Energetics of Si(001)

H. J. W. Zandvliet

Faculty of Applied Physics and MESA+ Research Institute, University of Twente, 7500 AE Enschede, The Netherlands

A classical thermodynamic description of a surface requires the introduction of a number of energetic parameters related to the surface steps. These parameters are the step free energy, the kink creation energy, and the energetic and entropic interactions between steps. This review will demonstrate how a statistical analysis of scanning tunneling microscopy images of stepped surfaces can provide appropriate values of these fundamental energetic parameters. The Si(001) surface is used as a model system. In order to illustrate the significance of these energetic parameters, two morphological surface phase transitions are discussed, namely, the thermal roughening transition and the orientational phase diagram.

CONTENTS

I. INTRODUCTION

Surface steps play a key role in many equilibrium surface processes (e.g., faceting and thermal roughening) as well as nonequilibrium processes (e.g., crystal growth and etching) which are very important from a technological point of view. With this in mind, it is not surprising that a detailed knowledge of the step-related parameters is a prerequisite for determination of the surface free energy of a vicinal surface. A primary objective of this review is to assess the information that can be extracted from surface steps encountered in scanning tunneling microscopy (STM) images. With the advent of the scanning tunneling microscope, developed in 1982 by Binnig and Rohrer (1982; Binnig *et al.* 1982a, 1982b), it became possible to image surfaces on the atomic scale in real space. In the 1950s, atomic-scale imaging of sharp metal tips was already possible using the field ion microscope (FIM; Müller, 1951; Tsong, 1988). Nevertheless, when seeking to identify an unknown structure or making exact measurements of structural parameters, it is still necessary to rely on diffraction experiments. Direct imaging techniques such as STM and field ion microscopy are ideal for studying features that represent a disruption in the periodicity, such as defects and steps. As will be shown, observation of these disruptions or deviations from the ideal periodicity will provide the information needed to describe quantitatively the thermodynamic behavior of surfaces. The scanning tunneling microscope has many advantages relative to the FIM, notably, the ability to examine a much wider range of solids and adsorbates and to view relatively large regions of single-crystal planes with atomic resolution. As previously shown by a number of scientists, the energetic parameters can be derived from FIM data; however, the statistical error is generally large due to the small sample size. STM measurements are possible over a temperature range from well below room temperature to as high as 800–1000 K (which is still below the equilibration temperature of some materials (Zandvliet, van Loenen, and Elswijk, 1992).[1]

This paper briefly reviews how step properties are related to the thermodynamic properties of surfaces, and how step-related energetic parameters can be extracted from atomically resolved images of surfaces. The main focus is on the method rather than the specific surface under study. Because steps are also the fundamental units in surface mass transport, which is required for morphological phase transitions and kinetically limited processes (such as growth and etching), the importance of steps extends beyond the equilibrium properties discussed in this review.

II. THE Si(001) SURFACE

The technologically important Si(001) surface is at present the most extensively studied surface. The driving force for this interest is that currently no less than 95% of the total microelectronic market relies on Si(001). Although Si(001) has a relatively small surface unit mesh, the surface displays a wealth of fascinating phenomena, rendering it an ideal model system for study. A number of reviews on this surface have appeared (Schlüter, 1988; Griffith and Kochanski, 1990; Markov, 1995; Liu and Lagally, 1997); however, none of these reviews has focused on the determination of the key energetic parameters of Si(001) that govern its thermodynamic behavior.

When a silicon crystal is cut along the (001) plane, each surface atom is left with two dangling bonds. However, the Si(001) surface reconstructs to form rows of

[1]The room-temperature or near-room-temperature STM images of Si(001) surfaces, for example, represent structures equilibrated at a higher temperature (typically 750–850 K).

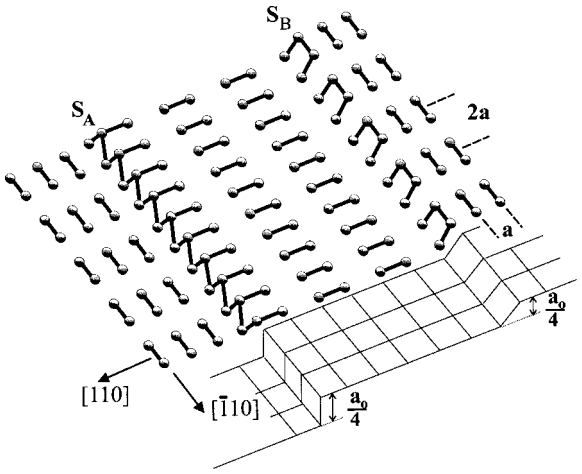

FIG. 1. Schematic diagram of a single-layer-stepped Si(001) surface. The dumbbells represent surface dimers. The two different types of single-layer steps are labeled S_A and S_B ($a_o = 5.43$ Å and $a = 3.84$ Å).

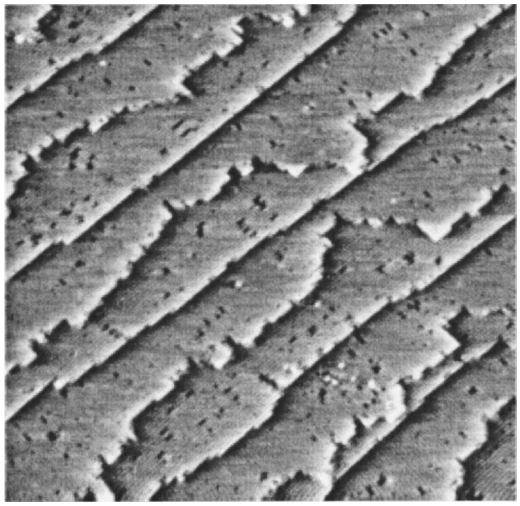

FIG. 2. Scanning tunneling microscope image (100×100 nm^2) of single-layer-stepped Si(001) 0.5° misoriented towards the [110] direction, obtained at -2-V sample bias and 0.5-nA tunneling current.

dimerized atoms, yielding a (2×1) unit cell. The driving force for reconstruction is the reduction in the number of dangling bonds from two per surface atom in the unreconstructed case to only one in the reconstructed case. This dimerization was first proposed by Schlier and Farnsworth (1959) and was first observed in real space by scanning tunneling microscopy in 1985 (Tromp *et al.*, 1985; Hamers *et al.*, 1986). Figure 1 shows a schematic view of the Si(001) surface. The (001) reconstruction establishes two characteristic directions on the surface, either parallel or perpendicular to the substrate dimer rows in orthogonal [110] directions, leading to many interesting anisotropies. For example, with dimerization an anisotropic surface stress tensor develops; the surface is under a tensile stress along the dimer bond and a compressive stress perpendicular to the dimer bond, i.e., along the substrate dimer row. According to *ab initio* calculations, the 2×1 dimerization reduces the energy by ~1 eV per surface atom. The Si(001) surface remains dimerized up to temperatures of at least 1475 K. Because of the tetrahedral bonding configuration in the diamond lattice, the dimer direction is orthogonal on terraces separated by an odd number of single-layer steps, giving rise to the existence of both 2×1 and 1×2 domains. There are two different types of single-layer steps: those that run along the dimer rows of the upper terrace (S_A steps) and those that run perpendicular to the dimer row direction of the upper terrace (S_B steps). On vicinal Si(001) surfaces misaligned along the [110] direction, the two types of steps alternate. The S_A step appears to be smooth, while the S_B step contains a high density of thermally excited kinks (Fig. 2). The kinks in S_A (S_B) consist of segments of S_B (S_A) steps. The surface periodicity rotates from 2×1 to 1×2 or vice versa at each single-layer step. Consider now a Si(001) surface slightly misoriented, by an angle θ, in the [110] direction. The surface misorientation can be accommodated by single-layer steps and double-layer steps, leading to surfaces that are different not only in the height of the step edges and width of the terraces, but also in lattice structure. The single-layer-stepped surface is a double domain surface, exhibiting both 2×1 and 1×2 domains, while the double-layer-stepped surface is a single domain surface. For miscut angles larger than about 1.5° toward [110], double-layer steps begin to form and their fraction increases with increasing miscut angle toward a maximum of nearly 100% at 5–6° (Fig. 3).

III. THE CLASSICAL THERMODYNAMIC MODEL

The classical thermodynamic formalism for describing surfaces was introduced by Gibbs (1957) and later extended by Herring (1951). In the 1960's Gruber and Mullins (1967) arrived at a particularly important conclusion, namely, that the surface free energy is governed

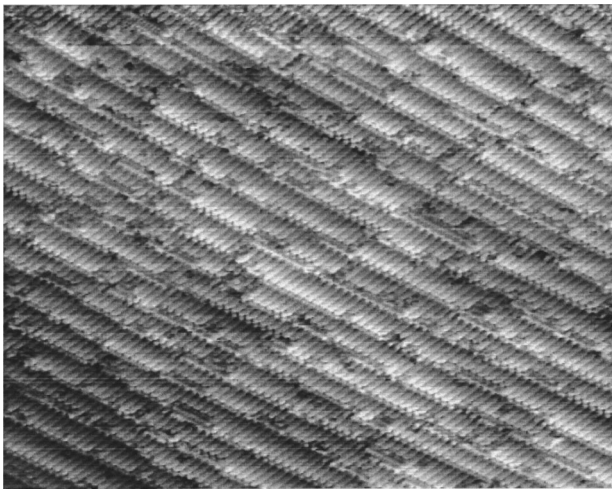

FIG. 3. Scanning tunneling microscope image (40×32 nm^2) of double-layer-stepped Si(001) 4.5° misoriented towards the [110] direction, obtained at -2-V sample bias and 0.5-nA tunneling current.

by the meandering of steps. More recently Williams and Bartelt (1991; Williams, 1994) developed statistical-mechanical theories describing the surface free energy in terms of step-related parameters. The interested reader is referred to particularly beautiful reviews by Williams and Bartelt (1991), Williams (1994), and Jeong and Williams (1999). To begin, it is appropriate to discuss the main expression for the surface free energy of a slightly misoriented low-Miller-index[2] surface. Low-Miller-index surfaces generally correspond to local minima in the surface free energy. At temperature T, the free energy per unit area $\gamma(\theta,T)$ of a surface having a small angle of misorientation θ with respect to a low-Miller-index surface can be calculated by considering the surface as consisting of terraces separated by steps of single-layer height. The surface free energy per unit area is given by

$$\gamma(\theta,T)=\gamma(0,T)+\frac{f(T)}{d}|\tan\theta|+q(T)|\tan\theta|^3, \qquad (1)$$

where $\gamma(0,T)$ refers to the surface free energy per unit area for the low-Miller-index surface, d is the step height, $f(T)$ is the free energy per unit length ($a_\parallel$) for forming a step of height d, and finally $q(T)|\tan\theta|^3$ corresponds to the energy cost per unit area due to step-step interactions. The average spacing between the steps is given by $L=d/\tan(\theta)$. Since steps cannot cross, there is a temperature-dependent entropic repulsion between them, decaying as L^{-2} (Gruber and Mullins, 1967; Fisher and Fisher, 1982). In general, the energetic step-step interaction exhibits a similar decay, i.e., $\sim L^{-2}$, but is not temperature dependent. In principle both $f(T)$ and $q(T)$ will depend on the polar angle ϕ (related to the averaged direction of the steps with respect to a high-symmetry direction). However, for the sake of simplicity I consider here only steps running along a high-symmetry direction, i.e., $\phi=0$.

IV. DETERMINATION OF STEP-EDGE-RELATED ENERGETIC PARAMETERS FROM REAL-SPACE OBSERVATIONS

At zero temperature, steps are always as straight as possible. The only kinks that are present in steps at zero temperature are the so-called forced kinks. These forced kinks are necessary in order to accommodate the misalignment of a step with respect to one of the principal crystal high-symmetry axes. At higher temperatures, thermally induced kink pairs are also generated. When viewed along the step edge, there are two types of kinks: those pointing toward the lower terrace (positive kinks) and those pointing toward the upper terrace (negative kinks). The symbol $n_{\pm k}$ will be used to denote the probability that there is a jump (or drop) of length $2ak$ at a given position in the step edge. A unit of two dimers has been chosen as the basic building block of the 2×1 surface, due to the observation that there are two types of single-layer S_B steps, depending on where the dimer row of the upper terrace ends with respect to the dimer rows of the lower terrace (rebonded or nonbonded S_B step). Both types of S_B single-layer steps have been observed (Tromp *et al.*, 1985), and hence occasionally kink lengths of an odd number in a do occur, but in general kinks are found on a (2×2) lattice (rebonded S_B steps occur much more frequently than nonbonded S_B steps). At any point on the step, the normalization condition holds:

$$n_0+\sum_{k=1}^{\infty}(n_{+k}+n_{-k})=1. \qquad (2)$$

In an elegant STM paper, Swartzentruber *et al.* (1990) analyzed the kink density along both types of single-layer step edges (S_A and S_B) on Si(001), using a Boltzmann probability for independent kink excitations, i.e., $n_{+k}+n_{-k}=2n_0\exp(-\varepsilon_k/k_bT)$, where ε_k is the creation energy of a kink with a length of k units. In retrospect this paper by Swartzentruber and co-workers (1990) marked the beginning of a period in which one attempted to extract the fundamental energetic parameters from atomically resolved real-space images.[3] An earlier, classic paper (Burton, Cabrera, and Frank, 1951) had already derived a number of thermodynamic relations for various kink densities of a step on a (001) surface of a Kossel crystal with isotropic nearest-neighbor and next-nearest-neighbor interactions. For an anisotropic Si(001) surface, the following thermodynamic relations for different single-layer S_A (S_B) steps can be derived (Zandvliet, Elswijk, *et al.*, 1992):

$$\varepsilon_{\parallel(\perp)}/k_bT=-\ln\left(\frac{n_{+1}n_{-1}}{n_0^2}\right), \qquad (3)$$

$$\delta/k_bT=-\ln\left(\frac{n_{\pm r}n_0}{n_{\pm(r-1)}n_{\pm1}}\right), \quad r\geqslant2. \qquad (4)$$

Hence $\varepsilon_{\parallel(\perp)}$ refers to the nearest-neighbor interaction energy between a unit building block ($2a\times2a$) along (perpendicular to) the substrate-dimer-row direction, and δ refers to the second-nearest-neighbor interaction.

[2]The relation between reciprocal lattice vectors and families of lattice planes provides a convenient way to specify the orientation of a lattice plane. In general one describes the orientation of a plane by giving a vector normal to the plane. The three primitive lattice vectors are often used to represent the normal. The Miller indices of a lattice plane are the coordinates of the shortest reciprocal lattice vector normal to that plane, with respect to a specified set of primitive reciprocal lattice vectors. The Miller indices of a lattice plane can be found in the following way: assume that the lattice plane intersects the coordinate axes in real space at values a, b, c, where each of these numbers refers to an integer multiple of the corresponding basis vector. One then takes the reciprocal values $1/a$, $1/b$, $1/c$ and multiplies this set by an integer m in such a way that m/a, m/b, and m/c are all integers. $h=m/a$, $k=m/b$, and $l=m/c$ are denoted as the Miller indices of the lattice plane (hkl).

[3]It should be noted that the application of the Boltzmann equation in the specific form proposed by Swartzentruber *et al.* (1990) leads to correct kink creation energies only in the absence of an azimuthal misalignment of the steps with respect to a high-symmetry direction ($\phi=0$).

TABLE I. Calculated step-edge formation energies of the Si(001) surface ($a=3.84$ Å).

	S_A (eV/2a)	S_B (eV/2a)	D_A (eV/2a)	D_B (eV/2a)
Chadi (1987)	0.02	0.30	1.08	0.10
Poon *et al.* (1990, 1992)	0.004	−0.022	0.321	−0.085
Bowler and Bowler (1998)	0.04	0.16		
Oshiyama (1995)	0.18	0.24	0.86	0.34

Swartzentruber *et al.* (1990) introduced a corner energy term rather than a second-nearest-neighbor interaction term to explain the kink length distribution. As has been discussed extensively by Zandvliet, Elswijk, *et al.* (1992), both interpretations, i.e., nearest-neighbor interactions and corner energy versus nearest-neighbor and next-nearest-neighbor interactions, are equally valid. Equations (3) and (4) can be understood as resulting from a comparison of several different situations that result in the same overall displacement of the step (the principle of detailed balancing). The density of thermally excited kinks in S_B steps turns out to be much higher than in S_A steps because the kink creation energy is much lower in S_B steps. For independent kink excitations, the probability of finding two kinks separated by r units must be $(1-n_0)n_0^{r-1}$. Swartzentruber *et al.* (1990) have shown that this assumption of independent kink excitations is indeed correct for Si(001). The step structure of Si(001) is static at room temperature; this means that the step roughness is frozen in at some temperature T_f that lies above room temperature. High-temperature STM measurements of thermally fluctuating steps reveal that the freeze-in temperature lies somewhere between 750 and 875 K (Zandvliet, van Loenen, and Elswijk, 1992; Kitamura *et al.*, 1993; Pearson *et al.*, 1995; Swartzentruber and Schact, 1995). Using the equations of Burton, Cabrera, and Frank, the following values for $\varepsilon_{\perp(\parallel)}$ and δ have been found:

$$\varepsilon_\perp=(3.6\pm0.2)k_bT_f, \tag{5a}$$

$$\varepsilon_\parallel=(5.7\pm0.3)k_bT_f, \tag{5b}$$

$$\delta=-(1.0\pm0.3)k_bT_f. \tag{5c}$$

The step-edge formation energy for S_A (S_B) step is

$$E_{Sa}=\frac{\varepsilon_\perp}{2}+\delta=(0.8\pm0.4)k_bT_f, \tag{6a}$$

$$E_{Sb}=\frac{\varepsilon_\parallel}{2}+\delta=(1.85\pm0.4)k_bT_f. \tag{6b}$$

The kink formation energy for a kink with a length of k units in an S_A (S_B) step is

$$E_{kink,Sa}=k\frac{\varepsilon_\parallel}{2}+(k-1)\delta=kE_{Sb}-\delta, \quad k\geqslant1, \tag{7a}$$

$$E_{kink,Sb}=k\frac{\varepsilon_\perp}{2}+(k-1)\delta=kE_{Sa}-\delta, \quad k\geqslant1. \tag{7b}$$

Double-layer step edges and their kink formation energies can be determined in a similar manner. One encounters, however, a fundamental problem, namely, that double-layer-stepped Si(001) exhibits only type-B double-layer (D_B) steps. Hence only the step-edge formation energy of the energetically unfavorable type-A double-layer (D_A) step edge can be determined. By counting the kinks in the D_B steps on Si(001) Zandvliet *et al.* (1996) have found a D_A step-edge formation energy of 4.5 k_bT_f. In order to determine the D_B step-edge formation energy, one must count the frequency with which a D_B step edge splits into a pair of single-layer steps separated by $2a$. The energy difference between a rebonded D_B step and two single-layer steps is $E_{Sa}+E_{Sb}+E_{strain}-E_{Db}$.[4] Hence the relative probability that this splitting will occur in a pair of single-layer steps is $\exp-(E_{Sa}+E_{Sb}+E_{strain}-E_{Db})/k_bT_f$. For double-layer-stepped Si(001), the strain term E_{strain} can be neglected. Using the above derived values for $E_{Sa,Sb}$ and the experimentally determined relative probability of step splitting (0.31), Zandvliet *et al.* (1996) found a value of $E_{Db}=1.5\ k_bT_f$. Eaglesham *et al.* (1993) measured the equilibrium shape of voids in Si formed by MeV He implantation and annealing. From this equilibrium shape they were able to extract the surface free energy $\gamma(\theta,T_f)$. The edge formation energy for the D_B step obtained from $d\gamma/d\theta$ was determined to be 92 meV/2a. This value is in good agreement with that extracted from STM experiments, assuming a reasonable freeze-in temperature of 775 K.

Before considering the step-step interactions, it is interesting to compare the experimentally obtained step-edge formation energies with available theoretical data. The first calculations of these energies for Si(001) were performed by Chadi (1987), who used semiempirical tight-binding total energy calculations and found the following energies: 0.02 eV/2a (S_A), 0.10 eV/2a (D_B), 0.30 eV/2a (S_B) to 1.08 eV/2a (D_A). Later Poon *et al.* (1990, 1992) used atomistic calculations with the Stillinger-Weber interatomic potential and found significantly different step-edge formation energies. The various theoretically calculated and experimentally determined energies are summarized in Tables I and II.

In addition to the step-edge and kink creation energies, step interaction energies can be extracted from a statistical analysis of STM images. Data from large numbers of images like Fig. 2 can be compiled to measure the strength as well as the form of the interactions between steps. When the steps are sufficiently far apart, they do not "feel" each other. However, as they come

[4]For the sake of simplicity, the force-dipole interaction between two single-layer steps has been omitted.

TABLE II. Measured step-edge formation energies of the Si(001) surface ($a=3.84$ Å).

	S_A (eV/2a)	S_B (eV/2a)	D_A (eV/2a)	D_B (eV/2a)
Swartzentruber *et al.* (1990)[a]	0.056	0.18		
Eaglesham *et al.* (1993)				0.092
Pearson *et al.* (1995)	0.064			
Bartelt and Tromp (1996)	0.056	0.14		
Laracuente and Whitman (2000)[b]	0.060	0.12		
Zandvliet, Elswijk, van Loenen, and Dijkkamp (1992)[b,c]	0.052	0.12	0.30	0.10

[a] Freeze-in temperature of 875 K.
[b] Freeze-in temperature of 775 K.
[c] Zandvliet, van Dijken, and Poelsema (1996).

closer, they begin to repel one another. Basically there are two types of step interactions. The first has a purely entropic origin and emerges because steps are not allowed to cross (Gruber and Mullins, 1967; Fisher and Fisher, 1982; Bartelt *et al.*, 1990, 1992a, 1992b). Each step-step collision decreases the number of configurations available per step by a factor of 2 and thus reduces the entropy by an amount $k_b \ln 2$. The increase in free energy per step-step collision is therefore $k_b T \ln 2$. The Gruber-Mullins (1967) model of a meandering step trapped between two hard walls, is well suited to the Si(001) surface, where the S_B and S_A steps correspond to the meandering step and hard wall, respectively (see Fig. 4). An estimate of the average spacing between successive step-step collisions can easily be obtained from this model by introducing the mean-square length $\langle k^2 \rangle$ of a kink in an S_B step (Zandvliet and Elswijk, 1993),

$$\langle k^2 \rangle = \frac{\sum_{k=-\infty}^{\infty} k^2 e^{-E_{kink,Sb}(k)/k_b T}}{\sum_{k=-\infty}^{\infty} e^{-E_{kink,Sb}(k)/k_b T}}. \tag{8}$$

The roughness of the steps is now most easily characterized by the correlation function $\langle (h_0 - h_r)^2 \rangle$, where r is the distance measured parallel to the step edge and h_r is the deviation measured perpendicular to the step edge with respect to a fixed, but arbitrarily chosen, reference line that runs parallel to the step edge. For an isolated step (or for steps spaced far apart), the mean-square displacement diverges linearly with r. This result follows immediately from the random distribution of kinks in a step at equilibrium (i.e., $\langle (h_0 - h_r)^2 \rangle = \langle k^2 \rangle r$). In Fig. 5, the mean-square displacement for the two types of single-layer steps extracted for a 0.5° misoriented Si(001) surface is shown. The slopes in Fig. 5 are consistent with experimental values of the interaction energy between the dimers, as listed in Eqs. (5a)–(5c). Up to values of approximately 20 dimer-dimer row spacings, the behavior is almost linear; however, at larger distances a deviation from this linearity arises due to the presence of step interactions (Zandvliet, Wormeester, *et al.*, 1993; Zandvliet, Louwsma, *et al.*, 1995). If L is the average spacing between neighboring steps, the average spacing between successive collisions is $L^2/\langle k^2 \rangle$. The increase in free energy per unit step-edge length due to entropic step repulsion is thus $\langle k^2 \rangle L^{-2} k_b T \ln 2$. If, for instance, $L = 20(2a)$ and $\langle k^2 \rangle_{Sb} = 1.7(2a)^2$, the spacing between successive step collisions is approximately 235 units or 1800 Å.

The second type of step interaction has an energetic origin and generally also decays as L^{-2}. From elasticity theory it can be shown that a crystal with degenerate phases and an anisotropic surface stress tensor can lower its energy with respect to a uniform single-domain surface by forming an ordered domain configuration

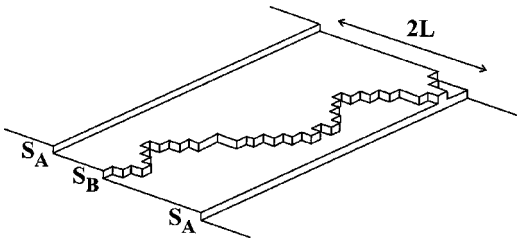

FIG. 4. Schematic drawing of a meandering step edge (S_B) trapped between two neighboring straight step edges (S_A).

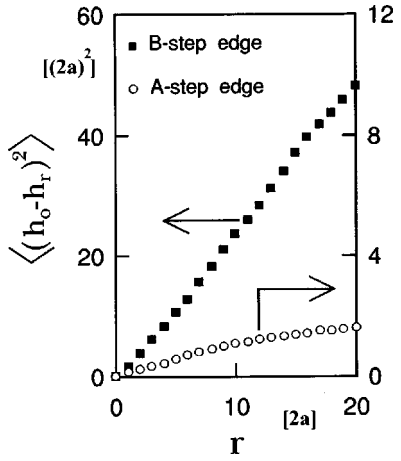

FIG. 5. Mean-square displacement, $\langle (h_0 - h_r)^2 \rangle$, measured in units of $(2a)^2$, of both types of single-layer steps vs the position r, measured in units of $2a$ (=7.7 Å) along the step. The slopes in this figure are directly related to the parameters $\varepsilon_\perp$, $\varepsilon_\parallel$, and δ.

(Marchenko, 1981; Alerhand *et al.*, 1988). This reduction in energy comes from a long-range elastic or strain relaxation, driven by the difference in the surface stress of both domains. In the particular case of Si(001), the surface stress is compressive perpendicular to the dimer bond ($\sigma_\perp$) and tensile along the dimer bond ($\sigma_\parallel$). The strain relaxation energy (per unit step length) for the single-layer stepped surface with terrace width L is (Alerhand *et al.*, 1988; 1990)

$$E_{strain}(L) = -\frac{(1-\nu)}{2\pi\mu}(\sigma_\parallel - \sigma_\perp)^2 \ln\left(\frac{L}{\pi a}\right) = -C\ln\left(\frac{L}{\pi a}\right), \tag{9}$$

where ν and μ are the Poisson ratio and bulk modulus of Si, respectively, and a is the microscopic cutoff length (i.e., the surface lattice constant). The potential $V(h,L)$ felt by a wandering step that is displaced h atomic positions out from the middle is then

$$V(h,L) = E_{strain}(L+h) + E_{strain}(L-h) \approx -2C\ln\left(\frac{L}{\pi a}\right) + C\left(\frac{h}{L}\right)^2. \tag{10}$$

The last term in Eq. (10), C/L^2, refers to the step-step interaction per lattice spacing. In general, the problem of interacting noncrossing steps can be mapped onto the problem of interacting, spinless fermions (Joós *et al.*, 1991). The step-distance distribution $P(h,L)$ is hence found simply by solving the 1D Schrödinger equation. Assuming no energetic step interactions, $P(h,L) = (1/L)\cos^2(\pi h/2L)$, whereas for the case of energetic step interactions behaving as $-C\ln L$ the step-distance distribution is given by

$$P(h,L) = \frac{1}{w\sqrt{2\pi}} e^{-(h^2/2w^2)}, \tag{11}$$

where w is the width of the Gaussian,

$$w(T,L) = \left(\frac{k_b T\langle k^2\rangle L^2}{8C}\right)^{1/4}. \tag{12}$$

The distribution of terrace lengths (h_i) determined from several STM images of adjacent areas is shown in Fig. 6. If the weak shoulders, located symmetrically around the main peak, are neglected, the data can be fitted with a Gaussian of width $w \approx 4-5$. Hence a surface stress anisotropy $\sigma_\parallel - \sigma_\perp$ of 0.9 ± 0.2 eV/a^2 is obtained (Zandvliet and Elswijk, 1993; Swartzentruber *et al.*, 1993). Besides the step interaction originating from the anisotropic surface stress tensor of Si(001), an additional direct step interaction decaying as L^{-2} (due to force-dipole interactions) is sometimes added (Swartzentruber *et al.*, 1993). However, for large terrace widths the force-monopole step interactions ($\ln L$) predominate over the direct step interaction (L^{-2}).[5]

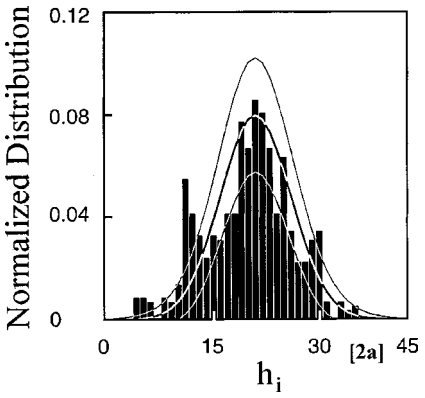

FIG. 6. Normalized terrace-width distribution $P(h,L)$ vs terrace width h in units of $2a$ ($=7.68$ Å).

V. MORPHOLOGICAL SURFACE PHASE TRANSITIONS

A. Thermal roughening of (001) facets of Si

Phase transitions occur because all systems in thermodynamic equilibrium seek to minimize their free energy $F = U - TS$. One phase will supplant another at a given temperature because different states partition their free energy between the U and TS terms in different ways. The idea that, in principle, there exists a roughening transition at a surface was put forward by Burton, Cabrera, and Frank (1951). Inspired by Onsager's (1944) solution of the two-dimensional Ising model, they conjectured that single-crystal surfaces in equilibrium with their vapor, melt, or solution, would become rough above a certain temperature. Macroscopically, the roughening transition of a surface is characterized by the disappearance of facets from the equilibrium shape. On the other hand, microscopically, this transition is defined by the free energy's becoming zero on a facet or, equivalently, by the appearance of strong fluctuations in the location of the facet. The formation energy of steps can be so high that the roughening temperature sometimes lies near or even above the melting point. On vicinal surfaces the thermal generation of kinks is responsible for roughening. Since thermal generation of a kink requires substantially less energy than the formation of a step on a facet, the roughening temperature of a vicinal surface is considerably lower than the roughening temperature of a facet.

[5]Adding the force-dipole interaction term $\lambda_d(a/L)^2$ results in a step interaction energy E_{step},

$$E_{step} = -C\ln\left(\frac{L}{\pi a}\right) + \lambda_d\left(\frac{a}{L}\right)^2.$$

A best fit of the terrace width distribution is obtained for $\lambda_d/C = 125$ (Swartzentruber *et al.*, 1993). Hence the force-dipole term plays a role only for very small terrace widths.

The most logical way to proceed is to model the Si surface by a compact structure of rigid elementary building blocks in a simple cubic arrangement. Because the Si(001) surface is dimerized even at temperatures as high as 1475 K (Tromp and Reuter, 1992; Bartelt *et al.*, 1994; Bartelt and Tromp, 1996) it seems most natural to take dimers as the elementary building blocks. However, since kink excitations in both monatomic step edges are mapped on a 2×2 lattice, it is more natural to take a unit of two surface dimers as an elementary building block. Because both the step-edge formation and the kink formation energies are known it is quite straightforward to determine the free energies of the A- and B-type step edges. As has been discussed above, one can easily determine these energies by simply measuring the Boltzmann factors for the various kink excitations and assuming a reasonable freeze-in temperature of step roughness. At first sight it might be expected that the step free energy of the two types of steps might vanish at different temperatures. Below it is shown that both step free energies vanish at precisely the same temperature T_r. The step free energy per unit length and per unit step height f is defined as

$$f=-k_bT\ln Z=-k_bT\ln\left(\sum_i e^{-E_i/k_bT}\right), \tag{13}$$

where the summation runs over all possible step configurations and E_i refers to the formation energy of the ith configuration.[6]

The partition function Z of an S_A (S_B) step edge is given by

$$Z_{Sa(Sb)}=e^{-E_{Sa(Sb)}/k_bT}\left(1+2\sum_{k=1}^{\infty}e^{-E_{kink,Sa(Sb)}(k)/k_bT}\right) \tag{14}$$

and hence

$$f_{Sa(Sb)}=\frac{\varepsilon_{\perp(\parallel)}}{2}+\delta-k_bT\ln\left(1+\frac{2e^{-(\varepsilon_{\parallel(\perp)}/2)/k_bT}}{1-e^{-(\varepsilon_{\parallel(\perp)}/2+\delta)/k_bT}}\right). \tag{15}$$

The roughening temperature can be found by solving

$$e^{-\varepsilon_{\parallel}/2k_bT_r}+e^{-\varepsilon_{\perp}/2k_bT_r}+e^{-(\varepsilon_{\parallel}+\varepsilon_{\perp})/2k_bT_r}(2-e^{-\delta/k_bT_r})+e^{\delta/k_bT_r}=0. \tag{16}$$

This expression is symmetric in $\varepsilon_{\perp(\parallel)}$ and therefore the step free energies of both types of steps vanish at the same temperature T_r (1430 K). At 1483 K low-energy electron microscopy (LEEM) images show that the steps become suddenly very rough and at slightly higher temperatures all contrast is lost (Tromp and Reuter, 1992). Near the roughening transition, entropic and energetic interactions between the steps also become important. These interactions are not taken into account here and it could be expected that the step free energies will approach zero more gradually than indicated by expression (15). As the entropic repulsion depends on the kink formation energy, it would be reasonable to expect that the degeneracy of the roughening temperature of both single-layer steps is lifted by the anisotropy in the entropic repulsion. However, the difference between S_A and S_B single-layer steps eventually disappears when one approaches the roughening temperature.

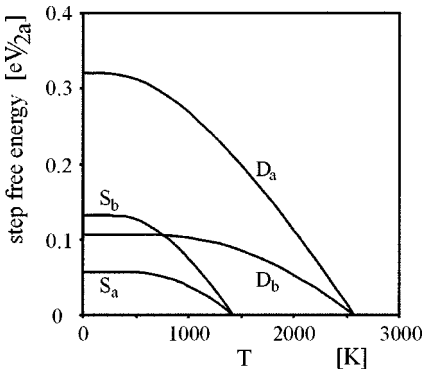

FIG. 7. Step free energies of single- and double-layer steps vs temperature T. Values for the single-layer steps: $\varepsilon_{\perp}=3.6\,k_bT_f$, $\varepsilon_{\parallel}=5.7\,k_bT_f$ and $\delta=-k_bT_f$. Values for the double-layer steps: $E_{Da}=4.5\,k_bT_f$ and $E_{Db}=1.5\,k_bT_f$ ($T_f=825$ K).

Because the step-edge formation and kink creation energies of both A- and B-type biatomic steps are known, similar expressions for the free energies of biatomic steps can also be derived (Zandvliet *et al.*, 1996). However, the step-edge free energy of the double steps[7] vanishes at a temperature of 2570 K, which greatly exceeds the melting temperature of 1683 K for Si.

A plot of the single- and double-layer step free energies versus temperature is presented in Fig. 7. It should be noted that Mètois and Wolf (1993; Pimpinelli and Mètois, 1994) found that sublimation holes appear in the terraces of Si(001) prior to roughening [indicating that the roughening of Si(001) has a kinetic rather than a thermodynamic origin]. However, the increasing step fluctuations observed by Bartelt *et al.* (1994) below 1375 K, where sublimation is very small, suggest that roughening has a thermodynamic origin. Although the dimerized Si(001) surface is clearly anisotropic, it should be realized that even unreconstructed Si(001) is anisotropic. The easiest way to demonstrate this is to recall

[6]For $\varepsilon_{\perp(\parallel)}\gg k_bT$ and $\delta=0$, simple expressions for the partition sum and step free energy are found,

$$Z_{Sa(Sb)}=e^{-\varepsilon_{\perp(\parallel)}/2k_bT}\left(1+2\sum_{k=1}^{\infty}e^{-k\varepsilon_{\parallel(\perp)}/2k_bT}\right)\approx e^{-\varepsilon_{\perp(\parallel)/2k_bT}}(1+\langle k^2\rangle_{Sa(Sb)}),$$

$$f_{Sa(Sb)}=-k_bT\ln Z\approx\frac{\varepsilon_{\perp(\parallel)}}{2}-k_bT\ln(1+\langle k^2\rangle_{Sa(Sb)})\approx\frac{\varepsilon_{\perp(\parallel)}}{2}-k_bT\langle k^2\rangle_{Sa(Sb)}.$$

[7]Here the breaking up of the double-layer steps into a pair of single-layer steps has not been taken into account. Including this excitation will lower the temperature at which the free energies of a double-layer step vanish.

that a diamond lattice basically consists of two interpenetrating fcc sublattices, which are rotated by 90° with respect to each other. A perfect (001) facet can only be terminated by one of the two fcc sublattices.

Thermal roughening of a vicinal surface is closely related to roughening of a low-index facet. On vicinal surfaces, where steps are already present, thermal formation of kinks is responsible for roughening. Since the creation of an adatom-vacancy pair is energetically much more expensive than the creation of a pair of kinks, it is obvious that the roughening temperature of a vicinal surface lies far below the roughening temperature of a facet. The creation of a kink pair depends not only on the kink creation energy but also on the strength of the (repulsive) interaction between the steps. If this step interaction is large, the extent of wandering of the steps will be limited and the roughening temperature will be relatively high. The terrace-ledge-kink model of Villain, Grempel, and Lapujoulade (1985) describes the roughening of vicinal surfaces. These authors demonstrated that the roughening temperature of a vicinal surface depends on two energetic parameters, namely, the kink creation energy and the step-step interaction. Using this theory it can be shown that vicinal Si(001) roughens at a temperature well below the freeze-in temperature T_f of the step roughness (Zandvliet, Wormeester, *et al.*, 1993; Hegeman *et al.*, 1995).

B. Orientational phase diagram of Si(001)

A vicinal crystal surface, i.e., one slightly misoriented with respect to a low-Miller-index facet, consists of terraces separated by steps that accommodate the misorientation. Vicinal surfaces can exhibit different structural phases, since steps of different types may be favored depending on temperature T or miscut angle θ. Consider now a Si(001) surface, slightly misoriented, by an angle θ, in the [110] direction. The surface misorientation can be accommodated by single-layer steps or double-layer steps, leading to surfaces that differ not only in the height of the step edges and the width of the terraces, but also in the lattice structure. The single-layer-stepped surface is a double domain exhibiting both 2×1 and 1×2 domains, while the double-layer-stepped surface is a single domain. Experimentally it has been observed that for θ larger than $4-5°$, the surface contains predominantly double-layer steps (Wierenga *et al.*, 1987; Aumann *et al.*, 1988), whereas for smaller miscut angles the surface is single-layer stepped. Also, for small miscut angles, Si(001) exhibits an interesting sequence of phase transitions. For miscut angles smaller than approximately 0.03°, a hill and valley structure with step loops is observed (Tromp, 1993), whereas a phase of wavy steps coexisting with a phase of relatively straight steps is observed in the range $0.03-0.1°$ (Tromp and Reuter, 1992). If the miscut angle lies in the range from 0.1° up to $1.5-2°$, two types, one much straighter than the other, of single-layer step edges are found (Dijkkamp *et al.*, 1990; Swartzentruber *et al.*, 1990). The most prominent transition is without any doubt the transition from a single-layer-stepped surface to a double-layer-stepped surface.

To understand this transition, let us compare the step-edge free energies of a pair of single-layer steps ($f_{Sa}+f_{Sb}$) with the double-layer step free energy (f_{Db}). As outlined above, there are two types of single-layer steps: S_A and S_B. Both types are necessary to accommodate the misorientation. On a double-layer-stepped surface, all terraces have the same orientation, and only one type of double-layer step is required. The D_B step is always found in STM experiments, as expected, since the D_A step has a much higher formation energy. Although the sum of the step-edge formation energies of pair of single-layer steps (172 meV/$2a$) is significantly higher than the formation energy of a D_B step (100 meV/$2a$), a single-layer-stepped surface is found experimentally for small misorientations. Because the entropy of the steps has not been taken into account here, this comparison is only appropriate at $T=0$ K. It is natural to assume that the entropy of a pair of single-layer steps is much larger than the entropy of a single wandering double-layer step (Dey *et al.*, 1996).[8]

With increasing temperature, the difference between the step free energy of a single-layer-stepped surface and a double-layer-stepped surface, i.e., $f_{SA}+f_{SA}-f_{DB}$, will decrease. However, this is not the only contribution that will favor single-layer steps for small misorientations. From elasticity theory it can be shown that a crystal with degenerate phases and an anisotropic surface stress tensor can lower its energy with respect to a uniform single-domain surface by forming an ordered domain configuration. This reduction in energy comes from long-range elastic or strain relaxation that is driven by the difference in surface stress between the two domains. In the particular case of Si(001), the surface stress is compressive perpendicular to the dimer bond ($\sigma_\perp$) and tensile along the dimer bond ($\sigma_\parallel$). The surface stress anisotropy is experimentally found to be $\sigma_\parallel-\sigma_\perp=0.9\pm0.2\,\mathrm{eV}/a^2$. The strain relaxation energy (per unit area) for a single-layer-stepped surface with terrace width L is (Alerhand *et al.*, 1988; 1990)

$$E'_{strain}(L)=-\frac{(1-\nu)}{2\pi\mu L}(\sigma_\parallel-\sigma_\perp)^2\ln\left(\frac{L}{\pi a_0}\right)=-\frac{C}{L}\ln\left(\frac{L}{\pi a_0}\right). \tag{17}$$

The double-layer-stepped surface is a single-domain structure in which all the terraces have the same orientation, and no strain relaxation occurs. Hence the free-energy difference between a single-layer- and double-layer-stepped surface with misorientation is

$$\Delta f=f_{Sa}+f_{Sb}-2C\ln\left(\frac{L}{\pi a_0}\right)-f_{Db}. \tag{18}$$

[8]The step-doubling transition found on $W(430)$ is an exception to the rule that two single wandering single-layer steps have more entropy than a single double-layer step.

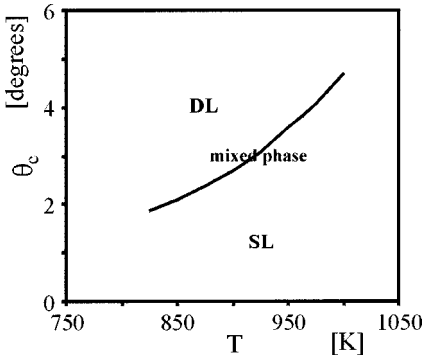

FIG. 8. Critical angle θ_c vs temperature T (T_f=825 K). SL and DL refer to single-layer- and double-layer-stepped Si(001), respectively. The mixed phase refers to a coexistence region where both SL and DL steps are found.

For sufficiently large terrace width L, the strain relaxation term stabilizes the single-layer-stepped surface. The condition $\Delta f=0$ defines a first-order transition at

$$\theta_c=\arctan\left(\frac{d}{\pi a_0}e^{(f_{Db}-f_{Sa}-f_{Sb}/2C)}\right), \tag{19}$$

where d is the single-layer step height (=1.36 Å). Above the temperature at which the steps are frozen in (here we have assumed T_f=825 K), the critical miscut angle θ_c depends on the temperature. A plot of calculated θ_c versus temperature ($T>T_f$) is shown in Fig. 8. The model predicts that for annealed Si(001) the transition between the single-layer- and double-layer-stepped surface takes place at $\sim 2°$. STM experiments have revealed that Si(001) is single-layer stepped for a miscut angle below 1° and double-layer stepped above 4°. In the range 1–4° single- and double-layer steps have been found experimentally to coexist.

Bartelt *et al.* (1991) pointed out that in principle there should be a range of θ over which the surface facets into regions with a low density of single-layer steps and small θ and regions with a high density of double-layer steps and large θ. By plotting the free energy per projected area in the (001) plane for single-layer and double-layer phases of the model versus $\tan(\theta)$ (i.e., step density), one immediately sees that the free-energy function is concave near the point of intersection of the two curves. Thermodynamic stability requires that the free-energy function be convex as a function of $\tan(\theta)$. Construction of the appropriate tie lines (''Gibbs's construction'') shows that the surface is unstable with respect to the breakup into a phase with a low density of single-layer steps and a phase with a higher density of double-layer steps. However, Alerhand *et al.* (1991) argued that such a coexistence of phases would require faceting of the surface, and hence substantial mass transport, which might not be kinetically allowed. Somewhat later Pehlke and Tersoff (1991a) showed that there is neither an abrupt transition from single-layer to double-layer steps with a change in angle nor a coexistence between spatial regions of single- and double-layer steps. Instead, as the angle increases past a critical value, pairs of single-layer steps collapse into double-layer steps in a complex pattern, so that at zero temperature the surface undergoes a cascade of transitions resembling a ''devil's staircase'' (Bak and Bruinsma, 1982; Pehlke and Tersoff, 1991a). Finally, Pehlke and Tersoff (1991b) calculated the full temperature-angle phase diagram of Si(001) and pointed out that there is a critical point in the phase diagram above which there is no phase transition at all. In their model they included the breakup of D_B steps into pairs of single-layer steps. It is this excitation which blurs the distinction between single- and double-layer steps at high temperature, leading to a critical point in the phase diagram of Si(001). However, because of uncertainty in the values of the parameters that enter their model, they could not rule out the possibility that the critical temperature might be above the freeze-in temperature of the step structure.

VI. CONCLUSIONS

The thermodynamic behavior of surfaces is governed by the surface free energy. The surface free energy depends on four parameters, namely, the step free energy, the kink creation energy, the interaction energy between steps, and the free energy of the facet. In this review it has been shown that scanning tunneling microscopy images of a surface can provide, in principle, three of these four key parameters. The Si(001) surface is used as a model system. In order to illustrate the significance of these step-related energetic parameters, two morphological surface transitions of Si(001) have been discussed. The method proposed here is, in principle, generally applicable to other surfaces as well. However, the applicability is more satisfactory for semiconductor surfaces, for which step and kink energies can be associated with covalent bonds, than for metal surfaces.

ACKNOWLEDGMENTS

I am grateful to N. C. Bartelt, D. Dijkkamp, H. B. Elswijk, B. Poelsema, B. S. Swartzentruber, E. J. van Loenen, and A. van Silfhout, for many helpful discussions.

REFERENCES

Alerhand, O. L., D. Vanderbilt, R. D. Meade, and J. D. Joannopoulos, 1988, Phys. Rev. Lett. **61**, 1973.
Alerhand, O. L., A. Nihat Berker, J. D. Joannopoulos, D. Vanderbilt, R. J. Hamers, and J. E. Demuth, 1990, Phys. Rev. Lett. **64**, 2406.
Alerhand, O. L., A. Nihat Berker, J. D. Joannopoulos, D. Vanderbilt, R. J. Hamers, and J. E. Demuth, 1991, Phys. Rev. Lett. **66**, 962.
Aumann, C. E., D. E. Savage, R. Kariotis, and M. G. Lagally, 1988, J. Vac. Sci. Technol. A **6**, 1963.
Bak, P., and R. Bruinsma, 1982, Phys. Rev. Lett. **49**, 249.
Bartelt, N. C., T. L. Einstein, and E. D. Williams, 1990, Surf. Sci. **240**, L591.
Bartelt, N. C., T. L. Einstein, and C. Rottman, 1991, Phys. Rev. Lett. **66**, 961.

Bartelt, N. C., T. L. Einstein, and E. D. Williams, 1992a, Surf. Sci. **276**, 308.
Bartelt, N. C., T. L. Einstein, and E. D. Williams, 1992b, Surf. Sci. **273**, 252.
Bartelt, N. C., and R. M. Tromp, 1996, Phys. Rev. B **54**, 11731.
Bartelt, N. C., R. M. Tromp, and E. D. Williams, 1994, Phys. Rev. Lett. **73**, 1656.
Binnig, G., and H. Rohrer, 1982, Helv. Phys. Acta **55**, 726.
Binnig, G., H. Rohrer, H. Gerber, and E. Weibel, 1982a, Appl. Phys. Lett. **40**, 178.
Binnig, G., H. Rohrer, H. Gerber, and E. Weibel, 1982b, Phys. Rev. Lett. **49**, 57.
Bowler, D. R., and M. G. Bowler, 1998, Phys. Rev. B **57**, 15385.
Burton, W. K., N. Cabrera, and F. C. Frank, 1951, Philos. Trans. R. Soc. London, Ser. A **243**, 299.
Chadi, D. J., 1987, Phys. Rev. Lett. **59**, 1691.
Dey, S., S. Kiriukhin, J. West, and E. Conrad, 1996, Phys. Rev. Lett. **77**, 530.
Dijkkamp, D., A. J. Hoeven, E. J. van Loenen, J. M. Lenssinck, and J. Dieleman, 1990, Appl. Phys. Lett. **56**, 39.
Eaglesham, D. J., A. E. White, L. C. Feldman, N. Moriya, and D. C. Jacobson, 1993, Phys. Rev. Lett. **70**, 1643; 1994, **72**, 1392(E); 1994, **72**, 2975(E).
Fisher, M. E., and D. S. Fisher, 1982, Phys. Rev. B **25**, 3192.
Gibbs, J. W., 1957, *The Collected Works of J. Willard Gibbs* (Yale University, New Haven, CT), Vol. 1.
Griffith, J. E., and G. P. Kochanski, 1990, Crit. Rev. Solid State Mater. Sci. **16**, 255.
Gruber, E. E., and W. W. Mullins, 1967, J. Phys. Chem. Solids **28**, 875.
Hamers, R. J., R. M. Tromp, and J. E. Demuth, 1986, Phys. Rev. B **34**, 5343.
Hegeman, P. E., H. J. W. Zandvliet, G. A. M. Kip, B. A. G. Kersten, and B. Poelsema, 1995, Surf. Sci. **331-333**, 1110.
Herring, C., 1951, Phys. Rev. **82**, 87.
Jeong, H.-C., and E. D. Williams, 1999, Surf. Sci. Rep. **34**, 171–294.
Joós, B., T. L. Einstein, and N. C. Bartelt, 1991, Phys. Rev. B **43**, 8153.
Kitamura, N., B. S. Swartzentruber, M. G. Lagally, and M. B. Webb, 1993, Phys. Rev. B **48**, 7504.
Laracuente, A., and L. J. Whitman, 2000, unpublished.
Liu, F., and M. G. Lagally, 1997, in *The Chemical Physics of Solid Surfaces*, edited by D. A. King and D. P. Woodruff (Elsevier, Amsterdam), Vol. 8, p. 258.
Marchenko, V. I., 1981, Pis'ma Zh. Eksp. Teor. Fiz. **33**, 397 [JETP Lett. **33**, 381 (1981)].
Markov, I. V., 1995, *Crystal Growth for Beginners* (World Scientific, Singapore), p. 222.
Mètois, J. J., and D. E. Wolf, 1993, Surf. Sci. **298**, 71.
Müller, E. W., 1951, Z. Phys. **131**, 136.
Onsager, L., 1944, Phys. Rev. **65**, 117.
Oshiyama, A., 1995, Phys. Rev. Lett. **74**, 130.
Pearson, C., B. Borovsky, M. Krueger, R. Curtis, and E. Ganz, 1995, Phys. Rev. Lett. **74**, 2710.
Pehlke, E., and J. Tersoff, 1991a, Phys. Rev. Lett. **67**, 465.
Pehlke, E., and J. Tersoff, 1991b, Phys. Rev. Lett. **67**, 1290.
Pimpinelli, A., and J. J. Mètois, 1994, Phys. Rev. Lett. **72**, 3566.
Poon, T. Z., S. Yip, P. S. Ho, and F. F. Abraham, 1990, Phys. Rev. Lett. **65**, 2161.
Poon, T. Z., S. Yip, P. S. Ho, and F. F. Abraham, 1992, Phys. Rev. B **45**, 3521.
Schlier, R. E., and H. E. Farnsworth, 1959, J. Chem. Phys. **30**, 917.
Schlüter, M., 1988, in *The Chemical Physics of Solid Surfaces and Heterogeneous Catalysis*, edited by D. A. King and D. P. Woodruff (Elsevier, Amsterdam), Vol. 8, p. 37.
Swartzentruber, B. S., Y.-W. Mo, R. Kariotis, M. G. Lagally, and M. B. Webb, 1990, Phys. Rev. Lett. **65**, 1913.
Swartzentruber, B. S., N. Kitamura, M. G. Lagally, and M. B. Webb, 1993, Phys. Rev. B **47**, 13432.
Swartzentruber, B. S., and M. Schact, 1995, Surf. Sci. **322**, 83.
Tromp, R. M., 1993, Phys. Rev. B **47**, 7598.
Tromp, R. M., R. J. Hamers, and J. E. Demuth, 1985, Phys. Rev. Lett. **55**, 1303.
Tromp, R. M., and M. C. Reuter, 1992, Phys. Rev. Lett. **68**, 820.
Tsong, T. T., 1988, Rep. Prog. Phys. **51**, 759.
Villain, J., D. R. Grempel, and J. Lapujoulade, 1985, J. Phys. F **15**, 809.
Wierenga, P. E., J. A. Kubby, and J. E. Griffith, 1987, Phys. Rev. Lett. **59**, 2169.
Williams, E. D., 1994, Surf. Sci. **299/300**, 502.
Williams, E. D., and N. C. Bartelt, 1991, Science **251**, 393.
Zandvliet, H. J. W., and H. B. Elswijk, 1993, Phys. Rev. B **48**, 14269.
Zandvliet, H. J. W., H. B. Elswijk, E. J. van Loenen, and D. Dijkkamp, 1992, Phys. Rev. B **45**, 5965.
Zandvliet, H. J. W., H. K. Louwsma, P. E. Hegeman, and B. Poelsema, 1995, Phys. Rev. Lett. **75**, 3890.
Zandvliet, H. J. W., S. van Dijken, and B. Poelsema, 1996, Phys. Rev. B **53**, 15429.
Zandvliet, H. J. W., E. J. van Loenen, and H. B. Elswijk, 1992, Surf. Sci. **272**, 264.
Zandvliet, H. J. W., H. Wormeester, D. J. Wentink, A. van Silfhout, and H. B. Elswijk, 1993, Phys. Rev. Lett. **70**, 2122.