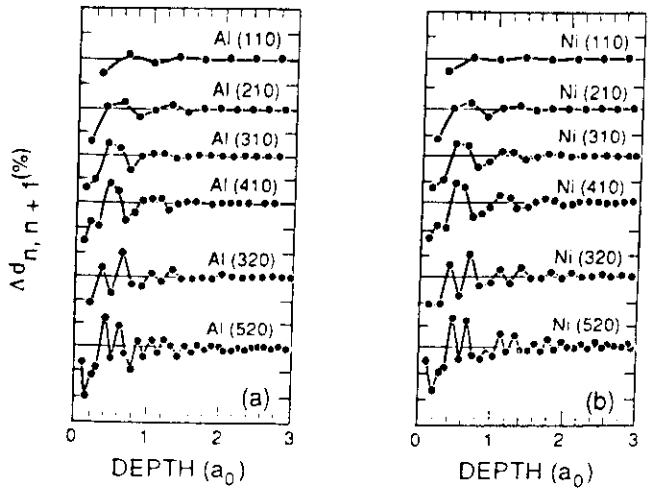


FIG. 5. The surface relaxation, $\Delta d_{n,n+1}$ (%), vs depth for six Al and Ni surfaces. Each tick on the vertical axis corresponds to 10% for the Al surfaces and 5% for the Ni surfaces.

and differences in defect energies. Again, the surface structure literature provides an excellent comparison. In some cases, the (110) surfaces of fcc metals are known to undergo a missing-row 1 × 2 reconstruction.[43] Using local volume potentials of the type described in Sec. II, the difference in energy between the reconstructed and unreconstructed (110) surfaces in Al, Ni, Cu, Pd, Ag, and Au were calculated.[24] The simulations show surface energies ranging between approximately 1000 and 2000 mJ/m^2, with differences in energy between unreconstructed and reconstructed surface typically of order 20 mJ/m^2. Although these differences are only about one percent of the surface energy, only the Au surface is found to reconstruct, in agreement with experiment.[49] Taken together with the data presented in Tables I and II, this result shows the accuracy of these potentials in describing defect energies.

V. EXCESS GRAIN BOUNDARY VOLUME

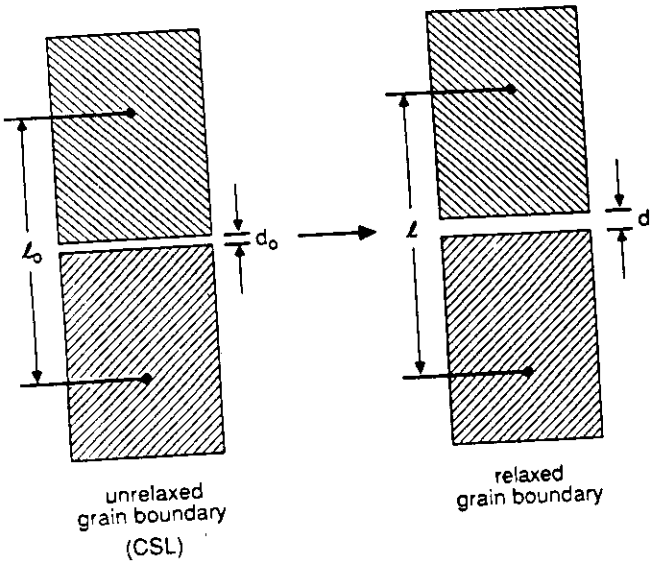
Associated with the atomic rearrangement that occurs upon minimization of the grain boundary energy is a relative displacement of one grain with respect to the other. The relative displacement of the two grains in the direction perpendicular to the grain boundary (i.e., the z-direction) is one measure of the grain boundary expansion or excess volume per unit grain boundary area. This measurement, if made far away from the grain boundary, is not sensitive to the grain boundary strain field, which decays away from the boundary as ze^{-z}.[20,50] We denoted this displacement or shift as Δz (see Fig. 6). Another measure of the grain boundary expansion is the relative z displacement of the two atomic planes closest to the grain boundary (i.e., the pair of planes with the largest spacing after the relaxation). This is a local measure and is denoted as Δz^*. For low angle grain boundaries, the dependence of Δz on relative grain misorientation, θ, may be determined in the framework of the dislocation model. Assuming that there is a net expansion, β, at the dislocation core due to the anharmonicity in the atomic interactions, the dislocation model predicts

$$\Delta z = \beta\rho = 2(\beta/b)\sin(\theta/2) \qquad (10)$$

where ρ is the inverse of the spacing between dislocations in the boundary and b is the magnitude of the dislocation Burgers vector. Unfortunately, we do not have sufficient data at small θ to determine β from the simulations.

These two measures of the grain boundary expansion are plotted vs the misorientation angle, θ, in Figs. 7 and 8 for Ni and Al, respectively. While Δz is seen to vary relatively smoothly with tilt angle, the local expansion, Δz^*, appears erratic. For high angle grain boundaries, Δz is of order one-tenth of a lattice parameter (a_0), while the local expansion is typically around twice that large.

The translation of the upper grain with respect to the lower grain has been measured for grain boundaries in Al by transmission electron microscopy[51] and has been calculated via atomistic (pair potential) simulation techniques by Pond and Vitek.[51] While agreement between simulation and experiment was found for the absolute value of the in-plane translation vector, the magnitude of the translation in the direction perpendicular to the grain boundary differed by a factor of between two and three. This apparent discrepancy is probably attributable to the fact that the simulations measured the local expansion (i.e., Δz^*) while the experiment measured the net expansion of the two grains (Δz). Though experiments were not performed on any



$$\Delta z = \mathcal{L} - \mathcal{L}_0$$
$$\Delta z^* = d - d_0$$

FIG. 6. Schematic drawings of the definition of the separation, d, total expansion, Δz, and grain boundary local expansion, Δz^*.

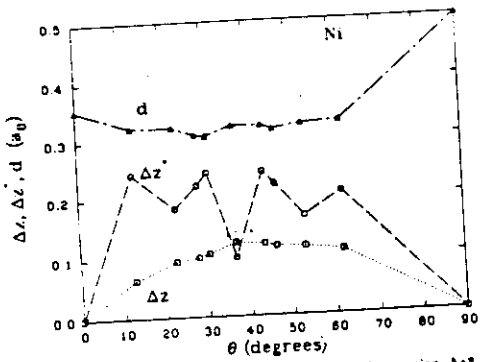


FIG. 7. The separation, d, total expansion, Δz, and local expansion, Δz^*, of Ni [001] tilt grain boundaries.

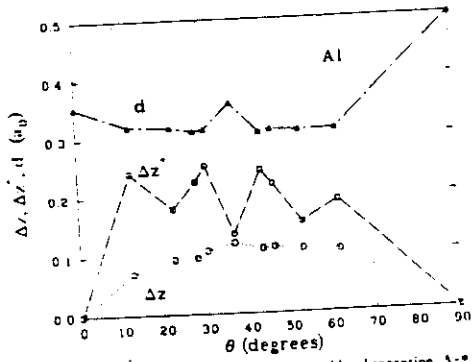


FIG. 8. The separation, d, total expansion, Δz, and local expansion, Δz^*, of Al [001] tilt grain boundaries.

[001] tilt boundaries, Smith, Vitek, and Pond[52] did calculate Δz^* for this series, and a comparison with the present results is shown in Table V. The local volume potential and pair potential values for Δz^* agree quite well, with a typical deviation of less than 5% and a maximum deviation of 19% for the $\Sigma 5(210)$ boundary. The ratio between our Δz^* and Δz values is of order 2, in agreement with the magnitude of the discrepancy between the pair potential simulation and experiment[51] reported above. Table V indicates that the small volume expansion at the grain boundary does not severely violate the constant volume assumption of the pair potentials.[51]

TABLE V. Δz and Δz^* in units of the lattice parameter (a_0) for Ni and Al in the present study and Δz^* from previous calculations on Al by Smith *et al.*[52] Θ is measured from the (110) plane of the top grain to the (110) plane of the bottom grain. The boundary plane indices are also given.

			Present study						Previous simulation[52]
			Ni			Al			Al
Θ	Index	d_0	d	Δz^*	Δz	d	Δz^*	Δz	Δz^*
0.00	(110)	0.354	0.354	0.000	0.000	0.354	0.000	0.000	...
12.68	(540)	0.087	0.322	0.235	0.063	0.321	0.234	0.081	0.23
22.62	(320)	0.139	0.321	0.182	0.094	0.318	0.179	0.094	0.20
28.07	(530)	0.086	0.308	0.222	0.100	0.311	0.225	0.096	0.23
30.51	(740)	0.062	0.306	0.244	0.106	0.314	0.252	0.107	...
36.87	(210)	0.224	0.322	0.098	0.125	0.358	0.134	0.119	0.16
43.60	(730)	0.066	0.308	0.242	0.212	0.308	0.242	0.109	0.24
46.40	(520)	0.093	0.315	0.222	0.117	0.312	0.219	0.111	0.21
53.13	(310)	0.158	0.323	0.165	0.115	0.310	0.152	0.107	0.16
61.93	(410)	0.121	0.327	0.206	0.108	0.311	0.198	0.104	0.20
90.00	(100)	0.500	0.500	0.000	0.000	0.500	0.000	0.000	...

It is also of interest to examine the interplanar spacing at the grain boundary. This interplanar spacing, denoted as d, is plotted in Figs. 7 and 8 and tabulated in Table V. While the spacing between unperturbed atomic planes parallel to the grain boundary (d_0) shows a relatively large variation with θ, d is relatively constant except for the (100) case at 90°.

VI. THE STRAIN FIELD NORMAL TO THE BOUNDARY PLANE

As shown in Fig. 5, the interlayer spacings near a relaxed free surface show oscillations that decay into the bulk. This kind of oscillatory behavior can be shown (Fig. 9) to arise from smoothing of the surface steps and steric interactions between layers.[18,20] Similar atomic relaxation occurs

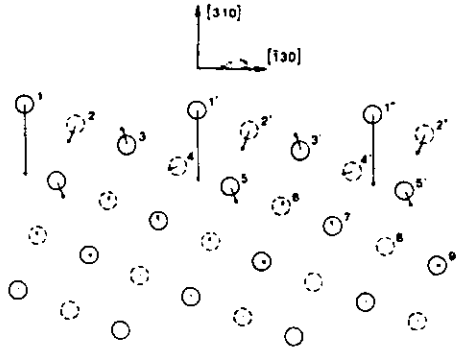


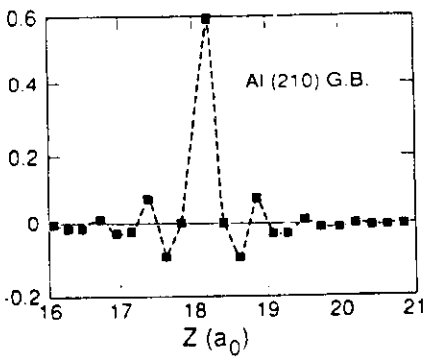
FIG. 9. Schematic indication of the surface smoothing and the steric repulsion model which explains why the surface relaxation occurs in an oscillatory fashion.[18]

in the vicinity of grain boundaries. Figure 10(a) shows the normal strain component (ϵ_{zz}) as a function of distance (z) from the grain boundary plane for Al $\Sigma 5(210)/36.87°$. This plot shows a symmetric oscillatory strain profile that has a maximum at the boundary plane and decays into the bulk. Careful comparisons of strain profiles near the symmetric tilt boundaries and the corresponding surfaces [Al(210) in this case, Fig. 10(b)] show that they share the same oscillation period.[18,20] The decay length of the oscillations also shows excellent correspondence (Figs. 10 and 11). However, the amplitudes of the oscillations are approximately twice as large in the grain boundary case. While in the (210) case, the grain boundary and surface oscillations are nearly 180° out of phase, examination of many grain boundaries shows that this is not a general feature of the relaxation. In fact, the phase of the oscillations in the grain boundary strain profile depends on which local minimum the grain boundary relaxes into (these different minima correspond to different relative shifts of the two grains).

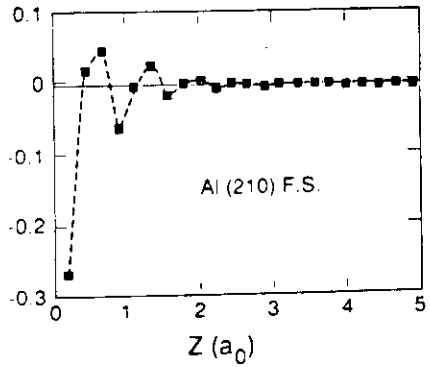
The oscillatory nature of the strain profile may be seen analytically for low-angle grain boundaries. Symmetric low-angle tilt boundaries may be described as an ensemble of parallel coplanar edge dislocations with Burgers vectors along the boundary normal.[4] The strain field associated with such an array of dislocations[30] is simply

$$\epsilon_{zz} = \frac{-b \sin(X)}{4D[\cosh(Z) - \cos(X)]^2} \cdot \left\{ \frac{\cosh(Z) - \cos(X)}{1+\nu} + \frac{Z \sinh(Z)}{1-\nu} \right\} \quad (11)$$

where D is the spacing between dislocations, $Z = 2\pi z/D$, $X = 2\pi x/D$, ν is the Poisson's ratio, and b is the Burgers vector. For low angle boundaries $D \approx b/\theta$. When the z-coordinates of the atoms are tabulated and ordered to measure ϵ_{zz}, we find that as z increases so does the dis-



(a)



(b)

10. The relative deviation of interlayer spacing in z-direction, ϵ_{zz}, function of distance from the (210) tilt grain boundary (a) and from (210) free surface (b) in Al. Note that the oscillation period is three rs in each case.

cement parallel to the boundary, x. This is simply a nsequence of the fact that the atoms are arranged in nes that are not strictly parallel to the grain boundary. nsequently, z is measured along a crystal plane and Δx proportional to Δz. For low angle boundaries, $\Delta x = /\tan(\theta/2)$. Inserting this relation into Eq. (11), we find t ϵ_{zz} is an oscillatory function of z since ϵ_{zz} is an oscilory function of y. The rapid decay of the amplitude of strain oscillations is seen by taking the large Z limit in . (11):

$$\epsilon_{zz} = \frac{-\theta \sin[Z \cot(\theta/2)]}{4(1-\nu)} Z e^{-z} \qquad (12)$$

is expression clearly shows the nature of the observed cillatory behavior and the decay of the amplitude.

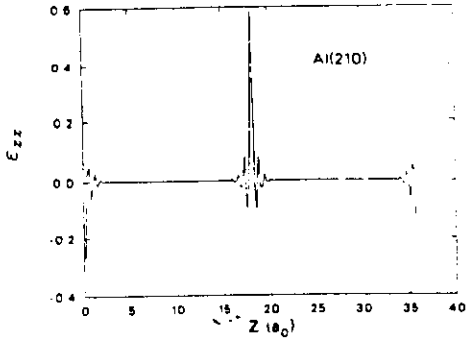


FIG. 11. The composite picture of Fig. 10(a) and (b). Note that the oscillation has an initial opposite sign and the oscillation at the GB is about twice as large as that of the free surface.

While this analysis of the strain oscillations was based on a dislocation picture of low angle boundaries, we expect it also to be valid for high angle boundaries. The difference between the low and high angle cases is in the identity of the dislocations (e.g., Burgers vector) and the dislocation spacing. Furthermore, since the oscillatory strain profile for grain boundaries is very similar to that for surfaces, where dislocations do not necessarily exist, we conclude that the oscillations are simply due to the atomic nature of these crystalline defects.

VII. GRAIN BOUNDARY STRUCTURE AND THE STRUCTURAL UNIT MODEL

A. Grain boundaries in pure Ni and pure Al

Historically, models of grain boundary structure have included the amorphous model[53] dislocation model,[4] CSL model,[5-8] polyhedral model,[11] O-lattice model,[9] and the structural unit model.[10,13] The most recent of these, the structural unit model,[10,13] has been very successful in describing the atomistic structures and energies of grain boundaries. We summarize the main features of this model here, restricting discussion to [001] symmetric tilt grain boundaries in fcc materials for simplicity. For the [001] tilt boundary, the period in the y-direction is only two layers, while in the x-direction it is longer. The structural unit model states[12] that this x-period can be completely characterized by a sequence of structural units. While, in general, more than one type of structural unit appears in this sequence, there are four special angles for which only one unit appears. The angles are 0° (perfect crystal), 36.87°, 53.13°, and 90° (perfect crystal also) and the associated structural units are called A, B, C, and D, respectively, as shown for the relaxed grain boundaries in Fig. 12. These are known as delimiting angles, because any [001] tilt boundary can be completely described using only the structural units corresponding to the two special angles that bound the grain boundary angle in question. For example, the (320) grain boundary at 22.62° is composed of alternating A and B units and the (530) grain boundary at 28.07° is composed of one A and two B units. The A unit ($\theta = 0°$) is in the majority (e.g., $\cdots$ AAAAAAB $\cdots$ at low angles), while close to 36.87° the B unit is dominant. Between 36.87° and 53.13°, the grain boundaries consist of only B and C units, and so on. For a particular angle, the ratio of the density of the two types of unit can be determined *a priori*[12] and is found to be in agreement with simulation results.[13]

In some cases, variations on the perfect A, B, C, and D units can occur; these are indicated by B', C', etc. For example, B can be transformed into B' by inserting an extra atom into the structural unit cell [see Fig. 13(a)]. This B' unit can be found in metastable grain boundary structures at the B delimiting angle (36.87°), and it sometimes occurs in the lowest-energy grain boundary at other angles. For [001] tilt boundaries, B' and C' units have been identified.[13] In principle, even more units of this type can exist (e.g., B", B"', ...). Note that because the B (or B') units begin and end on different atomic layers, they always appear in pairs (though not necessarily adjacent). This step ($\frac{1}{2}$ a_0 in magnitude) is indicated by a dot (e.g., B.B). This is also true for D units. Table VI shows the structural unit sequences for all the tilt boundaries we have studied, along with the sequences obtained by Wang and Vitek[13] using a pair potential description of Cu. The sequences for Ni and Al are the same in every case, and differ from Cu only in the interchange of B and B' in certain cases, including the delimiting angle 36.87°.

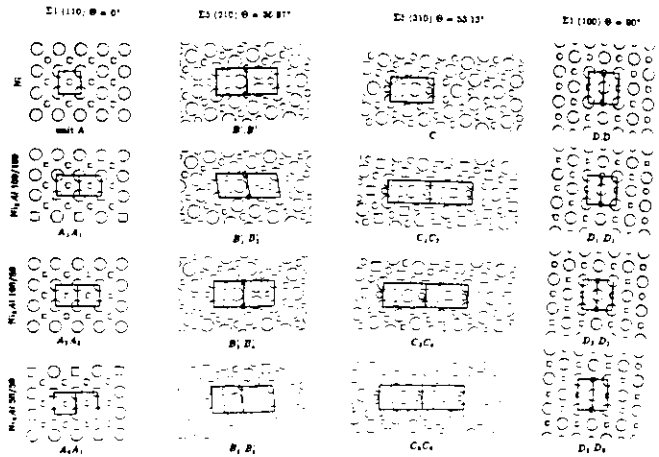


FIG. 12. Structural units observed in the present simulation for the lowest energy configuration in Ni and Ni_3Al.

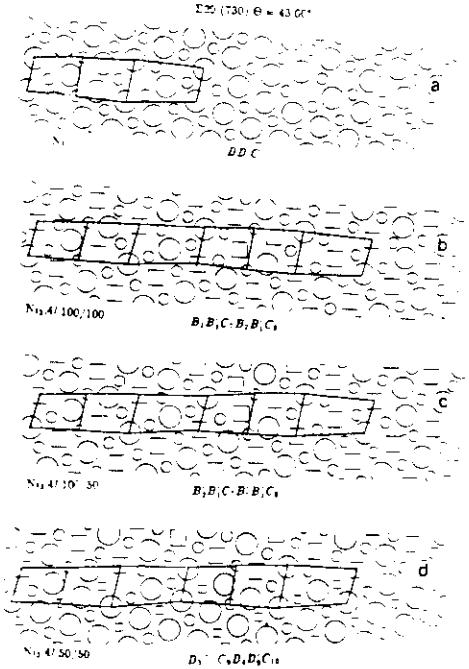


FIG. 13. Structural units in (730) GB for Ni (a), Ni_3Al 100/100 (b), 100/50 (c), and 50/50 (d). The Ni_3Al cells are twice as large as those of Ni. The generic units are not the same for Ni_3Al for different GB chemistry. The 100/100 case has B'B'CB'B'C generic units instead of BB'CBB'C as those of other systems and Ni cases.

TABLE VI. The lowest energy generic unit in the [001] symmetric tilt GB in Ni, Al, and Ni_3Al using "local volume" potentials in this study and the traditional pair potential approach for Cu.[15] Here " indicates the generic unit is the same as immediately to the left in the table and ··· means the data are not available.

					Ni_3Al		
Θ (°)	Index	Cu[15]	Ni	Al	50/50	100/50	100/100
30.51	(740)	ABBB'.ABBB'	AB'B'B.AB'B'B	Al	BB'CBB'C	BC.BC	B'B'CB'B'C
0	(110)	A	"	"	AA	"	"
12.68	(540)	AAAB'.AAAB'	"	"	"	"	"
22.62	(320)	AB.AB	AB'.AB'	"	"	"	"
28.07	(530)	ABB'	AB'B	"	AB'B	"	"
					AB'B		
30.51	(740)	ABBB'.ABBB'	AB'B'B.AB'B'B	"	"	"	"
36.87	(210)	B.B	B'.B	"	"	"	"
43.60	(730)	···	BB'C	"	BB'CBB'C	"	B'B'CB'B'C
46.40	(520)	B'C.B'C	"	"	"	BC.BC	"
53.13	(310)	C	"	"	CC	"	"
61.93	(410)	CD.CD	"	"	"	"	"
90.00	(100)	D.D	"	"	"	"	"

B. Structure of Ni_3Al grain boundary

Application of the structural unit model to the Ni_3Al grain boundaries is more complicated than for the pure elements. For each type of structural unit that can exist for the pure material, there is now an additional classification according to the chemical type of each atom in the unit; this is indicated by a numerical subscript. For example, the structural unit sequence at $\theta = 0°$ (110) is A_1A_1 for the 100/100 grain boundary stoichiometry, A_2A_3 for 100/50, and A_4A_4 for 50/50, as shown in Fig. 12. Units A_1, A_2, A_3, and A_4 differ from pure unit A only by the chemical identity of the atoms and minor geometric distortions (actually A_3 is the same as pure unit A); there are no topological differences. The 36.87° and 53.13° delimiting angles each show six chemical variations of the structural unit, and 90° shows three variations.

Note also that for the 0° and 53.13° (310) delimiting angles, the Ni_3Al grain boundary period is twice that of the pure elements. The level of variation is even more pronounced for the nondelimiting tilt angles. This is exemplified in Fig. 13, where the sequences for the 43.60° (730) boundaries are shown. In going from pure Ni to Ni_3Al, the "B-B-C" sequence is preserved with a doubled period, but the structure is otherwise much more complex. For the three stoichiometries, nine different "B" units ($B_1, B_2, B_3, B_4, B_1', B_2', B_3', B_4', B_5'$) and four "C" units ($C_7, C_8, C_9, C_{10}$) appear. Moreover, the BB'C sequence that occurs for pure Ni and the 50/100 and 50/50 grain boundaries changes to B'B'C for the 100/100 grain boundary.

In general, the structural sequences of grain boundaries in ordered alloys can show a large number of possible variations. For example, the simple Σ5(210)/36.87° boundary which consists of ten atoms in the B'B' sequence can have up to 2^{10} possible structures corresponding to occupying each of the ten atoms sites with either Ni or Al. (Note that this number was substantially smaller in the present study since we considered only perfect crystals on either side of the grain boundary.) This multiplicity becomes even more important as the period of the grain boundary increases. This result suggests that the importance of the degeneracy of grain boundary structures reported previously[34] for pure metals is potentially even more important in the case of alloys.

VIII. GRAIN BOUNDARY ENERGIES

We define the grain boundary energy as

$$\gamma_{gb} = \frac{1}{A}\sum_i \Delta E_i \tag{13}$$

where A is the cross-sectional area of the simulation cell parallel to the grain boundary, and ΔE_i is the change in energy of atom i due to the presence of the grain boundary:

$$\Delta E_i = E_i - E_{t_i}^{bulk} \tag{14}$$

(t_i = Ni or Al). The sum in Eq. (13) is over all atoms whose energy is affected by the proximity of the grain boundary (within the accuracy of the calculation). This includes atoms up to between 5 and 9 lattice parameters on each side of the grain boundary to obtain γ_{gb} to within 0.5 mJ/m^2. For pure Ni or Al, E^{bulk} is simply minus the fcc cohesive energy (−4.450 eV for Ni and −3.360 eV for Al), while for Ni_3Al the atom self-energies for the $L1_2$ crystal must be used (−4.509 eV for Ni and −4.821 eV for Al). The grain boundary energies for the tilt angles considered in the present study are shown in Table VII and Fig. 14. These values correspond to the lowest energy grain boundaries (i.e., the deepest energy minimum) found for each particular misorientation.

For the pure elements, Fig. 14 shows indications of cusps at 36.87° (210), 53.13° (310), and possibly at 28.07° (530). These first two cusp positions correspond to single structural unit, Σ5 boundaries (discussed above) and are in agreement with experimental results on grain boundaries in elemental Al[55] and Cu.[56] Previous pair potential simulations[55] on Al grain boundaries have indicated the presence of cusps at the three positions quoted above as well as 43.60° (730). There is no evidence for a cusp at this position in the present study. The average grain boundary energy (averaged over the data in Table VII, excluding $\theta = 0°$ and 90°) is determined to be 1210 mJ/m^2 for Ni and 341 mJ/m^2 for Al. The corresponding experimental values (with at least 10 ~ 15% errors) are 866 for Ni[57] and 325 for Al.[57] While the simulation values for Al are in excellent agreement with experiment, the discrepancy between the Ni values has not been accounted for. For the grain boundaries consisting of a single structural unit, the lowest energy boundary (there are multiple energy minima) was always found to have a mirror plane normal to the z-direction. However, for multiple structural unit boundaries, this was not always the case.

The most striking feature of the Ni_3Al grain boundary energies is that γ_{gb} varies substantially with the local grain boundary composition. The energy difference corresponding to different choices of the Al sublattice (see Fig. 14) varies from between 50 and 300 mJ/m^2 for the high angle boundaries with an average of 150 mJ/m^2. This variation is approximately 12% of the total grain boundary energy. It is interesting to note that the nonstoichiometric boundaries [i.e., (50/50) Al-rich and (100/100) Ni-rich] have nonzero energies at $\theta = 0°$ and 90°. The 0° (50/50) boundary corresponds to a complex stacking fault which is obtained by removing a 100% Ni (100) plane from the perfect crystal. This introduces Al–Al nearest neighbors which do not exist in the perfect $L1_2$ crystal. As the tilt angle increases from 0°, the atomic relaxations that occur act to lessen the severity of the Al–Al interactions. This can occur by local distortions or by relative shifts of the two grains. Indeed, the shift perpendicular to the grain boundary is found to be the largest for the (50/50) composition, and the excess energy (measured with respect to the (100/50) grain boundary energy) associated with these Al–Al interactions is reduced from 617 mJ/m^2 at 0° to an average of 119 mJ/m^2 for the other orientations. In general, we find that the Al-rich grain boundaries (50/50) have the highest energies and the Ni-rich and stoichiometric grain boundaries have very similar, lower energies (on average the Ni-rich boundaries have the lowest energies). The Ni_3Al grain boundary energies greatly exceed those for pure Al for a composition

TABLE VII. Symmetric tilt [001] grain boundary energies in Ni, Al, and Ni_3Al in (mJ/m^2). The (110) and (100) grain boundaries are not included in the average.

					Ni_3Al		
Θ (°)	Index	Σ	Al	Ni	50/50	100/50	100/100
0	(110)	1	0	0	441	0	17
12.68	(540)	41	278	866	1054	938	904
22.62	(320)	13	330	1109	1351	1155	1101
28.07	(530)	17	338	1198	1388	1303	1208
30.51	(740)	65	365	1253	1393	1321	1305
36.87	(210)	5	351	1278	1396	1213	1329
43.60	(730)	29	373	1353	1467	1434	1417
46.40	(560)	29	370	1347	1484	1397	1406
53.13	(310)	5	315	1221	1468	1247	1166
61.93	(410)	17	349	1261	1407	1334	1294
90.00	(100)	1	0	0	617	0	17
	grain boundary energy		341	1210	1379	1260	1237

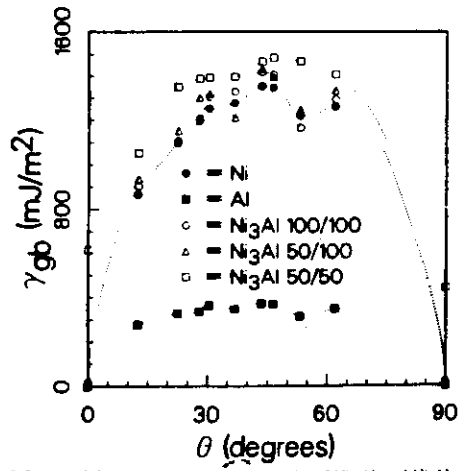


FIG. 14. [001] symmetric tilt boundary energies of Ni, Al, and Ni_3Al as a function of misorientation angle. The dashed lines are only guides for the eye.

weighted Ni–Al average, but are very close to those for pure Ni.

The cusps in the grain boundary energy vs misorientation plot show variations with grain boundary composition. At 36.87° [Σ5(210)], a cusp occurs for Ni, Al, and all three compositions of the Ni_3Al grain boundary, with the deepest cusp occurring for the stoichiometric boundary (100/50). On the other hand, at 53.13° [Σ5(310)] the Ni-rich boundary (100/100) has the deepest cusp, and the pure Al boundary shows no cusp at all. At each of these two angles, the largest cusp is on the order of 200 mJ/m^2 deep, which is approximately 13% of the total boundary energy. Based on the observations, we conclude that the cusps are extremely sensitive to local grain boundary composition.

As mentioned above, the Al-rich grain boundaries have both a larger energy and expansion than do the stoichiometric and Ni-rich boundaries. This relationship between grain boundary energy and grain boundary expansion (Δz) appears to be rather general in the present calculations. This is indicated in Fig. 15 where the grain boundary energies for Ni, Al [Fig. 15(a)] and the three Ni_3Al [Fig. 15(b)] grain boundary compositions are shown. To lowest order, there appears to be a linear relationship between γ_{gb} and Δz, with the proportionality factor depending on the material. The grain boundary energy rises faster with grain boundary expansion in Ni and in Ni_3Al than in Al. Since the expansions have been normalized by the appropriate lattice constants, this result is more indicative of the relatively weaker binding in pure Al than its larger atomic size.

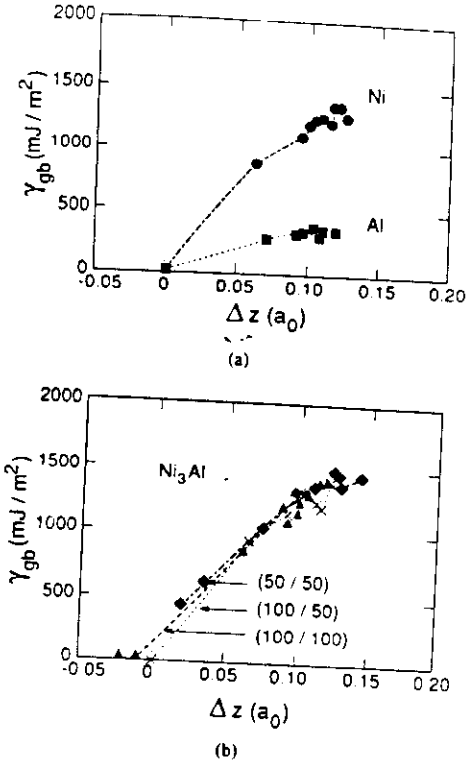


FIG. 15. The [001] tilt boundary energy as a function of total grain boundary expansion (Δz) for Ni, Al (a), and Ni_3Al (b).

IX. GRAIN BOUNDARY COHESION

The Griffith cohesive energy[58] is defined as the energy required to cleave a brittle material without plastic deformation. When the material is cleaved along the grain boundary, two free surfaces are created and the grain boundary is destroyed. Therefore, the grain boundary cohesive energy is defined as

$$\gamma_{coh} = \gamma_{s_1} + \gamma_{s_2} - \gamma_{gb} \quad (15)$$

Note that, even for the symmetric boundaries studied here, the two surface energies, γ_{s_1} and γ_{s_2}, may differ due to the different possible surface compositions. The cohesive energies of the grain boundaries discussed above are tabulated in Table VIII and Fig. 16.

In order to gauge the magnitude of the grain boundary cohesive energies, it is useful to know the cohesive energy of the (100), (110), and (111) planes. For Al, these values are 1710, 1918, and 1648 mJ/m^2, respectively. For Ni, the corresponding values are 3510, 3954, and 3248 mJ/m^2, respectively. Finally, for Ni_3Al these values are 3821, 4119, and 3510 mJ/m^2, respectively. As with the grain boundary energies, the cohesive energies of these perfect crystal planes are lowest for Al while the Ni and the Ni_3Al energies are comparable, with the Ni_3Al values slightly higher. At first glance, this suggests that Al would be more easily fractured than either Ni or Ni_3Al. However, the dominant term in the fracture energies of these materials comes from the plastic work associated with an advancing crack and hence the Griffith energies is not the same as fracture toughness. Nonetheless, there is a strong correlation between the Griffith cohesive energies and the fracture toughness, as shown by McMahon and Vitek.[59]

Consideration of Figs. 14 and 16 shows that the grain boundary energies and the cohesive energies are, crudely, inversely related. This may be most simply seen by the sign of the curvature in these two figures. Furthermore, the cusps in the grain boundary energy vs misorientation plots often correspond to inverted cusps (i.e., local maxima) in the cohesive energy plot. This is especially pronounced for the single structural unit boundaries: 53.13° and 36.87°. As for the cohesive energies of the perfect crystal planes, the Ni and Ni_3Al grain boundaries show higher cohesive energies (Ni_3Al slightly higher) than for pure Al. Not surprisingly then, the Al-rich boundaries generally show lower Griffith cohesive energies than the stoichiometric grain boundary and the Ni-rich boundary.

If the cohesive energies of grain boundaries in Ni_3Al exceed those for pure Ni and pure Al, why are the pure metals so ductile and Ni_3Al intergranularly brittle? Unlike for perfectly brittle materials, the cohesive energy of (microscopically) ductile materials represents only a small fraction of the fracture energy. The majority of the fracture energy in such materials comes from the plastic work associated with the advancing crack. Recent calculations by Hack, Chen, and Srolovitz[60] have demonstrated that the amount of plastic work associated with fracture is strongly controlled by the constitutive behavior of the material (especially the yield stress). This is a consequence of the fact that significantly higher stresses must be applied to Ni_3Al than to the pure materials to achieve a similar amount of deformation. Therefore, we believe that the relative intergranular fracture behavior of Ni and Ni_3Al is dictated more by the plastic properties of the matrix than by the nearly identical cohesive energies.

TABLE VIII. Grain boundary Griffith cohesive energies for symmetric tilt [001] boundary in Ni, Al, and Ni_3Al in (mJ/m^2). The ⟨110⟩ and ⟨100⟩ grain boundaries are not included in the average.

					Ni_3Al		
Θ (°)	Index	Σ	Al	Ni	50/50	100/50	100/100
0	(110)	1	1918	3954	3703	4119	4077
12.68	(540)	41	1690	3206	3186	3302	3336
22.62	(320)	13	1660	3023	2949	3146	3201
28.07	(530)	17	1658	2954	2934	3014	3104
30.51	(740)	65	1633	2905	2929	3002	3019
36.87	(210)	5	1647	2886	2968	3114	2961
43.60	(730)	29	1611	2777	2855	2881	2891
46.40	(520)	29	1604	2765	2818	2909	2904
53.13	(310)	5	1635	2835	2822	3027	3092
61.93	(410)	17	1557	2699	2815	2867	2886
90.00	(110)	1	1710	3510	3191	3821	3814
	average cohesive energy		1663	2894	2920	3029	3044

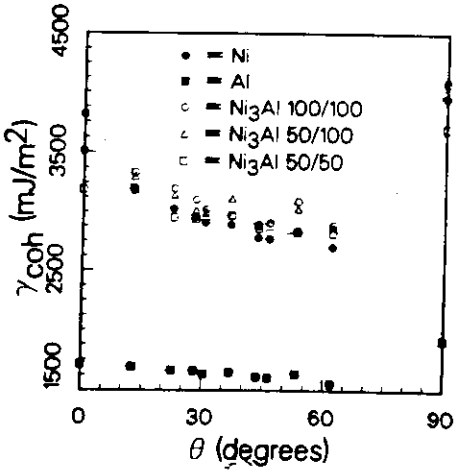


FIG. 16. Griffith cohesive energies as a function of misorientation angle for the [001] tilt GB in Ni, Al, and Ni_3Al. The dashed lines are only guides for the eye.

X. CONCLUSIONS

We have performed atomistic simulations of [001] symmetric tilt boundaries in Ni, Al, and Ni_3Al. The interactions among the atoms were described in terms of "local volume" potentials. For free surface and bulk properties, excellent correspondence with experimental data was obtained. The excess volume associated with grain boundaries was observed to scale with grain boundary energy and was in good agreement with experimental data on the pure metals. The strain in the direction perpendicular to the grain boundary was shown to oscillate and rapidly decay with increasing distance from the grain boundary. This result was explained in terms of an analytical dislocation model of grain boundaries. The structural unit model, which has been very successful in describing the structure of grain boundaries in pure metals, was shown to be of more limited utility in alloys. The energies of grain boundaries in Ni_3Al were shown to be of the same order as in Ni and much greater than in Al. The positions of the cusps in grain boundary energy vs misorientation plots were observed to depend sensitively on the local grain boundary composition. The Al-rich boundaries in Ni_3Al generally exhibited higher energies than the stoichiometric or Ni-rich boundaries. The misorientation dependence of the grain boundary cohesive energy was seen to be roughly

inversely proportional to the grain boundary energies (Al-rich Ni_3Al boundaries lowest). The intergranular brittleness of Ni_3Al was attributed to its plastic properties.

ACKNOWLEDGMENTS

We would like to thank V. Vitek and J. E. Hack for helpful discussions. This work was supported by the Energy Conversion and Utilization (ECUT) Program of the United States Department of Energy.

REFERENCES

[1]*The Nature and Behavior of Grain Boundaries*, edited by H. Hu (Plenum Press, New York, 1972).
[2]*Grain Boundary Structure and Properties*, edited by G. A. Chadwick and D. A. Smith (Academic Press, New York, 1976).
[3]*Grain Boundary Structure and Kinetics*, edited by R. W. Balluffi (ASM, Metals Park, OH, 1980).
[4]W. T. Reed and W. Shockley, Phys. Rev. **78**, 275 (1950).
[5]M. L. Kronberg and F. H. Wilson, Trans. AIME **185**, 501 (1949).
[6]F. C. Frank, *Conf. Plastic Def. of Cryst. Solids* (Mellon Inst., Pittsburgh, PA, 1950).
[7]D. G. Brandon, B. Ralph, S. Ranganathan, and M. S. Wald, Acta Metall. **12**, 813 (1964).
[8]S. Ranganathan, Acta Cryst. **21**, 197 (1966).
[9]W. Bollmann, Phil. Mag. **16**, 363 (1967).
[10]G. H. Bishop and B. Chalmers, Scripta Metall. **2**, 133 (1968).
[11]M. F. Ashby, F. Spaepen, and S. Williams, Acta Metall. **26**, 1647 (1978).
[12]A. P. Sutton and V. Vitek, Phil. Trans. Roy. Soc. (London) **A309**, 1, 37, 55 (1983).
[13]G. J. Wang and V. Vitek, Acta Metall. **34**, 951 (1986).
[14]W. Krakow, J. T. Wetzel, and D. A. Smith, Phil. Mag. **A53**, 739 (1986).
[15]Y. Oh and V. Vitek, Acta Metall. **34**, 1941 (1986).
[16]J. Budai, P. D. Bristowe, and S. L. Sass, Acta Metall. **31**, 699 (1983).
[17]A. F. Voter and S. P. Chen, Mat. Res. Soc. Symp. Proc. **82**, 175 (1987).
[18]S. P. Chen, A. F. Voter, and D. J. Srolovitz, Phys. Rev. Lett. **57**, 1308 (1986).
[19]S. P. Chen, A. F. Voter, and D. J. Srolovitz, Scripta Metall. **20**, 1389 (1986).
[20]S. P. Chen, A. F. Voter, and D. J. Srolovitz, Mat. Res. Soc. Symp. Proc. **82**, 515 (1987).
[21]S. P. Chen, A. F. Voter, and D. J. Srolovitz, Mat. Rcs. Soc. Symp. Proc. **81**, 45 (1987).
[22]K. Aoki and O. Izumi, Trans. JIM **19**, 203 (1978); C. T. Liu, C. L. White, and J. A. Horton, Acta Metall. **33**, 213 (1985).
[23]M. S. Daw and M. I. Baskes, Phys. Rev. Lett. **50**, 1285 (1983); Phys. Rev. B. **29**, 6443 (1984).
[24]A. F. Voter (to be published).
[25]J. K. Norskov and N. D. Lang, Phys. Rev. B **21**, 2131 (1980); M. J. Ston and E. Zaremba, Phys. Rev. B **22**, 1564 (1980).
[26]S. M. Foiles, M. I. Baskes, and M. S. Daw, Phys. Rev. B **33**, 7983 (1986).
[27]J. H. Rose, J. R. Smith, F. Guinea, and J. Ferrante, Phys. Rev. B **29**, 2963 (1984).
[28]C. Kittel, *Introduction to Solid State Physics*, 5th ed. (Wiley, New York, 1976).
[29]*Metal Reference Book*, 5th ed., edited by C. J. Smith (Butterworth's, London, 1976).
[30]*Handbook of Chemistry and Physics*, edited by R. C. Weast (CRC, Boca Raton, FL, 1984).
[31]G. Simons and H. Wang, *Single Crystal Elastic Constants and Calculated Aggregate Properties* (MIT Press, Cambridge, MA, 1977).
[32]R. W. Balluffi, J. Nucl. Materials **69**, 240 (1978).
[33]J. S. Koehler, in *Vacancies and Interstitials in Metals*, edited by A. Seeger, D. Schumacher, W. Shilling, and J. Diehl (North Holland, Amsterdam, 1970), p. 175.
[34]J. O. Noell, M. D. Newton, P. J. Hay, R. L. Martin, and F. W. Bobrowicz, J. Chem. Phys. **73**, 2360 (1980).
[35]K. P. Huber and G. Herzberg, *Constants of Diatomic Molecules* (Van Nostrand Reinhold, New York, 1979).
[36]S. Stassis, Phys. Stat. Soli. A **64**, 335 (1981).
[37]R. Hultgren, P. D. Desai, D. T. Hawkins, M. Gleiser, and K. K. Kelley, *Selected Values of the Thermodynamic Properties of Binary Alloys* (ASM, Metals Park, OH, 1973).
[38]M. H. Yoo, private communication, values from Ref. 33 were scaled to $T = 0$ K according to values in K. Ono and R. Stern, Trans. AIME **245**, 171 (1969).
[39]T. M. Wang, M. Shimotomai, and M. Doyama, J. Phys. F **14**, 37 (1984).
[40]P. Veyssiere, J. Doun, and P. Beauchamp, Phil. Mag. A **51**, 469 (1985).
[41]R. C. Pond and V. Vitek, Proc. Roy. Soc. (London) A **357**, 453 (1977).
[42]G. J. Wang, A. P. Sutton, and V. Vitek, Acta Metall. **32**, 1093 (1984).
[43]*The Structure of Surfaces*, edited by M. A. von Hove and S. Y. Tong, Springer Series in Surface Sciences (Springer, Berlin, 1985), Vol. 2.
[44]H. B. Nielsen, J. N. Anderson, L. Petersen, and D. L. Adams, J. Phys. C **14**, L1113 (1982), and J. Phys. C **17**, 173 (1985).
[45]J. R. Noonan and H. L. Davis, Phys. Rev. B **29**, 4349 (1984).
[46]D. L. Adams, L. E. Petersen, and C. S. Sorensen, J. Phys. C **18**, 1753 (1985).
[47]S. M. Yalisove, W. R. Graham, E. D. Adams, M. Copel, and T. Gustafsson, Surf. Sci. **171**, 400 (1986).
[48]D. Sondericker, F. Jona, and P. M. Marcus, Phys. Rev. B **33**, 900 (1986), and Bull. Am. Phys. Soc. **31**, 325 (1986).
[49]I. K. Robinson, Phys. Rev. Lett. **50**, 1145 (1983).
[50]J. P. Hirth and J. Lothe, *Theory of Dislocations* (McGraw-Hill, New York, 1968), p. 670.
[51]R. C. Pond and V. Vitek, Proc. Roy. Soc. (London), A **357**, 453 (1977).
[52]D. A. Smith, V. Vitek, and R. C. Pond, Acta Metall. **25**, 475 (1977).
[53]W. Rosenhain and J. C. W. Humphrey, J. Iron Steel Inst. **87**, 219 (1913).
[54]Y. Oh and V. Vitek, Acta Metall. **34**, 1941 (1986).
[55]G. C. Hasson and C. Goux, Scripta Metall. **5**, 889 (1971).
[56]N. A. Gjostein and F. N. Rhines, Acta Metall. **7**, 319 (1959).
[57]L. E. Murr, *Interfacial Phenomena in Metals and Alloys* (Addison-Wesley, Reading, MA, 1975).
[58]A. A. Griffith, Phil. Trans. Roy. Soc. A **221**, 163 (1920).
[59]C. J. McMahon and V. Vitek, Acta Metall. **27**, 507 (1979).
[60]J. E. Hack, S. P. Chen, and D. J. Srolovitz, Acta Metall. (to be published); J. E. Hack, D. J. Srolovitz, and S. P. Chen, Scripta Metall. **20**, 1699 (1986).

Journal of
MATERIALS
RESEARCH

Investigation of the effects of boron on Ni_3Al grain bound by atomistic simulations

S. P. Chen, A. F. Voter, R. C. Albers, A. M. Boring, and P. J. Hay
Los Alamos National Laboratory, Los Alamos, New Mexico 87545

pp. 955–970

Volume 5, Number 5
May 1990

Published by the Materials Research Society

Investigation of the effects of boron on Ni_3Al grain boundaries by atomistic simulations

S. P. Chen, A. F. Voter, R. C. Albers, A. M. Boring, and P. J. Hay
Los Alamos National Laboratory, Los Alamos, New Mexico 87545

(Received 19 June 1989; accepted 16 January 1990)

A series of simulations has been performed on grain boundaries in Ni and Ni_3Al with and without boron doping using embedded atom-style potentials. A new procedure of obtaining "reference" data for boron related properties from electronic band structure calculations has been employed. Good agreement with existing experimental structural and energetic determinations was obtained. Boron is found to segregate more strongly to grain boundaries than to free surfaces. Adding boron to grain boundaries in Ni and Ni_3Al increases their cohesive strength and the work required to pull apart the boundary. This effect is much more dramatic for Ni-rich boundaries than for stoichiometric or Al-rich boundaries. In some Ni-rich cases, adding boron increases the cohesive strength of the boundary to such an extent that the boundaries become stronger than the bulk. Bulk Ni_3Al samples that are Ni-rich produce Ni-rich grain boundaries. The best cohesive properties of Ni_3Al grain boundaries are obtained when the boundary is Ni saturated and also with boron present. Boron and nickel are found to cosegregate to the grain boundaries.

I. INTRODUCTION

The pronounced effect of grain boundaries on the physical properties of materials has motivated many studies with the aim of understanding the structure, energetics, and properties of grain boundaries. While significant experimental and theoretical progress[1,2] has been made in understanding boundaries in pure systems, the understanding in alloy systems is much less developed.[3,4]

In this paper we present our recent results on atomistic simulations of grain boundaries in the ordered intermetallic compound Ni_3Al. The brittle nature of this material and many other $L1_2$ compounds arises from the marked propensity toward intergranular fracture, which contrasts with the ductile nature of single crystals of Ni_3Al. While Ni_3Al precipitate has been widely used as the γ'-strengthener in Ni-based superalloys,[5] these brittle characteristics have precluded the use of Ni_3Al itself as a structural material, despite the many attractive properties of the single crystal, such as low diffusivity and enhanced yield strength at high temperatures.[6] The discovery that microalloying Ni_3Al with small amounts of boron (B) significantly reduces the tendency for intergranular fracture[7] has stimulated a renewed interest in intermetallic compounds.[8,9] Extensive recent studies by Liu and coworkers[10] have examined the role of boron in Ni_3Al in detail and have established that boron segregates preferentially to the grain boundaries where it has been detected by Auger analysis.[11] The influence of boron was also found to depend sensitively on the Ni–Al ratio. With stoichiometric or Al-rich Ni_3Al the boron is ineffective in improving ductility, while in Al-poor samples (between 23 and 25% Al) the boron dramatically improves the ductility of polycrystalline Ni_3Al.

In this study we examine the role of boron at grain boundaries in Ni_3Al and compare the results of the present calculations with our earlier studies of boundaries in Ni, Al, and Ni_3Al. Preliminary accounts have been given elsewhere.[12–14] The atomistic simulations employ potentials we have developed[15,16] using an approach based on the embedded atom method.[17] The potentials for Ni, Al, and Ni_3Al were determined by fitting to known thermodynamic and other physical properties of the bulk materials.[3,15,16] For B-containing systems, where there is considerably less experimental data, we have used the results of Linear Muffin Tin Orbital (LMTO) density functional calculations[18–20] to derive the potentials. The electronic structure calculations have been compared to available experimental results on Ni and Al-containing systems to assess the reliability of the LMTO approach.

The paper is organized as follows. The methodology and results of the LMTO calculations are discussed in Sec. II, and the procedure for deriving the interatomic potentials is described in Sec. III. The remainder of the paper is devoted to the atomistic simulations of grain boundaries with and without B atoms present.

II. ELECTRONIC STRUCTURE CALCULATIONS

All of the electronic calculations were done using the Linearized Muffin Tin Orbital (LMTO) method[18–20] with the combined correction terms added. The self-consistent calculations employed the local density approximation with a Barth–Hedin exchange-correlation potential and included all relativistic effects except spin orbit coupling. The final set of compounds included Ni_3Al, fcc Ni, fcc Al, fcc and bcc B, NiAl, NiB, AlB, Ni_3B, Al_3B, and Ni_3Al with B interstitials. All of the XY-type compounds were for the CsCl (B2) crystal structure and all of the X_3Y-type were for the Cu_3Au ($L1_2$) crystal structure.

In Table I some of the results of these calculations are given and compared with experiment (when available). As can be seen, the calculated lattice constant a_0, bulk modulus B, and cohesive energy E_{coh} are in very good agreement with the experimental data.[3,15,16] For all of the Ni systems, the basis set for the calculations included angular momentum components up to $f(l = 3)$ and for all the other systems up to $d(l = 2)$. The lattice constant was obtained by performing the calculations at several lattice spacings and finding the one with the lowest total energy (zero pressure). The cohesive energy was obtained from the total energy differences between the solid and the atom: $E_{coh} = E_{XTAL} - E_{ATOM}$. In this calculation the core wave functions in the crystal are frozen at their atomic values so that the total energy differences involve only the energy differences of the valence electrons. The calculated energy differences are average-of-configuration energies and we have corrected the atomic values to give the lowest term value since that is the true atomic ground state. It is these corrected values that are presented in Table I. Finally, the bulk modulus was a numerical approximation to the expression:

$$B = -\Omega \frac{dP}{d\Omega},$$

where Ω and P are atomic volume and pressure, respectively. These differences were calculated from points bracketing the calculated equilibrium lattice constant.

TABLE I. LMTO results.

Al (fcc)	LMTO	Expt.	Dev.
a_0 (Å)	4.0345	4.0498	0.4%
B (Mbar)	0.82	0.79	3.0%
E_{coh} (eV)	3.74	3.36	9.0%
Ni (fcc)	**LMTO**	**Expt.**	**Dev.**
a_0 (Å)	3.5141	3.5198	0.2%
B (Mbar)	1.99	1.80	9.5%
E_{coh} (eV)	4.45	4.45	0.0%
Ni_3Al (Cu_3Au)	**LMTO**	**Expt.**	**Dev.**
a_0 (Å)	3.571	3.5667	0.1%
B (Mbar)	1.78	1.64	7.9%
E_{coh} (eV)	4.79	4.57	4.6%
	B (fcc)	B (bcc)	
a_0 (Å)	2.9087	2.3423	
E_{coh} (eV)	5.33	4.95	
	Al_3B (Cu_3Au)	Ni_3B (Cu_3Au)	
a_0 (Å)	3.8048	3.3904	
B (Mbar)	0.93	2.56	
E_{coh} (eV)	4.0	4.8	
fcc Al or Ni + Octahedral B Interstitial (BI)			
	fcc Al + BI	fcc Ni + BI	
a_0 (Å)	4.52	3.79	
E_{coh} (eV)	2.5	3.7	

III. INTERATOMIC POTENTIALS

The interatomic potentials employed in the study of boron in Ni_3Al are based on the "local volume" potentials as described by Voter and Chen.[15,16] This type of potential consists of a pairwise interaction augmented by a term, for each atom, that depends on the local atomic density. The limitations of a constant-volume pair potential are thus circumvented, as each atom senses its own local "volume". Major defects, such as free surfaces, grain boundaries, and crack tips, can thus be modeled. This approach (also called the embedded atom method[17] or EAM) has been shown to provide a good description for a variety of systems[3,17,21–26] and is especially suited to the description of fcc metals. Potentials for the Ni_3Al system have also been presented by Foiles and Daw[27,28] using such a prescription.

We present here the potentials used for the Ni–Al–B system derived according to our own EAM approach. For the pure metals Ni and Al this approach is quite similar to the original method of Daw and Baskes,[17] while for Ni_3Al there are significant differences in the methods used to derive the potential compared to those of other investigators.[27,28] The potential for the Ni–Al subset is exactly as described previously,[3,15,16] and has been employed in studies of Ni, Al, and Ni_3Al grain boundaries[3,16,23] and surfaces.[16,24,25] We have augmented this Ni_3Al potential to include B–B, Ni–B, and Al–B interactions. It is important to note that the present potential does not explicitly take into account angular terms. However, we know of no other potential for the Ni–Al–B system, and we expect this potential to be significantly better than a pair potential. Moreover, the EAM description of boron at the grain boundaries in Ni_3Al may be a reasonable one because experimentally the boron is present at sufficiently low concentrations (~10% at GBs and ~1% in the bulk) that there will be relatively few B–B interactions.

Due to the complex structure of rhombohedral boron,[29] we simplify the fitting procedure by treating boron as an fcc material. Electronic band structure results using the linearized muffin tin orbital (LMTO)

approximation as described in the previous section are used instead of experimental data in fitting both pure B and the Ni–B and Al–B cross potentials. In this way, the fitting procedure for B closely parallels that of Ni and Al. We begin by describing the EAM potentials for pure Ni, Al, and B.

A. Potentials for pure Ni, Al, and B

The EAM energy of an n-particle homonuclear system is given by

$$E = \frac{1}{2}\sum_{i(\neq j)}^{n}\phi(r_{ij}) + \sum_{i}^{n}F(\bar{\rho}_i), \tag{1}$$

where r_{ij} is the distance between atoms i and j, ϕ is a pairwise interaction potential, F is the embedding function, and $\bar{\rho}_i$ is the density at atom i due to all its neighbors,

$$\bar{\rho}_i = \sum_{j(\neq i)}^{n}\rho(r_{ij}). \tag{2}$$

$\bar{\rho}_i$ thus acts as a sensor for the local atomic density about atom i, and $F(\bar{\rho}_i)$ provides the volume-dependent, many-body contribution to the energy. The pairwise interaction is taken to be a Morse potential,

$$\phi(r) = D_M\{1 - \exp[-\alpha_M(r - R_M)]\}^2 - D_M, \tag{3}$$

where the three parameters, D_M, R_M, and α_M, define the depth, distance to the minimum, and a measure of the curvature near the minimum, respectively. The density function, $\rho(r)$, is taken as

$$\rho(r) = r^6[e^{-\beta r} + 2^9 e^{-2\beta r}], \tag{4}$$

where β is an adjustable parameter. Physical justification for the forms in Eqs. (3) and (4) is given elsewhere.[15] Motivated by Rose *et al.*[30] and following Foiles *et al.*,[21,31] we define $F(\bar{\rho})$ by requiring that the energy of the fcc crystal is given by

$$E_{fcc}(a^*) = -E_{coh}(1 + a^*)e^{-a^*} \tag{5}$$

as the lattice constant is varied. Here E_{coh} is the fcc cohesive energy (per atom) and a^* is a reduced lattice constant defined by

$$a^* = (a/a_0 - 1)/(E_{coh}/9B\Omega)^{1/2}, \tag{6}$$

where a is the lattice constant, a_0 is the equilibrium lattice constant, B is the bulk modulus, and Ω is the equilibrium atomic volume. Thus, knowing E_{coh}, a_0, and B, the embedding function is defined by requiring that the crystal energy from Eq. (5) match the energy from Eq. (1) for all values of a^*. By fitting $F(\bar{\rho})$ in this way, the potential is appropriate for a large range of densities.

To ensure that the interatomic potential and its first derivatives are continuous, $\phi(r)$ and $\rho(r)$ are smoothly cut off at $r = r_{cut}$ by using

$$f_{smooth}(r) = f(r) - f(r_{cut}) + \left(\frac{r_{cut}}{m}\right)\left[1 - \left(\frac{r}{r_{cut}}\right)^m\right]\left(\frac{df}{dr}\right)_{r=r_{cut}} \tag{7}$$

where $f(r) = \phi(r)$ or $\rho(r)$ and $m = 20$. r_{cut} is used as a parameter in the fitting procedure, and $F(\bar{\rho})$ is also cut off smoothly.

We now describe the fitting procedure for fcc Ni, Al, and B. Because of the way $F(\bar{\rho})$ is determined, the potential always gives a perfect fit to the experimental values for a_0, E_{coh}, and B for any choice of $\phi(r)$ and $\rho(r)$ (for boron, the reference values are from LMTO calculations). The five parameters, R_m, D_m, α_M, β, and r_{cut}, are optimized by minimizing the root-mean-square deviation (χ_{rms}) between the calculated and reference properties of the material. For Ni and Al, these properties are the three cubic elastic constants (C_{11}, C_{12}, and C_{44}), the vacancy formation energy (ΔE_v^f), and the bond length (R_e) and bond energy (D_e) of the diatomic molecule. Also, the hcp and bcc crystal structures are required to be less stable than fcc. For boron, the properties are a_0, E_{coh}, and B for the fcc and bcc crystal (from LMTO calculations), a_0 and E_{coh} for an fcc crystal with every fourth atom missing (from LMTO calculations—in lieu of a vacancy formation energy), and R_e and D_e for diatomic boron (from experiment). The best fit for each element is shown in Table II, and the resulting parameter sets are shown in Table III.

B. The cross potentials

For a general alloy system, Eq. (1) is written as

$$E = \frac{1}{2}\sum_{i\neq j}^{n}\phi_{t_i t_j}(r_{ij}) + \sum_{i}^{n}F_{t_i}(\bar{\rho}_i), \tag{8}$$

with

$$\bar{\rho}_i = \sum_{j(\neq i)}^{n}\rho_{t_j}(r_{ij}), \tag{9}$$

where the subscripts t_i and t_j indicate atom types. The functions $\phi_{NiNi}(r)$, $\phi_{AlAl}(r)$, $\rho_{Ni}(r)$, $\rho_{Al}(r)$, $\rho_B(r)$, $F_{Ni}(\bar{\rho})$, $F_{Al}(\bar{\rho})$, and $F_B(\bar{\rho})$ are known from the fits to pure Ni, pure Al, and pure B. To treat the Ni–Al–B alloy, $\phi_{NiAl}(r)$, $\phi_{NiB}(r)$, and $\phi_{AlB}(r)$ need to be determined. These are each taken as Morse functions, with a variable cutoff distance, just as for the pure elements. In addition to these ($4 \times 3 = 12$) parameters, five other parameters are available to aid in fitting the experimental alloy properties. These arise from the invariance of Eq. (1) with respect to scaling of $\rho(r)$, and the invariance of Eq. (1) under addition of a linear term to $F(\bar{\rho})$. These two types of transformation affect the alloy energy, but leave the pure-element energy unchanged. Using pure boron as an example, the scaling transformation can be written as

$$\rho_B(r) = S_B\rho_B(r) \tag{10}$$

$$F_B(\bar{\rho}) = F_B(\bar{\rho}/S_B), \tag{11}$$

where the prime indicates the transformed function. The other transformation is

$$F_B(\bar{\rho}) = F_B(\bar{\rho}) + g_B\bar{\rho} \tag{12}$$

$$\phi_{B\text{-}B}(r) = \phi_{BB}(r) - 2g_B\rho_B(r). \tag{13}$$

[The definition of g_B depends on which transformation is applied first. We apply transformations (12) and (13) *before* the scaling of (10) and (11).] Recognizing that only two of the three $\rho(r)$ functions can be independently scaled, we obtain a total of five transformation parameters: S_{Al}, S_B, g_{Ni}, g_{Al}, and g_B.

The fits are performed in the following way. First, S_{Al}, g_{Ni}, and g_{Al} and the four parameters defining $\phi_{NiAl}(r)$ are optimized by fitting to a variety of experimental data on $L1_2$ Ni_3Al and B2 NiAl.[3,15] These quantities are the Ni_3Al lattice constant, cohesive energy, elastic constants, ordering energy (ΔE_{ord}), vacancy formation energy, (111) and (100) antiphase boundary (APB) and superlattice intrinsic stacking fault (SISF) energies,

TABLE III. Potential parameters for fcc Ni, Al, and B, optimized from data in Table II.

	Ni	Al	B
D_M (eV)	1.5335	3.7760	0.7182
R_M (Å)	2.2053	2.1176	1.6517
α_M (Å^{-1})	1.7728	1.4859	3.1915
β (Å^{-1})	3.6408	3.3232	2.0108
r_{cut} (Å)	4.7895	5.5550	4.3716

TABLE II. Properties used in fit to fcc Ni, Al, and B. Calculated values of a_0, E_{coh}, and B match experiment due to the way $F(\rho)$ is determined.

	Ni		Al		B	
Property	Expt.	Calc.	Expt.	Calc.	LMTO	Calc.
a_0 (Å)	3.52	[3.52]	4.05	[4.05]	2.91	[2.91]
E_{coh} (eV)	4.45	[4.45]	3.36	[3.36]	5.33	[5.33]
B (10^{12} erg/cm^3)	1.81	[1.81]	0.79	[0.79]	2.02	[2.02]
C_{11} (10^{12} erg/cm^3)	2.47	2.44	1.14	1.07		
C_{12} (10^{12} erg/cm^3)	1.47	1.49	0.619	0.652		
C_{44} (10^{12} erg/cm^3)	1.25	1.26	0.316	0.322		
ΔE_v^f (eV)	1.60	1.60	0.75	0.73		
D_e (eV)	1.95	1.94	1.60	1.54	1.59	1.62
R_e (Å)	2.2	2.23	2.47	2.45	3.08	3.27
bcc a_0 (Å)					2.34	2.33
bcc E_{coh} (eV)					4.95	5.31
bcc B (10^{12} erg/cm^3)				2.11		
fcc (3/4) a_0 (Å)					2.12	
fcc (3/4) E_{coh} (eV)					2.78	2.65
χ_{rms} (%)	0.75			3.85	5.78	5.66
						4.12

and the lattice constant and cohesive energy of B2 phase NiAl (CsCl structure). Next, S_B and g_B are optimized by fitting to LMTO data on $L1_2$ phase Ni_3B and B2 phase NiB, and $\phi_{AlB}(r)$ is optimized using LMTO data on $L1_2$ phase Al_3B and B2 phase AlB. In each case, variable fitting strengths are employed in an effort to best overall fit. For the Al–B fit, r_{cut} was fixed at 3.0 Å and $D_e = 1.60$ A was included to help control the fit. Tables IV and V show the reference and calculated properties for Ni–Al, Ni–B, and Al–B, and Table VI gives the optimized parameters.

C. Other experimental comparisons

Some additional comparisons of the Ni, Al, and B potentials derived according to the above proce-

TABLE IV. Metal properties used to fit the Ni–Al cross potential. Superscripts are the experimental references.

Ni_3Al properties	Expt.	Calc.
a_0 (Å)	3.567	3.573
E_{coh} (eV)	4.57	4.59
C_{11} (10^{12} erg/cm^3)	2.30	2.46
C_{12} (10^{12} erg/cm^3)	1.50	1.37
C_{44} (10^{12} erg/cm^3)	1.31	1.23
ΔE_v^f (eV)	1.6 ± 0.2	1.64(Ni)
		1.87(Al)
SISF(111) (mJ/m^2)	10 ± 5	13
APB(100) (mJ/m^2)	140 ± 14	83
APB(111) (mJ/m^2)	180 ± 30	142
B2 NiAl properties		
a_0 (Å)	2.88	2.87
E_{coh} (eV)	4.51	4.38

TABLE V. Properties used to fit the Ni-B and Al-B cross potentials.

	Property	LMTO	Calculated
Ni_3B ($L1_2$)	a_0 (Å)	3.39	3.39
	E_{coh} (eV)	4.80	4.80
	B (10^{12} erg/cm^3)	2.56	2.56
NiB (B2)	a_0 (Å)	2.55	2.58
	E_{coh} (eV)	5.30	5.30
Al_3B ($L1_2$)	a_0 (Å)	3.805	3.909
	E_{coh} (eV)	3.72	3.92
	B (10^{12} erg/cm^3)	9.27	9.22
AlB (B2)	a_0 (Å)	2.96	2.96
	E_{coh} (eV)	4.24	4.27

ure were made with two known experimental borides, exagonal AlB_2 and orthorhombic Ni_4B_3. The pre-icted lattice constants and cohesive energies for each oride compound are compared with the observed xperimental values in Table VII. It should be em-hasized that the experimental information on these ompounds was *not* used in the actual determination f any of the potentials.

In general, the agreement between the theoretically redicted values with experiment is reasonably good ypically about 5 to 7% error) with the exception of the lattice constant corresponding to the interlayer separa-on in AlB_2, which is overestimated in the simulations 39%. The predicted cohesive energies are within 7% the experimentally measured values. Considering at no angular dependence has been incorporated into e form of the potential for B, these results indicate at the potentials are providing a reasonably realistic scription of Ni–Al–B materials, provided that the ncentration of B remains relatively low.

ATOMISTIC SIMULATIONS

Intermetallic compounds generally suffer brittle acture under tensile stress at room temperature.[32] his phenomenon is generally thought to be related to insufficient number of slip systems (as in B2, DO_{19},

BLE VI. Optimized cross potential parameters.

	ϕ_{Ni-Al}	ϕ_{Ni-B}	ϕ_{Al-B}
(eV)	3.0322	0.28223	0.12949
(Å)	2.0896	2.3149	2.8876
(Å^{-1})	1.6277	2.4852	1.3904
(Å)	5.4639	2.8181	3.0000
	Ni	Al	B
	1.0	0.61723	3.4936×10^{-1}
eV Å^3)	6.51451	−0.22050	−0.078785

TABLE VII. Comparison of unit cell structural parameters predicted from potentials derived in this work with experimental values. These crystal properties were not included in the determination of the potentials.

	Property	Expt.[a,b]	Calculated
AlB_2	a (Å)	3.006	2.85
	c (Å)	3.252	4.52
	E_{coh}	5.298	5.23
Ni_4B_3	a (Å)	11.954	12.382
	b (Å)	2.9815	2.969
	c (Å)	6.5684	7.044
	E_{coh}	5.54	5.152

[a]*Pearson's Handbook of Crystallographic Data for Intermetallic Phases*, P. Villars and L. D. Calvert (ASM, Metals Park, OH, 1985).
[b]*Smithells Metal Reference Book*, edited by E. A. Brandes (Butterworth's, London, 1983).

DO_{22}, and other complicated structures),[6] or due to the weak cohesion of the grain boundary (GB) as in $L1_2$ Ni_3X cases, where X = Al, Ga, Si, and Ge.[33] The former classes of materials tend to fail by cleavage and the latter fail intergranularly, but all of them have very limited elongation before fracture. This brittleness prevents practical applications of these alloys despite their high strength, corrosion resistance, low diffusivity, and other desirable high-temperature properties.[6,8,9] These intermetallic compounds have attracted intensive attention recently, stimulated by the recent works of Aoki and Izumi,[7] and Liu *et al.*,[10] showing that Ni_3Al can be ductilized by adding small amount of borons in Ni-rich samples. Addition of boron results in larger elongation to fracture and transition from intergranular failure to transgranular failure. This microalloying effect of boron works for Ni-rich samples for Ni_3Al, Ni_3Ga, and Ni_3Si, but not for Ni_3Ge.[33] It has also been found that addition of carbon helps in Ni_3Si but not in Ni_3Al.[34] Later studies showed beryllium in Ni_3Al may also help to ductilize the GB.[35] Therefore, understanding the basic mechanism underlying the beneficial boron and stoichiometric effect may be useful in improving other systems which suffer brittle fracture problems.

The calculation cell used in the present study extended at least 18 lattice parameters on either side of the grain boundary. In the z-direction perpendicular to the grain boundary, the cells were terminated by free surfaces, and periodic boundary conditions were employed in the grain boundary plane (x, y directions). Using an energy gradient method, the structure of the grain boundary was relaxed to the minimum energy configuration until the maximum forces on any atom were less than 5×10^{-3} eV/A.

A. Grain boundaries in pure Ni and Ni_3Al

In order to understand the boron effect in Ni_3Al, we have to study theoretically why pure polycrystalline

Ni_3Al is intrinsically intergranularly brittle.[36] Two possible reasons have been suggested for the intrinsic intergranular brittleness of the GB in Ni_3Al. The first possibility[10] is that the GB is less cohesive in Ni_3Al than in Ni. The second possibility is that GB is a stronger obstacle for the dislocations to pass through, due to the long-range order of the $L1_2$ structure.[37]

Intrinsic brittleness in these materials has been correlated by various investigators with such properties as valence difference,[38] electronegativity difference,[39] size difference,[40] ordering temperature,[41,42] and difference in yield strength.[42] We have performed atomistic simulations to determine some of the basic mechanisms of the brittleness of the Ni_3Al GB. Here we summarize the major findings of these studies[3,4,16] together with some new results.

The first important result is that the cohesive energies of Ni_3Al and Ni grain boundaries are basically the same. The ratio of the cohesive energy of the GB with respect to the cohesive energy of the (111) surface is about 0.9, as shown in Table VIII. From this, we conclude that the difference between the mechanical behavior (ductile versus intergranular brittleness) between Ni and Ni_3Al grain boundaries is related to the plastic response of the materials. Initial steps to estimate the difference in plastic work are presented elsewhere.[42,43]

To study the mechanical properties of the grain boundaries under applied stress, we have calculated the maximum stress needed to cleave the crystal along perfect planes, e.g., (100), (110), and (210), and then compared it to the calculated maximum stress needed to pull apart the grain boundaries. This maximum stress is a quantity which has been shown recently by Hack, Chen, and Srolovitz[43] to control the plastic work. The *frozen maximum stress* is defined as the maximum stress (Fig. 1) required to separate the two grains by cleaving between the two layers closest to the grain boundary mid-plane without letting atoms on each side relax. The value calculated in this way is presumably the upper bound for the true maximum stress. The

TABLE VIII. Bulk cohesive energies and average GB cohesive energies (γ_{coh}) in Ni, Al, and Ni_3Al. The cohesive energies are defined in the Griffith sense, as $\gamma_{coh} = 2\gamma_s - \gamma_{gb}$, where γ_{gb} is zero for bulk cohesion, and are reported in mJ/m^2. The ratio of the average GB cohesive energy to the bulk (111) cohesive energy is also given.

Interface	Ni	Al	Ni_3Al
Bulk (100)	3510	1712	3821
Bulk (110)	3954	1918	4119
Bulk (210)	4164	1998	4327
Bulk (111)	3246	1648	3510
Average GB	2894	1633	2998
Average GB/bulk (111)	0.89	0.99	0.85

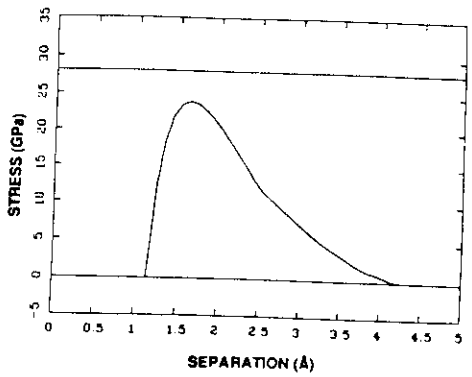


FIG. 1. Frozen stress versus separation curve for pure Ni (210)/[001] tilt GB. The value of σ_{max} for the (210) bulk is shown as a straight line.

maximum stress [Fig. 2(a)] of different cases for the perfect (110) crystallographic plane, the (210) GB without boron and the (210) GB with borons at various locations (discussed below) are shown in Fig. 2(b). The planes on each side of the GB can terminate in layers containing either 100% Ni atoms or 50% Ni and 50% Al atoms. The GBs are denoted as (100/100), or Ni-rich, when both sides of the GB are comprised of 100% Ni layers. The other two cases are denoted as (100/50), or stoichiometric, and (50/50), or Al-rich. Clearly, the maximum stress of the GB (22 GPa) is about 80% of the perfect (210) plane (27 GPa). The GB in pure Ni has the same ratio of maximum stress, with 24 GPa for the GB and 28 GPa for the (210) bulk.

We can also study the response of the GB as the system is pulled slowly apart. To do this we clamped the system 4 lattice parameters away from the GB, and then applied successive 2% strain steps, allowing the system to relax to the minimum energy (zero force) state[3] at each step. The process was repeated until total cleavage occurred. The stress strain curve for the (210) and (310) cases in Ni_3Al and Ni are shown in Figs. 3 and 4. A strong difference is seen between the bulk and GB results. If we take the *relaxed maximum stress* and work to pull apart the (210) [or (310)] plane as unity, then we can define the percentage stress and work needed to break the GB. We find that the Ni GB has 99% stress and 86% of the work, while for Ni_3Al, the stresses are in the range from 87% to 95% and the required work ranges from 42% to 55%, as shown in Table IX. The clear distinction lies in the relaxed work ratio of the GB, since the Ni_3Al GBs have only half of the value of the bulk. The value of the cohesive energy calculated in this way is about twice as large as that calculated in the Griffith sense,[3,4,16] because in the

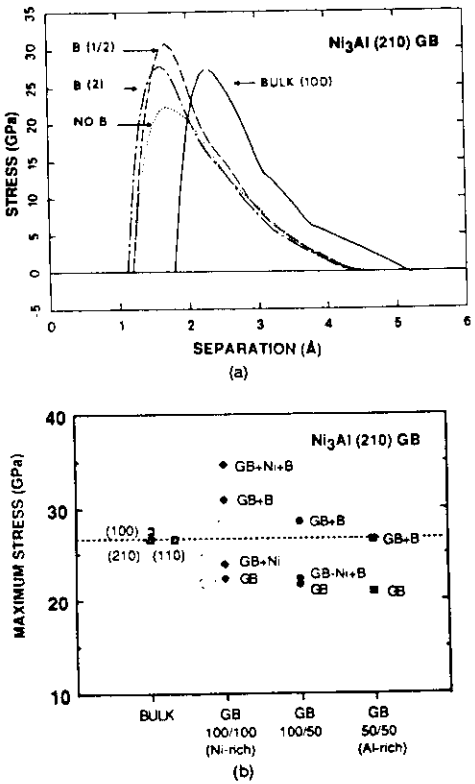


FIG. 2. The frozen stress versus separation curve for the "Ni-rich" (see text) Ni_3Al (210) GB (a), and the maximum frozen stress values for various bulk planes and GB (b).

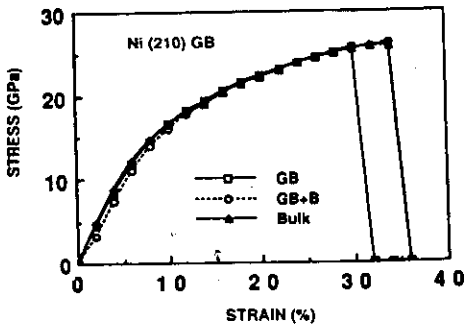


FIG. 3. Relaxed stress versus strain curves for pure Ni (210) GBs.

slow relaxation calculations, part of the plastic response is also calculated in the work. From McLean's data,[44] pure metals (Cu, Al) have work ratios of 0.8, while ionic crystals such as NaCl have ratios of 0.4 to 0.5. In terms of ideal work for fracture, the range of the ductile-brittle transition region is about 0.5 to 0.6. This correlation seems to work for oxides and other ionic crystals. Thus the work ratio results from the slow relaxation calculations are consistent with Ni being more ductile than Ni_3Al, in agreement with experiment. In contrast, Griffith energies incorrectly predict that both Ni and Ni_3Al should be in the ductile region.

B. Grain boundaries of Ni and Ni_3Al with boron

There are two major speculations about what the boron does in the GB to help to ductilize the Ni_3Al grain boundary: (1) boron increases the cohesive strength of the boundary,[10] (2) boron improves plasticity at the boundaries,[45] either directly or by promoting chemical or structural disorder.

The first step in studying the effect of boron is to understand where boron resides in the bulk and whether boron will segregate to the GB. This study serves to identify a reference point for boron position and energy and also tests the validity of the potentials used for the boron. Boron has been inserted in the three possible bulk interstitial sites: the tetrahedral site, the 4Ni–2Al octahedral site, and the 6Ni octahedral site leading to relaxed energies of −2.99, −3.65, and −4.59 eV, respectively. That the boron prefers an octahedral site is consistent with channeling/nuclear reaction analysis,[46] though the experiment cannot distinguish between the 6Ni and 4Ni–2Al sites.

Starting with relaxed Ni_3Al GBs,[3,4,12] boron was either inserted into the lowest density regions in the grain boundary or substituted for Ni atoms. Subsequently, the grain boundaries were relaxed to the lowest energy.[3,12–14] For the [001] symmetric tilt (210) and (310) GBs we have studied, the boron prefers low density "interstitial" sites at the GB to the "substitutional" sites.

If boron is placed in the site of the (210) GB shown by the square with X in Fig. 5(a), during relaxation the boron moves to the site shown in Fig. 5(b) without introducing large strains in the environment. (This is labeled as site 2 in Table X.) Placing B at an alternate site [denoted site 1 in Table X and not shown in Fig. 5(a)] lying midway between sites 2, results in a separate structure upon relaxation. The distortion caused by the insertion of B at the boundary is a small local dilation with a strain in the first shell of about 5%, whereas insertion in the bulk leads to larger strains of about 10%. The boron causes a smaller dilation at the boundaries because of the larger free volume[12,16] at the boundaries where the structure can more easily adjust to an external disturbance.

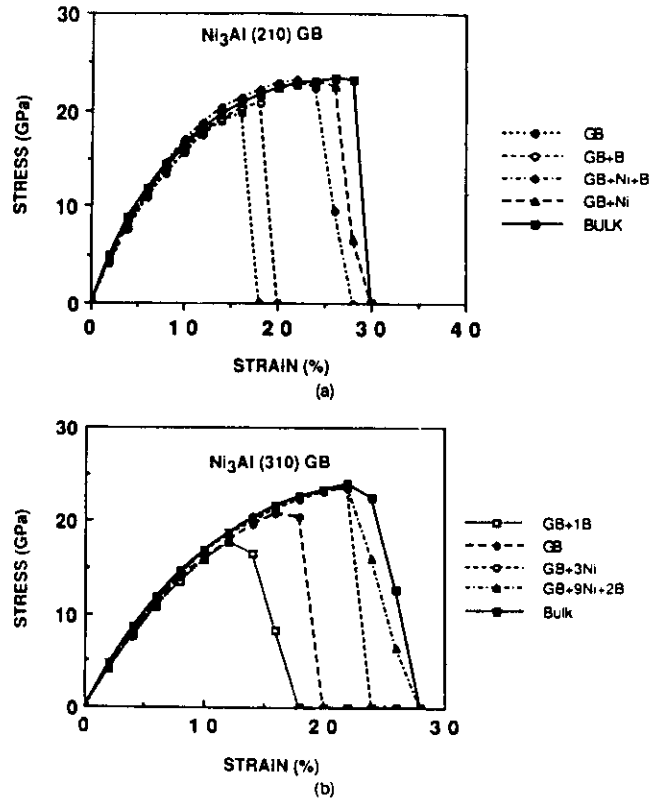


FIG. 4. Relaxed stress versus strain curve for (a) (210) GB in Ni_3Al with extra Ni, extra B, and extra B and Ni simultaneously, and (b) same as (a) for (310) GB.

In the (310) boundary, if the boron is inserted into a capped trigonal prism environment similar to the (210) case, as denoted by the square with X in Fig. 6(a), the boron induces a large lateral shift (25% of the lattice spacing) of the top grain relative to the bottom grain. The relaxed boron environment is that of a distorted pentagonal bipyramid, with the atoms at the apices moving sideways, as shown in Fig. 6(b). If the boron is first inserted at the alternate site, denoted by the square with cross in Fig. 6(a), having a pentagonal bipyramidal local environment, the boundary relaxes to the same structure shown in Fig. 6(a).

The topological environment and the size of the local cluster are seen to have a pronounced effect on the final structure of the grain boundary. These effects are usually not considered in electronic structure calculations[47,48] on small clusters with high symmetry as models for the local environment in grain boundaries.

The change in energy caused by the segregation of boron to various free surface or grain boundary sites is shown in Table X. The results show that the boron segregates preferentially to the GB rather than to the free surface, with a difference of about 1 eV. This preference for segregation of boron to GBs is in agreement with Auger studies of free surfaces and fractured grain boundaries,[11] which show B segregates to the GB but not to the free surface. In recent studies[49] B and N were observed to segregate to surfaces damaged by sputter ion etching, but they did not segregate strongly when the surfaces had been well annealed. As far as the relative preference for bulk to surface sites is concerned, however, the results in Table X do not necessarily indicate a tendency for surface segregation of B from the bulk. The current calculations correspond only to the relative enthalpy of such processes, and entropy effects—which favor bulk over interfacial sites—would

TABLE IX. Calculated maximum stress and total work involved in cleavage in Ni and Ni_3Al, expressed as a fraction of the value required to cleave the perfect crystal along the corresponding plane.

	Stress ratio	Work ratio
(210) pure Ni	0.99	0.86
(+2B)	(1.00)	(0.99)
(210) Ni_3Al		
Al-rich	0.88	0.42
Stoichiometric	0.95	0.55
Ni-rich	0.87	0.46
+2B	0.89	0.58
+extra 2Ni*	0.99	0.98
+extra 2Ni and 2B*	0.98	0.97
+extra 10Ni*	0.99	0.97
+extra 10Ni and 2B*	0.99	0.90
(310) Ni_3Al		
Ni-rich	0.87	0.59
+1B	0.74	0.42
+3Ni	0.98	0.92
+9Ni	0.96	0.72
+3Ni and 2B*	0.92	0.71
+9Ni and 2B*	0.99	0.93

*Breaks away from GB.

also need to be included in the overall free energy calculations needed for such predictions.

In addition to the segregation energy of boron, Table X also shows the boron-induced change in GB energy and surface energy. The energies of the interface are calculated by taking the reference energy of boron as the energy at the 6Ni octahedral site. Based on this calculation, we find that the boron segregates to the GB and lowers the GB energy. However, when present on surfaces, boron also reduces the surface energy. The change in the GB energy, however, is at least twice the change of the free surface energy. According to the classical thermodynamics argument given by Rice,[50] boron thus increases the Griffith cohesive energy ($\gamma_{coh} = \gamma_{s1} + \gamma_{s2} - \gamma_{gb}$) and therefore the cohesive strength of the boundaries, as first suggested by Liu *et al.*[10]

We examine more closely what boron does to the boundaries in equilibrium and under load conditions. Substitution of boron for Ni is only slightly beneficial in increasing the frozen maximum stress. The maximum stress, σ_{max}, of the GB is about 80% of the bulk, which is about the same for pure Ni. The substitution of boron for Ni atoms at the boundary (in the 100/50 cases) in Fig. 2 only slightly increases σ_{max}. Insertion of a boron atom into interstitial sites[13,14] [Fig. 5(b)] has a much more pronounced effect on σ_{max}, and the beneficial effect of boron doping (i.e., raising σ_{max}) increases with increasing Ni concentration at the boundary (from

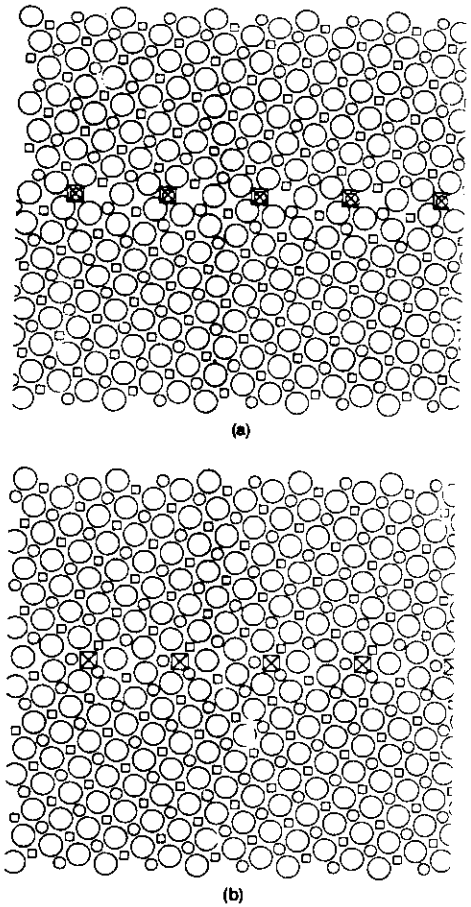


FIG. 5. Ni_3Al (210) symmetric tilt GB; the circles are Ni, the squares are Al, and the size indicates the distance from the viewer. In (a), the pure Ni_3Al has been relaxed, and a possible site for B is shown (square with X) before further relaxation. Relaxing this geometry leads to (b), in which the B has moved to the center of a capped trigonal prism. This is denoted as site 2 in Table X.

66% Ni, to 75% Ni, to 82% Ni in the GB structured unit). In the case of the Ni-rich boundary, adding boron to a certain site or certain combinations of sites raises σ_{max} of the GB above that for perfect crystal planes. This implies that boron segregation can effectively make the Ni_3Al grain boundaries *stronger* than the

TABLE X. Energy at B in various bulk, GB, and surface sites. For all interface sites, the planar density of B corresponds to one B per (210) GB unit cell (see Fig. 5). GB sites 2 are indicated in Fig. 5(b). GB sites 1 (not shown) lie midway between sites 2 in the figure.

	Boron energy (eV/atom)	Energy relative to 6Ni site (eV/atom)	$\Delta\gamma$ interface (mJ/m^2)
Bulk octahedral sites			
6Ni	−4.59	0	
4Ni, 2Al	−3.65	+0.94	
Ni-rich (210) GB			
boron site 1	−6.34	−1.75	−982
boron site 2	−6.83	−1.94	−1089
boron site 1 & 2	−13.05	−3.87	−2172
Stoichiometric (210) GB			
boron site 1	−6.92	−2.33	−1308
boron site 2	−5.70	−1.11	−623
boron site 1 & 2	−12.72	−3.54	−1987
Al-rich (210) Gb			
boror site 1	−5.67	−1.08	−606
boron site 2	−6.96	−2.27	−1274
boron site 1 & 2	−12.63	−3.45	−1936
Subsurface octahedral sites			
6Ni	−4.96	−0.37	−208
4Ni, 2Al	−5.30	−0.71	−398
(210) surface			
Ni-term. sites	−5.62	−1.03	

*The changes of GB energy and surface energy are calculated by taking isolated boron at 6Ni octrahedral site as the reference state.

bulk. Therefore other mechanisms, such as dislocation generation in the second grain, are made possible by this increase in σ_{max} effected by the boron and to a greater extent by the combination of extra Ni and boron together at the GB. σ_{max} increases from 22.8 GPa to 34.8 GPa when doped with B and Ni, as shown in Fig. 2. A similar beneficial effect of B segregation and extra Ni in raising σ_{max} is found for the case of the (310) GB (Fig. 2). The smaller increase of the maximum stress is due to the lower boron concentration at the (310) GB, which has a longer period. Also, boron was found to increase σ_{max} of the pure Ni GBs. From the frozen result we know that σ_{max} increases above the value of the bulk. We would expect that the bulk will fail before the GB. This indicates that the dislocation reaction mechanism proposed by King and Yoo[51] for the case of boron-doped Ni alloys[52] cannot fully account for the boron effect. The boron effect on the motion or emission of dislocations at the GB is currently under study.

A pure Ni GB doped with boron is shown in Figs. 7(a) and 7(b). The sample fails at the GB when

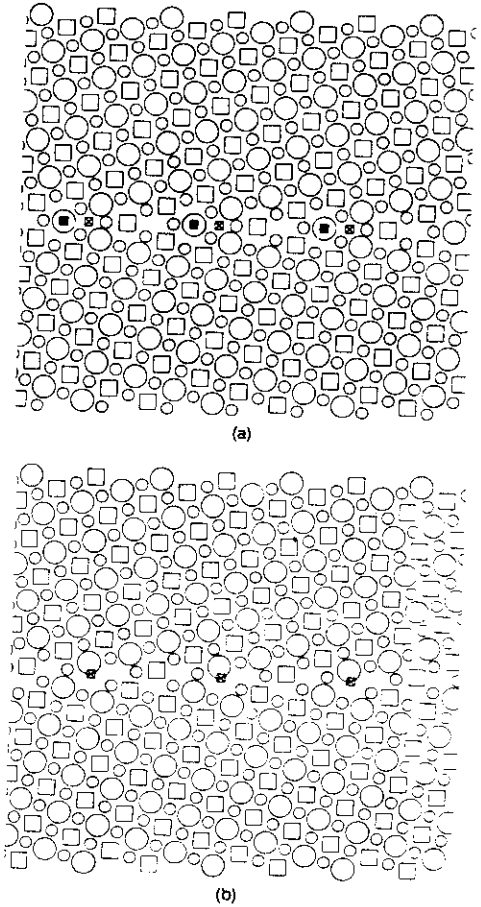


FIG. 6. Ni_3Al (310) symmetric tilt GB (symbols as in Fig. 5). In (a), the pure Ni_3Al GB has been relaxed, and two possible sites for B are shown (square with X and square with cross). Relaxing either of these two geometries leads to the one structure shown in (b), in which B is at the center of a pentagonal pyramid.

there is no boron, as expected, but when doped with boron, the sample fails away from the GB with almost 100% of the maximum stress and 100% of the work, as shown in Table IX.

For the Ni_3Al 100/100 (Ni-rich) GB, the results in Table IX and Fig. 3 show that the Σ5 (210) GB with one monolayer of boron increases the relaxed maximum

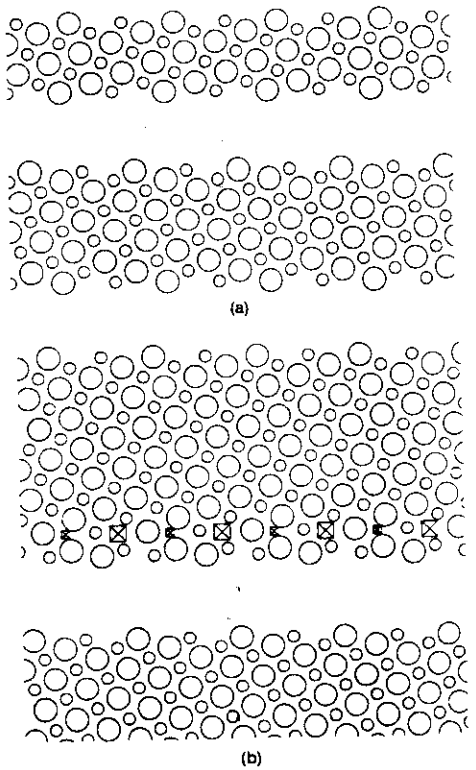


FIG. 7. The structure after slow deformation for the Ni (210) GB without B (a), and doped with one monolayer of B (b). The GB with boron breaks away from the original GB plane.

stress slightly, from 87% to 90%. One monolayer of boron corresponds to two B atoms per unit cell. However, the work ratio shows a much larger increase, from 46% to 56%. In stoichiometric samples, the exchange of Ni and Al at the GB is not allowed because the energy penalty for this exchange is 0.2 eV to 0.8 eV, depending on the initial GB chemistry. The substitution of Ni into Al sites at the GB is possible when the samples are Ni-rich; this is favored by −0.14 eV. In this case, the system shows an increase in stress to 99% (98% in work) and undergoes cleavage away from the GB [Fig. 8(a)]. A similar increase in cohesive properties occurs when the GB is Ni-rich and has segregated boron, as shown in Table IX. The fracture path is away from the GB [Figs. 8(b) and 8(c)]. *Cosegregation of Ni and B to the GB has an energy advantage of −0.46 eV per boron.* The GB with extra Ni and boron has a stress of 99% and a work value of 97%. This is a considerable improvement in the cohesive properties from 87% in stress and 46% in work.

The process of breaking can be visualized by the snapshot at various stages of the straining process, as shown in Figs. 9(a) and 9(b), for the cases without boron and with boron, respectively. (This particular simulation is done with the external clamps 2 a_0 from the boundary.) In the initial stages, for 0% to 14% applied strain, no major change occurs. At 16% strain, the GBs without and with boron start to develop cavities at different positions. At 22% the GB without boron has fractured [Fig. 9(a)] while the GB with boron still holds together with some cavities between the boron positions [Fig. 9(b)]. As the strain level increases, the GB with boron undergoes further necking and a change in direction of the cavities until it breaks at a strain level of 34%. The cleavage path is different from that observed in the case of the GB without boron. When the GB with boron breaks, the boron takes the surrounding atoms with it. These neighboring atoms have actually gone through a large migration before the GB breaks [see atom A in Figs. 9(b) and 9(c)]. This large migration is part of the plastic response of the GB. One could view this result in the context of the McMahon and Vitek,[53] who conclude that it is the change of the GB cohesive energy which controls the plasticity. The (310) GB with Ni saturation (4.5 monolayers of extra Ni at the GB) breaks at larger strain than the pure GB but still breaks at the boundaries. A large change occurs when the GBs are filled with one monolayer of boron and (at least) one layer of extra Ni. The fracture path does not go through the GB where the boron is, but runs close to the first interface of the saturated Ni and the bulk Ni_3Al layer, as shown in Figs. 8(c) and 10(c). The results clearly show that the GB in Ni_3Al with the best overall properties is the one with boron and extra Ni simultaneously at the grain boundaries. The GB becomes much stronger than the bulk. Therefore, the samples fail in the bulk transgranularly instead of intergranularly, as shown in Fig. 10.

As shown in Figs. 2–4, the local tensile cohesive strength of the GB is 25–30 GPa, which is close to the theoretical limit of $\sim E/10$. This is not the same as the applied stress when the sample finally breaks, because we are studying only the response of the region local to the GB (10–20 Å). The stress close to the theoretical strength is expected when dislocation pile-ups are impeded by the GB or other defect before the microcrack can be initiated. We note also that the stress before fracture in the slow straining approach is ~90% that of the frozen method. This is expected, because the system will fail at the weakest link instead of failing at an idealized plane. The effect of extra Ni, B, and S on the slip of a dislocation across the GB will be published elsewhere.

The chemical disordering that occurs when extra Ni is present at the Ni_3Al GB causes the GB region to be like an fcc (gamma) phase with ~2% structural distortion—in effect, an fcc Ni GB with larger lattice spacing (Fig. 11). In the present study, we find that this chemical disordering (or the existence of a second phase at GB) strengthens the GB. This result seems to have a bearing on recent experiments[54,55] on Ni_3Al doped with a few atomic percent of Fe or Mn, in which they found that the sample is ductile only when the GB has a second phase. If the GB is single phase, then the polycrystals doped with Fe or Mn are still intergranu-

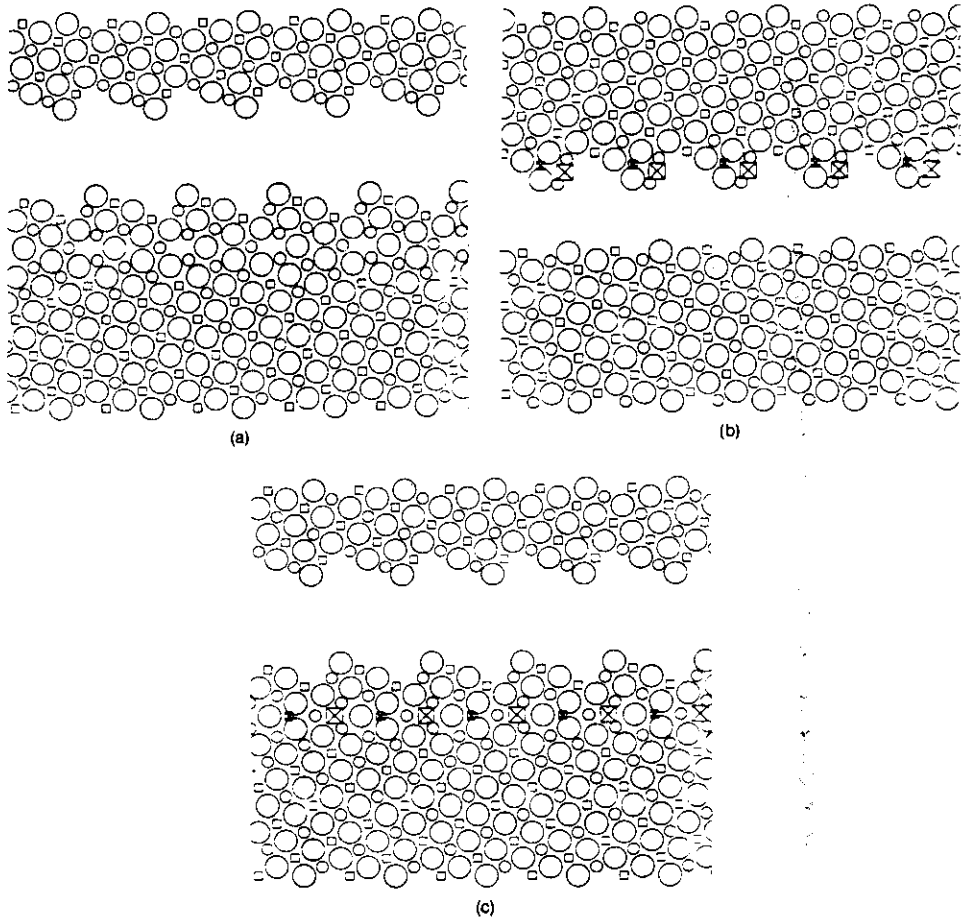


FIG. 8. The structure after slow deformation for the Ni_3Al 100/100 (210) GB with extra Ni (a), with 1 monolayer boron (b), with extra Ni and boron (c).

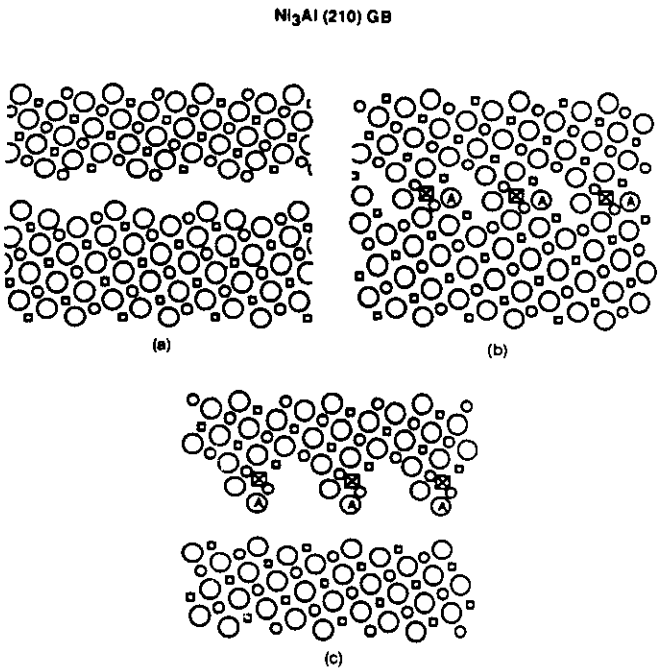


FIG. 9. The slow strain breaking process of Ni_3Al 100/100 (210) GB: (a) 22% strain without boron, (b) 22% strain with half monolayer boron, and (c) same as (b) with 34% strain. Significant atomic movement is observed near the boron. The B appears to be holding the sample together while cavities form between the borons.

urly brittle.[55] The chemically disordered GB attains the highest cohesive strength when boron segregates to the GB, as the GB then behaves very much like a pure Ni GB doped with boron, as shown in Figs. 7, 10, and 11. It has been shown that boron helps to strengthen the GB even with sulfur at the GB.[12,52] Therefore, the Ni_3Al case is like that shown in Fig. 11(b); the second phase (fcc) at the GB plus boron strengthens the GB as well as improves the plasticity in the GB region. Consequently, the sample breaks at the fcc-$L1_2$ interface in the bulk.

This mechanism for ductilizing the GB with a second phase and a dopant can be viewed in another way. If the fcc second phase GB has stronger cohesive properties than the corresponding $L1_2$ phase in the matrix, then it will be desirable to make the GB fcc phase, if possible. One can then look for elements as dopants for the second phase at the GB which are known to help the intrinsic cohesive strength of this second phase. In our case, it is the boron in pure fcc Ni. Similar mechanisms may operate in other systems, such as FeAl (B2), where an Fe bcc second phase at the GB and doped with carbon or boron could make it ductile. Based on this mechanism, we conclude it is important to understand pure metals with other dopants to be able to understand intermetallic compounds.

These calculations on the effects of B on grain boundaries in Ni_3Al are generally in accord with many recent experimental studies, although there are conflicting reports concerning certain aspects. The calculations indicated that B will tend to enhance cosegregation of Ni to the boundaries. Atom probe field ion microscopy by Sieloff *et al.*[56,57] found GBs of boron-doped Ni-rich Ni_3Al were, in fact, Ni-enriched relative to the bulk, while boron free samples did not show any measurable Ni enrichment. In contrast, George *et al.*[58] used Auger analysis of hydrogen charged GBs to look for Ni enrichment and found only a slight effect (0.6 to 0.8% more Ni at the boundary compared to bulk). This was regarded as negligible when compared to their experimental accuracy of 2%.

With regard to the issue of disordering at the boundary, Mackenzie and Sass[59] in lattice imaging studies found a 50 A disordered region on two grain boundaries in B-doped Ni-rich Ni_3Al. Baker and Schulson[60] also found a second phase present at some, but not all, B-free Ni-rich grain boundaries which could

6-0

Ni_3Al (310) GB WITH EXTRA Ni AND B

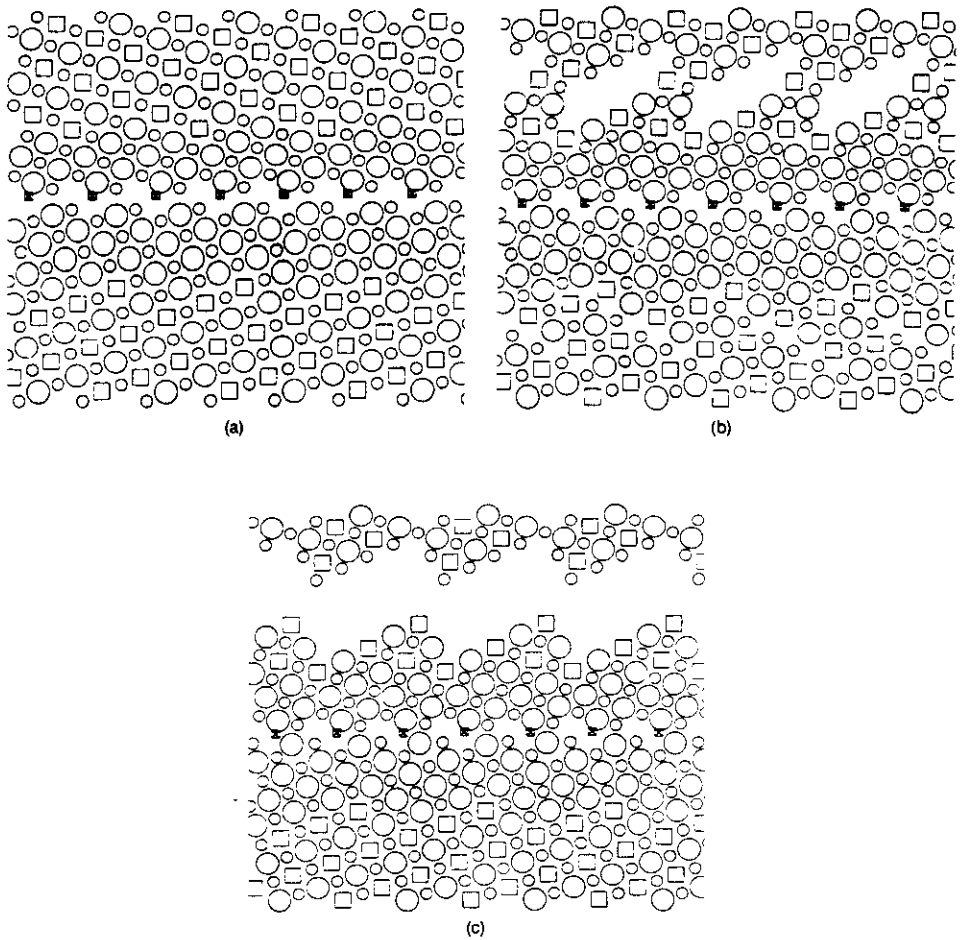


FIG. 10. The relaxed structure of the (310) GB in Ni_3Al with extra Ni and B: (a) 22% strain, (b) 24% strain (just before fracture), and (c) 26% strain (just after fracture).

extend up to 200 A in width and possessed a disordered fcc structure. When B was present, this disordered fcc phase was observed at all boundaries. Krzanowski[61] employed x-ray microanalysis to study GBs in Ni_3Al and found that superlattice images persisted all the way up to the GBs.

V. SUMMARY

Embedded-atom style potentials have been generated for boron–boron, boron–nickel, and boron–aluminum interactions based on data obtained from electronic structure calculations. This procedure is use-

6-1

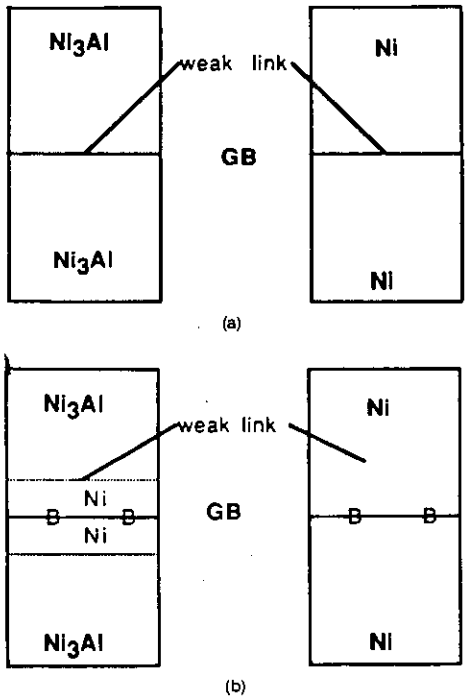


FIG. 11. Schematic drawing of the fracture path in a Ni_3Al GB: (a) before B addition, and (b) after adding B and/or extra Ni. The relationship to a Ni GB is shown.

ful for simulating complicated ternary systems for which experimental data are not available. The results show that the interstitial site and segregation behavior of boron to the GB are correctly predicted. We also find that boron segregates preferentially to the GB instead of the free surface. Furthermore, in Ni-rich samples, Ni substitutes into Al sites at the GB and, together with boron, makes the GB stronger than the bulk. Therefore, the boron and stoichiometric effects in ductilizing Ni_3Al GBs are largely understood. Calculations of the actual fracture toughness,[42,43] dislocation generation,[62] the sulfur embrittling effect,[12] and other intermetallic systems are currently underway. The best criterion for the cohesive properties seems to be the work under the stress strain curve, but the maximum GB cleavage stress obtained using either "frozen" or "relaxed" conditions is also very indicative of the boron effect. The maximum stress, however, does not delineate the intrinsic differences between Ni_3Al and Ni. Rather, these differences are clearly demonstrated only by the work ratio of GB to bulk under relaxed conditions.

The success of these simulations in explaining many of the phenomena involved in the effects of B on Ni_3Al grain boundaries lends credence to the approach taken here to obtain interatomic potentials from properties calculated by electronic structure band calculations on artificial structures. In cases of other types of materials where there may be limited or no experimental information, this approach offers a promising route for theoretical simulations to aid experiments in altering the properties of materials in desired ways.

ACKNOWLEDGMENTS

We would like to thank D.J. Srolovitz, J.E. Hack, F.M. Mueller, C.T. Liu, and V. Vitek for helpful discussions. This work has been supported by the Energy Conversion and Utilization Technologies (ECUT) Division of the United States Department of Energy.

REFERENCES

[1] *Grain Boundary Structure and Properties*, edited by G.A. Chadwick and D.A. Smith (Academic Press, New York, 1976).
[2] *Grain Boundary Structure and Kinetics*, edited by R.W. Balluffi (ASM, Metals Park, OH, 1980).
[3] S.P. Chen, A.F. Voter, and D.J. Srolovitz, Scripta Metall. **20**, 1389 (1986).
[4] S.P. Chen, A.F. Voter, and D.J. Srolovitz (Proc. Mater. Res. Soc. Symp.) (Materials Research Society, Pittsburgh, PA, 1987), Vol. 81, p. 45.
[5] H.R. Merrick, Metall. Trans. A **7A**, 505 (1976).
[6] *Intermetallic Compounds*, edited by J.H. Westbrook (John Wiley and Sons, New York, 1967); N.S. Stoloff and R.G. Davis, Prog. In Mater. Sci. **13**, 1 (1966).
[7] K. Aoki and O. Izumi, Nippon Kinzoku Gakkaishi **43**, 1190 (1979).
[8] *High-Temperature Ordered Intermetallic Alloys* (Proc. Mater. Res. Soc. Symp.), edited by C.C. Koch, C.T. Liu, and N.S. Stoloff (Materials Research Society, Pittsburgh, PA, 1985), Vol. 39.
[9] *High-Temperature Ordered Intermetallic Alloys II* (Proc. Mater. Res. Soc. Symp.), edited by N.S. Stoloff, C.C. Koch, C.T. Liu, and O. Izumi (Materials Research Society, Pittsburgh, PA, 1987), Vol. 81.
[10] C.T. Liu, C.L. White, and J.A. Horton, Acta Metall. **33**, 213 (1985).
[11] C.L. White, P.A. Padgett, C.T. Liu, and S.M. Yalisov, Scripta Metall. **18**, 1417 (1985).
[12] S.P. Chen, A.F. Voter, R.C. Albers, A.M. Boring, and P.J. Hay, Scripta Metall. **23**, 217 (1989).
[13] S.P. Chen, A.F. Voter, R.C. Albers, A.M. Boring, and P.J. Hay (Proc. Mater. Res. Soc. Symp.) (Materials Research Society, Pittsburgh, PA, 1988), Vol. 122, p. 355.
[14] S.P. Chen, A.F. Voter, and D.J. Srolovitz, J. de Phys. **C5**, 157 (1988).
[15] A.F. Voter and S.P. Chen (Proc. Mater. Res. Soc. Symp.) (Materials Research Society, Pittsburgh, PA, 1987), Vol. 82, p. 175.
[16] S.P. Chen, D.J. Srolovitz, and A.F. Voter, J. Mater. Res. **4**, 62 (1989).
[17] M.S. Daw and M.I. Baskes, Phys. Rev. B **29**, 6443 (1984).
[18] O.K. Andersen, Phys. Rev. B **12**, 3060 (1975).

[19] H.L. Skriver, *The LMTO Method* (Springer, Berlin, 1984).
[20] N.E. Christensen, Int. J. Quant. Chem. **25**, 233 (1984).
[21] S.M. Foiles, M.I. Baskes, and M.S. Daw, Phys. Rev. B **33**, 7983 (1986), and references therein.
[22] A.F. Voter (to be published).
[23] S.P. Chen, A.F. Voter, and D.J. Srolovitz (Proc. Mater. Res. Soc. Symp.) (Materials Research Society, Pittsburgh, PA, 1987), Vol. 81, p. 45.
[24] S.P. Chen, A.F. Voter, and D.J. Srolovitz, Phys. Rev. Lett. **57**, 1308 (1986).
[25] S.P. Chen, A.F. Voter, and D.J. Srolovitz, *Characterization of Defects in Materials* (Proc. Mater. Res. Soc. Symp.), edited by R.W. Siegel, R. Sinclair, and J.R. Weertman (Materials Research Society, Pittsburgh, PA, 1987), Vol. 82, p. 175.
[26] M.W. Finnis and J.E. Sinclair, Phil. Mag. **A50**, 45 (1984).
[27] S.M. Foiles and M.S. Daw, J. Mater. Res. **2**, 5 (1987).
[28] S.M. Foiles and M.S. Daw, J. of Metals **39**, 39 (1987).
[29] D. B. Sullenger and C.H.L. Kennard, Sci. Am. **215**, 96 (1966).
[30] J.H. Rose, J.R. Smith, F. Guinea, and J. Ferrante, Phys. Rev. B **29**, 2963 (1984).
[31] S.M. Foiles, Phys. Rev. B **32**, 7685 (1985).
[32] N.S. Stoloff, p. 3 in Ref. 8.
[33] A.I. Taub, C.L. Briant, S.C. Huang, K-M. Chang, and M.R. Jackson, Scripta Metall. **20**, 129 (1986).
[34] S.C. Huang, C.L. Briant, K-M. Chang, A.I. Taub, and E.L. Hall, J. Mater. Res. **1**, 60 (1986).
[35] T. Takasugi, N. Masahashi, and O. Izumi, Scripta Metall. **20**, 1317 (1986).
[36] T. Ogura, S. Hanada, T. Masumoto, and O. Izumi, Metall. Trans. A **16A**, 441 (1985); T. Tagasugi, E.P. George, D.P. Pope, and O. Izumi, Scripta Metall. **19**, 551 (1985).
[37] E.M. Schulson, T.P. Weihs, D.V. Viens, and I. Baker, Acta Metall. **33**, 1587 (1985).
[38] T. Tagasugi and O. Izumi, Acta Metall. **33**, 1247 (1985).
[39] A.I. Taub and C.L. Briant, p. 343 in Ref. 9.
[40] D. Farkas, Scripta Metall. **19**, 467 (1985).
[41] R.W. Cahn (private communication).
[42] J.E. Hack, D.J. Srolovitz, and S.P. Chen, Scripta Metall. **20**, 1699 (1986).
[43] J.E. Hack, S.P. Chen, and D.J. Srolovitz, Acta Metall. (to be published).
[44] D. McLean, J. de Phys. **C4**, 273 (1975).
[45] E.M. Schulson, I. Baker, and H.J. Frost (Proc. Mater. Res. Soc. Symp.) (Materials Research Society, Pittsburgh, PA, 1987), Vol. 81, p. 195.
[46] H.G. Bohn, J.M. Williams, J.H. Barrett, and C.T. Liu (Proc. Mater. Res. Soc. Symp.) (Materials Research Society, Pittsburgh, PA, 1987), Vol. 81, p. 127.
[47] C.L. Briant and R.P. Messmer, Phil. Mag. **42B**, 569 (1980); Acta Metall. **30**, 1811 (1982); R.P. Messmer and C.L. Briant, Acta Metall. **30**, 457 (1982).
[48] M.E. Eberhart and D.D. Vvedensky, Phys. Rev. Lett. **58**, 61 (1987).
[49] C.L. White, C.T. Liu, and R.A. Padgett, Jr., Acta Metall. **36**, 2229 (1988).
[50] J.R. Rice, in *Effect of Hydrogen on Behavior of Materials*, edited by A.W. Thompson and I.M. Bernstein (AIME, New York, 1976).
[51] A.H. King and M.H. Yoo (Proc. Mater. Res. Soc. Symp.) (Materials Research Society, Pittsburgh, PA, 1987), Vol. 81, p. 99.
[52] R.A. Mulford, in "Embrittlement of Engineering Alloys," *Treatise on Materials Science and Technology*, edited by C.L. Briant and S.K. Banerji (Academic Press, New York, 1983), Vol. 25, p. 1 and references cited therein.
[53] C.J. McMahon and V. Vitek, Acta Metall. **27**, 507 (1979); M.L. Jokl, V. Vitek, and C.J. McMahon, Acta Metall. **28**, 2479 (1980).
[54] T. Tagasugi, O. Izumi, and N. Masahashi, Acta Metall. **33**, 1259 (1985).
[55] D.M. Dimiduk, V.L. Weddington, and H.A. Lipsitt (Proc. Mater. Res. Soc. Symp.) (Materials Research Society, Pittsburgh, PA, 1987), Vol. 81, p. 221.
[56] D.D. Sieloff, S.S. Brenner, and M.G. Burke (Proc. Mater. Res. Soc. Symp.) (Materials Research Society, Pittsburgh, PA, 1987), Vol. 83, p. 87.
[57] D.D. Sieloff, S.S. Brenner, and H. Ming-Jian (Proc. Mater. Res. Soc. Symp.) (Materials Research Society, Pittsburgh, PA, 1989), Vol. 133, p. 155.
[58] E.P. George, C.T. Liu, and R.A. Padgett, Jr., Scripta Metall. **23**, 979 (1989).
[59] R.D. Mackenzie and S.L. Sass, Scripta Metall. **22**, 1807 (1988).
[60] I. Baker and E.M. Schulson, Scripta Metall. **23**, 1883 (1989).
[61] J.E. Krzanowski, Scripta Metall. **23**, 1219 (1989).
[62] J.E. Hack and S.P. Chen (unpublished research).