Encyclopedia of Applied Physics

Surfaces and Interfaces of Solids, Structure of

Feng Liu, M. Hohage, and M. G. Lagally

University of Wisconsin-Madison, Madison, WI 53706

This encyclopedia article gives an overview of basic concepts and fundamental principles underlying the structure of solid surfaces and interfaces. A brief discussion of surface thermodynamics is provided in the context of the Gibbs model, and the relationship between surface stress and surface tension for a solid surface is established. Basic definitions and notations of surface crystallography are introduced for the description of structures of single-crystal surfaces. Surface relaxation and surface reconstruction are discussed in detail and illustrated with examples of semiconductor and metal surfaces. Underlying physical mechanisms relating the atomic structure to the electronic structure are summarized. A qualitative description of the morphology of real surfaces, interfaces, and thin films is provided.

Surfaces and Interfaces of Solids, Structure of

Feng Liu, M.Hohage and M.G.Lagally, University of Wisconsin-Madison, Madison, Wisconsin, U.S.A

INTRODUCTION	1
1. SURFACE THERMODYNAMICS	2
 1.1 GIBBS SURFACE MODEL AND SURFACE TENSION 1.2 SURFACE STRESS AND ITS RELATION TO SURFACE TENSION 	
2. SURFACE CRYSTALLOGRAPHY	5
2.1 Orientation of crystalline surface	6
3. SURFACE RELAXATION AND SURFACE RECONSTRUCTION	10
 3.1. A GENERAL DISCUSSION 3.2 SEMICONDUCTOR SURFACES. 3.2.1 The {100} surfaces 3.2.1.1. Dimer buckling in the {100} surfaces of elemental semiconductors 3.2.1.2. The {100} surfaces of binary compound semiconductors and electron counting 3.2.2 The {111} surfaces 3.2.2.1. The (2 × 1) reconstruction on cleaved Si(111) and Ge(111) and on clean C(111). 3.2.2.2. The Si(111)-(7 × 7) and Ge(111)-c(2 × 8) reconstructions 3.2.3. The {111} surfaces of compound semiconductors 3.2.3. The {110} surfaces 3.3. The {110} surfaces 3.3.1 The fcc (110) surface 3.3.2 The fcc (100)-surface 3.3.3 fcc (111) surface. 	$ \begin{array}{c} 11 \\ 12 \\ 13 \\ 14 \\ 16 \\ 16 \\ 18 \\ 20 \\ 20 \\ 20 \\ 21 \\ 23 \\ 25 \\ 26 \end{array} $
4. MOPHOLOGICAL FEATURES OF REAL SURFACES, INTERFACES, AND THIN FILMS	
 4.1. THE TERRACE-STEP-KINK (TSK) MODEL	
GLOSSARY	
LIST OF WORKS CITED	
FURTHER READING	
FIGURE CAPTIONS	40

Introduction

Among the properties of surfaces and interfaces, structure is fundamental, as it determines or influences chemical, electrical, mechanical, and even magnetic and optical behavior. In describing structure, it is essential that a scale length be defined. For example, a surface may feel smooth or rough or it may scatter light specularly or diffusely, which are clearly macroscopic or mesoscopic measures of surface structure. On the other extreme, the atomic structure of the surface describes where the atoms are located relative to their positions in the bulk of the crystal. In the past half century, remarkable progress has been made in determining the atomic structure of the surface, driven initially by the development of ultra-high vacuum, which permitted the preparation of clean surfaces and maintenance of the cleanliness for sufficient time to make detailed structural measurements. The subsequent invention of numerous structural characterization methods and the development of supercomputers and advanced computational algorithms have led to a broad understanding of surface atomic structure.

"Surface structure" is not just "where are the atoms". There are also larger-scale features for whose description the detailed positions of the atoms need not be known. These features are typically affected or determined by extended structural entities such as steps, terraces, dislocations, grain boundaries, and voids. The term "surface morphology" has come into use to describe such larger-range features. Their influence on surface properties can be significantly greater than that of the atomic structure. The analogous term for the bulk of materials is "microstructure". Looking down on a surface to observe its form and topology in the way one looks at the ground from an airplane illustrates why "morphology" is appropriate.

This article discusses surface atomic structure and surface morphology under the umbrella of "surface structure". It aims to provide an overview of basic concepts and fundamental principles related to structural properties of solid surfaces and interfaces and to illustrate systematic trends in structures of surfaces and interfaces in different classes of materials. Section 1 reviews surface thermodynamics, as a basis and a driving force for formation of surface structure. Section 2 introduces concepts and notations required to describe the structures of single-crystal surfaces. Section 3 overviews intrinsic structures of single-crystal surfaces, and discusses surface relaxation and reconstruction, with examples chosen from semiconductor and metal surfaces. Section 4 deals with the morphology of real surfaces and interfaces, and the development of morphology through kinetic limitations thermodynamic driving or forces encountered during thin-film growth or surface treatment.

1. Surface Thermodynamics

1.1 Gibbs surface model and surface tension

J. W. Gibbs (1928) laid the groundwork for surface thermodynamics.

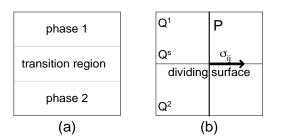


Fig 1. The Gibbs surface model, (a) the real system: (b) the idealized system. The thin horizontal line in the middle marks the dividing surface: the thick vertical line marks the imaginary plane P used to derive the surface stress tensor σ_{ijij} .

He introduced a notion of a dividing surface and derived all surface thermodynamic properties by an excess procedure. A real system consisting of two homogeneous phases separated by a finite inhomogeneous transition region (Fig. 1a) is idealized as two homogeneous phases separated by a hypothetical geometrical surface, the dividing surface (Fig. 1b). The dividing surface is chosen to be near the macroscopic observable physical interface and everywhere parallel to it; different choices of dividing surface may be used that are parallel to each other but displaced from each other in the direction of surface normal. The excess procedure expresses the thermodynamic properties of the real system in terms of the thermodynamic properties of the ideal system and excess properties, which define the difference between the real and ideal values. Thus, all extensive surface thermodynamic functions can be expressed in terms of the dividing surface through the excess For the two-phase system that is procedure. represented by two homogeneous phases, phase 1 and

2, and a dividing surface, as shown in Fig. 1(b), any extensive property can be divided unambiguously into contributions from phase 1, phase 2, and the surface as

$$Q = Q^{1} + Q^{2} + Q^{s} . (1)$$

Q is the extensive property of the whole system, which is the integration of the specific value of Q over the whole system including the inhomogeneous transition region; Q¹ and Q² are respectively the same extensive property for phase 1 and 2 that would apply as if they are homogeneous up to the (infinitely narrow) dividing surface; Q^s is the excess extensive property, i.e., the difference between Q and Q¹ + Q², assigned to the surface. One intriguing problem with the excess extensive property is that it is usually not unique depending on the choice (or location) of the dividing surface (see discussion below).

One defines surface tension, γ , as the reversible work required to create unit area of surface (by cleavage), at constant temperature (T) and pressure (p), this function is the partial derivative of Gibbs free energy (G) of the whole system with respect to area of surface formed (A), at constant T, p, and mole concentration of each component (n_i),

$$\gamma = \left(\frac{\partial G}{\partial A}\right)_{T,p,n_i}.$$
(2)

As $G = G(T,p,n_i,A)$, the differential of Gibbs free energy can be written as

$$dG = -SdT + Vdp + \sum \mu_i dn_i + \gamma dA, \quad (3)$$

where S is the entropy, V the volume, and μ_i the chemical potential of molecular species i.

Alternatively, γ can be defined as the partial derivative of Helmholtz free energy (F) of the whole system with respect to A, at constant T, V, and n_i,

$$\gamma = \left(\frac{\partial F}{\partial A}\right)_{T,V,n_i} \tag{4}$$

and the differential of Helmholtz free energy, F = G - pV, is

$$dF = -SdT - pdV + \sum \mu_i dn_i + \gamma dA.$$
 (5)

(Similarly, γ can be defined in terms of internal energy, E = F + TS, and grand canonical potential, $\Omega = F - \sum \mu_i n_i$.)

From eqns. (1) and (3) and applying the first and second law of thermodynamics to the two homogeneous phases 1 and 2, the differential of surface Gibbs free energy is derived as

$$dG^{s} = \sum \mu_{i} dn_{i}^{s} + \gamma dA.$$
 (6)

Here n_i^s is the excess mole concentration of the *i*th species assigned to the surface. According to Euler's theorem, the surface Gibbs free energy is

$$G^{s} = \sum \mu_{i} n_{i}^{s} + \gamma A, \qquad (7)$$

and similarly, the surface Helmholtz free energy is

$$F^{s} = pV^{s} + \sum \mu_{i}n_{i}^{s} + \gamma A \approx \sum \mu_{i}n_{i}^{s} + \gamma A, \text{ for } V^{s} \rightarrow 0.$$
(8)

Eqns. (7) and (8) give another definition of surface tension,

$$\gamma = g^{s} - \sum \mu_{i} \Gamma_{i}^{s} \approx f^{s} - \sum \mu_{i} \Gamma_{i}^{s}, \qquad (9)$$

where $g^{s}(f^{s})$ is the Gibbs (Helmholtz) free energy per unit area of surface, i.e., the specific surface Gibbs (Helmholtz) free energy, and Γ_{i}^{s} is the surface density n_{i}^{s}/A of the i-th species.

In general, surface tension, γ , which measures the change in Gibbs or Helmholtz free energy for the whole system in creating the surface, has two contributions: the change in Gibbs or Helmholtz free energy per unit area for the surface "phase", g^s or f^s, and the change per unit area of surface formed for the surrounding bulk phases. Note that g^s , f^s , and Γ_i^s are all excess surface properties that are dependent on the position of the dividing surface, but γ , a quantity associated with the whole system, is not, because the position of the dividing surface affects the Gibbs (or Helmholtz) free energy in such a way that the change in the surface term, g^s or (f^s), is always balanced by the change in the bulk term, $\sum \mu_i \Gamma_i^s$. For a one-component system, it is possible to choose the dividing surface so that n_1^{s} , and hence Γ_1^{s} vanishes; the specific surface Gibbs or Helmholtz free energy then becomes equal to surface tension. Consider a one-component system with solid and gas phases; the atomic density assumes a constant value of n_{bulk} in the solid phase and a value of zero in the gas phase. In the interface (surface) region, the atomic density decreases gradually from the bulk value to zero going from the bulk phase toward the gas phase. If the dividing surface is located inside the transition region, in the vicinity of the surface, there will be a thin region of "extra" atomic density (> 0) in the gas phase and a thin region of "lost" atomic density ($< n_{bulk}$) in the solid phase. By choosing the position of the dividing surface so that the extra density in the gas phase equals exactly the lost density in the solid phase, there will be no excess surface density, i.e., n_1^{s} and hence Γ_1^{s} vanishes. For a multi-component system, however, one choice of the dividing surface that makes n_i^s vanish will not make other $n_{j \star i}^{s}$ vanish. Therefore, in general, surface tension, γ , should not be confused with the specific Gibbs (or Helmholtz) surface free energy, g^{s} (or f^{s}).

Surface tension is a very difficult quantity to measure experimentally, but it can in principle be calculated with a reliable energy functional, such as those based on first-principles (or *ab initio*) computational techniques. In an empirical way, surface tension (surface energy) of a solid surface has two contributions: the formation energy and the relaxation energy. The former reflects the breaking of bonds to make a solid surface (at the ideal bulk terminations); the latter reflects the tendency of a solid surface to distort because it is a quasi-twodimensional (2D) system and hence would like to assume an atomic structure and bonding configuration different from that of the bulk. To create a surface, it costs energy to break bonds, and the cost is partly recovered by the relaxation process involving rearrangement of atoms and bonds at the surface. The formation energy (positive) dominates the relaxation energy (negative). The energy of a single-crystal surface scales approximately with the cohesive energy of the bulk crystal. The greater the bulk cohesion, the stronger the interatomic bonds, the higher the surface energy. The weakly van-der-Waals bonded rare gas-solids have the lowest surface energy; the strongly bonded semiconductors and metals have high surface energy. In contrast to that of an isotropic liquid surface, the energy of a solid surface is anisotropic, depending on the orientation of the surface because a different number of bonds is broken to create different surface orientations.

1.2 Surface stress and its relation to surface tension

Imagine that there exists a plane P normal to the dividing surface (Fig. 1b). One may

derive the surface force acting on P by the same excess procedure discussed above. The total force acting across P, from the material on one side of P to the material on the other side, can be divided into contributions from the two homogeneous bulk phases and from the surface phase. The surface force per unit length of the line intersection of the plane P with the dividing surface plane is then defined as the surface stress.

Gibbs (1928) first pointed out the distinction, for the case of a solid, between surface tension and surface stress: the former measures the energy cost to create unit area of new surface; the latter measures the energy cost to deform the surface. The relationship between surface stress and surface tension can be derived as following:

Suppose we deliberately deform the surface, at constant temperature and total amount of surface species, introducing an infinitesimal strain, ε_{ij} , to the surface. The work required to do so is

$$\delta W = A \Sigma \sigma_{ii} \varepsilon_{ii}, (i, j = 1, 2). \tag{10}$$

Here σ_{ij} denotes the surface stress tensor. This work is equal to the infinitesimal change of surface Gibbs or Helmholtz free energy, G^s or F^s , at constant T and n_i^s ,

$$\delta G^{s} = \delta F^{s} = \delta(\gamma A) = \gamma \delta A + A \delta \gamma$$

$$= \gamma A \Sigma \varepsilon_{ii} + A \Sigma \frac{\partial \gamma}{\partial \varepsilon_{ii}} \varepsilon_{ij}, (i, j = 1, 2)$$
(11)

Combining eqns. (10) and (11), we have

$$\sigma_{ij} = \gamma \delta_{ij} + \frac{\partial \gamma}{\partial \varepsilon_{ii}}, (i, j = 1, 2).$$
 (12)

where the Kronecker delta $\delta_{ij} = 1$ if i = j and $\delta_{ij} = 0$ if $i \neq j$. J. Surface stress, σ_{ij} , is independent of the location of dividing surface, although it is an excess surface property, as indicated by its relation to surface tension, γ , in eqn. (12).

Equation (12) also implies that the surface stress tensor (σ_{ij}) is defined as the derivative of surface Gibbs or Helmholtz free energy (G^s or F^s) with respect to surface strain (ϵ_{ij}), at constant temperature and total amount of surface species,

$$\sigma_{ij} = \frac{1}{A} \left(\frac{dG^s}{d\varepsilon_{ij}} \right)_{T,n_i^s} = \frac{1}{A} \left(\frac{dF^s}{d\varepsilon_{ij}} \right)_{T,n_i^s}, (i, j = 1, 2)$$
(13)

Equivalently, surface stress may be defined in terms of other thermodynamic energy functions.

For a liquid, surface stress is equal to surface tension because liquid is a purely plastic medium and its surface tension is independent of strain, i.e., the second strain-derivative term on the right-hand side of eqn. (12) vanishes. (For a onecomponent liquid, surface tension also equals the specific Gibbs or Helmholtz surface free energy. So, for this special case, the values of all three quantities, surface stress, surface tension, and specific surface free energy, are the same, which has often caused confusion.) When a liquid surface is stretched, atoms rapidly enter the surface from the bulk; the existing surface atoms stay where they are and the new surface is formed of atoms from the bulk now occupying surface positions. Thus, expanding a liquid surface is equivalent to creating new liquid surface. In contrast, solids are a purely elastic medium, at least at low deformations. When a solid surface is stretched, surface atoms are displaced from their minimum-energy positions, sitting at new sites for which the surface as a whole has a different energy. Thus, the surface tension of a solid surface depends on surface strain. While a liquid surface is always under tensile stress for γ to be positive, a solid surface can be under an overall stress that is either tensile or compressive because the second term in eqn. (12), the strain-derivative of tension, can be either positive or negative and may have a magnitude larger than the first term, the tension contribution. Also, the strain-derivative term often makes the surface stress of a solid surface anisotropic, while the surface stress of a liquid surface is always isotropic.

2. Surface Crystallography

2.1 Orientation of crystalline surface

Given a single crystal of a particular Bravais lattice, a lattice plane is defined to be any plane containing at least three noncollinear Bravais lattice points; it is denoted by Miller indices, which are the coordinates of the shortest reciprocal-lattice vector normal to the plane. A common way of creating a crystalline surface is to cut a bulk single crystal parallel to one of its low-index atomic planes (e.g., $\{100\}$, $\{110\}$, and $\{111\}$ planes in a cubic single crystal). An ideal surface (one that has all atoms in bulk positions) is made by removing all the atoms lying on one side of a chosen atomic plane and keeping the positions of all remaining atoms intact. The orientation of this surface is then well defined. with its surface normal specified by the directional indices, [hkl], defined in the conventional way with respect to the three-dimensional (3D) unit cell of the bulk crystal (Ashcroft and Mermin, 1976). The surface is denoted as an (hkl) surface, using the Miller indices of the corresponding parallel set of atomic planes in the bulk crystal. The notation $\{hkl\}$ denotes a set of equivalent planes or surfaces.

Low-Miller-index surfaces are widely used in surface science research and epitaxial growth of thin films because of their relatively greater stability and high symmetry. A real surface generally consists of a number of terraces of low-index surface planes of the same orientation separated by steps. In *nominal* {hkl} surfaces the normal to the average surface orientation coincides with the normal to the *singular* {hkl}terraces that constitute the surface. To achieve this condition, there must be on the average an equal number of "up" steps and "down" steps in a nominal {hkl} surface (see Fig. 2a).

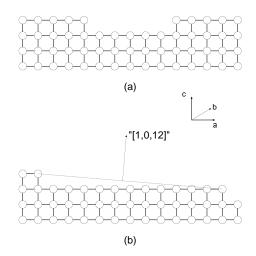


Fig 2. Side views ([010] projection of (001) surfaces of a simple cubic lattice. (a) A nominal surface consisting of three singular (001) terraces separated by one up and one down single-atomic-height step. (b) A vicinal surface miscut by $\approx 4.8^{\circ}$ toward [100]. The normal (arrow) to the average surface orientation (dashed line) is [1 0 12].

In practice, it is very difficult to cut a crystal exactly parallel to a chosen low-index plane. Even if there is no average miscut, the uncertainties in cutting and polishing lead to a wavy surface, giving only a nominally low-index surface. A *vicinal* surface is made by a deliberate cut up to a few degrees away from a low-index plane toward a specific direction. For example, Figure 2b shows a vicinal surface of a simple cubic lattice formed by a miscut ~ 4.8° away from [001] toward [100]. A vicinal surface may be specified either by the normal to its average surface orientation, which generally produces unreasonably large Miller indices [e.g., the surface in Fig. 2b would be labeled as $(1 \ 0 \ 12)$], or by the indices of the constituent terraces plus the miscut angle and miscut orientation (e.g., Si(001) miscut 2⁰ toward [110]). The latter choice makes it convenient to understand and discuss the properties of a small-miscut vicinal surface in terms of its low-index singular terraces and the effects of defects (steps). If the miscut angle is not too small and the index of the normal to the average surface orientation does not contain numbers that are too large the surface is named with its appropriate Miller indices.

2.2 Surface periodicity and symmetry

The structure of a single-crystal solid is uniquely defined by its intrinsic symmetry (Ashcroft and Mermin, 1976), which includes translational symmetry (periodicity), symmetry of the Bravais lattice (primitive unit cell), and symmetry of the unit cell basis (atoms or molecules) in the unit cell. The translational symmetry in 3D is realized by 7 different crystal systems: cubic, tetragonal, orthorombic, monoclinic, triclinic, trigonal, and hexagonal. Allowing more than one grid point per unit cell leads to 14 Bravais lattices, including types of primitive (p), body-centered (bc), face-centered (fc), and basecentered unit cells. Each lattice point may have more than one atom associated with it, reducing the symmetry of the original crystal system. There are 32 crystallographic point groups that a crystal structure can have, covering symmetry operations of rotation, reflection, and inversion, and combinations of them. The full symmetry of the crystal (including atomic arrangement) is described completely by 230 crystallographic space groups, with the addition of symmetry operations of screw axes and glide planes.

As a crystalline solid is truncated to make a surface, the translational symmetry of the solid perpendicular to the surface is removed, while the periodicity parallel to the surface remains. Because in the surface region each atom layer becomes intrinsically in-inequivalent to other layers, all the symmetry properties of a solid surface are twodimensional. Consequently, surface structures, in analogy to 3D crystal structures, are classified by the 2D Bravais lattice (net) and symmetry groups of the net and of the atomic basis associated with the net. There are 5 different 2D Bravais nets: oblique, primitive (p) rectangle, centered (c) rectangle, square, and hexagonal, as depicted in Fig. 3, 10 2D point groups, and 17 possible 2D space groups (Woodruff, 1981).

The 2D Bravais net consists of a grid of equivalent points; the atomic basis represents the arrangements of atoms at each net point.

The translational symmetry of a real surface usually differs from that of an ideal bulk termination of the solid, i.e., that of the corresponding atomic plane in the bulk, because of surface reconstruction (see discussion below).

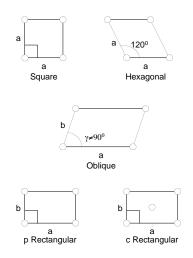


Fig 3. The five two-dimentional Bravais nets: oblique, square, hexagonal, and primitive (p) or centered (c) rectangle.

If (a_s, b_s) and (a_i, b_i) are the unit net vectors (the primitive translational vectors) of the real surface and

of the ideal surface, respectively, the general translational vectors are

$$\boldsymbol{T}_{\mathbf{s}} = n \, \boldsymbol{a}_{\mathbf{s}} + m \, \boldsymbol{b}_{\mathbf{s}} \tag{14}$$

in the real surface and

$$\boldsymbol{T}_{\mathbf{i}} = n \, \boldsymbol{a}_{\mathbf{i}} + m \, \boldsymbol{b}_{\mathbf{i}} \tag{15}$$

in the ideal surface. n and m are integers. The relationship between the real surface net and ideal surface net is uniquely defined by a (2×2) matrix, G (Park and Madden, 1968),

$$\begin{pmatrix} \boldsymbol{a}_s \\ \boldsymbol{b}_s \end{pmatrix} = \boldsymbol{G} \begin{pmatrix} \boldsymbol{a}_i \\ \boldsymbol{b}_i \end{pmatrix} \,. \tag{16}$$

The determinant of G, |G|, equals the ratio of the areas of the real unit net and the ideal unit net. If |G| is an integer, the nets are *simply* related; if |G| is a rational number, the two nets are *rationally* related. When |G|is irrational, the real surface has an incommensurate structure relative to the ideal surface (substrate).

Although the matrix notation is exact, it is seldom used. Instead, a notation introduced by

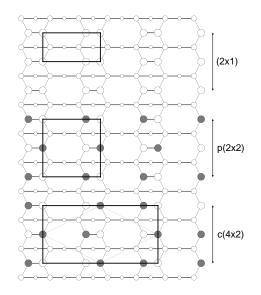


Fig 4. Schematic top view of three possible reconstructions on Si(001). Top region: (2×1) ; middle: $p(2 \times 2)$: bottom: $c(4 \times 2)$. Open

and solid circles mark atom positions: their size indicates different layers of atoms with the largest circles in the outermost layer. In the middle and bottom regions, the slightly different heights of the two atoms in a dimmer due to buckling are indicated by solid and open circles. The unit mesh of each reconstruction is depicted by a dark rectangle. For the c(4x2)structure, the primitive unit mesh is depicted by a dashed-line rhombus.

Wood (1964) has been widely adopted and is convenient for small commensurate unit meshes. This simple notation uses the ratio of the lengths of the real primitive net vectors to those of ideal primitive vectors plus an angle through which the real net is rotated with respect to the ideal net. Figure 4 illustrates the use of this notation to describe several possible reconstructions on Si(001). A detailed discussion of these reconstructions will be presented later.

2.3 Reciprocal-space representation and diffraction

Conventionally, surface structures are detected through their representations in reciprocal space using diffraction-based surface probes. An important consequence of periodicity is the modification of the law of momentum conservation. As a wave (e.g., electrons or photons) interacts with a periodic structure, its momentum is conserved plus or minus any reciprocal-lattice vector (lattice momentum). For a surface, the periodicity only exists parallel to the surface, so the modified law of momentum conservation applies only to the components of the wave vectors parallel to the surface,

$$\boldsymbol{k}_{\parallel} = \boldsymbol{k}_{0\parallel} + \boldsymbol{g}_{\mathrm{hk}}.$$
 (17)

 $\mathbf{k}_{0\parallel}$ and \mathbf{k}_{\parallel} are parallel to surface components of incident and diffracted-wave vectors. \mathbf{g}_{hk} is a surface reciprocalnet vector (i.e., a general translational vector in *k*-space), which is expressed as

$$\boldsymbol{g}_{\rm hk} = h\boldsymbol{a}^{*} + k\boldsymbol{b}^{*}, \qquad (18)$$

where *h* and *k* are integers. a^* and b^* are primitive unit vectors of the reciprocal net defined relative to the real-space net as

$$a^* = (2 \pi b_s \times n)/A, \quad b^* = (2 \pi a_s \times n)/A, \quad A = a_s \cdot b_s \times n,$$
(19)

where n is the unit vector normal to the surface.

The conditions for diffraction are totally defined by the momentum conservation expressed in eqn. (17) together with an energy conservation equation,

$$k^{2} = k_{0}^{2}$$
 or $k_{\parallel}^{2} + k_{\perp}^{2} = k_{0\parallel}^{2} + k_{0\perp}^{2}$. (20)

These conditions can be conveniently visualized using the Ewald sphere construction in k-space, as illustrated in Fig. 5.

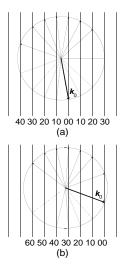


Fig 5. Ewald sphere consctruction for a two-dimensional system, e.g. a single layer. The incident beam is denoted by the vector k_0 and all other radial vectors ending on the Ewald sphere correspond to diffracted beams, k. (a) Near-normal incidence for LEED: (b) low-angle incidence for RHEED. Dashed lines indicate those diffracted beams penetrating into the bulk. For RHEED, k_0 is generally much larger, relative to the spacing of the reciprocal-lattice rods, than is shown in (b): for LEED, the relative magnitudes shown in (a) are appropriate.

One first draws a vector of the incident wave, \mathbf{k}_0 , to the origin of the reciprocal space, and then draws a sphere of radius k_0 about the tail of the vector \mathbf{k}_0 . In three dimensions, this (Ewald) sphere passes through certain reciprocal- lattice points in addition to the origin and the diffracted beam \mathbf{k} is simply given by a vector from the tail of \mathbf{k}_0 to those reciprocal-lattice points intersecting the Ewald sphere. A surface can be approximated by a 2D net.

It is straightforward to show that for a 2D net, the complete reciprocal-space consists of a set of lines or 'rods' that are infinite in extent, are perpendicular to the 2D net and pass through the reciprocal-net points. Because those rods inside the Ewald sphere will always intersect the sphere, the conditions for the appearance of diffracted beams are greatly relaxed, relative to those for a 3D crystal, for which the reciprocal lattice consists of points. Therefore, a small change in energy of the incident wave, i.e., in the length of k_0 and the radius of Ewald sphere, will totally extinguish one set of diffracted beams in three dimensions. The same change for a surface will only shift slightly the direction of the diffracted beams.

Electron beams and x-rays are typically used in diffraction-based surface probes. Figure 5 illustrates two common setups for electron diffraction: low-energy electron diffraction (LEED), generally performed at near-normal incidence, and reflection high-energy electron diffraction (RHEED), at glancing incidence. Most forms of radiation used for diffraction (electrons, x-rays) have a mean free path in solids sufficiently large to penetrate some distance into the crystal. For example, for electrons this mean free path varies from approximately 5 Angstroms up to thousands of Angstroms, as the energy is changed. The actual penetration into the crystal can be limited by changing the angle of incidence (compare LEED and RHEED). The finite penetration of radiation into the crystal implies that a 2D net and perfect reciprocal-lattice rods to describe scattering from a surface is only a limiting approximation, i.e., there is some 3D character to the reciprocal lattice. Thermal He and other noble-gas atoms, on the other hand, are scattered above the surface without penetrating the crystal because the energies for the wavelengths appropriate for crystal diffraction are much lower. For He scattering, a 2D net and pure reciprocal-lattice rods are a good approximation; consequently, of course, only the structure of the outer surface layer can be detected. Helium scattering is also very useful in studying "unstable" surface structures such as adsorbate layers because of the very low energy of the probe. In addition to determining the periodicity (i.e., long-range order) of the surface structure, which all diffraction probes used in the paper generally do well, diffraction is in principle able to determine the positions of the atoms in the structure. This sensitivity to the atomic basis associated with the lattice occurs through the intensities of the diffracted beams. For electron diffraction, determination of atomic positions is complicated by the strong interaction of electrons wit atoms, causing multiple scattering. The use, for example, of x rays (for which the interaction is much weaker) at grazing incidence avoids this problem.

As real-space surface structures are derived indirectly from their representations in reciprocal space

with diffraction techniques, controversies often arise when different structural models are fit to the diffraction data, especially when a large atomic basis is involved or when the relaxation from the ideal (bulk-termination) surface is small. The advance of a new generation of surface probes using tunneling and force microscopies has allowed the direct observation of surfaces in real space, often with atomic resolution. In many ways, the "old" diffraction probes and the new scanned probes compensate each other: the former, working in reciprocal space, are sensitive to the long-range order and the latter, working in real space, give local order with atomic detail. It is frequently much easier to construct a correct model to fit the diffraction data using scanned-probe real-space images as input. In modern laboratories, diffraction and real-space techniques are often combined to monitor and determine surface structures in situ and in real time.

3. Surface Relaxation and Surface Reconstruction

3.1. A general discussion

Real surfaces do not retain the ideal bulk termination. Atoms of surface and near-surface layers are generally displaced from their ideal bulk positions, residing at new positions of minimum energy in response to a change of atomic and electronic environments in the surface regions. The atomic rearrangements can take place a few layers deep into the bulk, but the most dramatic structural changes usually occur at only the true surface layer. The structure of this outer layer dominates most surface properties. Surface *relaxation* involves only atomic displacements that do not change the translational symmetry of the surface (i.e., the periodicity of the surface), e.g., collective equal displacements of all atoms in the surface layer; surface *reconstruction* describes atomic displacements that change the periodicity of surface. Surface relaxation should not be confused with atomic relaxations at defects (vacancies, adatoms, steps, etc.), at which the periodicity is already disturbed. A change in surface periodicity can also be caused by a change of atom number density in the surface layers. Because atoms in the surface layer are undercoordinated, missing all neighbors on the vacuum side of the surface, atoms may be added or deleted from the top layer, in order to minimize the surface free energy, as has been discussed in section 1.

The driving force for surface relaxation and reconstruction is minimization of surface (free) energy. Surfaces with high surface energy have a stronger tendency to relax and/or reconstruct than surfaces with low surface energy. For a given solid, higher-energy (generally higher-index) orientations of a surface usually relax and/or reconstruct more than lower-energy (lower-index) orientations (see, e.g., the discussion in section 3.3 of fcc metal (110), (100), and (111) surfaces). It is, however, the reduction of surface energy rather than the surface energy itself that controls the tendency for a surface to relax and/or reconstruct. The nature of interatomic interactions in different classes of solids is the key factor controlling their respective surface structural properties. The surfaces of rare-gas solids, in which atoms are bonded with the weak van der Waals interaction, are most inert, exhibiting only small relaxations perpendicular to the surface; surfaces of metals that have isotropic metallic bonds most often do not reconstruct but display large relaxations; semiconductor surfaces with broken highly directional covalent bonds have the strongest tendency to reconstruct.

The lowering of surface energy can be achieved by reducing surface chemical energy and/or surface strain energy; the two often compete. The reduction of chemical energy, which dominates at the atomic level by optimizing bonding and/or electron density, often occurs at the expense of strain energy, which is caused by bond distortion (see, e.g., the discussion of reconstruction on Si(001) in section 3.2). The reduction of strain energy becomes dominant in determining morphologies at the mesoscopic scale because elastic interactions extend to a much longer range than do chemical interactions.

The fundamental mechanisms for reducing surface chemical energy through surface relaxation/reconstruction depend strongly on the nature of chemical bonding in the underlying solid. In semiconductors, atoms bond with each other by sharing electron pairs in covalent bonds. The major effect of creating a surface is to break covalent bonds, introducing a large number of dangling bonds in the surface. Consequently, there exists a strong driving force for reconstructing semiconductor surfaces to remove dangling bonds on surface atoms. In metals, atoms bond with each other through the effective medium of an electron gas, and the major effect of creating a surface is the depletion of electron density in the vicinity of the surface. Surface atoms tend to rearrange themselves to re-optimize their surrounding electron density. Such atomic rearrangement in both semiconductor and metal surfaces involves inevitably change of interatomic spacings and distortion of bond angles, so the gain in chemical energy is usually achieved at the expense of strain energy.

The surface strain energy can be reduced by morphological changes at the mesoscopic scale, such as creation and modulation of steps and dislocations (see section 4.2). Such structural and morphological changes give rise to a variation of surface stress over the surface, which generates a distribution of elastic force density. The strain relaxation energy equals the integral of force density times displacement over the whole surface area. Thus, the strain relaxation process is always accompanied by a redistribution of the surface stress field.

A large portion of research in surface science, especially in earlier years, focused on determination of the intrinsic surface structure, in particular, surface relaxation and surface reconstruction of clean crystalline solid surfaces. In the following we discuss some general features of relaxation and reconstruction in semiconductor and metal surfaces and present a few typical examples.

3.2 Semiconductor surfaces

Semiconductors are generally made of atoms with s and p valence orbitals each containing on average two electrons (the half-filled s and p shells of average atomic configuration yields the filled valence and empty conduction band of band the semiconducting bulk). In those semiconductors that crystallize in diamond (elemental semiconductors) and zincblende (binary compound semiconductors) structures, the s and p orbitals at each atomic site hybridize into an sp³ configuration, which is characterized by four equivalent hybrid orbitals pointing to the corners of a tetrahedron with the atom at the center. Therefore, each atom is tetrahedrally coordinated, forming four covalent bonds with its four nearest neighbors, and each bond shares an electron pair, amounting effectively to a total of eight electrons, a very stable closed s and p shell, on each atom. In a group-IV elemental semiconductor, each atom contributes one electron per covalent bond. In a binary compound semiconductor, each atom contributes $f = n_v/4$ electrons per covalent bond, where n_v is the number of valence electrons in a single atom; $f_c + f_a = 2$, where subscripts *c* and *a* denotes cations and anions.

In creating a surface, some covalent bonds must be broken, leaving behind dangling bonds, hybrid orbitals with a single electron. Because of the high energy cost of dangling bonds, the surface energy of semiconductor surfaces is usually high, and there exists a strong tendency for semiconductor surfaces to relax and/or to reconstruct to reduce the number of dangling bonds or to reconfigure the dangling bonds into a lower-energy state.

3.2.1 The {100} surfaces

The {100}-oriented surfaces of elemental and compound semiconductors have been the most extensively studied surfaces because of their technological importance in real applications. In particular, the Si(100) surface serves as the foundation of modern semiconductor technology. In an ideal {100} surface of a diamond or zincblende structure, each atom has two dangling bonds, caused by the loss of two of its nearest neighbors (see Fig. 6a). The primary feature of a reconstructed {100} surface is dimerization, the rebonding of two neighboring surface atoms to form a dimer to eliminate one dangling bond per atom (Fig. 6b). The rebonding, which bonds two second-nearest-neighbor atoms into a distance almost equal to the firstnearest-neighbor separation, requires a rather large distortion of bond angles. Thus, the decrease in electronic energy by rebonding is partly offset by an increase in strain energy. The minimization of strain energy drives the dimers to order into parallel rows, leading to a (2×1) reconstruction (see top region of Fig. 4). The dimerization on Si(001) and Ge(001) was first proposed by Schlier and Farnsworth (1959) as an explanation of their (2×1) LEED patterns observed at room temperature. The dimer rows have been 'seen' directly by scanning tunneling microscopy (STM) on both Si(001) (Tromp et al., 1985, Hamers et al., 1986a) and Ge(001) (Kubby et al., 1987), as well as on (001) surfaces of compound semiconductors, such as GaAs(001) (Biegelsen et al. 1990a, 1990b).

Using a typical dimer bond length of 2.5 Å,

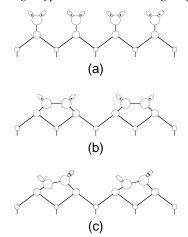


Fig 6. Dangling-bond configurations on (100) surfaces of elemental semiconductors: (a) two dangling bonds (hybrids with a single electron) on each surface atom in an ideal bulk-terminated surface; (b) one dangling bond on each atom in a dimmer. Each dangling bond is half-filled, leading to a metallic surface layer. (c) Charge transfer from downward atom to upward atom in a tilted (buckled) dimmer. The dangling bond on the lower atom is empty and the one on the higher atom is filled, leading to a semiconducting surface layer.

it is estimated that the gain in electronic energy is about 4eV per dimer based on a tight-binding model (Harrison, 1981), and the cost in strain energy is about 2 eV per dimer based on the Keating model (Keating, 1966); so, overall, the dimerization, or more precisely the (2×1) reconstruction, lowers the surface energy by ~ 2 eV/dimer. This estimate is roughly the same for all the elemental and binary compound semiconductor {100} surfaces, and it is in very good agreement with ab initio calculations (Ramstad et al., 1995) for Si(001).

3.2.1.1. Dimer buckling in the {100} surfaces of elemental semiconductors

The energies of two dangling bonds are degenerate in a dimer reconstruction if the dimers consist of two identical atoms in identical positions. Such dimers are called symmetric or "unbuckled" dimers. In the symmetric configuration, the two degenerate dangling bonds are both half-filled, resulting in a metallic surface. Because such energy degeneracy is not a spin-degeneracy, the symmetric structure is unstable against Jahn-Teller distortion, which will lower the free energy by lowering the symmetry of the dimer, i.e., by removing the degeneracy of the dangling bonds in the dimer and changing the surface from metallic to semiconducting. An asymmetric dimer can be most easily achieved by a simple tilt, or buckling, as shown in Fig. 6c.

The buckling dimers induces of rehybridization of surface bonds: the atom moving downward changes its back bonds more to an sp^2 -like bonding (about equally bonded to its three nearest neighbors within a plane) and hence its dangling bond to more p-like; the atoms moving upward changes its back bonds to more p-like (bond angles are reduced toward 90°) and its dangling bonds to more s-like. The two reconfigured dangling-bond states result in a charge transfer from the downward atom to the upward atom and a lowering of electronic energy. In other words, the buckling of a dimer splits two degenerate dangling-bond energy levels. The dangling-bond bonding level, of s character, lies below the maximum of the bulk valence band and will then be occupied; the dangling-bond anti-bonding level, of p character, lies above the maximum of the valence band and will be а metallic surface empty, making layer semiconducting. Chadi (1979), using tight-binding calculations, first showed that the symmetric dimers are unstable against buckling on Si(001). Ab initio

calculations (for a list of these calculations, see, e.g., Ramstad et al., 1995) later confirmed this conclusion. The energy gain from buckling is ~ 0.2 eV/dimer (Ramstad et al., 1995). Similar results are obtained for Ge(001) (Needels et al., 1987).

The buckling, like the dimerization, decreases electronic energy at the expense of increasing strain The minimization of strain energy causes energy. periodic arrays of buckling patterns that give rise to $c(4 \times 1)^{-1}$ 2) and $p(2 \times 2)$ reconstructions, which have been observed at low temperatures on Si(001) and Ge(001) by LEED (Kevan and Stoffel, 1984; Tabata et al., 1987) and by x-ray diffraction (Lucas et al., 1993). Lowtemperature STM (Wolkow, 1992) of Si(001) shows that surface dimer buckling begins to 'lock in' below ~ 200 K, the number of buckled dimers that are locked into the buckled position increases with decreasing temperature, and the ordering of buckled dimers forms local $c(4 \times 2)$ and $p(2 \times 2)$ domains, with the $c(4 \times 2)$ domains dominating.

Although buckled (tilted) dimers are more stable than unbuckled (untilted) dimers, there seems to be a rather low energy (kinetic) barrier for a buckled dimer to switch from one buckling configuration (e.g., with the left-side atom up) to another (with the left-side atom down). At sufficiently high temperature, surface dimers can switch their orientations so rapidly and independently of each other that the buckling configurations are in a dynamic disorder, leading to an averaged symmetric appearance in STM and to a (2×1) diffraction pattern. LEED (Kevan and Stoffel, 1984; Tabata et al., 1987) and x-ray diffraction (Lucas et al., 1993) have revealed a structural phase transition from the (2×1) to the $c(4 \times 2)$ phase as the temperature is decreased, corresponding to the freezing in of rocking dimers, in the temperature range between 150 K to 250 K. Theoretical calculations (Ihm et al., 1983; Saxena et al., 1985; Zubkus et al., 1991) predict a second-order phase transition from an ordered $c(4 \times 2)$ or $p(2 \times 2)$ structure to a disordered (2 × 1) structure at approximately 200 K to 250 K, in good agreement with experiments.

3.2.1.2. The {100} surfaces of binary compound semiconductors and electron counting

Covalent bonds in a binary compound semiconductor are partially ionic because the electronegativities are smaller for the metal atoms (cations) than for the nonmetal atoms (anions). In a zincblende structure, bulk {100} planes are alternately occupied by cations and anions, so an ideal {001} surface will be completely terminated by either cations or anions, leading to a polar surface. Polar surfaces, especially the {100}-oriented ones, have been widely used in homo- and hetero-epitaxial growth of compound semiconductors because the sticking coefficient of anion molecules depends strongly on the cation surface concentration. For example, the sticking coefficient of As₂ molecules on a GaAs(001) surface is unity if the surface is Gaterminated but zero if it is As-terminated (Arthur, 1974; Foxon and Joyce, 1977). During MBE growth, the surface composition can be controlled from anion- to cation-rich by adjusting growth conditions, e.g., the substrate temperature, the growth rate, and the ratio of the respective fluxes. All {100} surfaces of III-V and II-VI compound semiconductors display reconstructions depending on surface composition.

The GaAs(001) surface has attracted much attention because of its potential importance in optoelectronic devices. As the As-rich GaAs(001) surface is heated from room temperature to above 450^{0} C, the following sequence of reconstructions is

observed (Cho, 1971; Drathen et al. 1978; Massies et al., 1980):

$$c(2 \times 8) \xrightarrow{=475^{\circ}C} 1 \times 6 \rightarrow c(6 \times 4)$$

$$\xrightarrow{=575^{\circ}C} 3 \times 1 \rightarrow 4 \times 1$$

$$\xrightarrow{=650^{\circ}C} c(8 \times 2)$$

(21)

The sequence is correlated with the continuous decrease of the As concentration in the surface due to thermal desorption of As_2 molecules without evaporation of Ga (Arthur, 1974). A reverse sequence is observed (van Bommel et al., 1978) when cooling down the Ga-rich surface:

$$c(8 \times 2) \xrightarrow{\sim 450^{\circ}C} c(6 \times 4) \xrightarrow{\sim 350^{\circ}C} c(2 \times 8)$$
(22)

which is attributed to surface segregation of As from inside the bulk (Arthur, 1974; Neave and Joyce, 1978). In the As-rich surfaces, three different forms of (2×4) and/or $c(2 \times 8)$ reconstructions may occur depending on growth conditions, and they are denoted as (2×4) - α , - β , and - γ structures (Farrell and Palmstrøm, 1990). If an excess amount of As is present on the surface beyond one monolayer (ML) coverage, the surface displays a $c(4 \times 4)$ reconstruction (Neave and Joyce, 1978) and eventually becomes (1×1) when a thin film of As of a few monolayers thick is formed. All {100} surfaces of III-V and most of II-VI compound semiconductors except ZnSe behave like GaAs(001) (Cornelissen et al., 1988).

STM images of As-rich GaAs(001) surfaces (Biegelsen et al. 1990a, 1990b) reveal that the basic features of the surface reconstructions are dimers and dimer vacancies (missing dimers). At different coverages, different reconstructions arise from the specific density, the arrangement, and the orientation of dimers and dimer vacancies. The three forms of (2) \times 4)/c(2 \times 8) reconstruction on the As-rich surface correspond to a coverage of 0.5, 0.75, and 1 ML of As respectively; the c(4 \times 4) reconstruction corresponds to a As coverage of 1.75 ML.

The formation of dimers and dimer vacancies, and hence the reconstructions, can be understood by an electron counting model (Chadi, 1987a; Pashley, 1989). As for elemental semiconductors, the dimerization on {100} surfaces of compound semiconductors reduces the number of dangling bonds. The energy of dangling-bond states can be reduced by charge transfer between two dangling bonds to fill one dangling bond with a lone electron pair while emptying the other one. On elemental {100} surfaces, this energy redistribution is achieved by intra-dimer charge transfer through dimer buckling, while on compound {100} surfaces, it is achieved by charge transfer between the secondlayer atoms and surface dimers through the formation of dimer vacancy rows. Because danglingbond energies are lower on anions than on cations, charge transfer from dangling bonds on cations to those on anions is greatly favored over the reverse. On a cation-terminated surface, all the dangling bonds on surface cation dimers can be emptied by transferring their electrons to the second-layer anions that are exposed at the dimer-vacancy sites; on an anion-terminated surface, all the dangling bonds on surface anion dimers can be filled by obtaining extra electrons from the second-layer cations that are exposed at the dimer-vacancy sites. Consequently, all compound-semiconductor {100} surfaces exhibit a general $(2 \times N)$ reconstruction; the 2-fold periodicity results from the dimerization and the N-fold periodicity results from the dimer vacancy row formation. The periodicity, N, defining the optimal density of dimer vacancies, is determined by an electron counting rule (Mönch, 1995).

Consider an anion-stabilized $(2 \times N)$ structure with M_a anion dimers and $N-M_a$ missing dimers per unit mesh. Each dimer will contain six electrons; two in the dimer bond itself and two each in the dangling bonds at each of the dimer atoms. These electrons are supplied by the surface anions as well as by the second-layer cations that are exposed to the surface at the vacancy sites. Because each cation and anion contributes respectively f_c and f_a electrons per bond, counting the total number of electrons in the unit mesh, we have

$$6 M_a = 4 M_a f_a + 4 (N - M_a) f_c$$
(23)

Since $f_c = 3/4$ and $f_a = 5/4$ for III-V semiconductors, and $f_c = 2/4$ and $f_a = 6/4$ for II-VI semiconductors, we obtain from eqn. (23)

$$M_a(III-V) = 3N/4; \ M_a(II-VI) = N.$$
 (24)

Because the value of M_{a} , i.e., the number of dimers in the unit cell, must be an integer, the relations in eqn. (24) predict that the smallest unit mesh on anionstabilized {100} surfaces is (2 × 4) for III-V systems and (2 × 1) for II-VI systems. Figure 7 shows a schematic diagram of the (2 × 4) reconstruction on As-terminated GaAs(001). Similarly, for a cationstabilized (N × 2) structure with M_c cation dimers, we have

$$2 M_c + 8 (N - M_c) = 4 M_c f_c + 4 (N - M_c) f_a,$$
(25)

and

$$M_c(III-V) = 3N/4; M_c(II-VI) = N.$$
 (26)

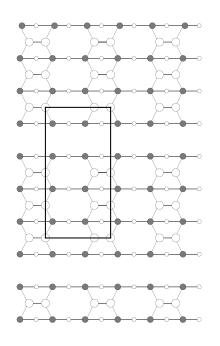


Fig. 7. Schematic top view of the (2x4)- β re-construction on As-stablized GaAs(001), induced by As vacancies. The large and small open circles denote As atoms in the first and third layers, respectively. The solid circles denote Ga atoms in the second layer. The solid rectangle marks the unit mesh.

The smallest unit mesh on cation-stabilized $\{100\}$ surfaces is (4×2) for III-V systems and (1×2) for II-VI systems. The acceptable values of N clearly depend on surface composition, which alters the number of electrons in the surface. So far, all the known reconstructions on the $\{100\}$ surfaces of binary compound semiconductors satisfy the electron counting rule (Mönch, 1995).

3.2.2 The {111} surfaces

The {111} surfaces of semiconductors, like the {100) surfaces, have also attracted much attention, the interest stemming from their complex structures, such as the well-known Si(111)-(7 \times 7) reconstruction. In an ideally bulk-terminated {111} surface of diamond or zincblende structures, surface atoms have three nearest neighbors in the second layer and one half-filled dangling bond perpendicular

the surface (see Fig. 8a). The elemental to semiconductors Si and Ge cleave along {111} planes and both as-cleaved surfaces exhibit a metastable (2×1) reconstruction. The ground-state structure of Si(111) is a (7×7) reconstruction and that of Ge(111) is a c(2 × 8) reconstruction. A simple rebonding (dimerization) mechanism like the one on Si(001) to reduce the number of dangling bonds is unlikely here because there is no easy direction for surface bonds to rotate. However, an adatom can effectively reduce the number of dangling bonds, two per adatom, by saturating three dangling bonds but only introducing one, a key ingredient on both the Si(111)- (7×7) and Ge(111)-c(2 × 8) reconstructions. For binary compound semiconductors, the {111} surfaces are fully terminated by either cations or anions, and the same (2×2) reconstructions are observed on different terminations, but with different structural origins.

3.2.2.1. The (2×1) reconstruction on cleaved Si(111) and Ge(111) and on clean C(111)

Both Si{111} and Ge{111} display a (2×1) reconstruction upon cleavage. The doubling of surface periodicity relative to the ideal surface (bulk {111} planes) occurs in <211> directions, and the three equivalent <211> directions in the diamond lattice may give rise to three different orientations of (2×1) domains. In an early model (Haneman, 1961), the doubling of periodicity was thought to result from alternating rows of depressed and raised surface atoms, which was expected to lower surface energy by rehybridization of sp³ orbitals on the surface atoms, inducing a charge transfer from the downward atoms with p-like dangling bonds to the upward atoms with s-like dangling bonds.

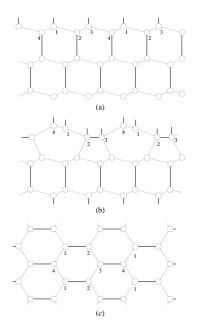


Fig 8. π -bonded chain model of the (2x1) reconstruction on {111} surfaces of solids with the diamond structure. (a) Side view of ideally terminated surface. (b) Top view of the reconstructed surface.

This simple model, accepted until 1980, was found to be inconsistent with the measurements of surface electronic properties, in particular, the core-level shifts of surface atoms measured by photoemission (Himpsel et al., 1980; Brennan et al., 1980). A model proposed by Pandey (1982), which consists of zigzag π -bonded chains along (110) directions (Figs. 8b and 8c), was found to be consistent with these data. These chains are joined to the underlying bulk by five- and seven-member rings instead of the sixmember rings that are characteristic of the diamond structure (see Fig. 8b). The arrangement of surface atoms into such zigzag chains reduces surface energy because it conserves the total number of dangling bonds in the surface but makes them nearest neighbors and hence promotes some *n*-bonding among them rather than having them as secondnearest neighbors, as in the ideal or the earlier model surfaces. Surface energy is further reduced by tilting (or buckling) of chains, similar to buckling of dimers on Si(001) and Ge(001), as shown by both ab initio theories and various experiments (for a collection of theories and experiments, see e.g., Schlüter, 1988; Mönch, 1995).

The chains can be easily formed by simultaneously depressing a surface atom (atom 3 in Fig. 8b) into the second layer and raising a second-layer atom (atom 4 in Fig. 8b) into the surface layer, with a very small barrier of ≤ 0.03 eV (Northrup and Cohen, 1982). The reconstructed surface is about 0.3~0.4eV/atom more stable than the ideal surface (Pandey, 1982; Northrup and Cohen, 1982). The small formation barrier explains why the chain structure can be readily formed during cleavage. The formation process may be mediated by generation of stacking faults that involves formation of five- and sevenmember rings (Reichardt, 1991).

Clean diamond {111} surfaces, which are prepared by annealing polished samples in ultrahigh vacuum in the temperature range of 950°C to 1200°C, show a (2×2) LEED pattern (Marsh and Farnsworth, 1964; Lander and Morrison, 1966). The (2×2) pattern is interpreted as the superposition of diffraction patterns from more than one of the three possible (2×1) domains present in the surface. [Occasionally, a (2×2) pattern also occurs on Si(111) and Ge(111)]. So, similar to the {111} surfaces of Si and Ge, the {111} surfaces of C display a (2×1) reconstruction, which can be explained by Pandey's chain model (Pandey, 1982). However, differences exist between the {111} surfaces of C and those of Si or Ge. In C(111), the chains are flat, as shown by total-energy calculations (for a list of theories, see Mönch, 1995), and are purely stabilized by strong π bonding between C dangling bonds. In Si(111) [or Ge(111)], however, the π -bonding between dangling bonds is much weaker, and it alone is not sufficient to stabilize the chain structure. Consequently, the chains are buckled and the structure is primarily stabilized by the buckling-induced charge transfer between surface atoms rather than π -bonding (Badziag and Verwoerd, 1988).

3.2.2.2. The Si(111)- (7×7) and Ge(111)- $c(2 \times 8)$ reconstructions

The (2×1) reconstruction on the as-cleaved Si(111) and Ge(111) surfaces transform irreversibly into (7×7) and $c(2 \times 8)$ structures, respectively, after annealing at elevated temperature (Lander et al., 1963), confirming the metastability of the (2×1) reconstruction. The calculated surface energy of Si(111)-(7 \times 7) is indeed lower than that of Si(111)- (2×1) (Brommer *et al.*, 1992). The determination of atomic structure involving large surface unit meshes is generally difficult. It took almost thirty years to solve the Si(111)- (7×7) reconstruction after it was first observed (Schlier and Farnsworth, 1959). After many different experimental techniques had been applied and a variety of models have been proposed, a complete description of the structure was finally achieved by Takayanagi et al. (1985a; 1985b), who used a comprehensive analysis of transmission electron diffraction patterns and all the prior knowledge developed for this surface. They introduced the dimer-adatom-stacking fault (DAS) model, showing how three key ingredients involved in the reconstruction that had each been independently suggested earlier, the adatom, the stacking fault, and the dimer, contributes to the overall structure shown in Fig. 9.

The (7×7) unit cell consists of two triangular subunits; one of them (the left half in Fig. 9) contains a stacking fault producing wurtzite-type stacking of the outermost two double layers. There are six adatoms in each of the two subunits, in a (2×2) arrangement. They are located directly over the second-layer atoms and bind to three first-layer atoms. The faulted and unfaulted subunits are separated by a triangular network of partial dislocations, and dangling bonds along these dislocations are partially saturated by the formation of dimers (three dimers at each triangle edge), which are linked by 8-member rings. At the corners of the unit cell, i.e., the crossing points of the dislocations, there are holes (corner holes) exposing large portions of the second double layer. Around the corner holes, atoms are connected by 12-member rings.

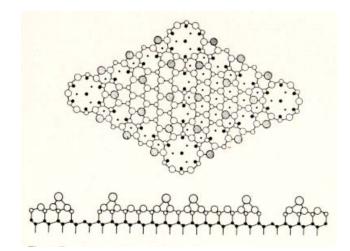


Fig 9. (Top) Top view and (bottom) side view taken along the long diagonal of the unit cell, of the dimer-adatom-stacking fault (DAS) model of the Si(111)-(7x7) reconstruction. The adatom layer (large shaded circles) and two double layers (open and solid circles) are indicated. The stacking fault is in the left half of the unit cell. (After Takayanagi et al. 1985b.)

The DAS model combines all the structural ingredients that had emerged from previous experimental studies. The adatom is the most effective way to reduce the dangling-bond density on a {111} surface of diamond structures (Harrison, 1976). However, incorporation of adatoms also introduces large strain in the surface. *Ab initio* theories (Northrup and Cohen, 1984; Northrup, 1986; Meade and Vanderbilt, 1989) show that the balance between the gain in electronic energy and the cost in strain energy produces, as the most stable arrangement of adatoms on Si(111), a (2 × 2) lattice in which adatoms sit on top of the second-layer atoms (in T₄ sites) with one rest atom (unsaturated surface atom) per (2 × 2) unit mesh. This arrangement has been seen directly by STM

(Binnig *et al.*, 1983; Becker *et al.*, 1985; Hamers *et al.*, 1986b) in both triangle subunits of (7×7) unit cell of Si(111). However, for Si(111), adatoms alone are unable to stabilize the surface.

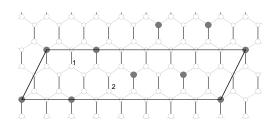


Fig. 10. Adatom model of the $Ge(111)-c(2 \ge 8)$ re-construction. Large and small open circles mark the first-layer and second-layer atoms respectively. Shaded circles are Ge adatoms sitting on top of the second-layer atoms. Numbers 1 and 2 denote two inequivalent sites of rest atoms in the first layer. (After Becker et al., 1989.)

Additional structural changes must be incorporated, in particular, stacking faults (Bennett et al., 1983) are introduced to create dimers at the boundaries between faulted and unfaulted subunits (McRae, 1983; Himpsel, 1983) to reduce further dangling-bond density. STM, equipped with spectroscopy (Tromp et al., 1986) also reveals details of other structural features, such as the rest atoms, the corner holes, and the asymmetry between the two triangle subunits. Consequently, in the (7×7) structure, the overall electronic-energy gain (from the reduction of dangling bonds by a factor of 2.6 by creation of adatoms and dimers) offsets the strain-energy cost associated with creation of adatoms and with the formation of dimers and stacking faults, as shown by theoretical calculations (Qian and Chadi, 1987a; Štich et al., 1992; Brommer et al., 1992).

The structure of the Ge(111)-c(2×8) surface, which involves only adatoms, is much simpler than that of the Si(111)-(7×7) surface. STM images (Becker *et al.*, 1985;1989) show that there are four adatoms and four rest atoms in the unit cell, as

shown in Fig. 10. All adatoms are located at T₄ sites (van Silfhout et al., 1990) while the rest atoms occupy two inequivalent positions (Hirschorn et al., 1991). There is one dangling bond on each adatom and on each rest atom, and the dangling bonds on adatoms are empty while those on rest atoms are filled (Becker et al., 1989), indicating charge transfer from adatoms to rest atoms to reduce surface energy. The Ge dangling-bond energy is lower than that of Si. Also, Ge adatoms introduce less surface strain on Ge(111) than do Si adatoms on Si(111) because Ge is softer than Si (the elastic constants are approximately 40% larger for Si than for Ge). Consequently, for Ge(111), the introduction of adatoms alone is energetically favorable because the energy gain by reducing dangling-bond density by a factor of 2 more than offsets the adatom-induced energy cost in strain. The balance between the electronic-energy gain and the strain-energy cost is rather delicate in the (7×7) and c(2) \times 8) reconstructions. Applying a compressive strain to a Ge(111) surface will convert the $c(2 \times 8)$ structure into the (7×7) structure, as demonstrated by the growth of thin Ge films on Si(111) (Gossmann et al., 1985). Also, there exists a family of $[(2n-1) \times (2n-1)]$ DAS structures that employ the same underlying principle of energy balance as in the (7×7) structure. Depending on sample conditions and annealing temperatures different domains with (5×5) , (7×7) , and (9×9) unit cells may appear on Si(111) (Becker et al., 1986). The detailed mode of reconstruction can be controlled by varying the lateral stress (strain) in the surface. For example, the (5×5) reconstruction appears in the SiGe alloy surface under compressive stress as a thin SiGe alloy film is grown on a Si(111)-(7×7) substrate (Gossmann et al., 1984), as well as in the Si(111) surface under tensile stress, obtained from a Si-Ge thin-film sandwich structures (Ourmazd et al., 1986).

3.2.2.3. The {111} surfaces of compound semiconductors

The ideal $\{111\}$ - (1×1) surfaces of binary compound semiconductors terminated by either cations or anions are conventionally denoted as (111) and $(\overline{111})$ surfaces, respectively. For III-V as well as II-VI compound semiconductors, both cation-terminated (111) and anion-terminated $(\overline{111})$ surfaces exhibit a (2×2) reconstruction (Ebina and Takahashi, 1982), but with different structural origins. The (2×2) structure of (111) surfaces arises from the formation of cation vacancies while that of $(\overline{111})$ surfaces arises from the formation of regular array of anion trimers on top of the anion-termination layer.

Most studies have been carried out on GaAs (111) and (111) surfaces. In GaAs(111), each (2×2) unit mesh contains one Ga vacancy, exposing three As atoms in the second layer below the surface. Consequently, there are equal numbers (three) of Ga and As atoms in the unit mesh, and Ga dangling bonds are empty and As ones are filled due to the charge transfer from Ga to As. This structural and electronic configuration has been directly seen by STM (Haberern and Pashley, 1990). A LEED study (Tong et al., 1984) shows that the exposed secondlayer As atoms relax toward the center of Ga vacancies, leading to almost planar (on Ga atoms) and pyramidal (on As atoms) surface bonding configurations. The resultant atomic arrangement becomes similar to the tilted chains on nonpolar (110) surfaces of compound semiconductors (see discussion below). Total-energy calculations (Chadi, 1984) reveal that the formation of Ga vacancies on GaAs(111) is exothermic but the formation of As vacancies on GaAs(111) is endothermic. The vacancy-induced (2×2) reconstruction is also observed on other similar systems, such as GaP(111), InSb(111), and GaSb(111) (Xu *et al.*, 1985; Bohn *et al.*, 1985; Feidenhans'l, *et al.*, 1987).

Different models have been proposed for the GaAs($\overline{111}$)-(2 × 2) reconstruction. It can not originate from As vacancies because of their large positive formation energy (Chadi, 1984). An anion-cluster model that consists of As trimers sitting on top of a complete terminating As layer was suggested from STM (Biegelsen *et al.*, 1990c). This structural model satisfies the electron-counting rule (Pashley, 1989) discussed in section 3.2.1.2, and it possesses a lower surface energy than many other proposed models, as shown by an *ab initio* calculation (Kaxiras *et al.*, 1986). The anion trimer model is also supported by a TEM study of InSb($\overline{111}$) (Nakada and Osaka, 1991).

3.2.3 The {110} surfaces

The {110} surfaces of elemental semiconductors have complex structures, most notably superstructures involving steps. For example, a (16 × 2) reconstruction, which consists of 25 Å wide stripes of singular terraces separated by up and down monatomic-height steps running along $\langle 112 \rangle$ directions, has been observed on both Si(110) (Yamamoto, 1994) and Ge(110) (Ichikawa *et al.*, 1995). A similar reconstruction with all the steps in the same sense (i.e., all steps up or down) also forms on Si(110) (Yamamoto, 1994), leading to a vicinal facet of (17,15,1) structure.

Both III-V and II-VI semiconductors that crystallize in the zincblende structure cleave naturally along {110} planes, because these planes contain an equal number of cations and anions and thus are intrinsically neutral.

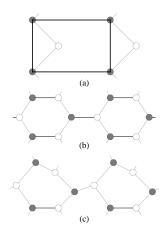


Fig 11. Schematic atomic arrangements at (110) surfaces of zinclende-stucture compound semi-conductors. Open and shaded circles denote cations and anions respectively. (a) Top view. (b) Side view of ideal bulk-terminated surface. (c) Side view of the relaxed surface.

The ideal surface consists of planar zigzag chains of alternating cations and anions along the [110] direction; each surface atom remains three-fold coordinated, with one dangling bond (Figs. 11a). The surface displays no reconstruction but sizable relaxations. The simplest relaxation at the $\{110\}$ surfaces of a zincblende crystal is the tilting (buckling) of the zigzag chains (MacRae and Gobeli, 1966; Levine and Freeman, 1970), as shown in Fig. 11b. Because bond energies change less with changing bond angle than with changing bond length, the buckling of the chain involves mainly bond rotations with bond length being approximately conserved, as shown by a collection of experimental results on {110} surfaces of III-V, II-VI, and even I-VII compounds (Mönch, 1995). Similar to buckled dimers on Si(100) or Ge(100), the buckled chains reduce surface energy by charge transfer from the downward atoms to the upward atoms because the dangling bond on the lower atoms becomes more p like (with backbonds more sp²-like) while that on the upper atoms becomes more s-like (with backbonds more p-like). Also, the dangling-bond energy is lower on anions than on cations (see section 3.2.2B), so all the cations on the {110} surfaces of compound semiconductors move downward and all the anions move upward (Fig. 11b), leading consistently a charge transfer from the downward cations to upward anions. A correlation exists between buckling angle and ionicity of the underlying compound. For compounds with large ionicity, substantial charge transfer can be induced by the large difference in electronegativities between surface cations and anions, without a need for much buckling of chains. Thus, in general, the angle of buckling is found to decrease with increasing ionicity (Mönch, 1995).

3.3 Metal surfaces

The nature of chemical bonding in metals differs from that in semiconductors. Covalent bonds are directional, while metallic bonds are more or less isotropic. Consequently, atoms in a semiconductor have to bond with each other in specific orientations, leading to a tetrahedral network and the diamond and zincblende structures: atoms in a metal, in contrast, bond as much as possible closely packed, forming face centered cubic (fcc), hexagonal close-packed (hcp), or body centered cubic (bcc) crystal structures, the three simple 3D lattices with highest atom number density. The fcc and hcp structures represent the closest packing of spheres in three dimensions, both having 12 nearest neighbors, while the bcc structure has 8 nearest neighbors but 6 nearly as close second-nearest neighbors, with a separation only slightly larger than the first-nearest-neighbor distance. Such close packing maximizes the electron density overlap. Different packing sequences result from different electronic structures of the constituent metal atoms.

The difference in bonding between semiconductors and metals also strongly influences the structural properties of the respective surfaces. As just described, in semiconductor surfaces, removal of atoms at the vacuum side produces a great number of dangling bonds, and thus a strong tendency for reconstruction to reduce their number. This driving force for reconstruction disappears in metal surfaces because there are no dangling bonds. Metal surfaces generally exhibit less tendency to reconstruct than semiconductor surfaces.

An empirical estimate of the magnitude of surface relaxation and the tendency to reconstruct for systems of close-packed spheres can be obtained by comparing the interatomic spacing to the diameter of a free atom. For example, rare-gas crystals can be approximated as the close packing of 'hard' spheres. The atoms bond with each other via the weak van der Waals interaction, and the atomic radii hardly change in the solid. Consequently, there is very little relaxation (less than a few percent) and no reconstruction of rare-gas crystal surfaces. Metallic bonds are much stronger than the van der Waals interaction; consequently interatomic spacings in a metal are less than the diameters of free metal atoms and metal surfaces have a stronger tendency to relax. Some metal surfaces also reconstruct, optimizing the interatomic spacings and/or the coordination numbers of surface atoms in order to minimize the surface free energy.

A more detailed understanding of the structure of metal surfaces than the sphere packing model requires knowledge of the electronic properties of metals. They can be described in the simplest limit by a jellium model, in which a periodic array of positive ions is embedded in the uniform density of a free-electron gas. Two theoretical approaches in the spirit of the jellium model, the embedded-atom method (EAM; see e.g., Foiles, 1987) and the effective-medium theory (EMT; see e.g., Jacobson and Nørskov, 1988) have been widely used to determine many properties of metals and metal surfaces. At equilibrium metal atoms are located at sites of most favorable electron density to minimize their free energy. At the surface the electron density is reduced because of the removal of neighbors on the vacuum side. The most intuitive (but not always correct) conclusion one draws is that the atoms in the outermost surface layer can reduce their energy by relaxing inward to higher electron density. Such inward relaxation frequently does occur. For example, the spacing of the two outermost layers of unreconstructed fcc (110) metal surfaces is reduced between 5.1% in Pd(110) (Schottke *et al.*, 1987) and 17.1% in Pb(110) (Breuer *et al.*, 1991) relative to the bulk value.

A relaxation in the opposite direction, i.e., an increase in the separation of the two outer layers, is also possible, as observed, e.g., in the Pt(111) surface (Materer *et al.*, 1995). Real metals are frequently not well described by a jellium model, i.e., the electron density does not decrease uniformly at the surface. Electronic surface states can introduce a local maximum in the electron density at a position outside the regular 3-D lattice position of the outermost layer of atoms. The lowest-energy state of the surface may then involve an outwards relaxation.

The electron density at the surface can also be minimized by reducing the interatomic spacing within the surface layer or by increasing the coordination number of the surface atoms, either by incorporating extra atoms into the surface layer or by rearrangement of the atoms. These effects produce reconstructions. A reconstruction may be caused by large atomic displacements or by atomic displacements so small that the reconstruction is difficult to observe. We restrict ourselves here to reconstructions that are unambiguously identified.

The desire of the surface atoms to reduce their lateral spacing leads to tensile stress in not yet reconstructed surfaces (this means that the surface tension would be reduced if the crystal were compressed). For example, *ab initio* calculations (Feibelman, 1995) have shown an excess surface stress in Pt(111) of about 0.25 $eV/Å^2$. Whether the excess surface stress is strong enough to induce a reconstruction depends on its value and on the energy cost in creating the reconstruction. example, atomic For the rearrangement, including the incorporation of additional atoms within the surface layer, displaces surface atoms away from their regular lattice positions. Because atoms in the topmost layer will lose some of their neighbors in the second layer, the interaction between the topmost layer and the second layer is weakened. The appearance of reconstruction depends on the balance between the energy gain from increasing the atom number density in the topmost layer and the weakening of the interaction between the topmost and second layer.

Most metal surfaces exhibit large relaxation but only some surfaces, including some orientations of the bcc metals W and Mo and of the fcc metals Au, Pt, and Ir, reconstruct. Intuitively, the higher the atom number density in a surface layer, the easier is the likelihood that the surface reconstructs. A higher coordination implies a more bulk-like environment. The atom number density of a surface plane depends on the surface orientation (see Fig. 12). Nevertheless, there are surfaces with highest atom number densities [e.g. Au(111)] that do reconstruct. At room temperature the W(100) surface is unreconstructed, but during cooling to 250K W(100) reconstructs spontaneously to a $\sqrt{2} \times \sqrt{2}$ R45° structure (Fig. 13) (see e.g., Pendry et al., 1988). Figure 13b shows that the lateral displacement of the atoms in the first layer leads to zigzag chains. The atoms in the second layer are also slightly displaced, by about 20% of the displacement of the atoms in the topmost layer (Altmann et al., 1988). Additionally the topmost

layer shows an inward relaxation of about 6%. *Ab initio* calculations (Singh and Krakauer, 1988) have identified the gain in energy due to an increase in the effective coordination of the surface atoms as the driving force towards the reconstruction. Mo(100) exhibits a similar reconstruction at lower temperatures, but with a 7 times larger unit cell (Daily *et al.*, 1993).

The fcc 5d transition metals (Au, Pt, Ir) display a variety of reconstructions on different low-index surfaces. In the following we discuss, as an example, the general trend of surface relaxation and reconstruction on these 5d metals.

3.3.1 The fcc (110) surface

Many experiments (for a list of these experiments, see, e.g., Rous, 1995) have shown that the (110) surface of fcc 5d metals has a strong tendency to reconstruct. Figure 14a shows a schematic diagram of the unreconstructed surface. Atoms in the topmost layer have only 7 nearest neighbors, a great reduction from the bulk value of 12. Atoms in the second layer also are missing one nearest neighbor, which would sit on top of them. Even more importantly, an atom located in the (110) plane has only 2 in-plane nearest neighbors along the [110] direction. In the perpendicular direction there are no in-plane nearest neighbors, leading to a very weak interaction between atoms in different [110] chains. Chains of atoms in such a configuration can be rearranged easily without losing nearest- neighbor bonds.

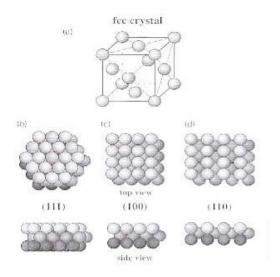


Fig. 12. Surfaces of fcc crystals. (a) Models of the fcc crystal. (b)-(d) Top and side views of the (111), (100), and (110) surfaces, respectively.

The large potential for reconstruction leads to the missing-row structure, observed, e.g., on Au(110) and Pt(110). Figure 14c shows a schematic view of the missing-row reconstruction. Counting all atoms exposed to vacuum in the top, second, and third layers as surface atoms, the reconstruction changes neither the average coordination of surface atoms nor the surface atom number density. However, the distribution of coordination of surface atoms has changed. In the unreconstructed surface (Fig. 14b), half of the surface atoms (in the topmost layer) lose 5 neighbors, and the other half (in the second layer) lose 1 neighbor; in the reconstructed surface (Fig. 14d), a fourth of surface atoms (in the topmost layer) lose 5 neighbors, one half (in the second layer) lose 3 neighbors, and another fourth (in the third layer) lose one neighbor.

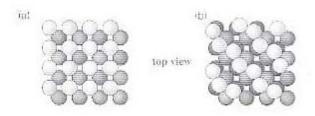


Fig. 13. Top view of the W(100) surface: (a) un-reconstructed and (b) $\sqrt{2} \times \sqrt{2}$ reconstructed.

The missing-row reconstruction transforms half of the atoms missing 5 or 1 neighbors into ones missing 3 neighbors by breaking some second-nearest-neighbor bonds between rows. Calculations have shown that this more uniform coordination leads to a reduction in free energy. The free-energy reduction for the missing-row reconstruction can be visualized by recognizing that the reconstruction is equivalent to forming small (111) facets (Binnig *et al.*, 1983), the most stable surface of fcc crystals.

Surfaces of fcc (110) metals exhibit a systematic correlation between formation of a missing-row reconstruction and electronic properties. Whereas all 5d fcc transition metals reconstruct, the fcc 4d metals and the 3d transition metals show no tendency to reconstruct. The trend can be understood by comparing the energy cost of forming the missing-row reconstruction with the energy gain.

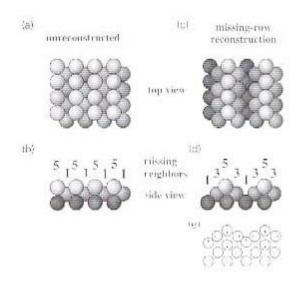


Fig 14. Reconstruction of fcc (110) surfaces. (a), (b) Top and side views of the unreconstructed surface same as Fig. 12(d). (c), (d) Top and side views of the 2 x 1 missing-row reconstructed surface. (e) A schematic side view of the relaxation of the Au(110)-(2 x 1) missing-row reconstruction as proposed by Moritz and Wolf (1985).

The cost of breaking the second-nearest-neighbor bonds between neighboring rows is balanced by the gain produced by a more uniform distribution of number of nearest neighbors. Thus, the gain is related to a more localized effect (the nearest-neighbor bonds) than the cost (the second-nearest-neighbor bonds). The electrons of the 5d metals are very localized, so that the first-nearest-neighbor effect dominates over the second-nearest-neighbor effect, favoring the reconstruction. As one moves from 5d to 4d and 3d metals, electrons become less and less localized. As a result, the second-nearest-neighbor effect becomes dominant, preempting the reconstruction. This trend is reproduced qualitatively by many calculations (see Woodruff, 1994 for a selection of theoretical publications).

An interesting question is why the (110) surface does not form larger (111) facets, leading to $(N \times 1)$ (N > 2) instead of (2×1) missing-row reconstructions. The simple model described above would lead to larger facets because the energy gain would be greater. The cause lies in the atomic relaxation. As shown in Fig. 14e the relaxation involves atomic displacements both perpendicular and parallel to the surface (Fery et al., 1988). For larger (111) facets the relaxation would have to be mainly a uniform displacement of atoms perpendicular to the (111) facet. In certain cases, such as Au(110), calculations have shown that there is only a slight energy advantage for the 2x1 missing-row reconstruction relative to $(N \times 1)$ structures with larger (111) facets [about 10 meV per atom (Ho and Bohnen, 1987)]. In agreement with these calculations, the Au(110) surface has an intrinsic amount of disorder involving (3×1) and (4×1) reconstructed regions (Binnig et al., 1983; Gritsch et al., 1991). On Ir(110) a (3 \times 1) reconstruction has been observed (Naumovets *et al.*, 1976).

3.3.2 The fcc (100)-surface

In the (100)-surface, atoms have 8 nearest neighbors, four of them within the surface plane. The (100) surface of metal crystals is expected to have a smaller tendency to reconstruct than the (110) surface. Nevertheless all three 5d transition metals have reconstructed (100) surfaces (for a list of these experiments, see, e.g., Rous, 1995). In contrast to the reconstruction of the (110) surface, the reconstruction of the (100) surface is controlled by gain and cost both related to nearest-neighbor bonds. The details of the underlying mechanisms are not yet known.

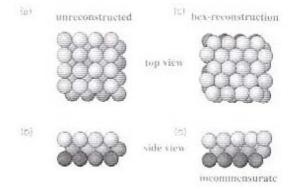


Fig. 15. Reconstruction of fcc (100) surfaces. (a), (b) Top and side views of the unreconstructed surface same as Fig. 12(c), (c), (d) Top and side views of the hexagonally reconstructed surface. The hexagonal layer has a periodicity incommensurate with that of the substrate lattice.

Both Pt(100) and Au(100) form a hexagonal surface layer on top of the square (100) lattice (Fig. 15) (Abernathy et al., 1992). The atom number density in this hexagonal outermost layer is 8 % higher than in a regular (111)-plane. This enhanced atom number density lowers the surface free energy. Such a layer is generally higher-order commensurate or incommensurate with the underlying bulk. The energy cost of this type of reconstruction is the reduced interaction with the second layer, because the incommensuration leads to a lower average coordination to the second layer.

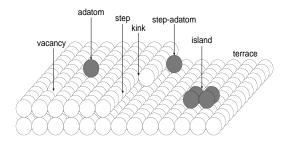


Fig 16. The TSK model of a surface defined for a simple cubic crystal. Open circles are substrate atoms. Shaded circles are adatoms deposited on the substrate. There are two terraces (the upper terrace is on the left) separated by a step containing a single-atom kink.

Ir(100) reconstructs by forming a quasihexagonal (1 × 5) reconstruction (Lang *et al.*, 1983). In one close-packed direction ($\langle 110 \rangle$) the interatomic spacing remains unchanged whereas along the other direction the structure is highly compressed. The layer remains commensurate with the second layer. Because more atoms are located at (or near) their regular lattice sites, the weakening of interactions between the topmost layer and the second layer is not as great as in the incommensurate structures of Pt(100) and Au(100). On the other hand, the bonds within the outermost layer are not as optimized as those in an incommensurate structure.

3.3.3 fcc (111) surface

The (111) surface is the most stable and inert surface of fcc crystals, in which surface atoms have 9 nearest neighbors, three fewer than in the bulk. The reconstructions on Au(100) and Pt(100), with their hexagonal outer layers having about 8% higher atom number density then a regular (111) plane, suggest that a (111) surface reconstruction that increases the outer layer atom number density would lower the free energy.

Au(111) reconstructs by forming the (22 × $\sqrt{3}$ herringbone-structure that allows the surface to increase the atom number density by about 5% (Perderau *et al.*, 1974; Barth *et al.*, 1990). Pt(111) reconstructs by forming a similar structure in equilibrium at high temperatures (~1300K) (Sandy *et al.*, 1992). The reconstruction of Pt(111) is also observed during homoepitaxial growth at lower temperatures (Hohage *et al.*, 1995). Calculations (Needs *et al.*, 1991) show that there is a general tendency for the fcc 5d metals to reconstruct in this manner, due to their excess surface stress.

4. Mophological features of real surfaces, interfaces, and thin films

4.1. The terrace-step-kink (TSK) model

Real surfaces are not perfect but contains defects, such as steps, kinks at steps, adatoms, and vacancies, which are generally described in terms of the terrace-step-kink (TSK) model (Burton *et al.*, 1951), as shown in Fig. 16. These morphological features not only influence surface properties but also play a fundamental role in thin-film growth. The STM has allowed the direct visualization of the TSK model (Swartzentruber, *et al.*, 1990; Wang *et al.*, 1990) and has made possible the quantitative determination of kinetic and thermodynamic properties of morphological features ranging from a single adatom to micrometer-sized terraces.

At thermal equilibrium, the concentration of surface defects is determined by minimization of free energy, i.e., the competition between their formation energy and their configurational entropy. The formation energy of a particular type of defect can be determined by measuring its concentration as a function of temperature. For example, surface steps meander by forming kinks. Step energies and kink energies on Si(001) have been determined by STM from the distribution function of kinks along the steps (Swartzentruber, et al., 1990). Beyond a critical temperature, the surface roughening temperature, surface steps form spontaneously, as the entropy term, which increases linearly with temperature, becomes equal to the step formation energy.

Surface defects may induce both local structural and long-range morphological changes. For example, steps and dimer vacancies on Si(001) "freeze" the motion of buckled dimers between the two equivalent positions at room temperature, as a means to release surface strain energy (Wolkow, 1992). Step-step interaction

(e.g., entropic repulsion) influences kink formation on steps; the meandering of an individual step is independent of that of other steps only when steps are far separated (Swartzentruber, *et al.*, 1990; Wang *et al.*, 1990).

4.2. Equilibrium step configurations

4.2.1 Stress-domain structure

Surface stress plays an important role in determining surface structure and morphology. Theories (Marchenko, 1981; Alerhand et al., 1988) predict that if the intrinsic surface stress is anisotropic, a single-domain surface is unstable against formation of alternating degenerate stress domains. For example, the (2×1) reconstruction on Si(001) introduces a highly anisotropic surface stress: the stress σ_{\parallel} along the dimer bond is tensile (i.e., surface atoms would like to be closer together along this direction than they are); the stress σ_{\perp} along the dimer row is consequently compressive (or at least less tensile than σ_{\parallel}). Because of the tetrahedral bonding configuration in the diamond structure, the dimer direction is orthogonal on terraces separated by monatomic steps. Thus, the intrinsic stress anisotropy, F $= \sigma_{\parallel} - \sigma_{\parallel}$, causes a morphological instability (Marchenko, 1981; Alerhand et al., 1988): a flat Si(001) surface is unstable against formation of monatomic steps as domain boundaries separating alternating (2×1) and (1×2) domains. Such stress domains are quite generally possible in systems that lower their surface energy by reconstruction.

The discontinuity of stress at the step introduces a force monopole, giving rise to an elastic step-step interaction that lowers the surface energy with a logarithmic dependence on step separation (terrace width, L) (Marchenko, 1981; Alerhand *et al.*, 1988), $E_s=C_1 - C_2$ ln(L/a), where C₁ denotes the step formation energy per unit length, C₂ reflects the strength of the interaction,

which is related to the stress anisotropy and elastic constants, and a is a microscopic cutoff length on the order of a lattice constant. The surface assumes the lowest-energy configuration with an optimal step separation of $L_0 = \pi a \exp(C_1/C_2 + 1)$. Therefore, a surface with sufficiently low step density could reduce its energy by introducing extra steps. However, the spontaneous formation of extra steps to reduce the size of the stress domain predicted by theory is not observed on nominally singular Si(001). Instead, the existing steps are observed to become wavy in the low-step-density regions (Tromp and Reuter, 1992). Step undulations lower the surface energy, just as creation of new steps does, by effectively reducing the size of the stress domains (Tersoff and Pehlke, 1992). Step undulations preempt the formation of extra steps because the undulations are kinetically preferred and are compatible with step flow.

4.2.2 Steps on vicinal surfaces

Vicinal surfaces, in particular, vicinal Si(001), play important roles in epitaxial growth and device fabrication, as templates for growing smooth thin films. The step morphology is frequently related to the vicinality of the surface. The behavior of steps on vicinal Si(001) demonstrates the range of morphological phenomena related to steps that can be encountered. On vicinal Si(001) surfaces with a small miscut ($<1.5^{\circ}$), the misorientation produces only monatomic steps, which separate alternating (2 \times 1) and (1 \times 2) domains constituting the stressdomain structure. If the surface is nominally singular (miscut by less than a few tenths of a degree), the equilibrium populations of these two domains are about equal (Men et al., 1988). When an external stress is applied to such a surface, the relative populations of the two domains can be changed by straining the surface at a temperature sufficiently high so that the step mobility is high (Men *et al.*, 1988). The domain compressed along the dimer bond is favored. The experiment is well understood by the theoretical model (Alerhand *et al.*, 1988) mentioned in the last section. The response of relative domain population to an external stress provides a unique way to determine the intrinsic surface stress anisotropy (Webb *et al.*, 1990;1991). For Si(001), the anisotropy is $0.07\pm0.01 \text{ eV/Å}$, as obtained from LEED and STM measurements (Webb *et al.*, 1990;1991). Recent *ab initio* calculations (Garcia and Northrup, 1993; Dabrowski *et al.*, 1994) have produced a value in good agreement with the experiment.

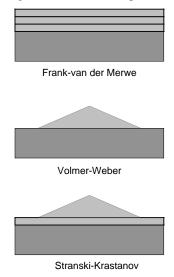


Fig. 17. Three equilibrium growth modes in heteroepitaxy.

Step configurations on high-miscut vicinal Si(001) surfaces are more complicated and their behavior is less well understood. As the miscut angle increases, the populations of (2×1) and (1×2) domains begin to differ, with the population of (1×2) domains increasing gradually with increasing miscut angle (Tong and Bennett, 1991; Pehlke and Tersoff, 1991; Swartzentruber *et al.*, 1992). At sufficiently high miscut (~ 4⁰) double-atomic-height steps start to form (Wierenga *et al.*, 1987; Swartzentruber *et al.*, 1992). The formation energy of double steps is lower than that of the single steps (Chadi,

1987b), but, at a sufficiently small miscut angle, single steps may be stabilized by their elastic stepstep interactions (Vanderbilt et al., 1989). Several model calculations (Alerhand et al., 1990; Poon et al., 1990; Pehlke and Tersoff, 1991) predict the existence of a first-order phase transition from the single-atomic-height-step surface to the doubleatomic-height-step surface as the miscut angle increases. Experimentally, however, no indication of a first-order phase transition has been observed. Instead, the step concentration changes continuously with both miscut angle (Swartzentruber *et al.*, 1992) and temperature (de Miguel et al., 1991). The experiments (de Miguel et al., 1992) also indicate that at high temperatures and/or large miscut angles, a simple 1D elastic model becomes unsatisfactory.

4.3 Surface/interface morphology of epitaxial thin films

4.3.1 Equilibrium growth modes

The surface morphology of a thin film grown on a substrate by vapor deposition is described, at thermal equilibrium, by three different growth modes, as depicted in Fig. 17. In the Frankvan der Merwe mode, growth proceeds layer-by-layer maintaining a smooth film; in the Volmer-Weber mode, 3D islands form immediately and the film does not cover the whole substrate surface; in the Stranski-Krastanov mode, the film first grows layer-by-layer, followed by the formation 3D islands. The realization of different growth modes, and hence different thin-film morphologies, is determined again by minimization of surface free energy, in this case for the combined substrate/film system. The surface energy of the substrate ($\gamma_{substrate}$), the surface energy of the film (γ_{film}), and the interface energy between the film and substrate ($\gamma_{interface}$) (Bauer, 1958) enter into this minimization. For heteroepitaxy, i.e., the

growth of material A on a different substrate B, the surface energy of the film (A) will differ from that of the substrate (B). If $\gamma_{substrate} < \gamma_{film} + \gamma_{interface}$, the deposited film tends to nucleate 3D islands, leaving the low-energy substrate exposed (the Volmer-Weber mode). The condition of $\gamma_{substrate} > \gamma_{film} + \gamma_{interface}$ defines the layer-bylayer growth of Frank-van der Merwe mode. However, γ_{film} increases with increasing film thickness because of misfit strain, which leads to Stranski-Krastanov growth: the film first grows layer-by-layer for $\gamma_{substrate} > \gamma_{film} +$ yinterface, forming a finite thickness of wetting layer, followed by formation of 3D islands as $\gamma_{substrate} < \gamma_{film} +$ $\gamma_{interface}$ to relieve the misfit strain built up in the film. In general, if A wets B, then B will not wet A. For homoepitaxy, i.e., the growth of material A on a substrate of the same material A, $\gamma_{substrate} = \gamma_{film} + \gamma_{interface}$. The equilibrium film grows layer-by-layer in Frank-van der Merwe mode.

Stranski-Krastanov growth has attracted renewed interest recently because of its potential for fabricating 3D spatially confined structures (Liu and Lagally, 1997a). One typical example is the growth of Ge or SiGe alloy on Si. Because Ge has a lower surface energy than Si, Ge initially grows layer-by-layer, wetting the Si substrate. As the film gets thicker, the strain energy in the film increases, and eventually 3D islands start to form to relieve the 4% lattice mismatch between Ge and Si. The equilibrium surface morphology of a thin Ge film grown on Si(001) is rather complex (Liu et al., 1997). For pure Ge films grown on Si(001), the thickness of wetting layer is extended up to 3 ML by strain relaxation mechanisms involving only 2D surface roughening processes before 3D islanding takes place (Mo et al., 1990; Tersoff, 1991). The misfit strain is partially relaxed by mostly the formation of dimer vacancy lines, the $(2 \times N)$ reconstruction (Chen et al., 1994). The reconstruction also induces changes in surface (step) morphology and stress field, reversing the relative roughness of the two types of steps on the surface (Wu *et al.*, 1995) as well as the surface stress anisotropy (Wu and Lagally, 1995; Liu and Lagally, 1996; Liu *et al.*, 1997). The 3D islands form in two stages. Initial islands are small and coherent with the substrate lattice, i.e., without dislocations (Eaglesham and Cerullo, 1990; Mo *et al.*, 1990). Larger dislocated islands form later.

4.3.2 Kinetics-limited growth morphology

Equilibrium thin-film morphologies for different film system can be achieved by growth at high temperature and low deposition rate. Very often, growth is carried out far from equilibrium, and consequently, the evolving surface morphology during thin-film growth is kinetically limited (Liu and Lagally, 1997b). The ability to control kinetic parameters to manipulate thin-film morphology allows us to grow novel artificial film structures with desirable properties. Surface diffusion is the most important kinetic parameter controlling growth morphology. Because the rate of diffusion is a exponential function of temperature, different regimes of growth that lead to different thin-film morphologies can be achieved by tuning the growth temperature.

Real surfaces have imperfections such as steps. As adatoms are deposited onto a surface, they are preferentially incorporated at steps, particularly at kink sites where they can bind to the most neighbors (Burton *et al.*, 1951). If all deposited adatoms are attached at steps, growth will proceed with advancement of steps, so called step flow growth. Step flow growth requires a high mobility of the adatoms so that they reach the existing steps before meeting with each other to nucleate islands. It is achieved by growing at high temperature, at low deposition rates, or on high-step-density surfaces. Step flow growth, representing a situation close to thermal equilibrium, produces smooth films. However, morphological instabilities, e.g., the bunching of substrate steps, may arise from the additional kinetic barrier for adatoms crossing the steps (Ehrlich and Hudda, 1966; Schwoebel and Shipsey, 1966) or from complex dynamics of step flow (Kandel and Weeks, 1995). For step-flow growth of a strained film, a step bunching instability caused by the elastic attractions between steps (Tersoff, 1995) or by strain modified adatom diffusion (Duport *et al.*, 1995) can occur.

At lower temperature, at high deposition rates, or on surfaces with a low step density, adatoms may meet each other in the middle of terraces before reaching the existing steps and nucleate stable 2D islands. As the deposition continues, the nucleated islands grow by incorporating new arriving adatoms, and more islands form by new nucleation events. Eventually, the 2D islands meet and coalesce. This regime is referred to as two-dimensional island growth. The 2D island growth mode is used for fundamental studies of kinetic mechanisms of growth and to calibrate deposition rates, but serves little technological function.

The nucleation and growth of 2D islands can be described by phenomenological rate equations (Venables *et al.*, 1984). The properties of islands (e.g., island number density, island compactness, and island shape) are largely determined by surface diffusion. The surface diffusion coefficient can be quantitatively determined by STM measurement of the island number density as a function of temperature at the initial stage of island growth (Mo *et al.*, 1991; Bott *et al.*, 1996; Liu and Lagally, 1997b). The compactness of islands (fractal-like vs. close-packed) is controlled by how fast an adatom diffuses along island edges and crosses corners where two edges meet (Hohage *et al.*, 1996; Zhang and Lagally, 1997). In general, growth at relatively lower

temperatures produces less compact (more fractallike) islands, because adatoms diffuse more slowly and thus do not reach their equilibrium positions before they become trapped by newly arriving atoms. On reconstructed surfaces, 2D island growth becomes more complex. For example, on (2×1) reconstructed Si(001), the large anisotropy in adatom sticking coefficients to two different edges (steps) of an island creates kinetics-limited islands with long and thin shapes. Reconstruction can also introduce additional roughness to the growth front. Two islands nucleating at different locations in the same terrace can form antiphase domain boundaries when they meet. For example, in Si(001) islands have a 50% chance to align their dimer rows out of phase (Hamers et al., 1989), roughening the growth front. The GaAs(001) (4×2) structure has four different possibilities of island positions with only 1/4 of the islands meeting without antiphase boundaries. The larger the unit mesh, the more possibilities exists.

Islands may nucleate not only on terraces but also on top of those islands that are already formed, when these become large enough so that atoms falling on them can no longer reach the edge. In addition to surface diffusion, the barrier for an adatom to jump off an island (i.e., crossing a step) now becomes a key kinetic parameter in determining the film roughness, even for homoepitaxial films. The roughness of the film increases as the film grows thicker, if atoms cannot cross the steps..

In heteroepitaxy, misfit strain can lead to various 2D morphological instabilities that roughen the surface, even in the absence of kinetic roughening, and are due to energetic (thermodynamic) driven forces. For example, as Ge is deposited on Si(001), the surface first roughens by formation of dimer vacancies, giving rise to the $(2 \times$ N) reconstruction. The dimer vacancy lines reduce adatom diffusion (Lagally, 1993), leading subsequently to additional roughness induced by new kinetic limitations (Mo and Lagally, 1991; Lagally, 1993; Wu and Lagally, 1996).

4.3.3 Interface morphology

The quality of the interface, in particular the smoothness of the interface of a thin-film structure, frequently influences other properties, and may determines the usefulness of the film, for example, for electronic and magnetic applications. A strained film is inherently unstable; strain energy increases as the film grows thicker. Upon reaching the "equilibrium critical thickness" (van der Merwe, 1963; Matthews and Blakeslee, 1976), a strained film will relax back equilibrium lattice constant, forming to its dislocations at the interface, if the kinetics permit it. Films can be grown beyond the equilibrium critical thickness by growing at low enough temperatures so that the kinetics are so small that dislocation formation is suppressed. For homogeneous nucleation of dislocations, the kinetic barrier for nucleating a dislocation decreases with increasing strain (LeGoues et al., 1993), that eventually, as the misfit strain increases, dislocations always will form. A structural defect at the surface or interface may activate the nucleation of dislocations.

It is difficult to create a smooth interface, even without forming dislocations. Interdiffusion roughens the interface by creating a transition region in which a compositional gradient exists. In general, interdiffusion lowers the interface energy by increasing the configurational entropy. In those cases when $2E_{A-B} > E_{A-}$ $A + E_{B-B}$, where E denotes the bond energy, there will also be a strong tendency for interdiffusion to form interface alloy compounds. Substantial interdiffusion can occur even at relatively low temperatures, at which bulk diffusion is negligible, through surface diffusion-assisted atom place exchange in the outer few layers, driven by surface segregation of the thermodynamically more stable component and/or by preferential occupation of different components at different lattice sites induced by surface reconstruction. For example, Ge surface segregation causes a decrease of Ge concentration at the interface when a SiGe alloy film is grown on Si and a non-abrupt interface when Si is grown on Ge. The (2×1) reconstruction on Si(001) introduces a nonuniform stress distribution in the surface layers (LeGoues *et al.*, 1990; Liu and Lagally, 1996), leading to preferential occupations of Si or Ge at different sites and formation of an ordered SiGe phase at the interface when SiGe alloy is deposited on Si(001) (LeGoues *et al.*, 1990).

Pre-existing surface roughness generally remains at the interface after overgrowth. In the growth of a multilayer thin film, the situation is more complex because the growth of B on A can be quite different from that of A on B. For example, if B wets A, A will not wet B. In general, the roughness of a surface or interface can be quantitatively characterized by its rms roughness, lateral correlation length, and fractal exponent (Sinha et al., 1988). A x-ray diffraction technique has been developed to measure the magnitude and correlation of interfacial roughness in a multilayer film (Savage et al, 1991; Phang et al, 1994). The interfacial roughness is frequently correlated among different layers (Savage et al, 1991; Phang et al, 1994). In particular, straininduced interaction between islands in different layers leads to a self-organization of islands, with progressively improved uniformity in islands size, shape, and spacing, as the multilayer thickness increases (Tersoff et al., 1996; Teichert et al., 1996), which provides an attractive route to fabricating arrays and superlattices of quantum dots (Liu and Lagally, 1997a).

Glossary

Adatom: A single atom adsorbed on a surface (terrace) bonded only to the surface atoms.

Basis: A group of atoms, molecules, or ions that constitute a single lattice point of the Bravais lattice.

Bravais lattice: An infinite array of discrete points with an arrangement and orientation that appears the same from whichever of the points the array is viewed.

Crystal Systems: Point groups of the Bravais lattice (the symmetry of the Basis is excluded).

crystallographic point groups: *Point groups* of the general crystal structure including the symmetry of the Basis.

crystallographic space groups: *Space groups* of the general crystal structure.

Dangling bonds: A hybridized electron orbital on an atom without bonding due to missing neighbors.

Dislocation: An extended line defect resulting from missing half of an atomic plane in the solid or from missing half of an atomic row in the surface.

Dividing surface: A hypothetical geometrical surface separating two homogeneous phases to replace the real physical transition region between the two phases.

Epitaxial growth: Growth of a film on a substrate with lattice structure of the film confined to that of the substrate.

Excess surface properties: The difference between the extensive thermodynamic property of a real system and that of an ideal system consisting of two homogeneous phases separated by the dividing surface.

Heteroepitaxy: Epitaxial growth of a material A on a substrate of different material B.

Homoepitaxy: Epitaxial growth of a material A on a substrate of the same material A.

Jahn-Teller distortion: Lattice distortion that lowers the symmetry, removing the degeneracy in electronic levels and hence lowering the total energy.

Kink: A kink is a jog of any length in a step terminated by one inside corner and one outside corner.

Miller indices: The indices of an atomic (lattice) plane that are coordinates of the shortest reciprocal lattice vector normal to the plane.

Misfit strain: Strain induced by different equilibrium lattice constants (misfit) between the film and the substrate during heteroepitaxial growth.

Primitive unit cell: A volume (area) of space that, when translated through all the vectors in a Bravais lattice, just fills all the space without either overlapping itself or leaving a void.

Reciprocal lattice: The set of all wave vectors K that yield plane waves with the periodicity of a given Bravais lattice.

Stacking fault: A single plane where the stacking differs from the usual stacking sequence of atomic planes inside the crystal.

Stacking fault region: A finite region of a stacking sequence of atomic planes different from that in the surrounding lattice.

Surface reconstruction: Atomic displacements at a surface that change the surface periodicity relative to that of an ideally bulk terminated surface.

Surface relaxation: Atomic displacements at a surface that do not change the surface periodicity relative to that of an ideally bulk terminated surface.

Surface stress: (I) The reversible work per unit area required to deform a surface. (II) The force per unit length on a line cut in the surface.

Surface tension: The reversible work required to create a new unit area of surface.

Vicinal surface: A crystal surface formed by a miscut a few degrees away from a low-index atomic plane.

List of Works Cited

Abernathy D.L., Mochrie S.G., Zehner D.M., Grübel G., and Gibbs D (1992), <u>Phys. Rev. B</u>, 45, p. 9272.

Alerhand O.L., Vanderbilt D., Meade R.D., and Joannopoulos J.D. (1988), <u>Phys. Rev. Lett.</u>, 61, p. 1973.

Alerhand O.L., Berker N.A., Joannopoulos J.D,

Vanderbilt D., Hamers R.J., and Demuth J.E. (1990), Phys. Rev. Lett., 64, p. 2406.

Altmann M.S., Estrup P.J., and Robinson I.K. (1988), Phys. Rev. B. 38, p. 5211.

Arthur J.R. (1974), Surf. Sci., 43, p. 449.

Ashcroft N.W. and Mermin N.D. (1976), <u>Solid State</u> Physics, Holt, Rinehart and Winston, London.

Badziag P. and Verwoerd W.S. (1988), <u>Surf. Sci.</u>, 201. p. 87.

Barth J.V., Brune H., Ertl G., Behm R. J. (1990), Phys. Rev. B, 42, p. 9307.

Bauer E. (1958), Z. Kristallogr., 110, p. 372.

Becker R.S., Golovchenko J.A., McRae E.G., and

Swartzentruber B.S. (1985), <u>Phys. Rev. Lett.</u>, 55, p. 2028.

Becker R.S., Golovchenko J.A., Higashi G.S., and Swartzentruber B.S. (1986), <u>Phys. Rev. Lett.</u>, 57, p. 1020.

Becker R.S., Golovchenko J.A., and Swartzentruber B.S. (1985), Phys. Rev. Lett., 54, p. 2678.

Becker R.S., Swartzentruber B.S., Vickers J.S., and Klitsner T. (1989), <u>Phys. Rev. B</u>, 39, p. 1633.

Bennett P.A., Feldman L.C., Kuk Y., McRae E.G., and Rowe J.E. (1983), <u>Phys. Rev. B</u>, 28, p. 3656.

Biegelsen D.K., Swartz L.-E., and Bringans R.D. (1990a), J. Vac. Sci. Technol., A8, p. 280.

Biegelsen D.K., Bringans R.D., Northrup J.E., and Swartz L.-E., (1990b), Phys. Rev. B, 41, p. 5701.

Biegelsen D.K., Bringans R.D., Northrup J.E., and Swartz L.-E., (1990c), Phys. Rev. Lett., 65, p. 452.

Binnig G., Rohrer H., Gerber Ch., and Weibel E. (1983), Phys. Rev. Lett., 50, p. 120. Binnig G., Rohrer H., Gerber Ch., and Weibel E. (1983), Surf. Sci., 131. p. L379. Bohr J., Feidenhans'l R., Nielsen M., Toney M., Johnson R.L., and Robinson I.K. (1985), Phys. Rev. Lett., 54, p. 1275. Bott M., Hohage M., Morgenstern M., Michely Th., and Comsa G. (1996), Phys. Rev. Lett., 76, p. 1304. Brennan S., Stöhr J., Jaeger R., and Rowe J.E. (1980), Phys. Rev. Lett., 45, p. 1414. Breuer U., Prince K. C., Bonzel H. P., Oed W., Heinz K., Schmidt G., and Müller K. (1990), Surf. Sci., 239, p. L493. Brommer K.D., Needels M., Larson B.E., and Joannopoulos J.D. (1992), Phys. Rev. Lett., 68, p. 1355. Burton W.K., Cabrera N., Frank F.C. (1951), Philos. Trans. R. Soc. London Ser., A 243, p. 299. Chadi D.J. (1979), Phys. Rev. Lett., 43, p. 43. Chadi D.J. (1987a), J. Vac. Sci. Technol., A5, p. 834. Chadi D.J. (1987b), Phys. Rev. Lett., 59, p. 1691. Chadi D.J. (1987a), J. Vac. Sci. Technol., A5, p. 834. Chen X., Wu F. Zhang Z.Y., and Lagally M.G. (1994), Phys. Rev. Lett., 73, p. 850. Cho A.Y. (1971), J. Appl. Phys., 42, p. 2074. Cornelissen H.J., Cammack D.A., and Dalby R.J. (1988), J. Vac. Sci. Technol., B6, p. 769. Dabrowski J., Pehlke E., and Scheffler M. (1994), <u>Phys. Rev. B</u>, 49, p. 4790. de Miguel J.J., Aumann C.E., Kariotis R., and Lagally M.G. (1991), Phys. Rev. Lett., 67, p. 2830. de Miguel J.J., Aumann C.E., S.G. Jaloviar, Kariotis R., and Lagally M.G. (1992), Phys. Rev. B, 46, p. 10257. Dovesi R., Pisani C., Roetti C., and Ricart J.M.

(1987), <u>Surf. Sci.</u>, 185, p. 120.

Drathen P., Ranke W., and Jacobi K. (1978), Surf.	Hamers R.J., Tromp R.J.
<u>Sci.</u> , 77, p. L162.	Phys. Rev. Lett., 56, p. 1
Duport C., Nozières P., and Villain J. (1995), Phys.	Haneman D. (1961), Phy
<u>Rev. Lett.</u> , 74, p. 134.	Harrison W.A. (1976), <u>S</u>
Eaglesham D.J. and Cerullo M. (1990), Phys. Rev.	Harrison W.A. (1981), <u>H</u>
<u>Lett.</u> , 64, p. 1943.	Himpsel F.J. (1983), Phy
Ebina A. and Takahashi T. (1982), J. Cryst. Growth,	Himpsel F.J., Heimann
59, p. 51.	D.E. (1980), Phys. Rev.
Ehrlich G. and Hudda F.G. (1966), J. Chem. Phys.,	Hirschorn E.S., Lin D.
66, p. 1039.	<u>Phys. Rev. B</u> , 44, p. 140
Farrell H.H. and Palmstrøm C.J. (1990), J. Vac. Sci.	Ho K.M. and Bohnen I
<u>Technol.</u> , B8, p. 903.	59, p. 1833.
Feibelman P.J. (1995), Phys. Rev. B, 51, p. 1786.	Hohage M., Bott M.,
Feidenhans'l R., Nielsen M., Grey F., Johnson R.L.,	Michely Th., and Coms
and Robinson I.K. (1987), Surf. Sci., 186, p. 499.	76, p. 2366.
Fery P., Moritz W., and Wolf D. (1988), Phys. Rev.	Hohage M., Michely Th
<u>B.</u> , 38, p. 7275.	<u>Sci.,</u> 337, p. 249.
Foiles S.M. (1987), Surf. Sci., 191, p. L779.	Ichikawa T., Sueyosi
Foxon C.T. and Joyce B.A. (1977), Surf. Sci., 64, p.	Udagawa F., and Sur
293.	<u>Commun.</u> , 93, p. 541.
Garcia A. and Northrup J.E. (1993), Phys. Rev. B.,	Ihm J., Lee D.H., Joan
48, p. 17350.	(1983), Phys. Rev. Lett.
Gibbs J.W. (1928), in Collected Works, New York:	Jacobson K.W. and Nø
Longmans, Green & Co., 1, p. 55.	Veen J. K. and Van How
Gossmann HJ., Bean J.C., Feldman L.C., and	Solid Surfaces II, Spring
Gibson W.M. (1984), <u>Surf. Sci.</u> , 138, p. L175.	Kandel D. and Weeks
Gossmann HJ., Bean J.C., Feldman L.C., McRae	74, p. 3632.
E.G., and Robinson I.K. (1985), Phys. Rev. Lett., 55,	Kaxiras E., Bar-Yam
p. 1106.	Pandey K.C. (1986), Phy
Gritsch T., Coulman D., Behm R. J., and Ertl G.	Keating P.N. (1966), Ph
(1991) <u>, Surf. Sci.</u> , 257, p.297.	Kevan S.D. and Stoffel
Haberern K.W. and Pashley M.D. (1990), Phys. Rev.	53, p. 702.
<u>B.</u> , 41, p. 3226.	Kubby J.A., Griffith J.
Hamers R.J., Köhler U.K., and Demuth J.E. (1989),	J.S. (1987), Phys. Rev. I
J. Vac. Sci. Technol., A8, p. 195.	Lagally M.G. (1993), Jp
Hamers R.J., Tromp R.M., and Demuth J.E. (1986a),	Lander J.J., Gobeli G.W
<u>Phys. Rev. B</u> , 34, p. 5343.	<u>Appl. Phys.</u> , 34, p.2298.

mp R.M., and Demuth J.E. (1986b), 56, p. 1972. 1), Phys. Rev., 121, p. 1093. 976), <u>Surf. Sci.</u>, 55, p. 1. 981), Phys. Rev. B, 24, p. 5838. 3), Phys. Rev. B, 27, p. 7782. mann P., Chiang T.-C., and Eastman <u>. Rev. Lett.</u>, 45, p. 1112. Lin D.S., and Leibsle F.M. (1991), p. 1403. hnen K.P. (1987), Phys. Rev. Lett., t M., Morgenstern M., Zhang Z., Comsa G. (1996), Phys. Rev. Lett., ely Th., and Comsa G. (1995), Surf. eyosi T., Sato T., Iwatsuki M., d Sumita I. (1995), Solid State 541. Joannopoulos J.D., and Xiong J.J. <u>. Lett.</u>, 51, p. 1872. nd Nørskov J.K. (1988) in van der an Hove M.A. (Ed.) The Structure of Springer, Berlin). Veeks J.D. (1995), Phys. Rev. Lett., Yam Y., Joannopoulos J.D., and 6), <u>Phys. Rev. Lett.</u>, 57, p. 106. 66), Phys. Rev., 145, p. 637. toffel N.G. (1984), Phys. Rev. Lett., fith J.E., Becker R.S., and Vickers <u>Rev. B</u>, 36, p. 6079. 93), Jpn. J. Appl. Phys., 32, p. 1493. eli G.W., and Morrison J. (1963), J. Lang E., Müller K., Heinz K., Van Hove M.A., Koestner R.J., and Somorjai G.A. (1983<u>), Surf. Sci.,</u> 240, p. 81.

LeGoues F.K., Kesan V.P., Iyer S.S., Tersoff J., and Tromp R. (1990), <u>Phys. Rev. Lett.</u>, 64, p. 2030.

LeGoues F.K., Mooney P.M., and Tersoff J. (1993), Phys. Rev. Lett., 71, p. 396.

Liu Feng and Lagally M.G. (1996), <u>Phys. Rev. Lett.</u>, 76, p. 3156.

Liu Feng and Lagally M.G. (1997a), <u>Surf. Sci.</u> 386, p.169.

Liu Feng and Lagally M.G. (1997b) in King D.A. and Woodruff D.P. (Eds.), <u>The Chemical Physics of Solid</u> Surfaces, Elsevier, Amsterdam, vol. 8, p. 258.

Liu Feng, Wu Fang, and Lagally M.G. (1997), Chemical Reviews, 97, p. 1045.

Lucas C.A., Dower C.S., McMorrow D.F., Wong G.C.L., Lamelas F.J., and Fuoss P.H. (1993), <u>Phys.</u> <u>Rev. B</u>, 47, p. 10375.

Massies J., Etienne P., Dezaly F., and Linh N.T. (1980), <u>Surf. Sci.</u>, 99, p. 121.

Materer N., Starke U., Barbieri A., Döll R., Heinz K., Van Hove M.A., and Somorjai G.A. (1995), <u>Surf.</u> <u>Sci.</u> 325, p. 207.

Marchenko V.I. (1981), JETP Lett., 33, p. 381.

Matthews J.W. and Blakeslee A.E. (1976), <u>J. Cryst.</u> <u>Growth</u>, 29, p. 273; 32, p. 265.

McRae E.G. (1983), Surf. Sci. 124, p. 106.

Meade R.D. and Vanderbilt D. (1989), <u>Phys. Rev. B</u>, 40, p. 3905.

Men F.K., Packard W.F., and Webb M.B., <u>Phys. Rev.</u> Lett., 61, p. 2469.

Mo Y.W. and Lagally M.G. (1991), <u>J. Cryst. Growth</u>, 111, p. 876.

Mo Y.W., Savage D.E., Swartzentruber B.S., and Lagally M.G. (1990), <u>Phys. Rev. Lett.</u>, 65, p. 1020.

Mo Y.W., Kleiner J., Webb M.B., and Lagally M.G. (1991), <u>Phys. Rev. Lett.</u>, 66, p. 1998.

Mönch W. (1995), Semiconductor Surfaces and Interfaces, Berlin; New York: Springer-Verlag, 2nd ed. chapter 7. Mullins W.W. and Hirth J.P. (1963), J. Phys. Chem. Solids, 24, p. 1391. Nakada T. and Osaka T. (1991), Phys. Rev. Lett., 67, p. 2834. Naumovets A.G. and Fedorus A.G. (1976), Sov. Phys.JETP, 41, p. 587. Neave J.H. and Joyce B.A. (1978), J. Cryst. Growth, 44, p. 387. Needels M., Payne M.C., and Joannopoulos J.D. (1987), Phys. Rev. Lett., 58, p. 1765. Needs R.J., Godfrey M.J., and Mansfield M. (1991), <u>Surf. Sci.</u>, 242, p. 215. Noonan J.R. and Davis H.L. (1984), Phys. Rev. B, 29, p. 4349. Northrup J.E. (1986), Phys. Rev. Lett., 57, p. 154. Northrup J.E. and Cohen M.L. (1982), Phys. Rev. Lett., 49, p. 1349. Northrup J.E. and Cohen M.L. (1984), Phys. Rev. B, 29, p. 1966. Ourmazd A., Taylor D.W., Bevk J., Davidson B.A., Feldman L.C., and Mannaerts J.P. (1986), Phys. Rev. Lett., 57, p. 1332. Pandey, K.C. (1982), Phys. Rev. Lett., 49, p. 223. Park R.L. and Madden H.H. (1968), J. Surf. Sci., 11, p. 188. Pashley M.D. (1989), Phys. Rev. B, 40, p. 10481. Pehlke E. and Tersoff J. (1991), Phys. Rev. Lett., 67, p. 465; p. 1290. Pendry J.B., Heinz K., Oed W., Landskron H., Müller K. and Schmidtlein G. (1988), Surf. Sci., 193, L1. Perderau J., Biberian J.P., Rhead G.E. (1974), J. Phys. <u>F</u>, 4, p. 798. Phang Y.H., Teichert C., Lagally M.G., Peticolas T.J.,

Bean J.C., and Kasper E., Phys. Rev. B, 50, p. 14435.

Poon T.W., Yip S., Ho P.S., and Abraham, Phys. <u>Rev. Lett.</u>, 65, p. 2161. 179, p. L63. Qian G.-X. and Chadi D.J. (1987), Phys. Rev. B, 35, p. 1288. Ramstad A., Brocks G., and Kelly P.J. (1995), Phys. 1502. Rev. B, 51, p. 14504. Reichardt D. (1991), Prog. Surf. Sci., 35, p. 63. Rous P.J. (1995) in de Boer F.R. and Pettifor D.G. (Eds.), Cohesion and Structure of Surfaces, Elsevier, Amsterdam, vol. 4, p.1. p. 816. Sandy A.R., Mochrie S.G.J., Zehner D.M., Grübel G., Hwang K.G., Gibbs D. (1992), Phys. Rev. Lett., 68, 2192. Savage D.E., Schimke N., Phang Y.H., and Lagally 67, p. 101. M.G. (1991), Phys. Rev. B, 71, p. 3283. Saxena A., Gawlinski E.T., and Gunton J.D. (1985), Phys. Rev. Lett., 55, p. 1303. Surf. Sci., 160, p. 618. Schlier R.E. and Farnsworth H.E. (1959), J. Chem. Science, 234, p. 304. Phys., 30, p. 917. Schlüter M. (1988) in King D.A. and Woodruff D.P. Lett., 68, p. 820. (Eds.), The Chemical Physics of Solid Surfaces, Elsevier, Amsterdam, vol. 5, p. 64. T.G.J. (1978), Surf. Sci., 72. p. 95. Schwoebel R.L. and Shipsey E.J. (1966), J. Appl. Phys., 37, p. 3682. p.123. Singh D. and Krakauer H. (1988), Phys. Rev. B. 38, p. 3999. Sinha D.T., Sirota E.B., Garoff S., and Stanley H.B. 89, p. 169. (1988), Phys. Rev. B, 38, p. 2297. Skottke M., Behm R.J., Ertl G., Penka V., and Moritz 1013. W. (1987), J. Chem. Phys., 87, p. 6191. Štich I., Payne M.C., King-Smith R.D., Lin J.-S., and Clarke L.J. (1992), Phys. Rev. Lett., 68, p. 1351. Swartzentruber B.S., Mo Y.-W., Kariotis R., Lagally M.G., and Webb M.B. (1990), Phys. Rev. Lett., 65, p. 1913. p. 2430. Swartzentruber B.S., Kitamura N., Lagally M.G., and Webb M.B. (1992), Phys. Rev. B, 47, p. 13432.

Tabata T., Aruga T., and Murata Y. (1987), Surf. Sci.

Takayanagi K., Tanishiro Y., Takahashi M., and Takahashi S. (1985a), J. Vac. Sci. Technol., A3, p.

Takayanagi K., Tanishiro Y., Takahashi M., and Takahashi S. (1985b), Surf. Sci. 164, p. 367.

Tersoff J. (1991), Phys. Rev. B, 43, p. 9377.

Tersoff J. and Pehlke E. (1992), Phys. Rev. Lett., 68,

Tersoff J., Phang Y.H., Zhang Z., and Lagally M.G. (1995), Phys. Rev. Lett., 75, p. 2730.

Tong X. and Bennett P.A. (1991), Phys. Rev. Lett.,

Tromp R.M., Hamers R.J., and Demuth J.E. (1986),

Tromp R.M., Hamers R.J., and Demuth J.E. (1986),

Tromp R.M. and Reuter M.C. (1992), Phys. Rev.

van Bommel, A.J., Crombeen J.E., and Dirschot

van der Merwe J.H. (1963), J. Appl. Phys., 34, p. 117;

van Silfhout R.G., van der Veen J.F., Norris C., and MacDonald J.E. (1990), Faraday Discuss. Chem. Soc.

Vanderbilt D., Alerhand O.L., Meade R.D., and Joannopoulos J.D. (1989). J. Vac. Sci. Tchnol., B7, p.

Venables J.A., Spiller G.D.T., and Hanbücken M. (1984), Rep. Prog. Phys., 47, p. 399.

Wang X.-S., Goldberg J.L., Bartelt N.C., Einstein T.L., and Williams E.D. (1990), Phys. Rev. Lett., 65,

Webb M.B., Men F.K., Swartzentruber B.S., Kariotis R., and Lagally M.G. (1990) in Lagally M.G. (Ed.), Kinetics of Ordering and Growth at Surfaces, Plenum, New York.

- Webb M.B., Men F.K., Swartzentruber B.S., Kariotis
- R., and Lagally M.G.(1991), Surf. Sci., 242, p. 23.

Wierenga P.E., Kubby J.A., and Griffith J.E. (1987), Phys. Rev. Lett., 59, p. 2169.

Wolkow R.A. (1992), Phys. Rev. Lett., 68, p. 2636.

Wood E.A. (1964), J. Appl. Phys., 35, p. 1306.

Woodruff D.P. (1981) in King D.A. and Woodruff

D.P. (Eds.), The Chemical Physics of Solid Surfaces

and Heterogeneous Catalysis, Elsevier, Amsterdam, vol. 1, p. 175.

Woodruff D.P. (1994) in King D.A. and Woodruff

D.P. (Eds.), <u>The Chemical Physics of Solid Surfaces</u>, Elsevier, Amsterdam, vol. 7, p. 513.

Wu F. and Lagally M.G. (1996), <u>Phys. Rev. Lett.</u>, 75, p. 2534.

Wu F., Chen X., Zhang Z.Y., and Lagally M.G. (1995), Phys. Rev. Lett., 74, p. 574.

Xu G., Hu W.Y., Puga M.W., Tong S.Y., Yeh J.L.,

Wang S.R., and Lee B.W. (1985), <u>Phys. Rev. B</u>, 32, p. 8473.

Yamamoto Y. (1994), Surf. Sci., 313, p. 155.

Zhang Z. and Lagally M.G. (1997), <u>Science</u>, vol. 276, p.377.

Zubkus V.E., Kundrotas P.J., Molotkov S.N., Zatarsi

V.V., and Tornau E.E. (1991), Surf. Sci., 243, p. 295.

Further Reading

Bauer E. (1958), Z. Kristallogr., 110, p. 372.

Burton W.K., Cabrera N., Frank F.C. (1951), Philos.

Trans. R. Soc. London Ser., A 243, p. 299.

Ertl G., Küppers J. (1974) Low Energy Electrons and

Surface Chemestry, Verlag Chemie, Weinheim.

Herring C. (1953), in Gomer R. and Smith C.S.

(eds.), Structure and Properties of Solid Surfaces, p.

5, The University of Chicago Press, Chicago.

King D.A. and Woodruff D.P. (Eds.) (1981), <u>The</u> <u>Chemical Physics of Solid Surfaces and</u> <u>Heterogeneous Catalysis</u>, Elsevier, Amsterdam, vol. 1.

King D.A. and Woodruff D.P. (Eds.) (1998), <u>The</u> <u>Chemical Physics of Solid Surfaces</u>, Elsevier, Amsterdam, vol. 8.

Linford R.G. (1973), in Green M. (ed.), <u>Solid State</u> <u>Surface Science</u>, Marcel Dekker, New York, vol. 2, p. 1.

Liu Feng, Wu Fang, and Lagally M.G. (1997), Chemical Reviews, 97, p. 1045.

Mönch W. (1995), <u>Semiconductor Surfaces and</u> <u>Interfaces</u>, Berlin;New York: Springer-Verlag, 2nd ed..

Zhang Z. and Lagally M.G. (1997), <u>Science</u>, vol. 276, p.377.

Figure captions

Fig. 1. The Gibbs surface model: (a) the real system; (b) the idealized system. The thin horizontal line in the middle marks the dividing surface; the thick vertical line marks the imaginary plane P used to derive the surface stress tensor σ_{ij} .

Fig. 2. Side views ($[0\ \overline{1}\ 0]$ projection) of (001) surfaces of a simple cubic lattice. (a) A nominal surface consisting of three singular (001) terraces separated by one up and one down single-atomic-height step. (b) A vicinal surface miscut by ~ 4.8^o toward [100]. The normal (arrow) to the average surface orientation (dashed line) is [1 0 12].

Fig. 3. The five two-dimensional Bravais nets: oblique, square, hexagonal, and primitive (p) or centered (c) rectangle.

Fig. 4. Schematic top view of three possible reconstructions on Si(001). Top region: (2×1) ; middle: $p(2 \times 2)$; bottom: $c(4 \times 2)$. Open and solid circles mark atom positions; their size indicates different layers of atoms, with the largest circles in the first layer. In the middle and bottom regions, the slightly different heights of the two atoms in a dimer due to buckling is indicated by solid and open circles. The unit mesh of each reconstruction is depicted by a dark rectangle. For the $c(4 \times 2)$ structure, the primitive unit mesh is depicted by a dashed-line rhombus.

Fig. 5. Ewald sphere construction for a twodimensional system, e.g., a single layer. The incident beam is denoted by the vector k_0 and all other radial vectors ending on the Ewald sphere correspond to diffracted beams, k. (a) near normal incidence for LEED. (b) low-angle incidence for RHEED. Dashed lines indicate those diffracted beams penetrating into the bulk.

Dangling-bond configurations on {100} Fig. 6. surfaces of elemental semiconductors. (a) two dangling bonds (hybrids with a single electron) on each surface atom in an ideal bulk-terminated surface. (b) one dangling bond on each atom in a dimer. Each dangling bond is half-filled, leading to a metallic surface layer. (c) charge transfer from downward atom to upward atom in a buckled dimer. The dangling bond on the lower atom is empty and the one on the higher atom is filled, leading to a semiconducting surface layer.Fig. 7. Schematic topview of the (2×4) - β reconstruction on As-stabilized GaAs(001), induced by As vacancies. The large and small open circles denote As atoms in the first and third layers, respectively. The solid circles denote Ga atoms in the second layer. The solid rectangle marks the unit mesh.

Fig. 8. π -bonded chain model of the (2 × 1) reconstruction on {111} surfaces of solids with the diamond structure. (a) Side view of ideally terminated surface. (b) Side view of reconstructed surface. (c) Top view of reconstructed surface.

Fig. 9. Top view (top) and side view, taken along the long diagonal of the unit cell (bottom), of the dimeradatom-stacking fault (DAS) model of the Si(111)-(7 \times 7) reconstruction. The adatom layer (large shaded circles) and two double layers (open and solid circles) are indicated. The stacking fault is in the left half of the unit cell. (After Takayanagi *et al.*, 1985b)

Fig. 10. Adatom model of the Ge(111)-c(2×8) reconstruction. Large and small open circles mark the first-layer and second-layers atoms, respectively. Shaded circles are Ge adatoms sitting on top of the second-layers atoms. Numbers 1 and 2 denotes two inequivalent sites of rest atoms in the first layer. (After Becker *et al.*, 1989)

Fig. 11. Schematic atomic arrangements at {110} surfaces of zincblende-structure compound

semiconductors. Open and shaded circles denote cations and anions, respectively. (a) Top view. (b) Side view of ideal bulk-terminated surface. (c) Side view of the relaxed surface.

Fig. 12. Surfaces of fcc-crystals. (a) model of the fcccrystal. (b) (c) (d) top and side views of the (111), the (100), and the (110) surfaces, respectively.

Fig 13. Top view of W(100) surfaces. (a) unreconstructed (b) $\sqrt{2} \times \sqrt{2}$ reconstructed.

Fig. 14. Reconstruction of fcc (110) surfaces. (a) (b) top and side views of unreconstructed surface. (c) (d) top and side views of the 2x1 missing-row reconstructed surface. (e) a schematic side view of the relaxation of the Au(110)-(2 × 1) missing-row reconstruction as proposed by W. Moritz and D. Wolf (1985).

Fig 15. Reconstruction of fcc (100) surfaces. (a) (b) top and side views of the unreconstructed surface. (c) (d) top and side views of the hexagonally reconstructed surface. The hexagonal layer has a periodicity incommensurate with that of the substrate lattice.

Fig. 16. The TSK model of a surface defined for a simple cubic crystal. Open circles are substrate atoms. Shaded circles are adatoms deposited on the substrate. There are two terraces (the upper terrace is on the left) separated by a step containing a single-atom kink.

Fig. 17. Three equilibrium growth modes in heteroepitaxy.