Nucleation theory and the early stages of thin film growth

C. Ratsch[a)]
Department of Mathematics, University of California, Los Angeles, California 90095-1555

J. A. Venables
Department of Physics and Astronomy, Arizona State University, Tempe, Arizona 85287-1504 and CPES, University of Sussex, Brighton BN1 9QJ, United Kingdom

(Received 9 June 2003; accepted 17 June 2003; published 2 September 2003)

A review is given of nucleation and growth models as applied to the earliest stages of thin film growth. Rate equations, kinetic Monte Carlo, and level set simulations are described in some detail, with discussion of remaining uncertainties, in particular the functional form of the so-called capture numbers in rate equations. Recent examples are given of sub-monolayer nucleation at surface defects, attachment-limited capture, and Ostwald ripening. The experimental literature is cited, and experiment–theory comparisons are made where possible. Emphasis is given to fast computational models that can span a large range of length and time scales, which might be further developed in the direction of on-line process control. © *2003 American Vacuum Society.*
[DOI: 10.1116/1.1600454]

I. INTRODUCTION

Nucleation and growth of thin films include processes on time and length scales that span many orders of magnitude. Atomic motion occurs on length scales of the order of Å, and time scales that reflect the typical atomic vibration frequencies (i.e., 10^{-13} s). On the other hand, a typical optoelectronic device might be up to several microns in size, and its growth can take minutes or even hours. Thus, modeling nucleation and growth of thin films presents a substantial challenge to theoretical physicists and material scientists. Moreover, some of the phenomena that occur are inherently stochastic in nature, and an ideal model would seamlessly combine the different time and length scales, but include only the necessary fluctuations.

The models typically used in nucleation theory are either completely stochastic or completely deterministic. Mean field rate equations (REs) are a set of coupled ordinary differential equations (ODEs), that were developed for this problem[1,2] more than 30 yr ago. They are easy to formulate and relatively easy to solve. Several results of nucleation theory have been successful in elucidating basic aspects of epitaxial growth. In particular, scaling results derived from RE nucleation models have, under the appropriate circumstances, been used to deduce microscopic parameters such as diffusion constants, adsorption and binding energies from comparison with experimental measurements. However, these equations contain no explicit spatial information, and thus do not readily yield information on surface morphology. One of the challenges is to include the spatial information properly into a model that is mean-field by construction; recent progress in this area and limitations are discussed in detail in this article.

Continuum models based on partial differential equations (PDEs) are appropriate mainly at large time and length scales.[3,4] By construction, features on the atomic scale are neglected, so they are poorly suited to describe growth on this scale, and we do not discuss continuum models in this article. However, we note that since continuum models, as well as rate equations, are based on differential equations, they are amenable to analytic treatments that can elucidate, e.g., asymptotic or stability properties.

An alternative to completely analytic approaches are atomistic models that explicitly take into account the stochastic nature of each microscopic process that may occur during nucleation and growth of thin films. They are typically implemented in the form of molecular dynamics (MD)[5] or kinetic Monte Carlo (KMC)[6] simulations. MD simulations are very useful for identifying relevant microscopic processes, such as the detailed steps during nucleation. But time and size limitations make them unfeasible for studying growth on technologically relevant time and length scales. KMC simulations, on the other hand, have been used successfully to study qualitative, and in limited cases, quantitative, behavior of growth. They allow for easy implementation of a large number of microscopic processes, whose rates are ideally obtained from first principles calculations.[7–9] However, the occurrence of very fast rates (which is particularly relevant at higher temperatures) ultimately limits the applicability of these methods to larger systems.

Recent work has attempted to develop new models that are hybrid models between continuum, PDE-based methods, and atomistic, stochastic methods. One approach, termed configurational continuum, which appears to be very promising, has been developed by Kandel and co-workers; we refer to Ref. 10 for further details. Another approach that has been developed by one of us (C.R.) in the past few years is an island dynamics model, based on the level-set method. This approach will be discussed in more detail in Sec. II D. This model allows us to describe thin film growth as continuous in the plane of the surface, yet each atomic layer is discretely resolved. Moreover, different sources of fluctuations can be isolated and studied individually.

[a)]Electronic mail: cratsch@math.ucla.edu

One generally distinguishes between growth on singular (nominally flat) surfaces, and vicinal, or stepped surfaces. A vicinal surface can be considered as a number of flat terraces that are separated by steps of atomic height. Growth on these surfaces proceeds either via step-flow, where atoms diffuse toward a step before meeting another atom, or via nucleation and growth, where islands nucleate and grow on the terrace. Growth on vicinal surfaces was first discussed in a seminal paper by Burton, Cabrera, and Frank (BCF).[11]

In this article we focus on modeling of growth on singular surfaces, where one distinguishes the following three growth modes: Frank–van der Merwe (FM) (layer-by-layer) growth, Volmer–Weber (VW), and Stranski–Krastanov (SK) growth.[12–14] During VW growth three-dimensional (3D) islands form on the surface. The competition between FM and VW growth can easily be understood based on energetic arguments as a competition between surface and interface energies. The case of SK growth is more complicated, and is intimately connected to (elastic) strain energy, which arises in most heteroepitaxial systems. During SK growth, one or more layers form initially (the so-called wetting layer), followed by the formation of 3D islands. Understanding the transition between growth modes is of great interest. The focus of this article is on submonolayer growth and will not address these issues, but the methods described are also applicable to multilayer growth.

The remainder of this article is organized as follows: In Sec. II B we discuss mean-field rate equations. In particular, we focus on scaling laws, and how these equations can properly describe mean-field quantities. In Sec. II C we describe the kinetic Monte Carlo method, and discuss some of its applications. The level-set method is described in Sec. II D. It is well known that rate equations do not properly predict the entire island size distribution, for the reason that rate equations do not contain any explicit spatial information. Recent attempts and progress in including this spatial information implicitly through particular forms of capture numbers is discussed in Sec. III A. Recent models of nucleation at defect sites are described in Sec. III B. In Sec. III C, we address the problem of attachment-limited, and time-dependent capture numbers. Section III D highlights issues related to coarsening and Ostwald ripening, and annealing in general. Finally, in Sec. IV we give an assessment of the relation of all the theoretical methods introduced, and a personal impression of future directions.

II. MATHEMATICAL AND COMPUTATIONAL METHODS

A. General considerations

Typical processes that may occur during epitaxial growth are illustrated schematically in Fig. 1. Atoms are deposited onto a perfect substrate surface with a deposition flux F (a); in the older literature, this same quantity has been termed the deposition rate R, both measured typically in monolayers per second. Once atoms are on the surface as adatoms, they can diffuse with a diffusion constant D (b). Adatoms can meet other adatoms to form a dimer (c), or attach to existing islands (d). Once adatoms are attached to an island, they can detach from the island edge (e) or diffuse along the island edge (f). Deposition of adatoms on top of islands and the corresponding processes have to be considered as well (g), (h); at high temperatures some adatoms can re-evaporate (i). In this section we describe different methods that model these processes at different levels of detail.

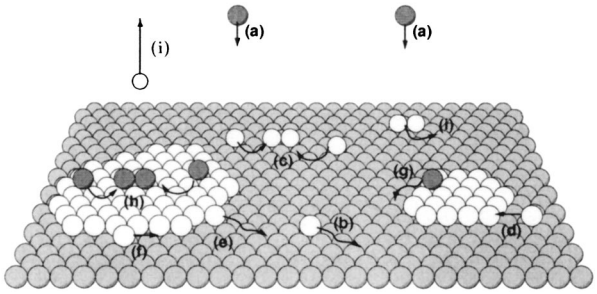

FIG. 1. Typical atomistic processes during epitaxial growth.

The most detailed description is a MD[5] simulation. In a MD simulation one calculates the forces on all atoms, and then moves the atoms according to the equations of motion. The most crucial aspect of an MD simulation is the knowledge of the correct potentials. The timestep is required to be small enough to resolve the vibrational frequency of the atoms, so that there is a natural limit on the time that can be simulated. Moreover, even with simple pairwise potentials the evaluation of forces is rather time-consuming, so that simulations of realistic system sizes on realistic time scales to describe nucleation and growth is currently not possible with MD simulations. This is still true, despite some impressive recent advances in speeding up MD simulations.[15] However, these simulations are very useful to identify some of the relevant microscopic pathways during nucleation and thin film growth.

One can make significant progress by using transition state theory (TST).[16,17] In TST rates are associated with microscopic events such as adatom diffusion, and all the irrelevant atomic vibrations are neglected. This way, the simulation timestep and thus the computational efficiency is increased by 5-10 orders of magnitude for realistic growth conditions. One can usually express the rates in the form $\nu = \nu_0 \exp(-\Delta E/kT)$, where ΔE is an activation energy barrier, k is the Boltzman constant, and T is the temperature. The prefactor ν_0 is typically of the order of the atomic vibration frequency, and is set to 10^{12}–10^{13} s^{-1} in many simulations. TST is the basis of RE, KMC, and many other approaches, which are the subject of the remainder of this section.

B. Mean-field rate equations

1. Basic concepts

The time evolution of mean-field quantities, such as the density of adatoms n_1 and of islands of size s, n_s, can be described by a set of coupled ODEs known as REs. Adatoms can either meet another adatom with a capture efficiency σ_1

to form a dimer, an island of size 2, or get captured by islands of size s with a capture efficiency σ_s.

At high substrate temperature, these adatoms can reevaporate. For example, evaporation is clearly required to establish equilibrium with the 3D vapor phase, and many models have been developed for the low supersaturation regime, following the pioneering work more than half a century ago by BCF,[11] much of which is now in the textbook and monograph literature.[18,19] If the adsorption energy E_a is low, then the evaporation time τ_a is short, and this evaporative quasiequilibrium is quickly established at low adatom density, with $n_1 = F\tau_a$. RE models for nucleation and growth in this situation were established in the 1960s,[1] and exact solutions for the capture efficiency, the so-called capture numbers, were obtained in this case.[2]

This high temperature problem is soluble in closed form, because the islands that form are isolated from each other, due to the short BCF diffusion length, $x_s = (D\tau_a)^{1/2}$, the mean adatom surface diffusion distance before evaporation. The approach becomes more complex if there are several competing processes at work. For this more general case, it has been argued[20,21] that one can develop the idea of competitive capture, in which characteristic times for different processes add inversely, and the shortest time dominates. In this formulation, the RE for n_1 at low coverage is essentially

$$dn_1/dt = F - n_1/\tau,$$

with

$$\tau^{-1} = \tau_a^{-1} + \tau_n^{-1} + \tau_c^{-1} + \ldots, \tag{1}$$

which has a steady state solution $n_1 = F\tau$. The composite term n_1/τ represents all the loss terms, adsorption (τ_a), nucleation (τ_n), capture by stable clusters (τ_c) and maybe other processes [... in Eq. (1)], all of which add like resistances in parallel. One clearly may envisage many other processes that might take place on, or close to, the substrate surface. However, as in any other modeling situation, completeness is bought at the price of loss of simplicity and clarity; thus one only adds new terms when compelled to do so by the (experimental) evidence. Some of these situations are developed in Sec. III. In addition to Eq. (1) for the single adatoms, we need a complete set of REs for the larger clusters, size $s \geqslant 2$. Assuming initially that we are concerned only with single adatom processes, these can most conveniently be written as

$$dn_s/dt = U_{s-1} - U_s, \tag{2}$$

where U_s is the net rate of conversion of clusters of size s into size $(s+1)$. As in Eq. (1), one can add other processes, most naturally coalescence between islands, U_c, which will reduce the island density at small sizes and increase the density at large sizes. This topic has been extensively considered, both in the early literature[2,22,23] and more recently.[24,25] Some comments are made below.

The above discussion implies that the models developed for the different growth modes are going to be different in detail. However, all RE models have certain features in common, and we discuss these features here. First, we need expressions for the net rate, U_{s-1} and U_s, in the set of Eqs. (2). Each of these is the difference between two terms, a capture and a decay term, such that when these terms are equal, we have local equilibrium. Ignoring direct impingement for the moment, the capture term is given by $\sigma_1 D n_1^2$ for forming a dimer, or is $\sigma_s D n_1 n_s$ for forming an $(s+1)$ cluster from an s cluster, where the capture numbers σ_s remain to be determined. The decay terms can be written in the form $n_s\Gamma_s$, where Γ_s is the decay rate of s clusters forming $(s-1)$ clusters. Details of the decay rate expressions are given in the Appendix.

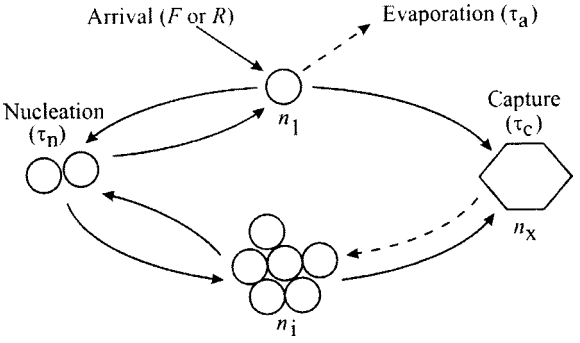

FIG. 2. Critical size in nucleation and growth models.

There are two main types of RE models in the literature. The first emphasizes the role of the critical nucleus of size i. The second approach does not include a critical island size explicitly. Rather, attachment and detachment rates for islands of all sizes are in principle included. Those two approaches are described in the following two subsections.

2. Rate equations with explicit critical island size

The idea of a critical nucleus has been explored in many papers in the literature, and has several related consequences. The main ideas are illustrated in Fig. 2. The left-hand side of this diagram indicates that, because of equal forward and back reaction rates (full lines), small clusters may be in local equilibrium with the instantaneous adatom density n_1. Using the detailed balance arguments set out in the Appendix, we are able to write that, for subcritical clusters $j \leqslant i$ the corresponding net rates U_j are zero, and the Walton relation[26] [Eq. (A1)] can be used to express the density of critical nuclei n_i, in terms of the adatom density n_1 and the energy of the critical nucleus E_i. The right-hand side indicates that 'stable' clusters eventually grow, and the back reaction rate becomes less important (dashed line). These stable clusters, size $s > i$, grow by diffusive capture and, maybe at a later stage, by direct impingement.

As developed by one of us (J.V.) and others, this approach can be combined with compacting the REs for all cluster sizes $s > i$ into one RE for the stable cluster density n_x, such that

$$dn_x/dt = \sum_{s>i} dn_s/dt = U_i - U_c, \tag{3}$$

where, from Eqs. (2), all the other U_s cancel out in pairs, and we have included the coalescence term U_c, which may limit n_x at higher coverage. Thus in this treatment we have reduced an infinite set of REs to just two, one for n_1 [essentially Eq. (1)] and one for n_x [Eq. (3)]. We have abandoned any serious attempt to calculate the island size distribution $n_s(s)$, and we need to assume specific island shapes to make further progress. But we can also calculate other mean-field quantities, e.g., mean island size (w_x), substrate coverage (Z), differential or integrated condensation coefficients (α or β) with this approach, using an auxiliary growth rate equation.[2,14,20,21] This equation for the number of atoms in stable clusters, $n_x w_x$, has the form

$$d(n_x w_x)/dt = n_1(\tau_n^{-1} + \tau_c^{-1}) + F\sum_{s>i} \kappa_s n_s a_s, \tag{4}$$

where all relevant growth processes [i.e., not including τ_a^{-1} in Eq. (1)] are included in the first term. The second term, included at this stage for illustration, corresponds to direct impingement onto all s clusters of area a_s, and $\kappa_s \geqslant 1$ is a geometrical correction factor, which may be needed to account for impingement immediately next to a cluster.[25] Without this factor, the sum $\Sigma_{s>i} n_s a_s$ is just equal to Z, the coverage of the substrate by stable and critical islands. Thus the expression $F(1-Z)$ has been used to correct Eq. (1) for n_1, for the effects of finite coverage.[20,21] The coalescence rate U_c at low coverage is simply proportional to $2n_x dZ/dt$;[2,23] high Z changes to this formula are given more recently.[25]

The above scheme is relatively simple, and using it we can make scaling predictions for all the mean field quantities, notably the densities n_1 and n_x, as a function of the various material parameters and the critical nucleus size i. In the version that the second author has developed, this involves scaling with the adsorption energy E_a (which governs τ_a), E_d (which governs the diffusion coefficient D), and the binding energy E_i of the critical cluster. The scaling relations in the complete condensation, initially incomplete, and extreme incomplete condensation regimes have been documented in tables for both 2D and 3D island shapes,[14,21] and some of these regimes, notably complete condensation, have been thoroughly tested, both by simulation and by experiment. These relationships, originally conceived to be useful for heteroepitaxy, have been revisited on occasion, and the 2D extreme incomplete condensation case has been modified[27] to make contact with the classic BCF case of homoepitaxy at large coverage.

When lateral bonds are suitably strong, at low temperatures, we have $i=1$, so that adatom pairs (dimers) are already stable nuclei, and E_i, or E_2 in this special case, is thereby also not relevant until higher temperatures. This simplest type of equation focuses on the formation of pairs and their subsequent growth as stable nuclei, a case that has been termed irreversible nucleation by the first author[28] and many others. In the complete condensation regime, with the only energy E_d left in the problem, in the form of the diffusion coefficient $D=(\nu_d/4)\exp(-E_d/kT)$, it has become customary to emphasize the scaling of n_x with the lumped parameter $(D/F)^{-\chi}$, with $\chi=1/3$. This scaling is the same as in the $i=1$ case for 2D islands, as discussed and tabulated previously, where the general form follows the power law with exponent $\chi=i/(i+2)$. For 3D islands, this is modified slightly to $\chi=i/(i+2.5)$[20,21] under steady state conditions. These expressions are valid when the island shape is compact; small corrections are needed for fractal islands.

Further progress cannot be made without specific assumptions about bonding, and in particular about how the energy E_i can be expressed in terms of lateral ''bonds'' of strength E_b within the critical cluster. With this assumption, one can compute both the densities and deduce the critical size i as an output of the calculation; this size is the size that (self-consistently) gives the lowest nucleation rate and density, for all possible sizes j considered.[20] With the adoption of the Einstein model of lattice vibrations, contact was also made with the equilibrium vapor pressure at low supersaturation (high temperature), and so this model spans the complete range of behavior from low to high supersaturation, using just the three energy parameters, E_a, E_d and E_b. This is the simplest three-parameter model to achieve this result, and as such it is valuable as a base for further exploration of more complex models, as described here in Sec. III.

3. Rate equations without explicit critical island size

The second type of RE model takes a somewhat different approach. At high supersaturation, re-evaporation is negligible, so E_a is irrelevant; this is the complete condensation regime. Within this regime, at higher temperatures, with weak lateral bonds, atoms can leave the clusters by detachment and subsequent diffusion. In this case, nucleation has been termed reversible, and we need a suitable formulation to describe the rate of such detachment processes. As an example, the first author[28] has used the set of equations, ignoring direct impingement and coalescence, as

$$\frac{dn_1}{dt} = F - 2D\sigma_1 n_1^2 - Dn_1\sum_{s>1} \sigma_s n_s + 2D_{\mathrm{det}}\gamma_2 n_2 + D_{\mathrm{det}}\sum_{s>2} \gamma_s n_s, \tag{5}$$

$$\frac{dn_s}{dt} = Dn_1(\sigma_{s-1}n_{s-1} - \sigma_s n_s) + D_{\mathrm{det}}(\gamma_{s+1}n_{s+1} - \gamma_s n_s) \quad \text{for all } s>1. \tag{6}$$

Here, the σ_s are the capture numbers as before, and the terms that involve $D_{\mathrm{det}}\gamma_s$ describe the rate of detachment of an atom from an island of size s. Note that these equations as written here are valid only in the submonolayer growth regime, but coverage effects, coalescence of islands or desorption of adatoms can also be included.

The second and third terms in Eq. (5) are explicit forms of Eq. (1), such that

$$\tau_n^{-1} + \tau_c^{-1} = 2D\sigma_1 n_1 + D\sum_{s>1} \sigma_s n_s, \tag{7}$$

while the brackets in Eq. (6) emphasize the possibility of steady-state rather than local equilibrium, which was highlighted in the discussion following Eq. (2). In both Eqs. (5) and (6) the terms involving $D_{\mathrm{det}}\gamma_s$ involve detachment from s clusters, as discussed in more detail in the Appendix.

The capture of adatoms by other adatoms and islands occurs with an efficiency that is given by the capture numbers σ_s. If chosen properly these σ_s do encode spatial information. Many recent studies have the goal of finding a form for the σ_s that properly accounts for the spatial correlations between adatoms and islands. Some recent progress will be discussed below in Sec. III A. There has been less focus on the detachment rates γ_s. One reason is that at low temperatures detachment is negligible, and often one assumes that $\gamma_s=0$ for all s. This is of course the regime that is termed irreversible aggregation. At higher temperatures, the detachment should be related to the number of (attached) edge atoms, and simple expressions for this have been derived analytically.[29] But we would like to point out that the real situation might be more complicated. For example, the product $D_{\mathrm{det}}\gamma_s$ might reflect the fact that detachment can be strain dependent, and thus this term also has some spatial dependence.

C. Atomistic kinetic Monte Carlo simulations

An alternative to completely analytic approaches are atomistic models that explicitly take into account the spatial information and stochastic nature of thin film growth. One way to implement such models is in the form of KMC simulations.[6,30,31] The main difficulty and challenge in a KMC simulation is to identify which, and how many, microscopic processes need to be included. Different philosophies exist to address the latter question. One philosophy is that such KMC models will never include all the microscopic details, and that one should keep the model as simple as possible. This approach has been used to a large extent in the past 20 yr, and has helped to study and understand many general trends during growth. In this section we will refer to such models as generic growth models. On the other end of the spectrum one might want to include all the relevant microscopic processes, with the goal of making detailed predictions for a specific material system. We will refer to such models as high resolution growth models.

In a typical KMC simulation, one first specifies the processes that are included, and associates each process with a rate. At every timestep one identifies all the sites where any of the processes might occur. One of these events is then executed with a probability that corresponds to the rates. In practice, this is implemented in a way that one always executes an event (so there are no rejected moves), and after every event the simulation clock is updated appropriately.

A model that has been used with considerable success to study general trends is a simple cubic, solid-on-solid model with nearest neighbor interactions. The processes that are allowed are (random) adatom deposition, and hops to a nearest neighbor site, where the rate for a hop is given by $\nu=\nu_0\exp(-(E_d+nE_n)/kT)$. E_d is the energy barrier for surface diffusion, and in fact for $n=0$ it is $\nu=4D$; in the KMC literature E_d is often referred to as the surface bond energy. E_n is a nearest neighbor bond energy, ν_0 is the relevant prefactor (typically 10^{-12}–10^{-13} s^{-1}), and $n=0,1,2,3,4$ is the number of in-plane nearest neighbors. Additional microscopic parameters such as enhanced diffusion along a step edge or reduced (or enhanced) diffusion over a step edge (the so-called Ehrlich–Schwoebel barrier) can easily be added. Generic KMC models have been used successfully to study qualitative trends during growth, such as scaling laws. In limited cases they have also been successful for quantitative predictions such as the occurrence and decay of the reflection high energy electron diffraction signal. Note that the rates for all possible events at every site only depend on the local environment; a dependence of the rates on more long range interactions is possible, but at a significant additional computational cost.

In the literature one often finds a distinction between models for irreversible and reversible aggregation. The model just described is a reversible model, as it allows for atoms with lateral nearest neighbors to move. The rate for an atom with one nearest neighbor $n=1$ to move (and detach form an island boundary) then is $D_{\mathrm{det}}=\nu_0\exp(-(E_d+E_n)/kT)$. In the limit that E_n is large, and/or T is small, this model then effectively becomes an irreversible growth model. Thus, a continuous change of the parameter E_n induces a continuous change from irreversible to reversible growth. It has been demonstrated by one of us (C.R.) that the scaled ISD changes its shape continuously as a function of the reversibility.[28,32] This is illustrated in Fig. 3, where the scaled ISD as obtained from a KMC simulation[32] is shown in comparison with the ISD as obtained from experiment[33] and also from the level-set method (discussion below). It is evident that the ISD sharpens as the reversibility increases (which is realized in the simulation by decreasing E_n), in excellent agreement with the experimental ISD that sharpens as the growth temperature increases. We stress that no value for a critical island size has to be specified. Rather, the interplay between attachment and detachment kinetics describes the degree of reversibility, which one might associate with a (noninteger) effective critical island size.[32] We note that a similar sharpening of the ISD has been obtained by Amar and Family[34] in a KMC model where islands larger than the critical size (that has to be specified in advance) are assumed to be stable against breakup.

We conclude this section with some remarks about detailed high resolution KMC models. The work of Madhukar and co-workers[31] spearheaded the development of growth models for III/V compound semiconductor systems. However, the large number of microscopic parameters were difficult to obtain, since for example not even the exact surface reconstructions were known 15 yr ago. But the situation has changed in the last few years with the advances in density functional theory (DFT) calculations, together with the more readily available comparison to detailed experimental data, in particular scanning tunneling microscopy (STM) images. We are aware of several detailed high resolution KMC mod-

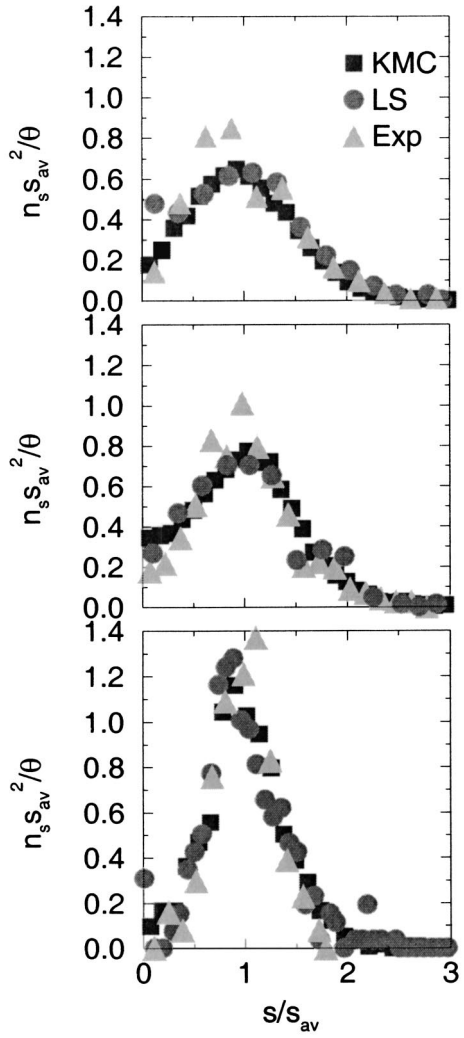

FIG. 3. Island size distribution, as given by KMC (squares) and LS (circles) methods, in comparison with STM experiments (triangles) on Fe/Fe(001) (Ref. 33). The reversibility increases from top to bottom.

els that have been published in the past few years.[35–37]

However, one problem with such high resolution growth models is that the range of magnitudes of all the different processes is rather large. Since the simulation timestep has to be defined by the fastest process, these models are often very slow. Thus, it has to be the goal to develop growth models that allow us to describe the fast processes in some averaged, mean-field approach. For example, repeated detachment and subsequent reattachment of atoms from and to island edges can be described by the overall net attachment (or detachment). One way to realize this is contained in a model that has been developed over the past few years, as described in the next section.

D. Level set (island dynamics) model

In the past few years, the island dynamics model, and a corresponding level set (LS) method for its numerical simulation have been developed.[38,39] This model is essentially continuum in the x-y-plane, but it resolves individual atomic layers. Within the model, the island boundaries for islands of height k are described as $\Gamma_k=\{\mathbf{x}:\phi(\mathbf{x})=k\}$ in which ϕ is the LS function that evolves according to

$$\frac{\partial\phi}{\partial t}+v|\nabla\phi|=0. \tag{8}$$

All the physical information is in the normal component v of the velocity function. Islands grow because atoms diffuse toward and attach to island boundaries, and shrink because they can detach from an island boundary. Thus, the velocity can be written as

$$v=v_{\text{att}}-v_{\text{shrink}}=D[\mathbf{n}\cdot\nabla\rho]-D_{\text{det}}p_{\text{esc}}\lambda. \tag{9}$$

In this equation, the first term on the right represents attachment of adatoms to the step from the upper and the lower terrace,[40] in which [ ] denotes the jump across the interface. The second term represents detachment of adatoms from island boundaries,[41] in which D_{det} is the microscopic detachment rate, p_{esc} is the probability for an atom to escape from the capture zone of the island it just detached from, and λ is the density of edge atoms that can detach. The adatom density $\rho(\mathbf{x},t)$ solves the diffusion equation

$$\frac{\partial n_1}{\partial t}-D\nabla^2 n_1=F-2\sigma_1 D\langle n_1^2\rangle \tag{10}$$

with the boundary condition $\rho=0$ for irreversible aggregation. The last term in Eq. (10) is the rate of nucleation of new islands, where $\langle\rangle$ denotes the spatial average.[42]

The model as described so far is almost completely deterministic. However, it was found that certain fluctuations are required to complete the model. The last term in Eq. (10) prescribes *when* a new island is to be nucleated. But the location of a new island has to be chosen with a probability weighted by the local value of n_1^2 as obtained from the solution of the diffusion equation.[43] Moreover, it was also found that randomness is essential in the thermal dissociation of small islands.[41] The idea of the LS method is illustrated in Fig. 4, where a snapshot of a typical LS simulation is shown. The stochastic prescriptions described above have been carefully validated by direct comparison with an atomistic KMC model. This is also illustrated in Fig. 3. Again, we see that the ISD becomes narrower and sharper as the detachment rate increases.

Significantly, and in contrast to atomistic simulation methods, the inclusion of fast microscopic processes comes at essentially no additional computational cost within this method. The reason for this is the following. In a typical atomistic simulation (where every event is resolved), the computational time increases dramatically when events with largely different rates are allowed. For example, frequent atomic detachment and reattachment at island boundaries is computationally expensive, even though the morphology might not change. In the LS approach, such fast events are accounted for in a quasi mean-field approach, without being explicitly resolved. This leaves the numerical time-step unchanged, resulting in essentially no increase in overall computational time.

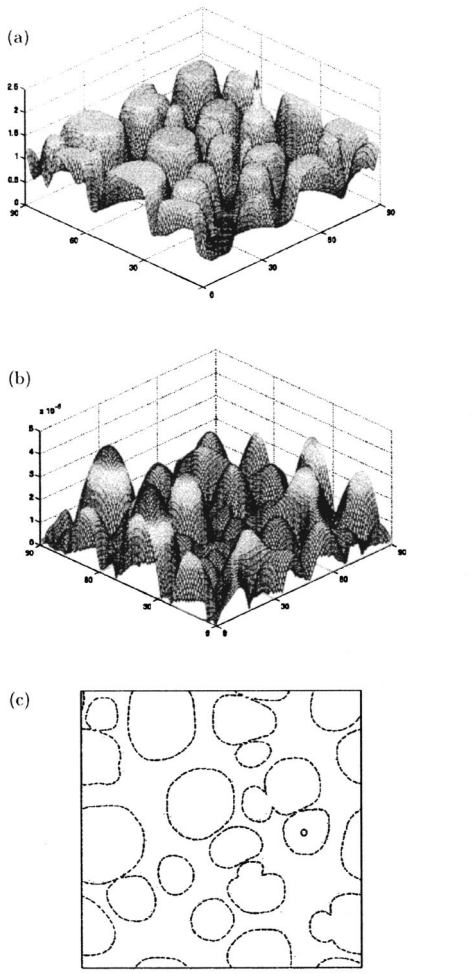

FIG. 4. Typical level set simulation. Shown is a snapshot of the level set function (a), and the corresponding adatom concentration (b). The island boundaries corresponding to this level set function are shown in (c).

This new method seems well-suited to model problems in epitaxial growth that may be difficult to describe with other methods. One current example is the inclusion of strain: the long range nature of the elastic field presents a challenge to include strain realistically in a KMC model (where the rates are typically determined by the local environment). However, recent progress in this area is intriguing and may present a way forward, as discussed in Secs. III C and III D. Moreover, solving the elastic equations is computationally very expensive, so that a method that allows for a large timestep is essential. As mentioned above, the level set method is also able to describe problems where the important events have vastly different rates. Work in both these areas is currently in progress.

III. RECENT PROGRESS: EXPERIMENTAL AND COMPUTATIONAL EXAMPLES

A. Capture numbers

The rate equations as formulated in Sec. II B are equations for mean field densities, and thus by construction do not include any explicit spatial information. Spatial information can however be included implicitly: consider the REs in the form of Eqs. (5) and (6). As mentioned already above, capture of adatoms by other adatoms and islands occurs with an efficiency that is given by the capture numbers σ_s. If chosen properly these σ_s do encode spatial information. For the purpose of the discussion in this section, we assume that there is no detachment from islands (i.e., $\gamma_s=0$). Many recent studies have the goal of finding a form for the σ_s that properly account for the spatial correlations between adatoms and islands. It has become customary to focus on the scaled ISDs to judge the validity of a certain choice for the capture numbers. In particular, simple RE approaches lead to a singularity of the ISD in the asymptotic scaling (large D/F) regime, and one test of any approximation of the σ_s is whether or not it leads to scaling of the ISD.

In this section we discuss the different choices that have been introduced for σ_s, and summarize some recent progress in understanding the form of the capture numbers. The simplest choice is $\sigma_s=1$ (i.e., a simple constant) for all s. This choice is often referred to as the point-island approximation. The only justification for this choice is that the equations are easier to analyze, and that certain analytic results can be easily obtained. Another simple choice is an s-dependence $\sigma_s=s^p$. In 2D, $p=0.5$ reflects the idea that capture is proportional to the perimeter of the island. Similarly, one can argue for the choice $p=1/3$ for 3D islands; both of these choices were assumed in the initial work on rate equations,[1] and have been made on occasion since then. However, it is well established that none of these choices for σ_s gives the correct scaling of the ISD; in particular, the scaled ISD becomes singular in the scaling limit.

There were several early attempts to account for the local environment of an island; this work was summarized and extended in Ref. 2. The uniform depletion approximation developed in that work is based on the mean-field assumption that at every point outside of an island the densities of islands of size s takes on an average value $n_s(t)$. This approach then gives an analytic formula for the σ_s in terms of modified Bessel functions. Bales and Chrzan[42] showed, for the irreversible nucleation ($i=1$) case, that the mean-field adatom and island densities n_1 and $N=\Sigma_{s\geq 2}n_s$ obtained from integrating the REs with these self-consistent capture numbers are in excellent agreement with the ones obtained from KMC simulations, for a wide range of D/F values. However, their approximation fails to reproduce properly the scaled ISD, which is a more stringent test for the spatial information in the σ_s.

A completely different approach was taken by Bartelt and Evans in Ref. 44. In this paper, the authors performed extensive KMC simulation of a point island model to *measure* numerically the capture numbers as a function of the island size. They suggest that σ_s depends linearly on s for $s>s_{\mathrm{av}}$, where s_{av} is the average island size, and is essentially a constant for $s<s_{\mathrm{av}}$. In a later article[45] they relaxed the point-island constraint and instead considered spatially extended islands, and obtained similar results (even though the

plateau is less pronounced in this case). The data is also in agreement with capture numbers that have been measured very carefully in an experiment where first Co islands are formed on Ru(001), and additional Cu is deposited subsequently. Because of the differences of Cu and Co in the STM, the additional (Cu) mass of each island can then be measured in a careful STM experiment. Similar data was obtained in a later experiment for Ag islands on Ag(001).[46]

Bartelt and Evans also established[44] that capture efficiency of islands, as described by the capture numbers σ_s, is intimately connected to the so-called capture areas A_s. The capture area of an island can be defined as the area surrounding an island, with the property that an adatom within a capture area, or zone, will on average diffuse toward the corresponding island (and will be captured by it). Geometrically, these capture areas A_s can be approximated by performing a Voronoi tessellation. In the complete condensation regime, the shape of the scaled capture area distribution A_s is similar to the shape of the scaled capture number distribution σ_s, a fact which has been known experimentally for a long time.[47,48] But in either case, capture numbers are only measured for one particular coverage, and these capture numbers are never used to integrate the REs. From more recent work[43,49] it is clear that these capture numbers can not produce the correct form of the ISD in the scaling limit.

Gibou *et al.*[49,50] determined capture numbers from simulations using the level set method. In contrast to the KMC approach, the dependence of the σ_s on s could be obtained at different times. The reason is that, due to the mean field treatment of the adatom density, a meaningful value for σ_s can be measured at any time, and growth of the islands does not have to be artificially suppressed. It was found that the functional form of the σ_s changes as a function of time. There is essentially no s-dependence in the nucleation phase, and an (almost[50]) linear dependence in the aggregation phase (cf. Fig. 5, upper panel). The rate equations can be integrated with these (time dependent) capture numbers, and the shape of the scaled ISD in fact does agree with the one obtained from LS or KMC simulations, and in particular does not display a singularity. This is shown in Fig. 5.

But there is no analytic form for σ_s that allows an integration of the REs from $t=0$, to get the proper ISD. The reason is that there is substantial cross-correlation between the capture numbers and related capture areas, which leads to significant numerical noise in the results in Refs. 49 and 50. It is not clear at the moment whether better simulation data will solve this problem, or whether there is a fundamental limitation within this approach.

The ultimate reasons why it has not been possible to date to find a functional form for all the $\sigma_s(t)$ are the following: Different islands of a particular size s have different capture areas. This is so because of spatial fluctuation during the nucleation of islands. As a result, these different islands of size s grow differently, which is hard to incorporate into a single σ_s. Moreover, the correlation between capture efficiency and the spatial distribution of islands is complicated due to nucleation. More precisely, the capture area (and thus the capture number) of a particular island may change (dramatically) if a new island is nucleated nearby.

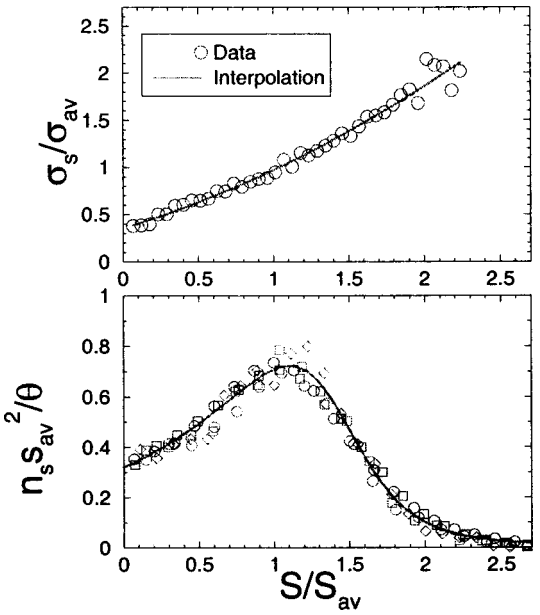

FIG. 5. Capture numbers in the aggregation phase as determined with the level set approach (upper panel). Shown in the lower panel is the scaled ISD obtained from the time dependent (measured) σ_s, in comparison with the original LS data (Ref. 50). The solid line in the lower panel is an analytic prediction from Ref. 45.

From the foregoing it is clear that any successful analytic treatment of the problem needs to consider the capture numbers σ_s as well as the capture areas A_s. There has been some progress in the past few years, but we believe that the problem has not yet been solved, and remains an area of active current research. Evans and Bartelt[51] developed a formalism where rate equations for the island densities n_s were complemented by rate equations for the average capture areas A_s. In a different approach it was suggested by Mulheran and Robbie[52] that the problem might be solved by considering a joint probability distribution (JPD) for islands and capture areas. Equations were introduced to describe the time evolution of the island density as a function of the island size s, and the size of capture areas, A. One complication that arises is: 'How are the capture areas treated in the case of a new nucleation event?' In Ref. 52 the evolution of the capture zones is modelled as a fragmentation process when new islands are nucleated. The JPD then obtained agrees with the one obtained from a KMC simulation.

Amar *et al.*[53,54] also utilized the idea of the JPD. They made some significant simplifications about the effect of nucleation on the capture areas. In essence they assumed that the capture area of the nth island is simply proportional to the average capture area at the time. For intermediate values of D/F their analytic model does indeed give better scaling of the scaled ISD. However, as mentioned in Ref. 55, and as is also evident from the simulation data presented by Popescu *et al.*,[54] the scaled ISD still exhibits singular behavior for larger values of D/F. It was pointed out by Evans and Bartelt[56] that the reason for this might be an unphysical delta

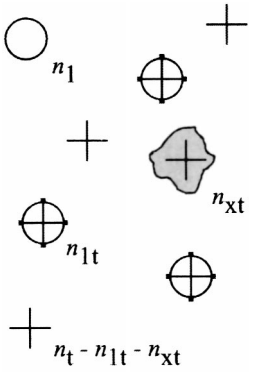

FIG. 6. Model for nucleation at attractive random point defects (density n_t), which can be occupied by adatoms (density n_{1t}), clusters (density n_{xt}) or can be empty (Refs. 19 and 67).

function scaling form for the JPD, which results from the fact that capture areas are not properly divided upon nucleation of new islands.

A recent, and perhaps more realistic treatment of the impact of nucleation on capture areas is given in Ref. 56. In this article, Evans and Bartelt study the JPD with a point island KMC simulation. Scaling of the JPD of the type $F(s/s_{\mathrm{av}}, A/A_{\mathrm{av}})$ is established. But whether or not this latest approach yields capture numbers that give proper scaling of the ISD in the asymptotic limit, as (D/F) tends to infinity, is currently still an open question that needs to be answered.

B. Nucleation on surface defects

All the above discussion has been concerned with nucleation and growth on a perfect substrate. But it is well known that many substrates are far from perfect, and indeed may contain impurities, surface point defects and/or steps, all of which may promote nucleation. Early examples, especially in island growth systems, which are particularly sensitive to such effects, are given in several reviews.[13,14,57] More recently, attention has turned to nucleation of small particles of Ag, Pd, and Pt on oxides such as MgO(001) and $Al_2O_3(0001)$, which are of interest as model catalysts. There are now several reviews of these systems that can be consulted for background information and detailed behavior.[58–61] There is an extensive theoretical literature, based on cluster chemical,[62] ionic pair potentials,[63,64] or density functional[65] calculations. For the most part, these calculations are not yet at the stage where definitive reviews can be given, largely because of their extreme sensitivity to charge imbalance at the oxide surface, and at surface defects, which may also be charged.

Here we discuss how defects have been incorporated into the framework of RE, KMC, and LS models. In all of these models, the introduction of defect traps requires that we specify their density, n_t, and the energy E_t with which diffusing adatoms are trapped. Within RE models, we can consider either irreversible trapping, such that adatoms cannot leave the traps (so that E_t is irrelevant), or reversible trapping as illustrated in Fig. 6, where we can qualitatively identify three regimes. This three-regime behavior has been reproduced in a steady state rate equation model developed by the second author,[64,66,67] which focuses on the rate equation for the trapped adatoms, density n_{1t}, and considers nucleation on both trap and normal sites.

The rate equation for n_{1t} is approximately

$$dn_{1t}/dt = \sigma_{1t} D n_1 n_{te} - n_{1t} \nu_d \exp(-(E_t + E_d)/kT), \quad (11)$$

where $n_{te} = n_t - n_{1t} - n_{xt}$ is the number of empty traps. For simplicity we have assumed that the capture numbers for empty and filled traps are the same. In steady state, this equation is zero, and inserting the usual form for $D \sim \nu_d \exp \times(-E_d/kT)$, we deduce

$$n_{1t}/(n_t - n_{xt}) = A/(1+A),$$

with

$$A = n_1 C_t \exp(E_t/kT), \quad (12)$$

where C_t is an entropic constant, which has been put equal to 1 in the illustrative calculations performed to date. Equation (12) shows that the traps are full ($n_{1t} = n_t - n_{xt}$) in the strong trapping limit, whereas they depend exponentially on E_t/kT in the weak-trapping limit, as expected. This equation is a Langmuir-type isotherm for the occupation of traps; the trapping time constant [τ_t, in analogy to Eq. (1)] to reach this steady state is very short, unless E_t is very large; but if E_t is large, then all the traps are full anyway.

The total nucleation rate is the sum of the nucleation rate on the terraces and at the defects. The nucleation rate equation without coalescence, analogous to Eq. (2), is

$$dn_x/dt = \sigma_i D n_1 n_i + \sigma_{\mathrm{it}} D n_1 n_{\mathrm{it}}, \quad (13)$$

where the second term is the nucleation rate on defects, and n_{it} is the density of critical clusters attached to defects, σ_{it} being the corresponding capture number. In the simplest case where the traps only act on the first atom which joins them, and entropic effects are ignored, we have

$$A_t = n_{1t}/n_1 = (n_t - n_{xt})A/[n_1(1+A)]. \quad (14)$$

Typically, there are three regions: a high-temperature region where adatoms visit the traps but can become detached from them; a low-temperature region where the traps are full, but the nucleation density is largely unaffected, since $n_x > n_t$. In between, there is a plateau region where $n_x = n_t$; this plateau is longer if E_t is higher and E_d lower. The first requirement is obvious, and the latter is required so that adatoms reach the traps before finding each other. This steady state model calculation, originally intended for Fe/CaF_2,[66,67] is shown for particular energy parameters in Fig. 7.

This defect nucleation model contains several subcases, depending on values of the parameters. An interesting example is Pd/MgO(001), studied with atomic force microscopy,[64,68] where a single set of experiments has been analyzed to put bounds on four energies; these data require a high trapping energy E_t and a low value of E_d, while also being sensitive to E_b and E_a. In this case, the high temperature portion of the data corresponds to the transition to $i = 3$, so that individual adatoms remain attached to traps, but

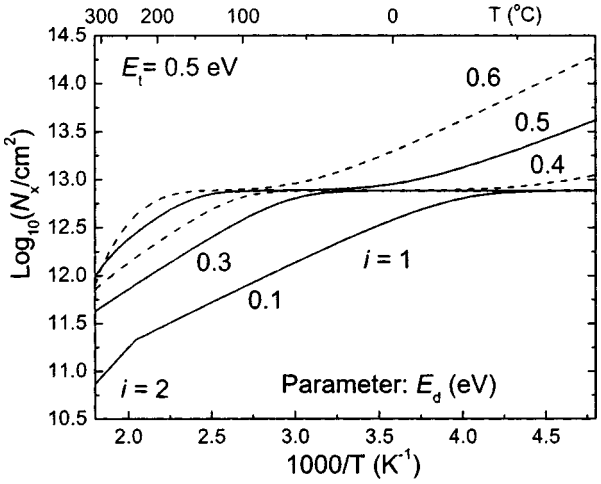

FIG. 7. Algebraic solution to rate equations for trapping energy $E_t=0.5$, $E_a=1.16$, $E_b=1.04$, and a range of E_d values between 0.1 and 0.6 eV. Recalculated (Ref. 67), after original model for $Fe/CaF_2(111)$ (Ref. 66).

subsequent adatoms can become detached. These features are in semi-quantitative agreement with calculations[62,64] for trapping of Pd in oxygen ion vacancies. The role of surface charges in stabilizing both surface vacancies on insulator surfaces, and small clusters attached to such point defects, is very marked. Currently, different calculations agree that such effects are strong, but disagree on their exact magnitude;[62,64,65] more comparative work is needed in this area.

The RE model described above predicts the number densities of islands, but (by construction) does not yield any spatial information, as for example the shape of the ISD. For this purpose, nucleation on defects has also been investigated using KMC and LS simulations. An example is the work of Lee and Barabási,[69] who showed in a KMC simulation that an ordered array of defect trapping centers can lead to a regular array of islands on the surface. A recent LS simulation[70] showed that the scaled ISD in the case of regularly spaced defect is a very sharp function (essentially a δ function). In the same study it was shown that in the opposite limit, when the defects are randomly distributed on the surface, the ISD has the form of a Γ distribution.

The ISD is controlled by the spatial distribution of defects when the mean diffusion length is comparable with or greater than the distance between defect traps. This corresponds to the upper end of the plateau regime, shown here in Fig. 7, where the adatom catchment area is roughly the same as the regular Voronoi polyhedron around each defect site. When the diffusion length is (much) less than the average distance between defects, both the nucleation density and the ISD converges to that obtained during nucleation on defect-free terraces. The narrow ISD suggests that if one can control experimentally the distribution of defect traps on the surface, one can control the ISD. The goal of a uniform ISD is clearly desirable for applications; we note that this may also be aided by stress and diffusion fields; this topic is also discussed in the next section.

C. Diffusion versus attachment-limited capture

The problem of determining capture numbers, and recent progress, has been described in Sec. III A, especially for determining size and spatial distributions in complete condensation. Most of this effort has gone into finding diffusion solutions for σ_1, σ_i and σ_x (or σ_s in general), especially for the case of irreversible nucleation ($i=1$), which is appropriate for STM experiments conducted at low temperatures.[24,25]

Low temperature deposition, even onto smooth close-packed metal surfaces, can be conducted in a regime where there is very little diffusion of adatoms during deposition. In such a case, essentially all motion occurs after deposition, and this regime, sometimes denoted by $i=0$, has been observed experimentally for the case of Cu/Ni(001),[71] see also Refs. 24 and 19. This regime implies no dependence or the nucleation density n_x on the flux F, but n_x increases and n_1 decreases following annealing at either higher temperatures and/or longer times.

Immediately following deposition there are no spatial correlations between adatoms on the surface, but diffusion during annealing establishes the spatial correlations, between adatoms and clusters, and between the adatoms themselves. There is an initial, transient, regime in which the capture number is time-dependent, before the diffusion solution becomes established. As discussed below, this transient is longer if there are (repulsive) interactions between adatoms, leading to attachment-limited behavior.

Two sets of STM experiments have been conducted on the deposition and annealing of Cu adatoms on Cu(111) at low temperatures.[72,73] After deposition and subsequent diffusion before observation, the spatial distribution is nonrandom, with a preferred spacing between adatoms. This feature has been analyzed quantitatively to determine the long-range oscillatory interaction between Cu adatoms as a function of radial separation. In the second of these experiments, Cu was deposited at 16.5 K, to sub-ML doses ($\sim 1.4\times 10^{-3}$ ML), followed by annealing at various temperatures around 20 K for times up to 20 min. At short distances, there is repulsion between adatoms, and this repulsion forms a barrier to ad dimer formation; but once formed, dimers are completely stable and do not diffuse.

These experiments test capture number models, as a repulsive barrier of height E_B, or a repulsive energy landscape of height V_0, changes the form of the diffusion field around adatoms and clusters, and reduces the capture number markedly if E_B/kT or $V_0/kT>0.2$. As shown recently,[74] the full time-dependent form of the capture numbers is required to obtain agreement between RE solutions and KMC simulations in the earliest stages of low coverage (sub-ML) annealing. The diffusion solution is almost sufficient when the barrier is zero, but for finite barriers the diffusion solution is quite wrong, and the attachment-limited (barrier) solution, $\sigma_s=2\pi(r_s+1)\exp(-E_B/kT)$ or $\sigma_s=2\pi(r_s+1)\times\exp(-V_0/kT)$ is much closer.

Surprisingly, this form of σ_s can be appropriate, even for barriers smaller than the diffusion energy. Note also that this capture solution is similar to the form originally used by

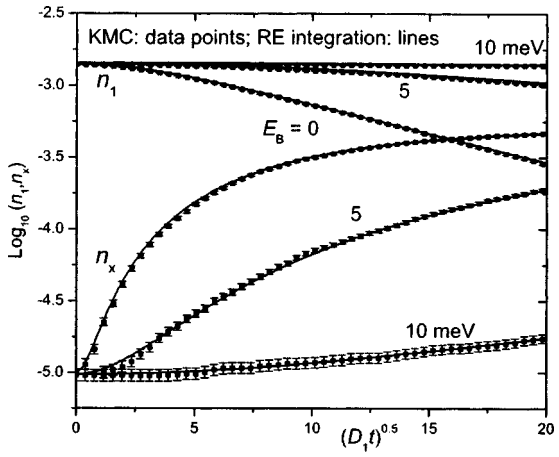

FIG. 8. RE solutions for n_1 and n_x annealing curves as a function of $(D_1 t)^{0.5}$, for annealing at 16.5 K with attachment barriers $E_B = 0$, 5, and 10 meV, compared to KMC simulations. The capture numbers used are based on an interpolation scheme between attachment barrier and diffusion solutions, showing essential agreement with the KMC simulations. See text for discussion of how these curves apply to STM experiments on Cu/Cu(111).

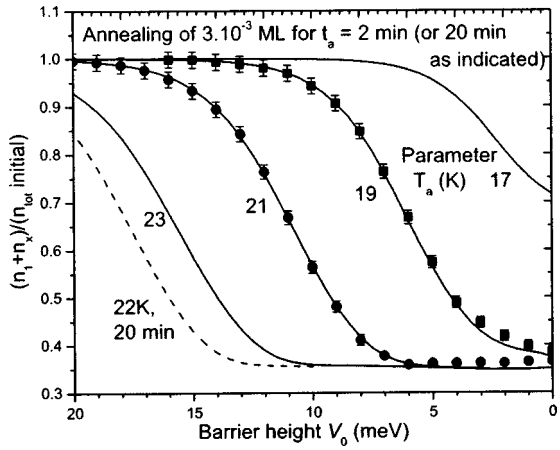

FIG. 9. RE solutions for annealing curves as a function of barrier height V_0, at temperatures 17 K$\leqslant T_a \leqslant$23 K. Plotted is the ratio (n_1+n_x) after a 2 min anneal, divided by the initial value $n_{\mathrm{tot}}=(n_1+n_x)$ after deposition. These curves use the time-dependent capture number expression as in Fig. 8. The curves for 19 and 21 K are also compared with the KMC simulations. Additionally a curve for annealing at 22 K for 20 min is given. See text for discussion of how these curves apply to STM experiments on Cu/Cu(111).

Zinsmeister,[1] but now reduced exponentially by the Boltzmann factor for the barrier. A related RE–KMC study by Ovesson[75] was applied to the deposition of Al/Al(111) as well as Cu/Cu(111). Fichthorn *et al.*[76] have found similar effects for Ag/Ag(111) and related systems. As emphasised above, these low temperature, smooth surface systems, correspond to the case where there are small, close-range, repulsive interactions prior to attachment, which determine the progress of nucleation. These repulsive interactions may be coupled with (even smaller) long-range oscillations, thought to be due to surface state interactions, which determine the preferred adatom spacings after deposition.

A full solution for annealing, appropriate to Cu/Cu(111), is shown in Fig. 8. As a result of having the agreement between the KMC and RE solutions, one can extrapolate with some confidence to other conditions, and compare with the experimental results.[72,73] These results showed no dimer formation during 20 min at $\sim$17 K and the completion of dimer formation after 20 min at 22 K. As a result, Venables and Brune[74] were able to deduce that the barrier height E_B, or alternatively the repulsive energy maximum V_0, for Cu/Cu(111), lies between 10 and 14 meV, as illustrated in Fig. 9. This figure is based on an integration of the REs for each V_0 value, up to the end of annealing (2 or 20 min) using the known E_d value, which is (40 ± 1) meV for Cu/Cu(111).[25,74] The comparison with KMC simulations is excellent, but the RE computation is much faster. This again points to a role for RE solutions to summarize large amounts of computation done by other methods.

D. Nucleation and Ostwald ripening

In 1900, Ostwald published a famous paper on the approach to equilibrium for a solution where a dense phase and a dilute phase coexist.[77] When the dense phase is present in the form of a distribution of compact clusters with different sizes, he argued that the Gibbs–Thomson effect[78] provides a thermodynamic driving force for large clusters to grow at the expense of small clusters. This phenomenon is called ripening or coarsening. The basic physics is the desire of the system to minimize the free energy associated with the interfaces between the two phases. Often ripening is a late-stage phenomenon, related to degradation of microstructures over long times, and thereby unrelated to nucleation and growth at early stages, the topic of this article. However, there are some cases where coarsening needs to be considered, and the relationship between nucleation, growth and ripening is of interest in its own right. Some comments on this topic are given below.

An authoritative review of the subject in the context of thin film growth is given by Zinke-Allmang *et al.*[79] In a series of articles, this group has studied island coarsening, both due to cluster mobility and to island instability. The term ripening is generally reserved for the latter phenomenon; it is usually formulated in terms of the (adatom) diffusion coefficient and the edge energies of islands.

In an atomistic context, the rate of Ostwald ripening for each island depends on the difference between attachment and detachment rates to/from the island. The first quantity depends on the number of adatoms on the surface and their mobility D, while the latter is related to the detachment rate D_{det} and density of atoms that can detach. Ostwald ripening can be incorporated into nucleation and growth models in several ways. For rate equations, there are at least two. In Sec. II B, we have not relied in Eqs. (5) or (6) on the idea of a critical nucleus size, and can consider the possibility that all clusters are to some extent unstable, both during deposition, and particularly during annealing. Then the difference between $\sigma_s D n_1$ and $D_{\mathrm{det}}\gamma_s$ is the relevant quantity to describe ripening. Calculations with specific forms of the latter term are described in Ref. 80.

An essentially equivalent RE approach, used by one of us

(J.V.), is to modify the capture time (τ_c) in Eq. (1), to allow for decay of 'stable' clusters. Integration of rate, or rate diffusion, equations containing such terms does allow for competition between nucleation, growth and ripening during deposition, and also gives a good overall description of annealing. This method has been applied to experiments on both Ag/Fe(001)[81] and Ag/Ge(111) and Ag/Si(111),[82] and relevant energies have been extracted. But for the semiconductor systems[82] there are details left to sort out, notably those concerned with small particle mobility and interdiffusion. So far this rate-diffusion equation approach has only been tested in one-dimensional geometries.

One might imagine that KMC simulations would be a powerful tool to study 2D Ostwald ripening. However, we are aware of only a single paper devoted to this subject.[83] The reason is, for realistic values of the physical parameters, island detachment processes are so slow that good statistics require an inordinate amount of computer time. Consequently, simulationists tend to focus on coarsening by island diffusion and coalescence,[84,85] a scenario that is known to occur for Ag(100).[86] Another approach that has been used to study Ostwald ripening is based on the boundary integral method.[87] This method is extremely efficient for solving the diffusion equation (and hence island growth velocities) for a surface morphology in the absence of nucleation and merger.

In recent studies the LS method described in Sec. II D has been used to simulate Ostwald ripening.[88,89] The LS method allows us to simulate nucleation and growth and subsequent coarsening within one unified approach, which is in contrast to the boundary integral method, where only the coarsening of islands can be simulated efficiently. It was shown in Ref. 88 that the predicted scaling behavior of the ISD, and the time evolution of the average island size, can be described very well within the LS approach.

One additional complication is that, in many systems of practical interest, we also would like to include the effects of stress in the islands, which gets larger during growth until limited by the introduction of misfit dislocations. We are presently quite a way from a fully quantitative model that includes all competing effects such as nucleation, growth (initial and stress-limited), ripening or coarsening, cluster shape fluctuations and transitions, stress-influenced interdiffusion, and so on. But many pieces of the argument are in place. For semiconductor systems such as Ge/Si(001) the high ad-dimer concentration and the small dimer–dimer interaction relates to a high critical nucleus size (small supersaturation) for nucleation and initial growth. The high density of mobile species makes both Ostwald ripening and cluster shape fluctuations relatively easy. For further discussions of this topic, and the role of stress and interdiffusion in the formation of quantum dots in semiconductor systems, we refer to recent reviews.[90–93]

IV. DISCUSSION AND FUTURE PROJECTION

Despite the large amount of work performed over the last 30 yr, even on submonolayer growth, there are many avenues still be to be explored. Some recent progress has been described in Sec. III. From these examples, it is clear that each new level of complexity requires new variables and material parameters for an adequate description of the model. There is a clear trend towards using at least two of the three mathematical methods, i.e., REs, KMC, and LS methods that we have described in Sec. II, for a comparison with experiment. There is also a clear trend, which we have not yet emphasized strongly in this article, to combine the results of these methods, or the parameters needed as input for such methods, with MD simulations, and with *ab initio* calculations such as DFT or cluster chemistry computations. But to construct all such models, and do all of these calculations, even for one system, represents a large amount of effort, which is typically well beyond the capacity of one research group. Thus the system has to be important enough, and the possible results decisive enough, to warrant the investment, both in money and time.

Our main argument for investing time and money in growth modeling is that we need to understand the basic processes in detail in order to be able to make worthwhile predictions. But, from an applied or industrial viewpoint, the argument in favor of modeling is simply that the cost of experiment by trial and error is rising even faster. There are now so many process steps in producing a device that engineers usually insist that experiments, even to introduce small changes, are performed under conditions essentially identical to those used in production. In the semiconductor industry, for example, this means that large wafers, few and infrequent experiments, and incremental change are the rule; modeling, provided it is well-grounded enough to extrapolate to new situations realistically, can be a much less expensive option. There are several calculations in the recent literature which have this as an aim, and many process simulators for which collaborations and web sites are available.[94]

The challenge for the scientists working on such topics is to know enough about the system to model it with sufficient confidence. It is well known, for example, that the widely used chemical vapor deposition (CVD) technique presents a particular challenge to modeling, since a whole sequence of processes occur, in the gas phase, on the surface, and during desorption of the reactants. By contrast molecular beam epitaxy (MBE), which is the context for much of the work described here, is much less used in production, and is only used for the most demanding applications. The reason is that it cannot match CVD for speed of growth and selectivity of reaction, and because it is more costly to install and maintain. Somehow, these gaps have to be overcome in the future by the modeling methods described here, as emphasized in recent articles[95] and funding initiatives.[96]

In this article we have not dealt in detail with coalescence[25] nor with subsequent nucleation and growth on higher layers, which can in principle be formulated.[97] There are also more recent papers on this topic[98–100] but the number of effective parameters becomes rather large, so that most such work has focused on the submonolayer regime, as described here. A particular current theoretical interest focuses on the adatom statistics on the second and higher lay-

ers being different from the first layer, because of the need to consider adatoms as belonging to a particular island, so that they cannot in general be considered to roam all over the substrate. We have not considered such effects in this article, and refer the reader to the above papers for ongoing discussion.

An ultimate goal for workers in this field would be to have models that are robust enough, and fast enough, so that production processes could be directly controlled from the model predictions. As far as we know, no atomistic model has actually been developed in this direction, and implemented to date. Clearly, measurement of system parameters, such as gas pressure, flow rate, source temperatures and time sequences are already used routinely for real time, in-situ, control of CVD and MBE growth. Of the three types of model discussed here, only REs are fast enough to accomplish real time prediction, and it seems worthwhile to consider further development of such tools in the direction of on-line control.

ACKNOWLEDGMENTS

The authors would like to thank R. Caflisch, D. Cahill, J. Drucker, E. Hall, and R. Vardavas for stimulating discussions. C.R. acknowledges financial support from the NSF focused research group Grant No. DMS-0074152.

APPENDIX: DECAY AND DETACHMENT RATES

This Appendix amplifies the remarks made in Sec. II B on decay and detachment rates in the various rate equation formulations. For example, we know that if single adatoms and dimers are in local equilibrium, we have $n_2 = C_2 n_1^2 \times \exp(-E_2/kT)$, a result originally derived by Walton,[26] and used in all models that contain the idea of a critical nucleus size i greater than one. The general form of the Walton equilibrium relation for size j is, in ML units

$$n_j = n_1^j \sum_m C_j(m) \exp(E_j(m)/kT), \tag{A1}$$

where we take account of the possibility of different configurations (m) at each size j, each with a statistical weight $C_j(m)$, which is an equilibrium property. Thus we can see that for small clusters $j \leqslant i$, if we make the equilibrium assumption, $U_j = 0$, and we have

$$\sigma_j D n_1 n_{j-1} = n_j \Gamma_j . \tag{A2}$$

Thus the complete expression for Γ_j must be

$$\Gamma_j = \sigma_j D \frac{\Sigma_{m-1} C_{j-1}(m) \exp(E_{j-1}(m)/kT)}{\Sigma_m C_j(m) \exp(E_j(m)/kT)}. \tag{A3}$$

If we restrict consideration to the configuration that has the highest binding energy, and drop the configuration label m, we obtain

$$\Gamma_j = (\sigma_j C_{j-1}/C_j) D \exp(-\Delta E_j/kT). \tag{A4}$$

This Eq. (A4) has the correct form as discussed above in the text for the general size s, but also has the correct preexponential $(\sigma_j C_{j-1}/C_j)$ appropriate for subcritical clusters, each of which has a dominant configuration. Assuming a single configuration is good for large supersaturation, but is very questionable at high temperatures, close to full (3D) equilibrium, when many fluctuations can be important. The reason for proceeding with such assumptions in practice is that, under typical growth conditions, one or more of the reaction rates are far from equilibrium, thus rendering the corresponding decay terms negligible.

The discussion in the text following Eq. (7) can be understood as follows. By comparing Eq. (6) with Eq. (A4), we see that local equilibrium for small clusters would correspond to

$$D_{\det} \gamma_s = (\sigma_s C_{s-1}/C_s) D \exp(-\Delta E_s/kT), \tag{A5}$$

where $\Delta E_s = E_s - E_{s-1}$ is the same definition as previously. Thus the two rate equation formulations both contain decay terms, which describe equivalent physical phenomena.

[1] G. Zinsmeister, Vacuum **16**, 529 (1966); Thin Solid Films **2**, 497 (1968); subsequent papers are cited in Ref. 14.
[2] J. A. Venables, Philos. Mag. **27**, 697 (1973).
[3] J. Villain, J. Phys. I **1**, 19 (1991).
[4] J. Krug, Adv. Phys. **46**, 139 (1997).
[5] M. Schneider, I. K. Schuller, and A. Rahman, Phys. Rev. B **36**, 1340 (1987); M. H. Grabow and G. H. Gilmer, Surf. Sci. **194**, 333 (1988); B. W. Dodson, CRC Crit. Rev. Solid State Mater. Sci. **16**, 115 (1990).
[6] J. D. Weeks and G. H. Gilmer, Adv. Chem. Phys. **40**, 157 (1979).
[7] R. Stumpf and M. Scheffler, Phys. Rev. B **53**, 4958 (1996).
[8] C. Ratsch, P. Ruggerone, and M. Scheffler, in *Surface Diffusion: Atomistic and Collective Processes*, edited by M. C. Tringides (Plenum, New York, 1997), pp 83–101.
[9] P. J. Feibelman, J. Vac. Sci. Technol. these proceedings.
[10] N. Israeli and D. Kandel, Phys. Rev. Lett. **88**, 116103 (2002).
[11] W. K. Burton, N. Cabrera, and F. C. Frank, Philos. Trans. R. Soc. London, Ser. A **243**, 299 (1951).
[12] E. Bauer, Z. Kristallogr. **110**, 372 (1958); E. Bauer and H. Poppa, Thin Solid Films **12**, 167 (1972).
[13] R. Kern, G. LeLay, and J. J. Métois, in Curr. Top. Mater. Sci. **3**, 139 (1979).
[14] J. A. Venables, G. D. T Spiller, and M. Hanbücken, Rep. Prog. Phys. **47**, 399 (1984).
[15] A. F. Voter, J. Chem. Phys. **106**, 4665 (1997); A. F. Voter, Phys. Rev. Lett. **78**, 3908 (1997).
[16] G. H. Vineyard, J. Phys. Chem. Solids **3**, 121 (1957).
[17] H. R. Glyde, Rev. Mod. Phys. **39**, 373 (1967).
[18] I. Markov, *Crystal Growth for Beginners* (World Scientific, 1995).
[19] J. A. Venables, *Introduction to Surface and Thin Film Processes* (Cambridge University Press, Cambridge, 2000).
[20] J. A. Venables, J. Vac. Sci. Technol. B **4**, 870 (1986); Phys. Rev. B **36**, 4153 (1987).
[21] J. A. Venables, Surf. Sci. **299–300**, 798 (1994).
[22] B. Lewis, Surf. Sci. **21**, 289 (1970); B. Lewis and G. J. Rees, Philos. Mag. **29**, 1253 (1974); B. Lewis and J. C. Anderson, *Nucleation and Growth of Thin Films* (Academic, New York, 1978).
[23] M. J. Stowell, Philos. Mag. **26**, 361 (1972); Thin Solid Films **21**, 91 (1974).
[24] H. Brune, Surf. Sci. Rep. **31**, 121 (1998).
[25] H. Brune, G. S. Bales, J. Jacobsen, C. Boragno, and K. Kern, Phys. Rev. B **60**, 5991 (1999).
[26] D. Walton, J. Chem. Phys. **37**, 2182 (1962).
[27] P. Jensen, H. Larralde, and A. Pimpinelli, Phys. Rev. B **55**, 2556 (1997).
[28] C. Ratsch, A. Zangwill, P. Šmilauer, and D. D. Vvedensky, Phys. Rev. Lett. **72**, 3194 (1994).
[29] R. E. Caflisch, W. E, M. F. Gyure, B. Merriman, and C. Ratsch, Phys. Rev. E **59**, 6879 (1999).
[30] S. Clarke and D. D. Vvedensky, Phys. Rev. Lett. **58**, 2235 (1987).
[31] A. Madhukar and S. V. Ghaisas, CRC Crit. Rev. Solid State Mater. Sci. **14**, 1 (1988).

[32]C. Ratsch, P. Šmilauer, A. Zangwill, and D. D. Vvedensky, Surf. Sci. Lett. **329**, L599 (1995).
[33]J. A. Stroscio and D. T. Pierce, Phys. Rev. B **50**, 6057 (1994).
[34]J. G. Amar and F. Family, Phys. Rev. Lett. **74**, 2066 (1995).
[35]M. Itoh, G. R. Bell, A. R. Avery, T. S. Jones, B. A. Joyce, and D. D. Vvedensky, Phys. Rev. Lett. **81**, 633 (1998).
[36]P. Kratzer and M. Scheffler, Phys. Rev. Lett. **88**, 036102 (2002).
[37]F. Grosse, W. Barvosa-Carter, J. Zinck, M. Wheeler, and M. F. Gyure, Phys. Rev. Lett. **89**, 116102 (2002).
[38]M. F. Gyure, C. Ratsch, B. Merriman, R. E. Caflisch, S. Osher, J. J. Zinck, and D. D. Vvedensky, Phys. Rev. E **58**, R6927 (1998).
[39]S. Chen, M. Kang, B. Merriman, R. E. Caflisch, C. Ratsch, R. P. Fedkiw, M. Gyure, and S. Osher, J. Comput. Phys. **167**, 475 (2001).
[40]C. Ratsch, M. F. Gyure, R. E. Caflisch, F. Gibou, M. Petersen, M. Kang, J. Garcia, and D. D. Vvedensky, Phys. Rev. B **65**, 19503 (2002).
[41]M. Petersen, C. Ratsch, R. E. Caflisch, and A. Zangwill, Phys. Rev. E **64**, 061602 (2001).
[42]G. S. Bales and D. C. Chrzan, Phys. Rev. B **50**, 6057 (1994).
[43]C. Ratsch, M. F. Gyure, S. Chen, M. Kang, and D. D. Vvedensky, Phys. Rev. B **61**, R10598 (2000).
[44]M. C. Bartelt and J. W. Evans, Phys. Rev. B **54**, R17359 (1996).
[45]M. C. Bartelt, A. K. Schmidt, J. W. Evans, and R. Q. Hwang, Phys. Rev. Lett. **81**, 1901 (1998).
[46]M. C. Bartelt, C. R. Stoldt, C. J. Jenks, P. A. Thiel, and J. W. Evans, Phys. Rev. B **59**, 3125 (1999).
[47]J. A. Venables and D. J. Ball, Proc. R. Soc. London, Ser. A **322**, 331 (1971).
[48]P. A. Mulheran and J. A. Blackman, Phys. Rev. B **53**, 10261 (1996).
[49]F. Gibou, C. Ratsch, S. Chen, M. F. Gyure, and R. E. Caflisch, Phys. Rev. B **63**, 115401 (2001).
[50]F. Gibou, C. Ratsch, and R. E. Caflisch, Phys. Rev. B **67**, 155403 (2003).
[51]J. W. Evans and M. C. Bartelt, Phys. Rev. B **63**, 234508 (2001).
[52]P. A. Mulheran and D. A. Robbie, Europhys. Lett. **49**, 617 (2000).
[53]J. G. Amar, M. N. Popescu, and F. Family, Phys. Rev. Lett. **86**, 3092 (2001).
[54]M. N. Popescu, J. G. Amar, and F. Family, Phys. Rev. B **64**, 205404 (2001).
[55]D. D. Vvedensky, C. Ratsch, F. Gibou, and R. Vardavas, Phys. Rev. Lett. **90**, 189601 (2003).
[56]J. W. Evans and M. C. Bartelt, Phys. Rev. B **66**, 235410 (2002).
[57]J. A. Venables, Mater. Res. Soc. Symp. Proc. **440**, 129 (1997).
[58]C. T. Campbell, Surf. Sci. Rep. **27**, 1 (1997).
[59]V. E. Henrich and P. A. Cox, *The Surface Science of Metal Oxides*, 2nd ed. (Cambridge University Press, Cambridge 1996); C. Noguera, *Physics and Chemistry of Oxide Surfaces* (Cambridge University Press, Cambridge, 1996).
[60]H. J. Freund, Rep. Prog. Phys. **59**, 283 (1996); Surf. Sci. **500**, 271 (2002).
[61]C. Henry, Surf. Sci. Rep. **31**, 231 (1998).
[62]A. M. Ferrari and G. Pacchioni, J. Phys. Chem. **100**, 9032 (1996).
[63]J. H. Harding, A. M Stoneham, and J. A. Venables, Phys. Rev. B **57**, 6715 (1998).
[64]J. A. Venables and J. H. Harding, J. Cryst. Growth **211**, 27 (2000).
[65]A. Bogicevic and D. R. Jennison, Surf. Sci. Lett. **515**, L481 (2002).
[66]K. R. Heim, S. T. Coyle, G. G. Hembree, J. A. Venables, and M. Scheinfein, J. Appl. Phys. **80**, 1161 (1996).
[67]J. A. Venables, in *Thin Films: Heteroepitaxial Systems*, edited by W. K. Liu and M. B. Santos (World Scientific, 1999), Chap. 1.
[68]G. Haas, A. Menck, H. Brune, J. V. Barth, J. A. Venables, and K. Kern, Phys. Rev. B **61**, 11105 (2000).
[69]C. Lee and A.-L. Barabási, Appl. Phys. Lett. **73**, 2651 (1998).
[70]R. Vardavas, C. Ratsch, and R. E. Caflisch (unpublished).
[71]B. Müller, L. Nedelmann, B. Fischer, H. Brune, and K. Kern, Phys. Rev. B **54**, 17858 (1996).
[72]N. Knorr, H. Brune, M. Epple, A. Hirstein, M. A. Schnieder, and K. Kern, Phys. Rev. B **65**, 115420 (2002).
[73]J. Repp, F. Moresco, G. Meyer, K.-H. Rieder, P. Hyldgaard, and M. Persson, Phys. Rev. Lett. **85**, 2981 (2000).
[74]J. A. Venables and H. Brune, Phys. Rev. B **66**, 195404 (2002).
[75]S. Ovesson, Phys. Rev. Lett. **88**, 116102 (2002).
[76]K. A. Fichthorn and M. Scheffler, Phys. Rev. Lett. **84**, 5371 (2000); K. A. Fichthorn, M. L. Merrick, and M. Scheffler, Appl. Phys. A: **75**, 17 (2002).
[77]W. Ostwald, Z. Phys. Chem. **34**, 495 (1900).
[78]P. Pelcé, *Dynamics of Curved Fronts* (Academic, San Diego, 1988); see also J. G. McLean, B. Krishnamachari, D. R. Peale, E. Chason, J. P. Sethna, and B. H. Cooper, Phys. Rev. B **55**, 1811 (1997).
[79]M. Zinke-Allmang, L. C. Feldman, and M. H. Grabow, Surf. Sci. Rep. **16**, 377 (1992).
[80]G. S. Bales and A. Zangwill, Phys. Rev. B **55**, R1973 (1997).
[81]H. Noro, R. Persaud, and J. A. Venables, Surf. Sci. **357–358**, 879 (1996).
[82]J. A. Venables, F. L. Metcalfe, and A. Sugawara, Surf. Sci. **371**, 420 (1997).
[83]P.-M. Lam, D. Bayayoko, and X.-Y. Hu, Surf. Sci. **429**, 161 (1999).
[84]T. R. Mattsson, G. Mills, and H. Metiu, J. Chem. Phys. **110**, 12151 (1999).
[85]A. Lo and R. T. Skodje, J. Chem. Phys. **112**, 1966 (2000).
[86]L. Bardotti, M. C. Bartelt, C. J. Jenks, C. R. Stoldt, J.-M. Wen, C.-M. Zhang, P. A. Thiel, and J. W. Evans, Langmuir **14**, 1487 (1998).
[87]N. Akaiwa and D. I. Meiron, Phys. Rev. E **51**, 5408 (1995).
[88]M. Petersen, A. Zangwill, and C. Ratsch, Surf. Sci. (in press, 2003).
[89]D. L. Chopp, J. Comput. Phys. **162**, 104 (2000).
[90]R. M. Tromp and J. B. Hannon, Surf. Rev. Lett. **9**, 1565 (2002).
[91]J. Drucker, IEEE J. Quantum Electron. **38**, 975 (2002).
[92]D. G. Cahill, J. Vac. Sci. Technol. A, these proceedings.
[93]J. A. Venables, P. A. Bennett, H. Brune, J. Drucker, and J. H. Harding, Philos. Trans. R. Soc. London, Ser. A **361**, 311 (2003).
[94]U. Hansen, S. Rodgers, and K. F. Jensen, Phys. Rev. B **62**, 2869 (2000); see for example the Semiconductor Subway web site at http://www-mtl.mit.edu/semisubway.html
[95]T. D. De la Rubia and V. V. Bulatov, MRS Bull. **26**, 169 (2001); see especially the article by F. H. Baumann *et al.*, *ibid.* **26**, 182 (2001).
[96]See, e.g., B. Cipra, Science **297**, 174 (2002).
[97]R. Kariotis and M. G. Lagally, Surf. Sci. **216**, 557 (1989).
[98]J. Krug, P. Politi, and T. Michely, Phys. Rev. B **61**, 14037 (2000).
[99]J. Rottler and P. Maas, Phys. Rev. Lett. **83**, 3490 (1999); S. Heinrichs, J. Rottler, and P. Maas, Phys. Rev. B **62**, 8338 (2000).
[100]P. Politi and C. Castellano, Phys. Rev. E **66**, 031605 (2002); **66**, 031606 (2002).