Modeling and Simulation of Strain-mediated Nanostructure Formation on Surface

Feng Liu

Department of Materials Science & Engineering
University of Utah
Salt Lake City, Utah 84112, USA

Phone: +1 801 587 7719
Fax: +1 801 581 4816
email: fliu@eng.utah.edu

Table of Contents

§.1 Introduction
§.2 Surface Elastic Theory
- **§.2.1 Definition and Physical Origin of Surface Elastic Forces**
  - **a. Surface Stress**
  - **b. Surface Elastic Forces: force monopoles vs. force dipoles**
- **§.2.2 Interaction Energy of Surface Elastic Forces**
  - **a. The Field of Elastic Displacements Induced by Surface Forces**
  - **b. The Interaction Energy between Point Forces on Surface**
  - **c. The Interaction Energy of Two Parallel Lines of Force Monopoles**
- **§.2.3 Elastic Energy of Strained 2D Islands on Surface**
  - **a. The Elastic Energy of a 2D square Isalnd**
  - **b. The Elastic Energy of a 2D circular Island**
  - **d. The Elastic Interaction Energy between Two Islands**

§.3 Equilibrium Shape of Stressed/Strained 2D islands
- **§.3.1 Equilibrium Shape of 2D Islands and Wulff Construction**
- **§.3.2 Stress/Strain Induced Size Dependence of Equilibrium Shape**
  - **a. Spontaneous Shape Instability**
  - **b. Dependence of Islands Shape on Island Size**
- **§.3.3 Experimental Growth of Si 2D Island on Si(001)**
- **§.3.4 Equilibrium Shape of Anisotropically Strained Islands: Growth of Nanowires**

§.4 Stress/Strain Induced Self-Assembly/Self-Organization of Surface Stress Domains and Lattice of 2D islands
- **§.4.1 Spontaneous Formation of Surace Stress Domains**
  - **a. Theoretical Prediction**
  - **b. Experimental Realization and Manipulation of Surface Stress-Domain Patterns**
  - **c. Self-Organized Pattern Formation via Sterss-Mediated Surface Faceting**
- **§.4.2 Self-Assembly of 2D islands ("Quantum Platelets") at Steps**
  - **a. 2D Analogs of 3D Equilibrium Growth Modes**
  - **b. Theoretical Formulation**
  - **c. Experimental Observation**
- **§.4.3 Self-Organization of 2D islands via Strain-Mediated Coarsening**

§.5 Concluding Remarks

ISBN: 1-58883-046-2

Copyright 2006 by American Scientific Publishers
All rights of reproduction in any form reserved.

Handbook of Theoretical and Computational Nanotechnology
Edited by Michael Rieth and Wolfram Schommers
Volume 4: Pages (577-625)

§.1 Introduction

The continued miniaturization of electronic and optoelectronic devices is leading us closer to the need for nanoscale structural assemblies to perform the function of computer, laser, data storage, sensor and satellite communication. As the size of the structural feature (building block) in a device is reduced to the nanometer scale, its electronic, optical, electrical, and magnetic properties begin to differ drastically from its bulk counterparts, displaying a strong size and shape dependence due to quantum size effects arising from the reduced dimensions and surface/interface effects arising from the increased surface-to-volume ratio. The new properties, in particular those arising from quantum size effects, lead to a wide range of potential applications in electronic, optoelectronic, and magnetic devices. Quantum well devices, in which one of the dimensions of the underlying materials is reduced to the nanometer range so that carriers (electrons and/or holes) are confined in this dimension, are already widespread. Recent effort has focused on designing and fabricating devices using quantum wires and quantum dots, respectively the one- and zero-dimensional analogs of a quantum well. Just as uniform thickness required for quantum wells, a high degree of uniformity in size, shape, and spacing are often required for quantum wires and quantum dots in real applications. However, fabrication of quantum wires and quantum dots with desirable uniformity remains, so far, a challenging problem. This is because our fundamental understanding of the physical and dynamical processes at the nanometer scale is still limited and it is very difficult to design and apply fabrication and characterization tools at the nanometer scale.

Much recent effort has been devoted to the controlled fabrication of nanostructures to achieve a high degree of order and uniformity. In general, two different routes have been taken: one is the so-called top-down approach by direct surface patterning or lithography with nanometer resolution, the other is the bottom-up approach by taking advantage of self-assembly and/or self-organization of surface structures occurring naturally at the nanometer scale. Surface structures with desirable size and shape can be made *artificially* by surface patterning and lithography, either patterning the adsorbates or growing on an *a priori* patterned substrate. A beautiful example of building nano-objects with atomic precision is the creation of quantum corrals [1] and quantum mirages [2] demonstrated by Eigler et al., using atom manipulation by scanning tunneling microscopy (STM). However, such an approach may not be suitable for mass production in device fabrication, because it is inherently a time-consuming serial process.

Alternatively, nanostructures, in particular quantum wires and quantum dots, can form *naturally* during the growth of thin films by self-assembly and/or self-organization. For example, coherent three-dimensional (3D) islands (free of dislocations) form spontaneously in strained thin films [3-18]. These 3D islands (clusters) form, in general, with crystalline perfection and nanometer dimensions, lending themselves naturally to use as quantum dots. The self-assembly and self-organization approach has shown great promise for device fabrication because it offers an economical parallel process with the added advantage of being compatible with the existing Si processing technology.

Self-assembly and self-organization processes are actually very common in nature. They lead to a variety of beautiful patterns in a wide range of length scales, from the cosmic array of galaxies to the atomic structure of crystals. The length scale of a self-assembled or self-organized pattern is governed by the typical range of interaction applied by the underlying forces. For example, galaxy arrays are governed by gravitational force acting at the celestial distance, while crystal structures are governed by chemical bonds acting at the angstrom (one tenth of nanometer) range. Therefore, for self-assembly and self-organization of nanostructures, forces acting at the nanometer range have to be identified. One force that falls into the nanometer range is the elastic interaction in solid materials. Coincidentally, devices are often made of layered structures of thin films, in which elastic interaction can be a dominant force arising from the misfit strain that is inherently present due to the lattice mismatch between the different layers of materials. The study of morphological instability and ordering in strained thin films has, therefore, attracted considerable recent interest, because strain-induced self-assembly and self-organization provides an attractive "*natural*" route for nanofabrication [3,19].

A strained thin film is inherently unstable. Upon reaching the "equilibrium critical thickness" [20,21], a strained film will relax back to its equilibrium lattice constant, forming dislocations at the interface. Below the critical thickness, strain is partially relieved by surface roughening, which may manifest itself in various forms [22]. To grow a quantum well, such roughening is highly undesirable, because it prevents the growth of smooth films and introduces nucleation centers for dislocations. Consequently, ealier work had focused on understanding the morphological instability, or better yet, finding ways to suppress it. Recently, however, it has been recognized that surface/interface roughening can also lead to potentially useful structures, in particular, by strain-induced self-assembly and self-organization of roughened surface/interface structures and morphologies. This has rapidly opened up a new horizon for the research of strained thin films, the "strain engineering of nanostructures" [3,19]. The objective is to promote and take advantage of strain-induced surface/interface roughening process, occuring with a typical nanometer length scale, to engineer low-dimensional nanostructures, such as quantum dots and quantum wires on surface.

Although naturally grown structures self-assemble/self-organize with crystalline perfection and sometimes exhibit surprising uniformity [3-18], they are in general not yet sufficiently uniform in size, shape and position to consider their use in practical applications. By manipulating growth kinetics, surface/interface thermodynamics, and their dependence on *strain*, it is possible to further manipulate the size, shape, and position of the growing structural and/or morphological features to improve their uniformity. Extensive experimental and theoretical work has been done in the last decade in an attempt to stimulate, guide, and control the self-assembly and self-organization processes to improve the nanostructure uniformity. In this chapter, I will provide an overview of the progress made in the last decade on theoretical modeling and computer simulation of strain-mediated formation of nanostructures on surface, focusing on strain-induced self-assembly and self-organization of two-dimensional (2D) patterns and structures. As part of a handbook, the main objective of the chapter is not to provide an extensive literature review on the topic. Instead, I will try to provide a general introduction and overview of the basic concepts and physical models along with some relatively detailed discussion of mathematical derivations and technical treatments so that readers (especially graduate students) who are interested in this topic can use this chapter as a guide and reference to start their own modeling and simulation.

§.2 Surface Elastic Theory

In this section, we begin with a preparation of some basic surface elastic theory to be used in modeling and simulation of strain-meadiated formation of 2D nanostructures on surface. In particular, we focus on the eleastic interactions underlying different surface structures and morphologies.

§.2.1 Definition and Physical Origin of Surface Elastic Forces

a. Surface Stress

We will confine our discussion to crystaliline solid surfaces. Crystalline surfaces used in surface science research and epitaxial growth of thin films are generally low-Miller-index surfaces because of their relatively greater stability and high symmetry [23]. A real surface consists of a number of terraces of low-index surface planes of the same orientation separated by steps. Thin films can be grown on either a *nominal* surface or a *vicinal* surface. A nominal surface has its normal of the average surface orientation coinciding with the normal of the singular terraces that constitute the 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, so that it consists of a staircase of steps separating the low-index terraces (see Fig. §.2.1.1f) . It is usually specified by the indices of the constituent terraces plus the miscut angle and miscut orientation.

One intrinsic property of a solid surface is surface stress, which represents the energy cost to *deform* a solid surface, to be distinguished from the energy cost to *form* a solid surface [23].

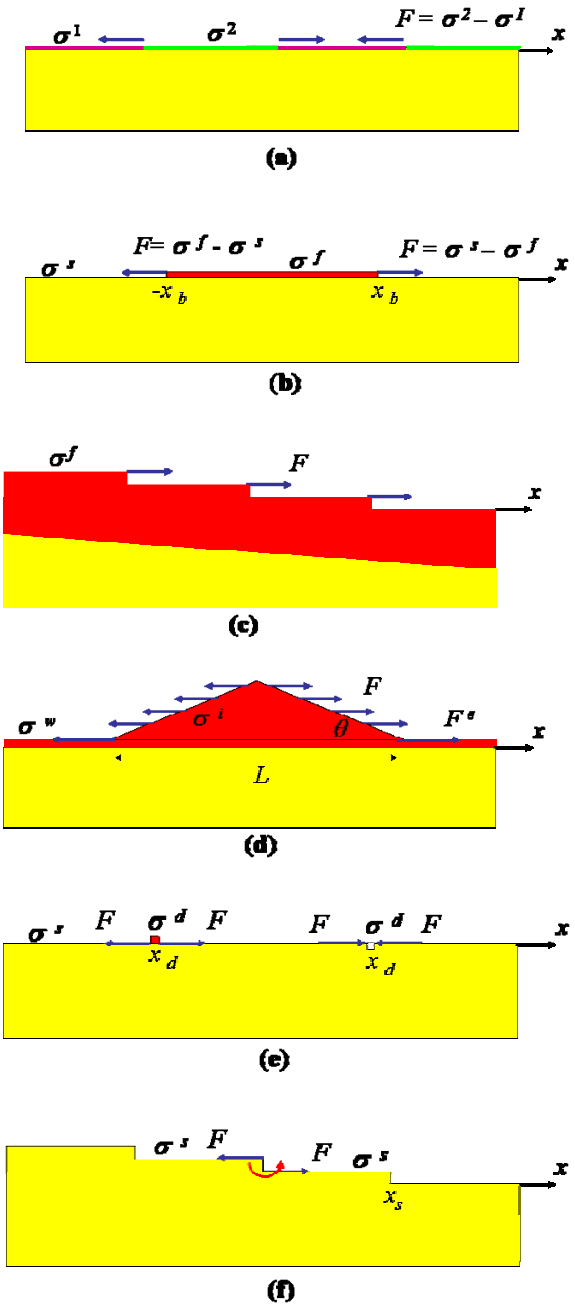

Figure §2.1.1. Schematic illustration of surface elastic forces resulted from (a) surface structure domains, (b) growth of a strain overlayer film, (c) growth of a strained film on vicinal surface, (d) SK growth of a faceted strained island, (e) line defects in the surface, and (f) a stepped surface. (a)-(d) are examples of elastic force monopoles; (e) and (f) are examples of elastic force dipoles. See text for detailed discussion.

The surface stress tensor of a solid surface is defined as

$$\sigma_{ij} = \frac{1}{A}\frac{dE}{d\varepsilon_{ij}} \qquad (\S.2.1.1a)$$

or

$$\sigma_{ij} = \gamma\delta_{ij} + \frac{d\gamma}{d\varepsilon_{ij}}, \quad (i, j = 1,2), \qquad (\S.2.1.1b)$$

where E is the surface energy of a solid surface of area A. γ is the surface tension, defined as the energy cost to create a unit area of surface, i.e., $E = \gamma A$. ε_{ij} is the two-dimensional surface strain tensor. The indices i and j denote two orthogonal directions in the plane of the surface. The Kronecker detla $\delta_{ij} = 1$ if $i = j$ and $\delta_{ij} = 0$ if $i \neq j$. For a crystal surface having a two-fold or higher symmetry, the stress tensor becomes diagonal, i.e., $\sigma_{i \neq j} = 0$. In contrast to solid surface, because there is no distinction between *forming* a liquid surface from *expanding* a liquid surface (both causing the same amount of energy), the surface stress of a one-component liquid equals its surface tension, i.e., the second term in Eq. (§.2.1.1b) vanishes [23].

b. Surface Elastic Forces: force monopoles vs. force dipoles

The surface elastic forces result from divergence of surface stress. The distribution of elastic forces in a surface can be calculated from the distribution of surface stress as [24]

$$F_i(\boldsymbol{x}) = \frac{\partial \sigma_{ij}(\boldsymbol{x})}{\partial x_j}. \qquad (\S.2.1.2)$$

Thus, an ideal homogeneous flat surface free of defects is free of elastic forces, because it has a uniform surface stress. When a surface becomes inhomogeneous, elastic forces will be introduced at the places where surface stress is discontinuous or varies with nonzero gradients. This can be caused by a variety of physical processes, such as formation of surface defects, surface adsorption, growth of an overlayer or film, and variation of surface structures and morphologies.

Figure §.2.1.1 illustrates several typical cases of elastic forces originated on the surface, which are relevant to our later discussion of the formation of nanostructures on surface. We first consider a surface consisting of alternating domains of different surface structures along the x-direction [25,26], as shown in Fig. §.2.1.1a. The difference in the surface structures may arise from different surface reconstructions (or the same reconstruction of different orientations) naturally or due to surface adsorption. Because the intrinsic surface stress tensor of each domain is different, there exists a discontinuity in surface stress at the domain boundary, giving rise to a distribution of surface elastic force (per unit length along the boundary) that can be expressed as [25,26]

$$F_x = \pm(\sigma_{xx}^1 - \sigma_{xx}^2)\,\delta(x - x_b), \qquad (\S.2.1.3)$$

where σ_{xx}^1 and σ_{xx}^2 are respectively the surface stress tensors in the domain 1 and 2, and x_b denotes the position of domain boundary.

Figure §.2.1.1b illustrates the case of epitaxial growth of an overlayer film of material B on a nominal surface of substrate of material A (heteroepitaxy [22,23]). The surface stress of substrate is σ^s, while it is σ^f in the film. In general, σ^f contains two parts: one is the intrinsic surface stress of the film surface (σ^{f_0}, being different from σ^s), and the other is misfit strain induced surface stress, which equals to the strain induced bulk stress times film thickness [27]. So, in general, $\sigma^f = \sigma^{f_0} + E\varepsilon h$, where $E\varepsilon$ is the strain induced bulk stress, E is elastic constant and ε is the misfit strain due to lattice mismatch between the film and substrate. h is the overlayer film thickness. Since the film thickness is very small, negligible compared to the thickness of substrate (a semi-infinite elastic medium), we treat the film as part of a plane surface (see discussion below in §.2.2a). Thus, the growing film introduces a single surface stress domain in the original substrate surface, similar to the case of domains of different surface structures in Fig. §.2.1.1a. Surface stress is uniform everywhere, but discontinuous at the film/substrate boundary. Consequently, surface elastic forces are introduced at the film/substrate boundary. Assuming the film forms an infinite stripe along the y-direction, with its x-boundaries at x_b and $-x_b$, as shown in Fig. §.2.1.1b, the surface elastic force (per unit length along the film boundary) can be expressed as

$$F_x = \pm(\sigma^s - \sigma^f)\,\delta(x \mp x_b). \qquad (\S.2.1.4)$$

Figure §.2.1.1c illustrates the case of epitaxial growth of an overlayer film of material B on a vicinal surface of substrate of materials A, consisting of a staircase of steps. The surface stress of the film, σ^f contains again two parts, the intrinsic surface stress of the film surface (σ^{f_0}) and the misfit strain induced surface stress (bulk stress times film thickness). The intrinsic surface stress of the film is the same everywhere, but there will still be a surface stress discontinuity at each step resulted from the misfit strain induced surface stress due to different film thickness. The surface elastic force (per unit length along the step) at each step is [27]

$$F = E\varepsilon h, \qquad (§.2.1.5)$$

i.e., the strain induced bulk stress times the step height (h). The surface stress discontinues at each step, because there is an extra layer of film being strained on the left of each step than on the right.

Figure §.2.1.1d shows a simplified 2D representation of Stranski-Krostonow (see section §.4.2a for discussion of equilibrium growth mode) growth of a faceted strained island on a surface. The surface stress of the film (island) induced by the misfit strain equals to the strain induced bulk stress times film thickness (island height), which varies continuously within the island, i.e., $\sigma^f = E\varepsilon h(x)$ [28]. Consequently, there will be a distribution of surface elastic forces on the island side facet that can be calculated as

$$F = E\varepsilon\partial_x h(x) = E\varepsilon \cdot tg\theta, \qquad (§.2.1.6)$$

as illustrated in §.2.1.1d. In addition, the intrinsic surface stress of the island facet (σ^i) is expected to be different from that of the wetting-layer surface (σ^w), due to their different surface structures. This gives rise to an additional elastic force distributed around the island edge [29], $F^e = E\varepsilon \cdot tg\theta + (\sigma^w - \sigma^i)$, as indicated in §.2.1.1d.

The type of distribution of elastic forces illustrated above in Figs. §.2.1.1a-d are also called force monopoles, as they are single forces defined at a given position. If there are two forces (monopoles) of the same magnitude pointing to the opposite direction and being separated by an infinitesimal distance at a given location, we call them together as a force dipole. A force dipole has zero total force, but may or may not have a zero total moment. Figure §.2.1.1e illustrates force dipoles induced by line defects (such as a single line of adatoms, adsorbates, or vacancies) with zero moment, whose distribution of forces can be expressed as [30]

$$F_x = (\sigma^d - \sigma^s)\frac{\partial}{\partial x}\delta(x - x_d). \qquad (§.2.1.7)$$

($\sigma^d - \sigma^s$) is the strength of the dipole given by the difference between the extra 'local" surface stress (σ^d) introduced by the defect and the intrinsic surface stress (σ^s). x_d denotes the position of the force dipoles (defects).

Figure §.2.1.1f illustrates force dipoles occurring at a surface step with nonzero moment, whose distribution of forces can be expressed as [30]

$$F_n = (\sigma^s h)\frac{\partial}{\partial x}\delta(x - x_s). \qquad (§.2.1.8)$$

σ^s is the intrinsic surface stress, h is the step height, x_s denotes the position of the steps. In addition, at every step, there will be also a force dipole of the type with zero moment, as expressed in Eqn. §.2.1.7, due to local structural distortion (or reconstruction) along the step edge. Comparing Fig. §.2.1.1c with Fig. §.2.1.1f, one sees that (strain induced) bulk stress induces a force monopole at a surface step, while (intrinsic) surface stress induces a force dipole at a surface step [27]. Because there are always local structural distortions associated with steps and surfaces always have nonzero surface stress, force dipoles are inherently present at a step. Therefore, in addition to the force monopoles indicated at each step in Figs. §.2.1.1b and c, there will be also force dipoles at each step.

§.2.2 Interaction Energy of Surface Elastic Forces

a. The Field of Elastic Displacements Induced by Surface Forces

We consider the growth of thin film on surface of an infinite thick substrate, which can be treated as a semi-infinite elastic medium bounded by a plane (surface). Since the film thickness and its variation are very small, negligible compared to the thickness of substrate, we neglect the height differences due to the structural and morphological variations in the surface and film, and treat the real surface of the film as part of a plane surface. In other words, all the elastic forces discussed above in Figs. §.2.1.1a-f are approximately considered to be lying in the plane surface of a semi-infinite elastic medium. This allows us to use the existing solutions of continuum theory of the elastic field in a semi-infinite medium induced by a point force on the surface [24,25]. Since in most cases, we are dealing with the interaction between the surface elastic forces, we need only the elastic field in the surface induced by the surface forces, which mediates their interaction. Throughout this chapter, we use bold letters to denote vectors. The vector field of displacement ($\boldsymbol{u}(\boldsymbol{r}) = u_x\boldsymbol{i} + u_y\boldsymbol{j} + u_z\boldsymbol{k}$, $\boldsymbol{r} = x\boldsymbol{i} + y\boldsymbol{j}$) in the surface induced by a single force ($\boldsymbol{F} = F_x\boldsymbol{i} + F_y\boldsymbol{j} + F_z\boldsymbol{k}$) at the center of the surface (chosen as the origin of the coordinates) is given as [24]

$$u_x = \frac{1+\nu}{2\pi Y}\{\frac{2(1-\nu)}{r}F_x + \frac{2\nu x}{r^3}(xF_x + yF_y) - \frac{(1-2\nu)x}{r^2}F_z\}, \qquad (\S.2.2.1a)$$

$$u_y = \frac{1+\nu}{2\pi Y}\{\frac{2(1-\nu)}{r}F_y + \frac{2\nu y}{r^3}(xF_x + yF_y) - \frac{(1-2\nu)y}{r^2}F_z\}, \qquad (\S.2.2.1b)$$

and

$$u_z = \frac{1+\nu}{2\pi Y}\{\frac{2(1-\nu)}{r}F_z + \frac{(1-2\nu)}{r^2}(xF_x + xF_y)\}. \qquad (\S.2.2.1c)$$

ν and Y are respectively Poisson's ratio and Young's modulus. If the elastic forces have only the in-plane components (i.e., $F_z = 0$), as the cases shown in Figs. §.2.1.1a-f, then only the in-plane components of the displacement is need, which can be written in a compact form as

$$\boldsymbol{u}(\boldsymbol{r}) = \frac{1-\nu^2}{\pi Y}\frac{1}{r}\boldsymbol{F} + \frac{\nu(1+\nu)}{\pi Y}\frac{\boldsymbol{r}}{r^3}(\boldsymbol{F}\cdot\boldsymbol{r}). \qquad (\S.2.2.2)$$

Physically, the displacement vector at $\boldsymbol{r}$ induced by a force at the origin has two components: one in the direction of the applied force $\boldsymbol{F}$ (the first term, proportional to $1-\nu^2$) and the other in the direction of the distance $\boldsymbol{r}$ [the second term, proportional to $\nu(1+\nu)$].

b. The Interaction Energy between Point Forces on Surface

When a point force $\boldsymbol{F}$ is applied at $\boldsymbol{r}$, it induces a displacement $\boldsymbol{u}[\boldsymbol{r}', \boldsymbol{F}(\boldsymbol{r})] = G(\boldsymbol{r}', \boldsymbol{r})\boldsymbol{F}(\boldsymbol{r})$ at $\boldsymbol{r}'$, where $G(\boldsymbol{r}', \boldsymbol{r})$ is called the elastic Green's function, given by Eqn. §.2.2.2. Now, if another force $\boldsymbol{F}'$ is applied at $\boldsymbol{r}'$, it will interact with the local displacement field induced by $\boldsymbol{F}$, doing the amount work $\boldsymbol{F}'(\boldsymbol{r}')\cdot\boldsymbol{u}[\boldsymbol{r}', \boldsymbol{F}(\boldsymbol{r})]$, i.e., work = force times displacement. In addition, $\boldsymbol{F}'$ at $\boldsymbol{r}'$ will also induce displacement at $\boldsymbol{r}$ interacting with force $\boldsymbol{F}$. Thus, the interacting energy between the two point forces on the surface can be calculated as

$$\begin{aligned} E &= -\frac{1}{2}\{\boldsymbol{F}'(\boldsymbol{r}')\cdot\boldsymbol{u}[\boldsymbol{r}', \boldsymbol{F}(\boldsymbol{r})] + \boldsymbol{F}(\boldsymbol{r})\cdot\boldsymbol{u}[\boldsymbol{r}, \boldsymbol{F}'(\boldsymbol{r}')]\} \\ &= -\frac{1+\nu}{\pi Y}\{(1-\nu)\frac{\boldsymbol{F}\cdot\boldsymbol{F}'}{|\boldsymbol{r}'-\boldsymbol{r}|} + \nu\frac{[\boldsymbol{F}\cdot(\boldsymbol{r}'-\boldsymbol{r})][\boldsymbol{F}'\cdot(\boldsymbol{r}'-\boldsymbol{r})]}{|\boldsymbol{r}'-\boldsymbol{r}|^3}\} \end{aligned} \qquad (\S.2.2.3)$$

Note that there are two terms in Eqn. §.2.2.3: the first arising from the force-induced displacement in the direction of the applied force and the second from the displacement in the direction of distance vector between the applied forces, $\boldsymbol{r}' - \boldsymbol{r}$. Consequently, the interaction

between two forces depends sensitively on their directions. Figure §.2.2.1a illustrates the interaction between several different point force configurations, which we discuss below.

(1) Two parallel forces lining along their distance vector. In this case, both terms in Eqn. §.2.2.3 contribute to the interaction energy:

$$E = -\frac{(1+\nu)}{\pi Y}\frac{F^2}{d} \qquad \text{(§.2.2.4a)}$$

(2) Two parallel forces lining perpendicular to their distance vector. In this case, because $\boldsymbol{F}\cdot(\boldsymbol{r}'-\boldsymbol{r})=0$ only the first terms in Eqn. §.2.2.3 contributes to the interaction energy:

$$E = -\frac{(1-\nu^2)}{\pi Y}\frac{F^2}{d} \qquad \text{(§.2.2.4b)}$$

(3) Two orthogonal forces with one lining along their distance vector. In this case, because $\boldsymbol{F}\cdot\boldsymbol{F}'=0$ and $\boldsymbol{F}\cdot(\boldsymbol{r}'-\boldsymbol{r})=0$ both terms in Eqn. §.2.2.3 vanishes and the interaction energy is zero.

(4) The above cases in (1)-(3) describe the interaction between force monopoles (i.e., single point forces). If there are pairs of point forces of the same magnitude pointing at the opposite direction that are separated by an infinitesimal distance, a, at the same location, they constitute force dipoles, as shown in Fig. §.2.2.1a(4). The interaction energy between the dipoles can be calculated by summing up all the monopole-monopole interactions as

$$\begin{aligned}
E = &-\frac{(1+\nu)F^2}{\pi Y}\left(\frac{2}{d}-\frac{1}{d+a}-\frac{1}{d-a}\right)-\frac{(1-\nu^2)F^2}{\pi Y}\left(\frac{2}{d}-\frac{2}{(d^2+a^2)^{1/2}}\right)\\
&-\frac{(1+\nu)\nu F^2}{\pi Y}\left(\frac{2}{d}-\frac{1}{d+a}-\frac{1}{d-a}\right)-\frac{(1+\nu)\nu F^2}{\pi Y}\frac{2a^2}{(d^2+a^2)^{3/2}}\\
&-\frac{(1+\nu)\nu F^2}{\pi Y}\frac{-4\cdot(d+a/2)(a/2)}{[(d+a/2)^2+(a/2)^2)]^{3/2}}\\
&-\frac{(1+\nu)\nu F^2}{\pi Y}\frac{4\cdot(d-a/2)(a/2)}{[(d-a/2)^2+(a/2)^2)]^{3/2}}\\
=&\frac{(1-\nu)^2F^2}{\pi Y}\frac{a^2}{d^3}+O(d^{-3})
\end{aligned} \qquad \text{(§.2.2.4c)}$$

From Equations. §.2.2.4a-b, we see that the interaction between two like (parallel pointing at the same direction) force *monopoles* is *attractive* (the energy decreases with decreasing distance). Conversely, the interaction between two dislike (anti-parallel pointing at the opposite directions) force monopoles is repulsive. From Eqns. §.2.2.4c, we see that the interaction between two like force *dipoles* [both with forces pointing outward as shown in Fig. §.2.2.1a(4)] is *repulsive*. Conversely, the interaction between two dislike force dipoles (one with forces pointing outward and the other with forces pointing inward) is attractive. The dipole-dipole interaction has been often studied in the context of surface adsorbates and point defects [31].

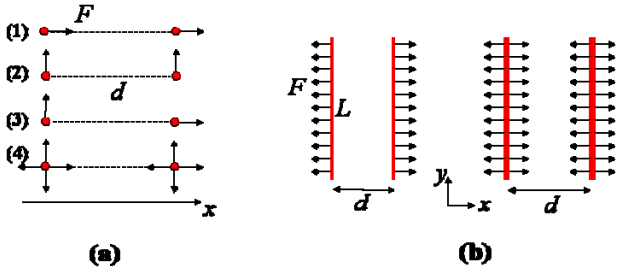

Figure §2.2.1. Schematic illustration of interaction between surface elastic forces. (a) Interaction between point force monopoles {(1), (2), (3)} and dipoles {(4)}. (b) Interaction between line force monopoles (left) and dipoles (right).

For example, an adatom may induce a "positive" dipole by introducing outward forces, while a vacancy may induce a "negative" dipole by introducing inward forces. If so, then an adatom will repel another adatom but attract a vacancy, and a vacancy will repel another vacancy but attract an adatom.

c. The Interaction Energy of two Parallel Lines of Force Monopoles

If there exists a continuous distribution of surface elastic forces, their interaction energy is calculated by integration over the force distribution as

$$E = -\frac{1}{2}\iint \boldsymbol{F}'(\boldsymbol{r}')\cdot\ \boldsymbol{u}[\boldsymbol{r}',\boldsymbol{F}(\boldsymbol{r})]\, d\boldsymbol{r}d\boldsymbol{r}'$$
$$= -\frac{1}{2}\iint \{\frac{1-\nu^2}{\pi Y}\frac{\boldsymbol{F}'\cdot\boldsymbol{F}}{|\boldsymbol{r}'-\boldsymbol{r}|}+\frac{\nu(1+\nu)}{\pi Y}\frac{[\boldsymbol{F}\cdot(\boldsymbol{r}'-\boldsymbol{r})][\boldsymbol{F}'\cdot(\boldsymbol{r}'-\boldsymbol{r})]}{|\boldsymbol{r}'-\boldsymbol{r}|^3}\} d\boldsymbol{r}d\boldsymbol{r}'. \quad (§.2.2.5)$$

One simple and important form of force distribution is the line distribution, which occurs at stress domain boundaries, edges of strained films and surface steps, as shown in Fig. §.2.1.1. Below, we calculate the interaction energy between two parallel line distributions of force monopoles pointing at opposite directions, as shown in §.2.2.1a. They may represent the interaction between two parallel domain boundaries (Fig. §.2.1.1a), between two opposite edges of a square or rectangular islands (Figs. §.2.1.1b and 2d), or between two steps (Fig. §.2.1.1c).

In general, we consider two lines of length L separated by a distance d, with a uniform distribution of force F (per unit length along the line), as shown in Fig. §.2.2.1b (left panel). The total elastic interaction energy includes two contributions: the self-interaction among monopoles on the same line and the cross-interaction between monopoles on the different lines, which can be calculated as

$$E_{elastic} = 2(E_{self} + E_{cross}) \quad (§.2.2.6a)$$

$$E_{self} = -\frac{(1-\nu^2)F^2}{2\pi Y}\{\int_0^{L-a_0} dy_1 \int_{y_1+a_0}^{L}\frac{dy_2}{y_2-y_1}+\int_{a_0}^{L} dy_1\int_0^{y_1-a_0}\frac{dy_2}{y_1-y_2}\}$$
$$= -\frac{(1-\nu^2)F^2}{\pi Y}[L\ln\frac{L}{ea_0}+a_0] \quad (§.2.2.6b)$$

$$E_{cross} = -\frac{1+\nu}{2\pi Y}\int_0^L dy_1\int_0^L \{\frac{(1-\nu)(-F^2)}{[(y_2-y_1)^2+d^2]^{1/2}}+\frac{\nu(-F^2d^2)}{[(y_2-y_1)^2+d^2]^{3/2}}\} dy_2$$
$$= -\frac{(1-\nu^2)F^2}{2\pi Y}[-2d+2\sqrt{L^2+d^2}+L\ln\frac{\sqrt{L^2+d^2}-L}{\sqrt{L^2+d^2}+L}] \quad (§.2.2.6c)$$
$$-\frac{(1+\nu)\nu F^2}{\pi Y}[d-\sqrt{L^2+d^2}]$$

In calculating self-interaction energy of forces on the same line, the forces are orthogonal to their displacement vector, i.e., $\boldsymbol{F}\cdot\boldsymbol{r}\equiv 0$, so it contains only the term proportional to $(1-\nu^2)$ from the first term of Eqn. §.2.2.3. In contrast, the cross-interaction energy has both the $(1-\nu^2)$ and the $\nu(1+\nu)$ terms. Also, integration of §.2.2.6b requires that $|y_2-y_1|\geq a_0$, or it diverges. a_0 is a cutoff length (usually in the range of lattice constant), below which the continuum elastic theory is no longer applicable. Similarly, integration of §.2.2.6c is carried out for $d\geq a_0$.

The total elastic interaction energy is

$$E_{elastic} = -\frac{2(1+\nu)F^2}{\pi Y}\{(1-\nu)[L\ln\frac{L}{a_0}+\frac{L}{2}\ln\frac{\sqrt{L^2+d^2}-L}{\sqrt{L^2+d^2}+L}-(L+d-a_0)+\sqrt{L^2+d^2}] \quad (§.2.2.5)$$
$$+\nu[d-\sqrt{L^2+d^2}]\}$$

The elastic interaction energy *per unit length* between two *infinite* long lines of force monopoles pointing at the opposite direction is

$$E_\infty = \lim_{L\to\infty} \frac{E_{elastic}}{2L} = -\frac{(1-\nu^2)F^2}{\pi Y} \ln \frac{d}{a_r}. \qquad (§.2.2.6)$$

where $a_r = (2e^{\nu/1-\nu})a_0$ is the renormalized cut off length. This interaction is repulsive, as the energy increases with decreasing separation, *d*. The repulsive force between them is

$$f = -\frac{\partial E_\infty}{\partial d} = \frac{(1-\nu^2)F^2}{\pi Y}\frac{1}{d}. \qquad (§.2.2.7)$$

Note that we use the upper case (*F*) to denote elastic force monopole and the lower case (*f*) to denote the interactive force between force monopoles (i.e., the energy derivative of force monopole-monopole interaction). Conversely, the interaction between the two lines of force monopoles pointing at the same direction is attractive. For example, the interaction between the two edges of a single-layer strained film in Fig. §.2.1.1b is repulsive, while the interaction between the steps in a vicinal strained film in Fig. §.2.1.1c is attractive.

From Eqn. §.2.2.6, we can easily calculate the interaction energy between two infinite long lines of force dipoles, as shown in Fig. §.2.2.1b (right panel), by summing over monopole-monopole interactions:

$$\begin{aligned} E &= -\frac{(1-\nu^2)F^2}{\pi Y}\{-2\ln\frac{d}{2a_0} + \ln\frac{d+a}{2a_0} + \ln\frac{d-a}{2a_0}\} \\ &= \frac{(1-\nu^2)}{\pi Y}\frac{(Fa)^2}{d^2} \end{aligned} \qquad (§.2.2.8)$$

This interaction between two like line dipoles is repulsive; the repulsive force between them is

$$F = \frac{2(1-\nu^2)}{\pi Y}\frac{(Fa)^2}{d^3}. \qquad (§.2.2.9)$$

Effectively, an infinite long line force in a two-dimensional (2D) surface can be considered as a point force in one 1D. The monopole-monopole interaction is long-ranged and scales as ln*d* in 1D (Eqn. §.2.2.6) and as 1/*d* in 2D (Eqn. §.2.2.4a,b), while the dipole-dipole interaction is short-ranged and scales as $1/d^2$ in 1D (Eqn. §.2.2.8) and as $1/d^3$ in 2D (Eqn. §.2.2.4c). The interaction is attractive (repulsive) between two like (dislike) monopoles but repulsive (attractive) between two like (dislike) dipoles.

§.2.3 Elastic Energy of Strained 2D Islands on Surface

a. The Elastic Energy of a 2D Square Island

Epitaxial growth starts with nucleation and growth of 2D islands of one-atomic-layer (monolayer) height in the submonolayer regime. The 2D island is generally strained in heteroepitaxy duo the lattice mismatch between the overlayer film and the substrate. The strain induces force monopoles, $F = E\varepsilon h$, around the island boundary, as described in Fig. §.2.1.1c. Thus, elastic energy of a strained 2D island can be calculated by integrating the monopole-monopole interaction around the island boundary. Let's consider a square island of side length *L*, as shown in Fig. 2.3.1a. The interaction energy between the two opposite sides can be calculated using Eqn. §.2.2.5 by setting *d*=*L*, i.e., the interaction between two line forces of length *L* separated by a distance *L*,

$$E_{\parallel} = -\frac{2(1+\nu)F^2}{\pi Y}\{(1-\nu)[L\ln\frac{L}{a_0} + \frac{L}{2}\ln\frac{\sqrt{2}-1}{\sqrt{2}+1} - (2L-a_0) + \sqrt{2}L] + \nu(1-\sqrt{2})L\} \quad (§.2.3.1)$$

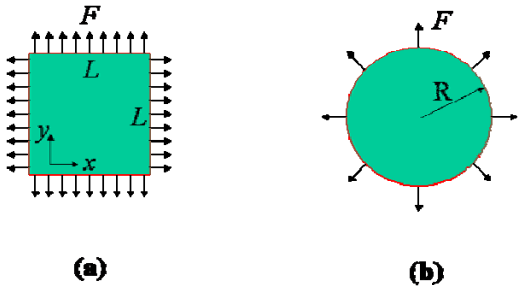

Figure §2.3.1 Schematic illustration of strained 2D islands with elastic force monopoles distributed along the island boundary. (a) A square island of size L x L. (b) A circular island of radius R.

The interaction between two neighboring sides can be calculated as

$$E_{\perp} = -\int_0^L \int_0^L \frac{1+\nu}{\pi Y}\left[\frac{\nu(-F^2 xy)}{(x^2+y^2)^{3/2}}\right]dxdy = \frac{(1+\nu)\nu F^2}{2\pi Y}\left[(2-\sqrt{2})L - \frac{a_0}{2}\right] \quad (\S.2.3.2)$$

Note that because the forces on the neighboring sides are perpendicular to each other, so only the second term [proportional to $\nu(1+\nu)$] in Eqn. §.2.2.2 left in Eqn. §.2.3.1. Also, integration of §.2.3.1 requires that $\sqrt{x^2+y^2} \geq a_0$, which gives an additional term proportional to a_0, reflecting the corner effect. The total elastic energy of the square island is

$$\begin{aligned} E &= 2E_{\parallel} + 4E_{\perp} \\ &= -\frac{4(1+\nu)F^2}{\pi Y}\{(1-\nu)[L\ln\frac{L}{a_0} + \frac{L}{2}\ln\frac{\sqrt{2}-1}{\sqrt{2}+1} - (2L-a_0) + \sqrt{2}L] + \nu(1-\sqrt{2})L\} \\ &\quad + \frac{2(1+\nu)\nu F^2}{\pi Y}\left[(2-\sqrt{2})L - \frac{a_0}{2}\right] \\ &= -\frac{(1-\nu^2)F^2}{\pi Y}[(4L)\ln\frac{L}{a_r} + LO(a_0)] \end{aligned} \quad ,(\S.2.3.3)$$

where the renormalized cutoff length, $a_r = [(\sqrt{2}+1)e^{g(\nu)}]a_0, g(\nu) = (2-\sqrt{2}+\frac{\sqrt{2}\nu}{2-2\nu})$. The corner-effect term $O(a_0) = \frac{4-3\nu}{4-4\nu}a_0$, which can be generally neglected [32] because $L >> a_0$.

Now, for island formation in real growth process, we need to consider the total energy of the island. For a square island of size $L \times L$, it can be expressed as

$$E_t = (4L)[\gamma - C\ln\frac{L}{a_r}], \quad (\S.2.3.4)$$

where γ is the island boundary (step) formation energy per unit length and $C = \frac{(1-\nu^2)F^2}{\pi Y}$. The energy per area is then

$$\frac{E_t}{L^2} = 4[\frac{\gamma}{L} - \frac{C}{L}\ln\frac{L}{a_r}], \quad (\S.2.3.5)$$

which has a minimum at $L_0 = a_r e^{(C/\gamma)+1}$. This has been used in some theoretical analyses on island stability (see section §.4.2b) . Another useful quantity is the island chemical potential, which governs island growth during coarsening process. It can be readily calculated as

$$\mu = \frac{dE_t}{dN} = 2s[\frac{\gamma - C}{L} - \frac{C}{L}\ln\frac{L}{a_r}], \quad (\S.2.3.6)$$

where $N=(L^2)/s$ is the total number of atoms in the island and *s* is the area occupied by one atom. The stable island size corresponding to the minim chemical potential is then given by $L_0 = a_r e^{(C/\gamma)}$.

b. The Elastic Energy of a 2D Circular Island

Next, we consider a circular island of radius *R*, as shown in Fig. 2.3.1b. Its elastic energy can be calculated by integration around its boundary as

$$E=-\frac{1+\nu}{2\pi Y}\int\int[\frac{1-\nu}{r}\boldsymbol{F}_1\cdot\boldsymbol{F}_2+\frac{\nu}{r^3}(\boldsymbol{F}_1\cdot\boldsymbol{r})(\boldsymbol{F}_2\cdot\boldsymbol{r})]ds_1 ds_2 . \qquad (§.2.3.7)$$

Using the polar coordinate system (R,θ), we have

$$r=2R\left|\sin\frac{\theta_2-\theta_1}{2}\right|, \qquad (§.2.3.8a)$$

$$\boldsymbol{F}_1\cdot\boldsymbol{F}_2=F^2\cos(\theta_2-\theta_1)=F^2(1-2\sin^2\frac{\theta_2-\theta_1}{2}), \qquad (§.2.3.8b)$$

$$\boldsymbol{F}_1\cdot\boldsymbol{r}=-2FR\sin^2\frac{\theta_2-\theta_1}{2},\ \boldsymbol{F}_2\cdot\boldsymbol{r}=2FR\sin^2\frac{\theta_2-\theta_1}{2}, \qquad (§.2.3.8c)$$

$$ds_1=Rd\theta_1,\ ds_2=Rd\theta_2 . \qquad (§.2.3.8d)$$

So,

$$E=-\frac{(1+\nu)F^2}{4\pi Y}\int_0^{2\pi}\int_{\theta_1+\theta_0}^{\theta_1+2\pi-\theta_0}[(1-\nu)(\frac{1}{\sin\frac{\theta_2-\theta_1}{2}}-2\sin\frac{\theta_2-\theta_1}{2})-\nu\sin\frac{\theta_2-\theta_1}{2}]d\theta_1 d\theta_2$$
$$=-\frac{(1-\nu^2)F^2}{\pi Y}(2\pi R)\ln\frac{R}{a_r} \qquad (§.2.3.8)$$

where the renormalized cutoff length, $a_r=(\frac{1}{4}e^{g(\nu)})a_0$, $g(\nu)=\frac{2-\nu}{1-\nu}$, $\theta_0=\frac{a_0}{R}$.

Similar to a square island, the total energy of a circular island can be expressed as

$$E_t=(2\pi R)[\gamma-C\ln\frac{R}{a_r}], \qquad (§.2.3.10)$$

where γ is again the island boundary (step) formation energy per unit length and $C=\frac{(1-\nu^2)F^2}{\pi Y}$. The energy per area is then

$$\frac{E_t}{\pi R^2}=2[\frac{\gamma}{R}-\frac{C}{R}\ln\frac{R}{a_r}], \qquad (§.2.3.11)$$

which has a minimum at $R_m=a_r e^{(C/\gamma)+1}$. The island chemical potential, which will be used later in section §.4.3, can be readily calculated as

$$\mu=\frac{dE_t}{dN}=s[\frac{\gamma-C}{R}-\frac{C}{R}\ln\frac{R}{a_r}] , \qquad (§.2.3.12)$$

where $N=(\pi R^2)/s$ is the total number of atoms in the island and s is the area occupied by one atom. The stable island size corresponding to the minim chemical potential is then given by $R_0=a_r e^{(C/\gamma)}$.

c. The Elastic Interaction Energy between Two Islands

We now consider the interaction between two islands as shown in Fig. §.2.3.2. For simplicity, we assume the size of islands is much smaller than their separation, i.e. $L_{1,2} << D$ for square islands in Fig. §.2.3.2a and $R_{1,2} << D$ for circular islands in Fig. §.2.3.2a. In this case, their interaction can be treated as interaction between two concentrated distributions of dipoles. By scaling the force F (per unit length) to L_1F and L_2F, similar to derivation of Eqn. §.2.2.4c, we sum up all the interactions between the line distributions of monopoles and obtain the interaction energy between two square islands (Fig. §.2.3.2a) as

$$E = \frac{(1-\nu)^2 F^2}{\pi Y} \frac{L_1^2 L_2^2}{D^3}. \qquad (§.2.3.13)$$

For calculating the interaction energy between two circular islands, we decompose the force monopole along the island boundary into x- and y-components, as shown in Fig. §.2.3.2b, and sum up all the pair interactions between the monopole components as

$$\begin{aligned}
E = &-\frac{(1-\nu^2)}{\pi Y}\int_0^{2\pi}\int_0^{2\pi} d\theta_1 d\theta_2 \left[\frac{F^2 R_1 R_2 \cos\theta_1 \cos\theta_2}{r}\right] \\
&-\frac{(1-\nu^2)}{\pi Y}\int_0^{2\pi}\int_0^{2\pi} d\theta_1 d\theta_2 \left[\frac{F^2 R_1 R_2 \sin\theta_1 \sin\theta_2}{r}\right] \\
&-\frac{(1+\nu)\nu}{\pi Y}\int_0^{2\pi}\int_0^{2\pi} d\theta_1 d\theta_2 \left[\frac{(F^2 R_1 R_2 \cos\theta_1 \cos\theta_2) r_x^2}{r^3}\right] \\
&-\frac{(1+\nu)\nu}{\pi Y}\int_0^{2\pi}\int_0^{2\pi} d\theta_1 d\theta_2 \left[\frac{(F^2 R_1 R_2 \sin\theta_1 \sin\theta_2) r_y^2}{r^3}\right] \\
&-\frac{(1+\nu)\nu}{\pi Y}\int_0^{2\pi}\int_0^{2\pi} d\theta_1 d\theta_2 \left[\frac{(F^2 R_1 R_2 \cos\theta_1 \sin\theta_2) r_x r_y}{r^3}\right] \\
&-\frac{(1+\nu)\nu}{\pi Y}\int_0^{2\pi}\int_0^{2\pi} d\theta_1 d\theta_2 \left[\frac{(F^2 R_1 R_2 \sin\theta_1 \cos\theta_2) r_x r_y}{r^3}\right] \\
&= \frac{(1-\nu^2) F^2}{\pi Y} \frac{(\pi R_1^2)(\pi R_2^2)}{D^3}
\end{aligned} \qquad (§.2.3.14)$$

where

$$r = [(D + R_2\cos\theta_2 - R_1\cos\theta_1)^2 + (R_2\sin\theta_2 - R_1\sin\theta_1)^2]^{1/2} \qquad (§.2.3.15a)$$

$$r_x = (D + R_2\cos\theta_2 - R_1\cos\theta_1) \qquad (§.2.3.15b)$$

$$r_y = (R_2\sin\theta_2 - R_1\sin\theta_1) \qquad (§.2.3.15c)$$

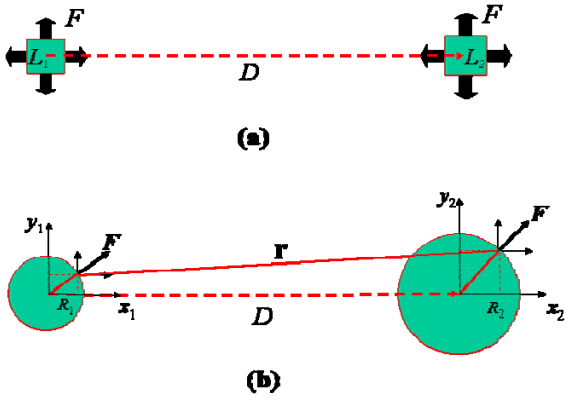

Fig. §.2.3.2. Schematic illustration of two interacting strained 2D islands: (a) two square islands; (b) two circular islands.

From equations (§.2.3.13) and (§.2.3.14), we can see that generally the island-island interaction is proportional to the product of their area and inversely proportional to their separation to the third power, as $E \sim \frac{S_1 S_2}{D^3}$, where S_1 and S_2 are respectively the area of island 1 and 2.

§.3 Equilibrium Shape of Stressed/Strained 2D islands

The properties of two-dimensional (2D) islands play a very important role in the study of epitaxial growth. Much of our understanding of the fundamental mechanisms of film growth originates from investigations of 2D islands at the very early stage of epitaxial growth. Both kinetic and thermodynamic mechanisms can be determined. Examples of kinetic parameters include the surface diffusion coefficient, extracted from the nucleation number density of 2D islands as a function of growth temperature [33]; the anisotropy in surface diffusion [34] and in adatom sticking to island edges [35], inferred from the anisotropic growth shape of islands; and the kinetics of island-edge diffusion [36] and corner-crossing [37], derived from the shape and compactness of the islands. Examples of thermodynamic parameters include step energies, determined from equilibrium island shapes and their thermal fluctuations [38-41]. Most early studies have ignored the effect of stress/strain on 2D island growth and morphology. Here, we review some recent theoretical studies [32,42,43] that have provided detailed analysis of the effect of stress/strain on equilibrium shape of 2D islands, from which additional thermodynamic parameters of surface stress can be derived.

§.3.1 Equilibrium Shape of 2D islands and Wulff Construction

A 2D island on a surface is bounded by surface steps of different orientations. The ratio of step free energies is commonly believed to define the equilibrium shape of a 2D island, which can be conveniently constructed by making a Wullf plot [44-46]. (This is in analogous to Wulff construction for constructing the equilibrium shape of a crystallite based on the ratio of surface free energies of different facets.) Let's first illustrate a simple example of equilibrium shape a rectangle island with length a and width b, as shown in Fig. §.3.1.1a.

Let E_a and E_b be the step free energy of the step oriented in the x and y direction, respectively. (The step orientation is denoted by its normal direction, in analogous to surface orientation denoted by the surface normal direction.) The total island step (boundary) energy is then

$$E_s = 2aE_a + 2bE_b . \qquad (§.3.1.1)$$

For a given island area $A = ab$, it can be readily shown that minimization of total step energy E_s with respect to island boundary length leads to the equilibrium island shape defined by

$$\frac{a}{b} = \frac{E_b}{E_a} . \qquad (§.3.1.2)$$

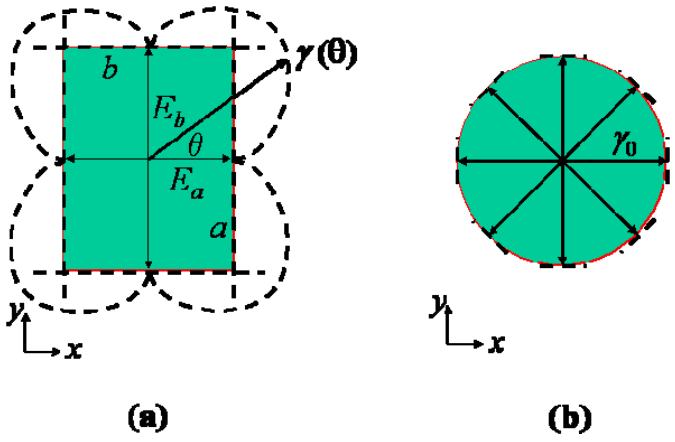

Fig. §.3.1.1 Shematic illustration of graphical Wulff construction of 2D equilibrium island shape. (a) A rectangular island constructed with highly anistropic step energy. (b) A circle island constructed with isotropic step energy $\gamma(\theta)$= constant.

Thus, the aspect ratio of the equilibrium rectangular island scales inversely with the ratio of step energy: the lower the step energy is, the longer is its length. Physically, this simply means that the island minimizes its energy by taking the step of lower energy with longer length. Alternatively, this can be viewed as that *for the equilibrium island shape, the distances of the steps to the center of the island are proportional to their step energies per unit length*, which forms the basis of Wulff construction [45]. Thus, to graphically construct the equilibrium rectangular island, one first draws two vectors of length E_a (step energy in the x direction) starting from the center in the $\pm x$ direction and two vectors of length E_b (step energy in the y direction) starting from the center in the $\pm y$ direction, as shown in Fig. §.3.1.1a, and then at the end of each vector draws a tangent line perpendicular to each vector (dashed lines in Fig. §.3.1.1a). The equilibrium shape is simply the area enclosed by those innermost tangent lines connected together.

The Wulff construction (plot) is also called $\gamma-$plot. In general, the orientational dependence of step energies can be expressed as $\gamma(\theta)$, with $0 \le \theta < 2\pi$ measured from the x-axis. The equilibrium island shape is constructed by drawing vectors of length $\gamma(\theta)$ along the direction of angle θ and then connecting the end tangent lines of all the vectors. For a rectangular island, it implies that step energies in all other orientations are much higher than those two in the x- and y-direction, i.e., $\gamma(\theta) >> E_a$; $\gamma(\theta) >> E_b$ for $\theta \neq 0, \pi/2, \pi,$ and $3\pi/2$, as indicated in Fig. §.3.1.1a (dot lines). These step orientations of high energy can't appear in the equilibrium shape are called forbidden angles, as discussed below. If $\gamma(\theta) = \gamma_0$ (i.e., the step energies are isotropic, same in all directions), then the equilibrium island shape is simply a circle, as constructed in Fig. §.3.1.1b.

For a crystal surface of N-fold symmetry, the step energies may take the following general form:

$$\gamma(\theta) = \gamma_0[1 + \varepsilon_N \cos(N\theta)], \qquad (§.3.1.3)$$

where $\varepsilon_N = [0,1)$ defines the degree of step energy anisotropy. In principle, one may always manually construct the equilibrium island shape as described above, but it would be rather inconvenient with a general form of $\gamma(\theta)$. Instead, one can draw the equilibrium island shape using the following analytical solution of the island boundary coordinates (x, y) at angle θ:

$$x = \gamma\cos\theta - \gamma_\theta \sin\theta, \qquad (§.3.1.4a)$$

$$y = \gamma\sin\theta - \gamma_\theta \cos\theta, \qquad (§.3.1.4b)$$

where γ_θ denotes the first derivative of γ with respect to θ. This solution is obtained by solving the Gibbs-Thomson equation, which expresses the chemical potential along the island boundary in terms of boundary (step) energy and curvature [45,46].

For $\varepsilon_N = 0$, $\gamma(\theta) = \gamma_0$, the equation (§.3.1.4) gives $x = \gamma_0 \cos\theta$ and $y = \gamma_0 \sin\theta$ corresponding to a circular equilibrium island shape, consistent with the graphical construction in Fig. §.3.1.1b. A circular island boundary consists evenly of steps of all orientations, because they all have the same energy for $\varepsilon_N = 0$ (isotropic step energy). As ε_N increases, the step energy becomes increasingly anisotropic, so does the island shape. If ε_N is sufficiently large, however, the step-energy anisotropy may become so large that steps in certain high-energy orientations will be excluded from the equilibrium island shape. These metastable and unstable step orientations, which are absent in the equilibrium shape, are so called forbidden angles in the Wulff plot.

The forbidden angles occur when $\varepsilon_N > 1/(N^2 - 1)$ [46]. For example, consider a 2D island on a crystal surface of two-fold symmetry with $\gamma(\theta) = \gamma_0[1 + \varepsilon_2 \cos(2\theta)]$, the forbidden angles will occur when $\varepsilon_2 > 1/3$. Figure §.3.1.2 shows the two equilibrium island shape calculated from Equation (§.3.1.4) using $\varepsilon_2 = 1/5$ and $\varepsilon_2 = 1/2$, respectively. For $\varepsilon_2 = 1/5$, the difference of step energy in the x- and y-direction is $\gamma(0) - \gamma(\pi/2) = 0.4\gamma_0$, a small step energy anisotropy, the resultant equilibrium island shape is continuous and smooth (without the forbidden angles), as shown in Fig. §.3.1.2a. For $\varepsilon_2 = 1/2$, the difference of step energy in the x- and y-direction is $\gamma(0) - \gamma(\pi/2) = \gamma_0$, a large step energy anisotropy, the resultant equilibrium island shape develops a shape corner (cusp) in the x-direction (see Fig. §.3.1.2b), corresponding to

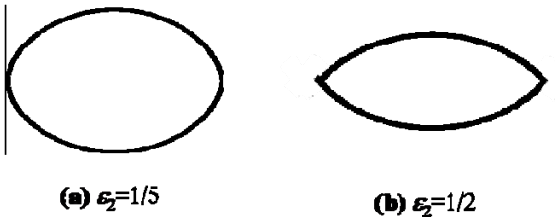

Fig. §.3.1.2 Equlibrium 2D island shape on a surface with twofold symmetry constructed using the mathematical equations (150) and (151). (a) A smooth elliptical shape obtained with small step energy anisotropy at ε_2= 1/5. (b) A cusped island shape (like an "American" football) obtained with large step energy anisotropy at ε_2 = ½.

the orientation of the highest step energy. A range of forbidden angles are associated with the cusp, reflecting the physical fact that the steps at and near the cusp orientation are metastable or unstable, having a much too high energy to enter the equilibrium island shape.

§.3.2 Stress/Strain Induced Size Dependence of Equilibrium Shape

The conventional Wulff construction of equilibrium 2D island shape discussed above includes only the island boundary (step) energy. Thus, most earlier studies of 2D island morphology have been limited to homoepitaxial systems, assuming the aspect ratio of equilibrium islands is defined by the ratio of step free energies [38-41]. However, in heteroepitaxy, where the growing material has a different lattice constant from that of the substrate, lattice-mismatch induced misfit strain is also expected to influence island equilibrium shape. Recently, a theory [32] has been developed to describe the effect of strain (lattice mismatch as well as intrinsic anisotropic surface stress) on the equilibrium shape of a 2D island. It is demonstrated that the conventional wisdom that the equilibrium shape of a 2D island is determined by the ratio of step free energies is in general incorrect, even for homoepitaxial systems if a surface stress anisotropy is present.

One of the main physical consequences of Wullf construction is that the equilibrium shape of a 2D island is size independent: no matter how big or small the island is, its shape is always the same. However, this is not true for a stressed/strained 2D island. In general, stress/strain drives islands to a great anisotropy as island size increases [32]. Below, we discuss the size dependence of the stressed/strained 2D island equilibrium shape by first demonstrating the strain induced spontaneous island shape instability.

a. Spontaneous Shape Instability

We apply the framework of continuum elastic theory presented in the last section to investigate the stability of a single 2D island under biaxial isotropic stress/strain on the surface of a semi-infinite substrate. We will limit our study to a single island isolated from other islands and steps, to eliminate possible complications of elastic island-island or island-step interactions on island shape. We show that minimization of strain energy for different island sizes leads to a complex evolution of island shape with increasing island size that depends on the relative strengths of step and strain energies and on the anisotropy of step energies. Biaxial isotropic stress induces a spontaneous shape instability: for isotropic step energies, an island adopts an isotropic shape at small sizes and transforms into an anisotropic shape beyond a critical size.

Let's consider a biaxially strained epitaxial 2D island on a surface with two-fold symmetry (e.g. the (001) surface of a material with the diamond structure). We assume it has a rectangular shape as shown in Fig. §.3.2.1. [Physically, the equilibrium shape can only be attained by growing and annealing at a sufficiently high temperature, and entropy will make the actual island elliptical (see discussion in section §.3.3 below). However, the elliptical integration is not analytically solvable. So, for simplicity, we use a rectangular shape to illustrate the effect of strain on island shape anisotropy, which is qualitatively equivalent.] The lattice mismatch between the island and the substrate introduces an elastic-force monopole along the island periphery proportional to the misfit strain and the height of the step that forms the edges of

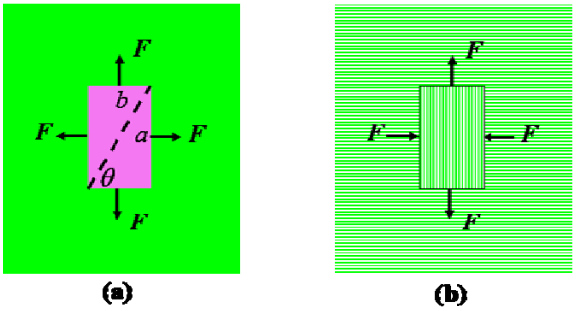

Fig. §.3.2.1. Schematic views of 2D islands grown on a surface of twofold symmetry, with a rectangular shape of length a and width b. $\theta=\tan^{-1}(a/b)$ defines the aspect ratio of the island. (a) Heteroepitaxial growth. F represents the elastic force monopole along the island periphery induced by the lattice mismatch between the island and substrate. (b) Homoepitaxial growth on a surface with anisotropic surface stress. Dashed lines indicate alternating stress domains arising from surface stress anisotropy. F represents the elastic force monopole induced by the surface stress anisotropy. Note that the force monopole on the two a sides in a direction opposite to the force monopole in heteroepitaxy, shown in (a).

the island, as discussed in section §.2.1b and shown in Fig. §.2.1.1b. Similar to the calculation of a square island in section §.2.3a, the strain energy of the rectangular island can be calculated as

$$E_{strain} = E_0\{2(1-\nu)(a\ln\frac{\sqrt{a^2+b^2}+a}{\sqrt{a^2+b^2}-a}+b\ln\frac{\sqrt{a^2+b^2}+b}{\sqrt{a^2+b^2}-b}-2a\ln\frac{a}{ea_0}-2b\ln\frac{b}{ea_0}) \quad .(\S.3.2.1)$$
$$+4[a+b-2(1-\nu)\sqrt{a^2+b^2}]\}+O(a_0)$$

Here, we define $E_0=\frac{1+\nu}{2\pi Y}F^2$ as the unit strain energy, representing the interaction energy between two parallel force monopoles at unit separation (see equation §.2.24a). Notations of F, ν, Y and a_0 are the same as in section §.2. We may rearrange equation §.3.2.1 into a generic compact form as

$$\frac{E_{strain}}{E_0}=PG(c)-P\cdot 2(1-\nu)\ln\frac{D}{a_0}+PO(\frac{a_0}{D}), \qquad (\S.3.2.2)$$

where $P=2(a+b)$ is the perimeter, $c^2=a/b$ is the shape aspect ratio, and $D=\sqrt{ab}$ is the diameter (size) of the island. The term $O(\frac{a_0}{D})=-4(1-\nu)\frac{a_0}{D}$ is very small and can be neglected [32], because generally $D >> a_0$. $G(c)$ is a dimensionless geometric factor that depends on the island aspect ratio c^2 as follows,

$$G(c)=\frac{1}{(c+c^{-1})}\{(1-\nu)(c\ln\frac{\sqrt{c^2+c^{-2}}+c}{\sqrt{c^2+c^{-2}}-c}+c^{-1}\ln\frac{\sqrt{c^2+c^{-2}}+c^{-1}}{\sqrt{c^2+c^{-2}}-c^{-1}}-2c\ln\frac{c}{e}-2c^{-1}\ln\frac{c^{-1}}{e}) \quad .(\S.3.2.3)$$
$$+2[c+c^{-1}-2(1-\nu)\sqrt{c^2+c^{-2}}]\}$$

For a square island, $G(c)$ reduces to a constant and equation §.3.2.1 reduces to §.2.3.3. From equations §.2.3.3, §.2.3.8, and §.3.2.2, we may deduce that the generic compact form of §.3.2.2 applies to all the island shapes, although we have not derived it rigorously from an arbitrary shape.

The island's strain energy, E_{strain}, has two contributions in Equation §.3.2.2: both are proportional to the perimeter P of the island but with opposite signs. The first term, which is positive, arises primarily from the elastic interactions between force monopoles along the same island edge (either a or b); the second term, which is negative, arises from the interactions between force monopoles on opposite island edges (a and a, or b and b) separated by the average island dimension $D=\sqrt{ab}$. The balance of these two terms defines the optimal island shape at a given size D, neglecting step energy contributions. (If the total energy is minimized with respect to both the island shape and size, there exists a solution of the most stable island

size [47] for which the island adopts an isotropic (square) shape. Here, we consider the case that the island continues to grow beyond its stable size.) For small island sizes, the first term dominates [i.e. $G(c) >> \ln(D/a_0)$], and the energy minimization requires minimizing P, favoring an isotropic island shape, a square; for large island sizes, the second term begins to dominate, and the energy minimization requires maximizing P, favoring an anisotropic shape, a rectangle.

The free energies of steps bounding the island of course also contribute in defining the island shape. Let E_a and E_b are respectively the free energies per unit length for island edges a and b, the total step energy is $E_{step} = 2aE_a + 2bE_b$, and the island total energy (step plus strain) is

$$\frac{E_{total}}{E_0} = P \cdot [\alpha \frac{\beta^2 c^2 + 1}{\beta c^2 + \beta} + G(c)] - P \cdot 2(1-\nu) \ln \frac{D}{a_0}, \qquad (\S.3.2.4)$$

where $\alpha = \sqrt{E_a E_b} / E_0$ defines the ratio of the average step energy to the unit strain energy and $\beta^2 = E_a / E_b$ denotes the ratio of the step free energy of edge a and b.

We consider the case of "isotropic" step energy, i.e., $\beta = 1; E_a = E_b$, to illustrate the strain-induced spontaneous shape instability. [Here, we use the term "isotropic" step energy in a restricted sense with respect to only two orientations of steps (x- and y-direction) present in the equilibrium island shape being equal of energy. All other high-energy step orientations are excluded from discussion, as they are excluded in the island shape, as illustrated in Fig. §.3.2.1a .] In Fig. §.3.2.2a, we plot the calculated total energy of an island as a function of θ= arctan(a/b) for different island sizes with isotropic step free energies ($E_a = E_b$). We choose to use θ instead of the aspect ratio a/b as the variable for island shape, because θ is a better choice of order parameter in describing the island shape here for the total energy being symmetric about θ= 45°, as shown in Fig. §.3.2.2a.

The island originally adopts an isotropic (square) shape, with an energy minimum at θ= 45°. As the island grows beyond a critical size D_c, strain induces a spontaneous shape instability: the island adopts an elongated rectangular shape in either of the two orthogonal directions with two degenerate energy minima at $\theta_m = 45^o \pm \Delta\theta$. $\Delta\theta$, and hence the aspect ratio of the elongated island, increases with increasing island size (D). The critical size D_c is defined by the condition

$$\frac{d^2}{d\theta^2} E_{total}\Big|_{\theta=45^o} = 0, \qquad (\S.3.2.5)$$

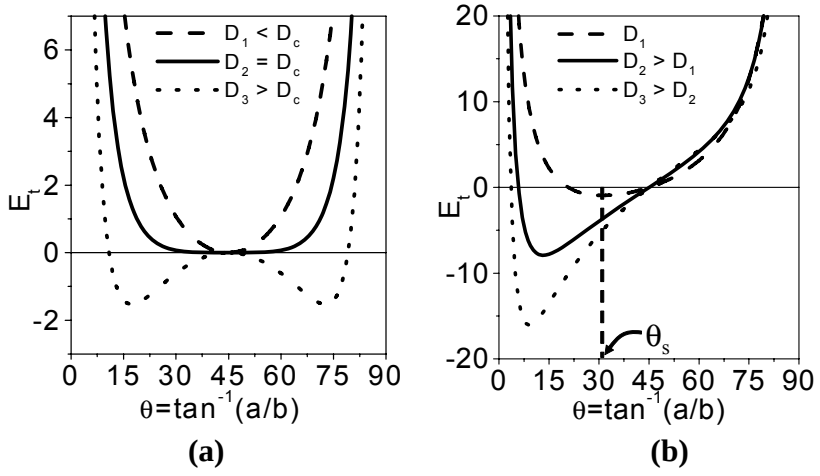

Fig. §.3.2.2. Total energy of strained 2D islands versus angle 0=tan^{-1} (a/b) demonstrating the strain-induced shape instability. The ratio of average step free energy to unit strain energy, α, is chosen to be 2. (a) 2D islands with isotropic step free energy (E_a=E_b). (b) 2D islands with anisotropic step free energies with β^2= E_a/E_b= 1.44. The vertical dashed line marks θ_s=tan-^{1} (E_a/E_b), defined by the step free-energy ratio.

which gives rise to

$$D_c = a_0 \exp[\frac{\alpha+2}{2(1-\nu)}+1.30]\cdot \qquad (\S.3.2.6)$$

The existence of the spontaneous shape instability originates from he strain relaxation energy. This is especially obvious that the instability remains even when the step free energy is zero (α = 0). The isotropic step energy shifts the critical size D_c to a larger value, because it would act to drive the island toward an isotropic shape for all sizes. The step energy becomes the dominant factor in defining the critical size when the step energy is much stronger than the strain energy ($\alpha >> 1$).

The strain-induced shape instability has been shown before for 3D inclusions in bulk materials [48,49]. However, on surface, such instability in principle exists only for 2D islands, because a 3D coherently strained island on surface can always lower its strain energy by increasing its height. However, if the increase in height of a 3D island is kinetically limited, the 3D island may grow only laterally behaving effectively like a quasi-2D island. It then can exhibit a shape instability [47] similar to the one we describe here for 2D island, driven now by the competition between strain energy and island surface (facet) energy.

b. Dependence of Island Shape on Island Size

The strain-induced shape instability redefines the traditionally assumed relationship between the equilibrium island shape and the step free-energy ratio [38-41], namely that the shape reflects the step free-energy anisotropy. Now, we extend the above analysis to the general case of anisotropic step energies. In this case, the island always has an anisotropic shape but its aspect ratio increases continuously with increasing island size, as the strain energy becomes a more significant contribution to the total free energy. Furthermore, the same behavior also occurs for a homoepitaxial 2D island growing under stress induced by substrate surface stress anisotropy [32].

In Fig. §.3.2.2b, we plot the calculated total energy of an island as a function of θ= arctan(a/b) for different island sizes with anisotropic step free energies ($E_a \neq E_b$). When the step free energy is anisotropic (referring to the x- and y-directional steps only), the symmetry between the two orthogonal directions of strain-induced island anisotropy is broken, i.e., the energy degeneracy between the strain-induced island elongation along x- or y-direction is lifted. The anisotropic step free energies cause islands to elongate only along the low-step-free-energy direction, as shown in Fig. §.3.2.2b, in which both the step free energy and strain energy are minimized. (The other direction becomes energetically unfavorable because the step free energy would not be optimized.) For any given island size, in general, the strain favors an optimal island aspect ratio different from what could have been defined solely by the step free-energy ratio; the total energy of the island has one deep minimum at $\theta_m = 45^o - \Delta\theta$ for $E_a > E_b$ (or at $\theta_m = 45^o + \Delta\theta$ if $E_a < E_b$). Figure §.3.2.2b shows that θ_m moves farther away from 45^o with increasing island size, indicating that the strain relaxation makes the aspect ratio of the islands increase continuously with increasing island size rather than stay equal to the step free energy ratio. At small size, strain drives the islands toward a more isotropic shape, making the island aspect ratio smaller than the step free-energy ratio; at large size, strain makes the island aspect ratio larger than the step free-energy ratio.

Figure §.3.2.3 demonstrates the relationship between the island shape (aspect ratio) and the step free-energy ratio under the influence of strain. Figure §.3.2.3a shows the dependence of the island aspect ratio, a/b, on the step free-energy ratio, E_a/E_b, for different island sizes, for a fixed value of the ratio of average step free energy and strain energy, α = 5. At a given value of E_a/E_b, a/b increases with increasing island size, D. Figure §.3.2.3b shows a/b as a function of E_a/E_b for different values of α for a fixed island size (D = 6). The horizontal dashed line (α = 0) marks the island's aspect ratio defined solely by strain energy relaxation, i.e., when the step energies are zero. The 45^o inclined dashed line corresponds to $\alpha = \infty$, i.e., there is no strain energy, and the island's aspect ratio equals the step free-energy ratio. All the curves pass through the same point, at which the island aspect ratio defined by minimizing the strain energy coincides with that defined by minimizing the step free energy. Below this point, strain relaxation drives the island toward an anisotropy higher than the

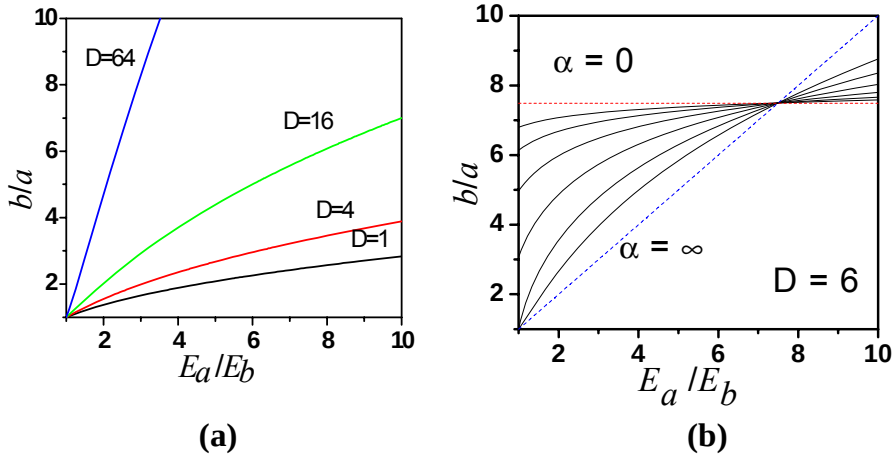

Fig. §.3.2.3. Island aspect ratio versus step free-energy ratio, demonstrating the strain effect on island aspect ratio. (a) At a given value of α for different island sizes. Note that for E_a=E_b, a/b =1 for small island sizes but becomes larger than 1 at D= 64, as the island becomes larger than the critical size D_c defined in Eq. (54) (see fig. 3.2.2). (b) At a given island size for different value of α.

step free-energy ratio; above this point, the reverse is true, although it is never possible for strain to force an isotropic island shape. This point will, of course, shift if the size D of the island is changed.

If a substrate surface stress anisotropy is present, as in the Si(001) surface, the above conclusion obtains even for homoepitaxy. The surface stress anisotropy introduces a force monopole [25,26,32] along the periphery of the 2D island similar to that introduced by misfit strain, the only difference being that the force monopoles on the two a sides point in directions opposite to those in heteroepitaxy (compare Fig. §.3.2.1b to Fig. §.3.2.1a), leading to a slightly different geometry factor $G(c)$. In Equation §.3.2.3, the last term is replaced by

$$4[(1-4\nu)(c+c^{-1})-2(1-3\nu)\sqrt{c^2+c^{-2}}]\,. \qquad (\S.3.2.7)$$

§.3.3 Experimental Growth of Si 2D Island on Si(001)

Experimentally, it has been a common practice to derive the step free-energy ratio on an anisotropic surface from the equilibrium aspect ratio of 2D islands at a given temperature [38-41]. However, the above theory shows that, for islands under stress, the aspect ratio of 2D islands does not simply equal the step free-energy ratio, but becomes dependent on the island size and on the ratio of the strengths of the step free energy and strain energy. Because both the ratio of step energy to unit strain energy α and the step free-energy ratio β vary with temperature, there is no way to determine the step free-energy ratio at different temperatures from only one data point of island shape at each temperature. In addition, the strain-induced island-island interaction may also influence island shape, further complicating the problem.

The theory discussed in the above sections provides a theoretical framework for extracting surface stress, steps energies, and their azimuthal dependence from measurements of island shape evolution with size. It is proposed [32] that to derive the step free-energy ratio from the shape of a strained 2D island, an experiment needs to be carried out to observe the changing shape of a single "isolated" island during growth at a *fixed* temperature. As the temperature is fixed, both the parameters α and β remain constant. By carefully measuring the increasing island aspect ratio with increasing equilibrium island size (i.e., very slow growth or interrupted growth), one can uniquely derive both the step free-energy ratio and the ratio of step energy to strain energy (for that particular temperature) with a best fit of the theoretical curve to experimental data. If the unit strain energy (E_o), i.e., the misfit strain or surface stress anisotropy, is also known, one can further determine the individual step free energies.

Such an analysis has been first carried out for Si 2D islands grown on Si(001) substrate [42]. It indeed confirms quantitatively the theoretical prediction [32], in which the equilibrium aspect ratio of a single isolated Si island grown on a large Si(001) surface was observed

to increase continuously with increasing size, using low-energy electron microscopy (LEEM). Here, we briefly review this experiment and show an example of analyzing the experimental data using the above theoretical model.

Surface stress plays an important role not only in determining the static structure and morphology of a surface, but also in controlling the evolution of surface structure and morphology during epitaxial growth. In general, surface stress is inevitably connected with surface reconstruction, because tilting or stretching of atomic bonds is necessary to reconstruct the surface. For the case of Si(001) surface, formation of dimer rows gives rise to a well-known (2x1) reconstruction [23]. The (2 × 1) reconstruction on Si(001) introduces a highly anisotropic surface stress [50]: the stress $\sigma_{\parallel}$ along the dimer bond is tensile (i.e., surface atoms would like to be closer together along this direction than they are); the stress $\sigma_{\perp}$ along the dimer row is compressive (or at least less tensile than $\sigma_{\parallel}$). Because of the tetrahedral bonding configuration in the diamond structure, the dimer direction is orthogonal on terraces separated by monatomic steps. Consequently, the dimer direction in a Si 2D island is orthogonal to that on Si(001) surface, as illustrated in Fig. §.3.2.1b, introducing force monopoles along the island boundary whose magnitude is proportional to the intrinsic surface stress anisotropy of Si(001)- (2 × 1) surface, i.e., $|F| = |\sigma_{\square} - \sigma_{\perp}|$. Therefore, in the case of homoepitaxial growth of Si 2D island on Si surface, the underlying driving force for the size dependence of island shape is the "surface stress anisotropy" rather than the misfit strain as in heteroexpitaxy. But the general physical principle and the framework of theoretical analysis is the same.

The interplay between the dimer-row formation and the anisotropic surface stress has led to a wide range of interesting and intriguing surface and growth phenomena on Si(001)-(2x1) surface [50]. For example, a stress-domain structure forms naturally on a vicinal (stepped) Si(001) surface, consisting of alternating (2x1) and (1x2) stress domains, due to the surface stress anisotropy [26,51]. It has been shown that applying a uniaxial stress to the Si(001) surface will drive step movement, favoring one domain over the other [51]. However, surface stress has generally been neglected in interpreting thermodynamic properties of homoepitaxial Si 2D islands [38-41] in the submonolayer growth regime. The theoretical proposal [32] had then stimulated an experimental study to investigate the effect of stress on Si 2D island growth [42].

The experiment [42] provided direct evidence that the equilibrium shape of 2D Si islands on Si(001) is *size dependent*. The island originally adopts an elliptical equilibrium shape, whose aspect ratio increases continuously with increasing island size, and then transforms into an "American football"-like shape beyond a critical size, as shown in Fig. §.3.3.1. Both the size-dependent aspect ratio and the elliptical-to-football shape transition are driven by surface stress relaxation. The changing equilibrium aspect ratio of an elliptical island with increasing size but formed at a fixed temperature was fitted nicely by the theoretical model

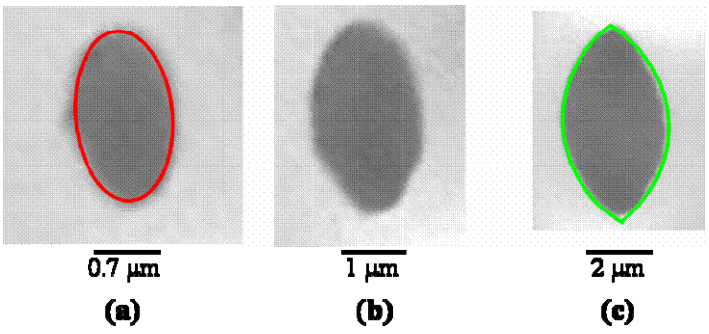

Fig. §.3.3.1. Island shape transition with island size depicted for three stages: (a) elliptical shape, (b) intermediate shape, (c) football shape. Images are scaled to compensate for island size. The elliptical outline enclosing the island in (a) is constructed using the elliptical function. The football outline enclosing the island in (c) is constructed as an intersection of two circles.

[32], and quantitative values of not only step free energies but also surface stress anisotropy at different temperatures were extracted. The study also showed, for the first time, that the intrinsic surface stress anisotropy of Si(001) is temperature dependent, decreasing with increasing temperature.

We refer readers to Ref. [42] for details of the experiments. One of key points was to use a large step-free terrace with a single island nucleating in the middle of the terrace far from the terrace edges, which not only allows to observe islands that grow up to several micrometers in diameter, but also to exclude possible stress-induced island-island and island-step (terrace edge) interactions that can influence island shapes. Earlier studies of Si island growth on Si(001) had either carried out at low-temperature in which island shapes were affected by slow kinetics [38,39] or limited the islands to less than ~100 nm in diameter before they interacted with terrace edges or neighboring islands [40,52]. Consequently, these early experiments are inappropriate for the analysis of stress induced size dependence.

To illustrate the island shape evolution, Figure §.3.3.1 displays three island shapes, scaled to compensate for island size, at different stages of a single island growing slowly on a rectangular 10x15 μm^2 large mesa terrace at 855°C, recorded by LEEM. The island appears as black and the base terrace as white [42]. At small size, e.g., 1.4 μm (major axis, Fig. §.3.3.1a), the island appears in an elliptical shape, as indicated by the perfect elliptical fit, the red contour enclosing the island. At large size, e.g., when the island size becomes three times larger (Fig. §.3.3.1c), the island develops sharp tips at the far ends of the island, changing into an American-football-like shape, as indicated by the fit of a football shape (the intersection of two circles), the green contour enclosing the island. Figure §.3.3.1b shows an intermediate state between the elliptical-to-football shape transition.

In general, the shape of 2D islands reflects the dependence of step free energy on step orientation, just as the shape of isolated 3D crystallites reflects the dependence of surface free energy on facet orientation [44-46]. For Si islands on Si(001), the anisotropic shape was conventionally attributed solely to the different step energies of two orthogonal orientations [38-40,53], namely S_A and S_B steps [54]. However, step energies are island size independent, which would lead to a size-independent island aspect ratio. Therefore, the size dependence of the island shape must originate from an additional free-energy contribution. The stress relaxation energy provides the appropriate behavior [32].

To analyze island shape evolution quantitatively, the island shape aspect ratios and island areas were determined using image analysis software [42]. High-resolution images had been used to capture island sizes over a wide range from 0.1 to 3.0 μm in diameter. Figure §.3.3.2 shows the island aspect ratio vs. island area at 855°C, with islands remaining in a

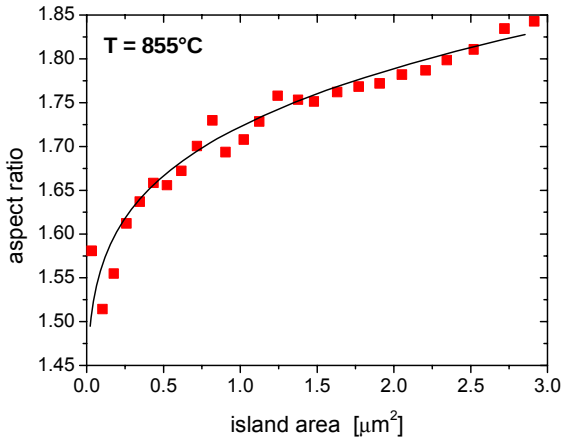

Fig. §.3.3.2 . Dependence of 2D island shape aspect ratio (major to minor axis) of a single 2D Si island grown on an extremely large single-domain Si(001) terrace. Squares are experimental data at a temperature of 885° C and the solid lines are model fits to the data as described in the text.

perfect elliptical shape. The experimental data was then analyzed using the theoretical model [32,42] that includes stress relaxation energy in addition to step free energy in determining the equilibrium shape of islands as a function of island size.

As a 2D island forms on Si(001), the dimer rows in the island are rotated by 90° with respect to the substrate dimer rows, as illustrated in Fig. §.3.2.1b. Consequently, the anisotropic surface stress introduces stress discontinuities, i.e., force monopoles along the island step edges. For a model rectangular island analyzed in the above section, island edges are straight, which greatly simplifies the problem. The real experimental island has an elliptical shape, which requires theoretical analysis to be generalized in two aspects: first, an general azimuthal dependence of the island step (boundary) energy must be included to account for curved island boundaries and their continuous evolution; second, the island stress relaxation energy has to be calculated numerically because integration along an elliptical boundary is not analytical solvable. Below, we illustrate such a general theoretical analysis for the Si(001) homepitaxial 2D island [42]. The total free energy of the elliptical island can be calculated as the sum of the step free energy and stress relaxation energy. The step free energy is integrated by

$$F_{step} = \int_0^{2\pi} F_s(\theta) r(\theta) d\theta \ , \qquad (\S.3.3.1)$$

where $F_s(\theta)$ is the step free energy at the angle θ in reference to the S_A step orientation (chosen along the x-direction) in Si(001) surface, $r(\theta)$ is the island radius. It can be calculated from the step energies (E_A and E_B) and corner energy (E_c) at zero temperature as

$$F_s(\theta) = F_A \cos\theta + F_B \sin\theta + F_C \sin 2\theta \ , \qquad (\S.3.3.2a)$$

$$F_{A,B} = E_{A,B} - TS_{A,B} \ , \qquad (\S.3.3.2b)$$

$$F_c = E_C - k_B T \ln 2 \ . \qquad (\S.3.3.2c)$$

T is the temperature and k_B is the Boltzmann constant. $S_{A,B}$, the step entropy is calculated from E_A, E_B, and E_C, using a solid-on-solid model [55,56]. The stress relaxation energy is again calculated by integrating the force monopole-monopole interaction around the island boundary, as has been done above for other island shapes, except now the integration has to be carried out numerically for an elliptical shape.

A trial-and-error procedure was used to fit the experimental data. For each set of trial values of step energies (E_A, E_B, E_C) and surface stress anisotropy, the optimal island aspect ratio for a given island size is determined by minimizing the island total energy, in which the step free energies are calculated analytically while the stress energies are calculated numerically. The calculated dependence of aspect ratio vs. island size (area) is in turn compared to the experimental data. The trial values of the energy and stress parameters are then adjusted until the theoretical curves converge to the experimental data with the smallest deviation, i.e., the sum of the differences between the two at each experimental data point. The fitting was done for island shapes at different temperatures and an example at 855°C was shown in Fig. §.3.3.2. It yields step energies E_A = 25 ± 3 meV, E_B = 53 ± 6 meV, and E_C = 56 ± 8 meV at zero temperature, and surface stress anisotropies of 80 ± 3 meV/Å at 695°C to 68 ± 4 meV/ Å at 855°C. These values agree very well with previous experiments and theories [40,55-59].

The island aspect ratio increases continuously with increasing island size. One particular interesting point is that the measured island aspect ratio is smaller than the step free-energy ratio in certain range of small island sizes. This can be explained readily by the equilibrium stress model [32], as shown qualitatively in Fig. §.3.3.2. Furthermore, the elliptical-to-football shape transition at large island size can also be explained within the same theory. From the point of Wulff construction of equilibrium island shape, the continuous increase of the island aspect ratio with increasing island size reflects that "effectively" the step energies have become more and more anisotropic with increasing island size. Especially, when the step energy anisotropy becomes so large that some step orientations will be forbidden from the equilibrium shape, as discussed in section $.3.1, leading to the tip development in the highest-energy step orientation and an elliptical-to-football shape transition.

Note that the observed elliptical-to-football shape transition is actually driven by the physical mechanism of size-dependent stress energy rather than step energies (the step energy ratio remains fixed at fixed temperature). One could combine the stress energy into "effective" step energies that would be island size dependent. In particular, the ratio of "effective" step energies would increase with increasing island size, leading to increasing island aspect ratio and eventually to the elliptical-to-football transition. It is possible to derive the "effective" step energies from the measured island shapes, using the inverse Wulff construction [38,60], which contain a stress contribution. But the stress energy can only be determined separately after knowing the island shape as a function of island size at constant temperature, as shown above; there is no mechanism to decompose stress energy into individual step energies of different orientations.

The above example demonstrates the influence of surface stress on island shape in Si homoepitaxy and illustrates the use of theoretical model [32,42] for analyzing the equilibrium 2D island shape under stress. The same theoretical analyses have also been applied in understanding the effect of stress/strain on 2D island shape in some other experimental systems. These include the surface stress and surface stress anisotropy induced by different surface reconstructions in homepitaxial systems, such as Ge(001) [61,62], GaAs(001) [63], Si(111) [64], and Cu(001) [65] surfaces, as well as surface stress and strain induced by surface adsorption and lattice mismatch in heteroepitaxial systems, such as Ge/Si(113) [66], Sb-Si/Ge(001) [67], Cl/Si(001) [68], Co/Cu(001) [69], and Ni/W(110) [70] surfaces. Furthermore, the effect of strain on equilibrium shape of 3D islands (implicitly with a fixed height) has been studied in the context of the similar theory formulated earlier for the 3D islands [47]. The readers are referred to these works for more detailed discussions.

§.3.4 Equilibrium Shape of Anisotropically Strained Islands: Growth of Nanowires

The study of equilibrium shape of stressed/strained islands on surface is not only of scientific interest but also of technological significance. As discussed above, equilibrium shape of a streesed/strained 2D island gives a direct manifestation of surface thermodynamic properties including surface step energies and surface stresses [32,42,64]. Furthermore, growth of strained 2D and 3D islands provides a potentially useful method for fabrication of nanowires on surface [47,71-77].

The above equilibrium theories of island shape have been established for 2D islands on surface under *isotropic* strain. Most noticeable, isotropic strain induces a spontaneous shape instability in growth of 2D island with isotropic step energies on surface [32]: there exists a critical size, controlled by competition between isotropic step energy and strain energy, below which the island adopts an isotropic shape (such as a square) and above which it adopts an anisotropic shape (such as a rectangle).

The strain induced island shape elongation has been applied for growing nanowires on surface [47,71-77]. One problem associated with growing wires using isotropically strained island is that it may elongate along either x- or y-direction with two energetically degenerate shapes (see Fig. §.3.2.2a). To make all the wires oriented along the same direction, an idea has been proposed [75,76] to grow island that is strained in one direction but strain-free in the other (orthogonal) direction so that it grows only along the strain-free direction. However, such a scenario is hard to find with a real materials combination between the islands and substrate.

Most wires have been grown with islands that are generally *anisotropically* strained on surface. A typical system is rare-earth metal silicide nanowires [74-76] grown on Si(001), which has attracted much attention because of its application in integrated circuits. However, existing theories [32,42,47] have focused mostly on *isotropically* strained islands. Although the effect of strain anisotropy has been speculated [47,75] and a special case of 2D island anisotropy with the same stress magnitude but opposite signs has been considered in the case of Si homoepitaxy [32,42] (see Fig. §.3.2.1b), a general theory of equilibrium shape of 2D or 3D islands under *anisotropic* stress/strain is still lacking. This creates a gap in the understanding of the

experimental growth of nanowires using anisotropically strained islands, such as those silicide wires [74-76].

Several pending questions are especially important with regard to growth of nanowires: (1) Will an anisotropically strained island, like an isotropically strained island, exhibit a shape instability? (2) When an island is anisotropically strained, will it be only stable along the less-strained direction with one energy minimum, or will it be also metastable along the other more-strained direction with a local energy minimum? (3) How does the island shape anisotropy depend on strain anisotropy? (4) For given strain anisotropy, does the sign of strain (i.e., tensile or compressive in both directions vs. tensile in one direction but compressive in the other direction or vise versa) matter?

Below, we review a recent theoretical analysis of equilibrium shape of 2D island under *anisotropic* strain (or stress) [43]. We demonstrate that the existence of strain anisotropy removes completely the island shape instability associated with isotropically strained island. As long as the strain is anisotropic, there will be only one energy minimum along the less-strained direction at all island sizes without the second local minimum in the orthogonal more-strained direction. Thus, thermodynamically, the island prefers always to grow along the less-strained direction, and its aspect ratio increases with increasing island size. For a given island size, the larger the strain anisotropy is, the larger the island aspect ratio is. The sign of strain makes only a quantitative difference, without changing the island shape behavior qualitatively.

We analyze the equilibrium shape of a 2D island under biaxial *anisotropic* strain with the same theoretical framework, as we discussed above for a 2D island under *isotropic* strain. Consider the general case of a rectangular island of length a and width b strained along the x- and y-direction respectively with ε_{xx} and ε_{yy}, as shown in Fig. §.3.4.1 (inset). The strain introduces elastic force monopoles (F_x and F_y) along the island boundary, whose magnitude is proportional to the magnitude of strain and the island height and whose direction follows the respective sign of strain (tensile vs. compressive) in each direction.

As discussed in section §.2, the island strain energy can be calculated by integrating the force monopole-monopole interaction along the island boundary expressed as

$$E_{strain} = -\frac{1}{2}\iint d\boldsymbol{r}d\boldsymbol{r}'\boldsymbol{F}_1(\boldsymbol{r})\cdot\boldsymbol{u}[\boldsymbol{r},\boldsymbol{F}_2(\boldsymbol{r}')], \qquad (\S.3.4.1)$$

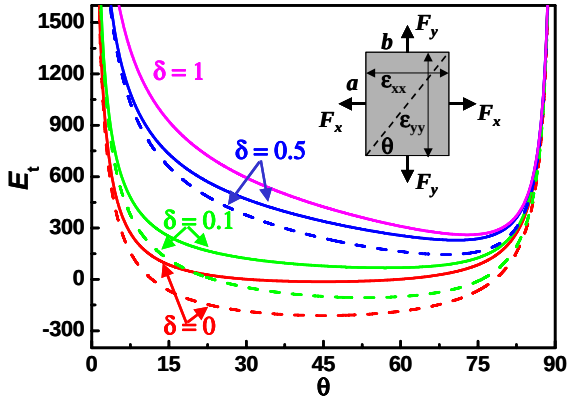

Fig. §.3.4.1. Total energies (E_t) of a strained 2D rectangular island (shown in the inset) versus $\theta = \tan^{-1}(a/b)$ for different strain anisotropies (δ). The island size is chosen to be smaller than the critical size associated with the isotropically strained island. Solid curves are for strains having the same sign in the two directions (both compressive or tensile); dashed curves are for strains having different signs (compressive along x- and tensile along y-direction; or vice versa).

where, $\boldsymbol{u}[\boldsymbol{r}, \boldsymbol{F}_2(\boldsymbol{r}')]$ is the displacement at $\boldsymbol{r}$ induced by a force at $\boldsymbol{r}'$, and

$$\boldsymbol{F}_1(\boldsymbol{r}) \cdot \boldsymbol{u}[\boldsymbol{r}, \boldsymbol{F}_2(\boldsymbol{r}')] = \frac{1-\nu^2}{\pi Y} \frac{\boldsymbol{F}_1 \cdot \boldsymbol{F}_2}{|\boldsymbol{r}' - \boldsymbol{r}|} + \frac{\nu(1+\nu)}{\pi Y} \frac{[\boldsymbol{F}_1 \cdot (\boldsymbol{r}' - \boldsymbol{r})][\boldsymbol{F}_2 \cdot (\boldsymbol{r}' - \boldsymbol{r})]}{|\boldsymbol{r}' - \boldsymbol{r}|^3}. \qquad (§.3.4.2)$$

Y is Young's modulus and ν is Poisson's ratio. For a rectangular island, the first term in Eqn. (§.3.4.2) arises only from the interaction between parallel or anti-parallel force monopoles on the two opposite island sides (a-to-a or b-to-b). The second term includes also the interaction between orthogonal force monopoles at the two neighboring sides (a-to-b or b-to-a), describing the Poisson effect that compression (tension) in one direction leads to tension (compression) in the orthogonal direction.

Integration of Eqn. (§.3.4.1) along the island boundary gives

$$\frac{E_{strain}}{E_0} = PG_1(c,\gamma) - P \times 2(1-\nu) G_2(c,\gamma) \ln \frac{D}{a_0} + O(\frac{a_0}{D}). \qquad (§.3.4.3)$$

$E_0 = \frac{1+\nu}{2\pi\mu} F_x^2$, is the unit interaction strain energy between two parallel force monopoles on the island two opposite a sides at unit separation in the x-direction. $\gamma = \frac{F_y}{F_x} = \frac{\varepsilon_{yy}}{\varepsilon_{xx}}$ is the ratio of strain in the two directions. All other notations are the same as in sections §.2 and §.3. The term $O(\frac{a_0}{D}) = 4P[(1-\gamma+\gamma^2)\nu - (1+\gamma^2)]\frac{a_0}{D}$ is very small and can be neglected as in [32], because generally $D >> a_0$. G_1 and G_2 are given by

$$G_1(c,\gamma) = \frac{1}{2(c+\frac{1}{c})}\{2(1-\nu)[c \ln \frac{\sqrt{c^2+\frac{1}{c^2}}+c}{\sqrt{c^2+\frac{1}{c^2}}-c} + \frac{\gamma^2}{c} \ln \frac{\sqrt{c^2+\frac{1}{c^2}}+\frac{1}{c}}{\sqrt{c^2+\frac{1}{c^2}}-\frac{1}{c}} - 2c \ln \frac{c}{e} - 2\frac{\gamma^2}{c} \ln \frac{1}{ce}] \qquad (§.3.4.4)$$

$$+ 4[c\gamma^2 + \frac{1}{c} - (1+\gamma^2 - 2(1-\gamma+\gamma^2)\nu)]\sqrt{c^2+\frac{1}{c^2}} + 2(1-\gamma)(c\gamma - \frac{1}{c})\}$$

and

$$G_2(c,\gamma) = \left(c + \frac{1}{c}\right)^{-1} \left(c + \frac{\gamma^2}{c}\right). \qquad (§.3.4.5)$$

Note that setting $\gamma = 1$, $G_2(c, \gamma) = 1$ and $G_1(c, \gamma)$ reduces to $G(c)$ in Equation (§.3.2.3); the equations (§.3.4.3) – (§.3.4.5) returns to solutions of isotropically strained island. Here, we consider only the case of isotropic step (boundary) energy, the total energy of the island is then $E_t = E_{strain} + PE_b$, where E_b is the boundary energy per unit length.

The equilibrium island shape is determined via total-energy minimization, as a function of island size (D) and of strain anisotropy (δ). To facilitate our discussion, we define strain anisotropy as $\delta = 1 - |\gamma|$, with $0 \le \delta \le 1.0$. For $\delta = 0$, strain is isotropic, i.e., zero or no strain anisotropy; for $\delta = 1.0$, strain anisotropy is the largest, as the y- direction is strain free. The strain in the x-direction, and hence E_0 is kept fixed.

Figure §.3.4.1 shows the total energy E_t as a function of $\theta = \tan^{-1}(a/b)$, calculated using the parameters of, $E_b / E_0 = 3.02$. The island size ($D/a_0 = 50$) is chosen to be smaller than the critical size for shape instability ($D_c/a_0 = 113$ for $\gamma > 0$, and $D_c/a_0 = 209$ for $\gamma < 0$) associated with the isotropically strained island, i.e., for $\delta = 0.0$, the island adopts an isotropic square shape with a single minimum at $\theta = 45^o$, as shown in Fig. §.3.4.1.

As long as $\delta \neq 0.0$ (no matter how small it is), the energy minimum always shifts away from $\theta = 45^o$ to a larger value. Thus, the existence of strain anisotropy leads always to an anisotropic shape, removing the shape instability (i.e. the existence of a critical size) possessed by an isotropically strained island. As the energy minimum shifts to the $\theta > 45^o$ side, the island will always elongate along the less-strained y-direction. The energy minimum continues to move further to larger θ with increasing δ. So, for a given island size, the larger the strain anisotropy, the larger the island aspect ratio, i.e. more elongated island along the less-strained direction.

For a given strain anisotropy, δ (except for δ=1.0 when strain is zero along the y-direction), there may be different signs of strains: both negative (compressive) or positive (tensile) along x- and y-direction vs. compressive along x- and tensile along y-direction or vise versa. Figure §.3.4.1 shows that for the same δ the energies of islands under strains of different signs (dashed curves) are always slightly lower than those under strains of the same sign (solid curves). This is resulted from the Poisson effect described by the second term in Equation (§.3.4.2), the interaction between orthogonal force monopoles on the neighboring island sides. (Without the second term in Equation (§.3.4.2), the energy would be the same for both cases and the solid and dashed curves in Fig. §.3.4.1 would merge together.) The overall strain is effectively smaller and hence the energy is lower when the island is under different signs of strain, because compression in one direction compensates tension in the other direction and vise versa. However, the sign of strain does not qualitatively change the physical picture.

Figure §.3.4.2 shows the same plot as Figure §.3.4.1 for the case of island size (D/a_0 = 300) chosen to be larger than the critical size of shape instability associated with the isotropically strained island. So, for δ = 0.0, there are two degenerate energy minima, one with θ > 45° and the other with θ < 45° (the two lowest curves in Fig. §.3.4.2). This means that at this size, the isotropically strained island would elongate along either x- or y-direction with equal probability.

For $\delta \neq$ 0.0, however, no matter how small it is, there is only one energy minimum at the far side of θ > 45°. This indicates that for any anisotropically strained island, it will only be thermodynamically stable by elongating along the less-strained y-direction. There does not exist a metastable shape along the more-strained x-direction. For large island size, the effect of elongation along the less-strained direction is even very pronounced with a very small strain anisotropy of δ = 0.1 with θ > 85°.

The fact that the island is only thermodynamically stable to elongate along one direction, the less-strained direction, without a metastable state in the orthogonal direction has an important practical implication. It implies that to align the nanowires along the same direction, from the thermodynamic point of view, one may not need to find a materials combination with very large strain anisotropy, in particular strained only in one direction and strain-free (lattice matching) in the other direction as suggested earlier [75,76]. A small strain anisotropy with 10% difference in the two directions may still be sufficient to drive wires to grow along one direction, at the thermodynamic limit. This provides a greater degree of flexibility in choice of materials for growing nanowires via elongation of anisotropically strained islands.

Same as Figure §.3.4.1, Figure §.3.4.2 shows that for the same strain anisotropy (δ) the energies of islands under strains of different signs (dashed curves) are always slightly lower than those under strains of the same sign (solid curves), with energy minima at slightly different positions. In general, as indicated by the positions of energy minima, the larger the strain

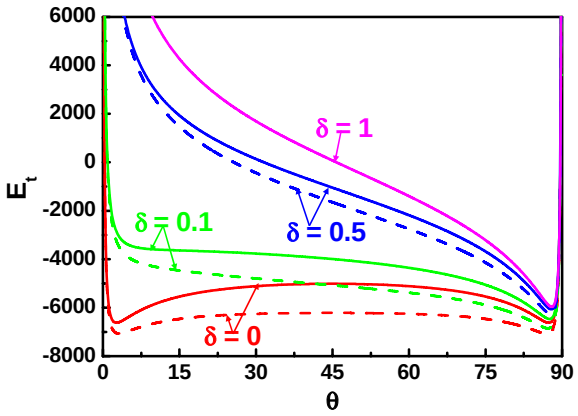

Fig. §.3.4.2 The same plot as Fig 3.4.1 except that the island is chosen to be larger than the critical size associated with the isotropically strained island.

anisotropy is, the larger the island aspect ratio is. This is more clearly illustrated in Fig. §.3.4.3, which shows island aspect ratio (a/b) as a function of strain anisotropy (δ) for several island sizes (D/a_0 = 10, 100, 200 and 300). For δ= 0, small islands, such as D/a_0 = 10 and 100, have a/b = 1.0, as they are smaller than the critical size of shape instability and maintain an isotropic shape; large islands above the critical size, such as D/a_0 = 300, have a/b > 1.0, becoming elongated. For $\delta \neq$ 0.0, islands of all sizes have a/b > 1.0, elongating along the less-strained y-direction.

For all strain anisotropy δ, the aspect ratio of those islands under the different signs of strains (dashed curves in Fig. §.3.4.3) is smaller than that of those islands under the same sign of strains (solid curves). This difference decreases with increasing δ. As δ increases, the island becomes more elongated so that the interaction between force monopoles on the two long opposite sides becomes dominant, while the interaction between orthogonal monopoles on the two neighboring sides, which gives rise to the difference, becomes negligible.

One interesting phenomenon occurs at δ = 0 for D/a_0 = 200. The island under the different signs of strains (y-intercept of D/a_0 = 200 dashed curve) has an isotropic shape with a/b = 1.0, while the island under the same sign of strains (y-intercept of D/a_0 = 200 solid curve) has an anisotropic elongated shape with a/b ~ 9.0. This reflects that the sign of strain changes the critical size of shape instability for the isotropically strained island, which again can be understood by the Poisson effect discussed above. Although the magnitude of strain is the same in two directions, the overall strain is effectively smaller for the island under different sings of strains, leading to a larger critical size, as the critical size increases with decreasing strain [32,47].

Previous studies [32,47] have shown that an isotropically strained island elongates only above the critical size, beyond which it continues to elongate with a fixed width. In contrast, an anisotropically strained island elongates at all sizes. In Fig. §.3.4.4, we show the evolution of equilibrium island shapes with increasing island size for strain anisotropy of δ = 0, 0.5 and 1.0. Under isotropic strain (δ = 0), the island first grows its size isotropically, then elongates along one-direction while shrinks in the other, later it elongates with a fixed width [47]. Under anisotropic strain (δ = 0.5 and 1.0), the island has initially an elongated shape with a narrow width. It then grows preferentially along the less-strained direction while its width converges quickly to a fixed value. In general, the larger the strain anisotropy is, the longer and narrower a nanowire will grow.

Although, for simplicity, analysis has been performed here only for 2D islands, the results can be readily applied to 3D islands for growth of nanowires. In fact, those 3D islands used for growing nanowires must be constrained kinetically with a fixed height, so effectively they can be treated as quasi-2D islands [32,47]. Without height constraint, a 3D island would always in principle grow its height to more effectively relax strain, rather than elongating laterally on surface [32].

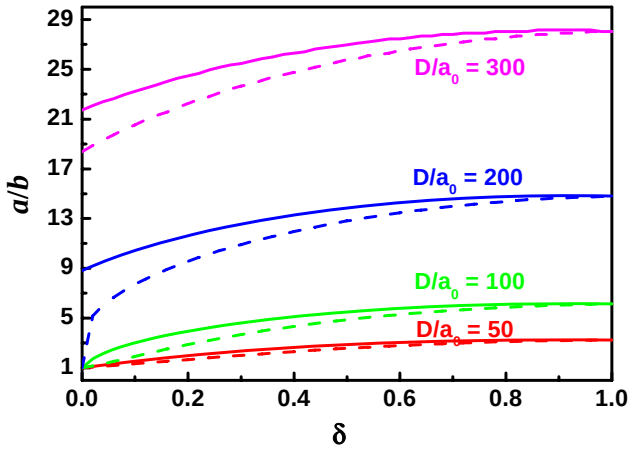

Figure §3.4.3. Calculated optimal (equilibrium) island aspect ratio (a/b) versus strain anisotropy (δ) for four different island sizes of D/a_0=10, 100, 200, and 300. The notations are the same as in figure 3.4.1.

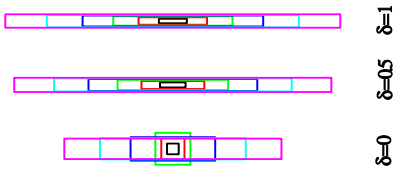

Fig. §.3.4.4 Evolution of equilibrium island shape with increasing island size for islands under different strain anisotropy of $\delta = 0$, 0.5, and 1.0 from bottom to top.

The above analysis is especially relevant to growth of rare-earth silicide nanowires [74-76], which are generally under anisotropic strain and are of great technological significance. We show that anisotropic strain may help improve orientational order of nanowires, driving them to grow along the less-strained direction. The wires may grow with uniform width distribution. The sign of strain makes only a quantitative difference. Longer and narrower nanowires can be grown by choosing a material combination with larger strain anisotropy. These are all in good qualitative agreement with existing experimental results [71-77].

In summary, the above theoretical analysis of equilibrium shape of anisotropically strained 2D islands establishes thermodynamic limits for growing nanowires on surface. We note that while thermodynamically strain anisotropy may work favorably for fabricating nanowires, kinetic factors have also to be taken into account. Future work is needed to investigate the competition between thermodynamics and kinetics, in order to obtain a more complete understanding of growing nanowires with anisotropically strained islands.

§.4 Stress/Strain Induced Self-Assembly/Self-organization of Surface Stress Domains and Lattice of 2D Islands

The formation of stress domains is a ubiquitous manifestation of surface stress, an intrinsic property of solid surfaces. A single isolated stressed/strained 2D island (or void, vacancy island) discussed above in section §.3 can be viewed generally as a typical representation of a single stress domain, i.e., a single domain of modified surface stress (e.g., due to lattice mismatch) embedded in the 2D matrix of a different surface stress (the intrinsic substrate surface stress). The equilibrium shape of an embedded single stress domain (a 2D island) is controlled by the balance between the domain (island) boundary free energy and the (island) stress relaxation energy [32]. If many 2D islands form at the same time, they represent a domain structure of multiple stress domains. In this case, the balance between the domains' boundary free energy and the overall stress relaxation energy would control the characteristic sizes and spatial distributions of all the domains. This physical intuition has been proved by theories [25,26] and confirmed by experiments [51,59]. Most remarkable, such a balance often induces self-assembly/self-organization of stress domains and 2D islands, leading to formation of ordered periodic arrays of domains and islands [78-81]. In this section, we will review a few examples of such self-assembly/self-organization process.

§.4.1 Spontaneous Formation of Surface Stress Domains

a. Theoretical Prediction

The spontaneous formation of periodic stress domains was first predicted by theory [25]. Let's consider a surface consists of 1D periodic array of alternating majority (1) and minority (2) stress domains of "equal surface energy" but different surface stress of σ^1 and σ^2, as shown in Fig. §.4.1.1a. Let L be the periodicity and L - d and d be the majority and minority domain size (width), respectively. (The length of domains is infinite.) There exists a density of elastic forces (force per unit length) along all the boundaries, $f = \pm(\sigma^1 - \sigma^2)$, as shown in section §.2.1b (Equation §.2.1.3 and Fig. §.2.1.1a). The total stress relaxation energy per period of

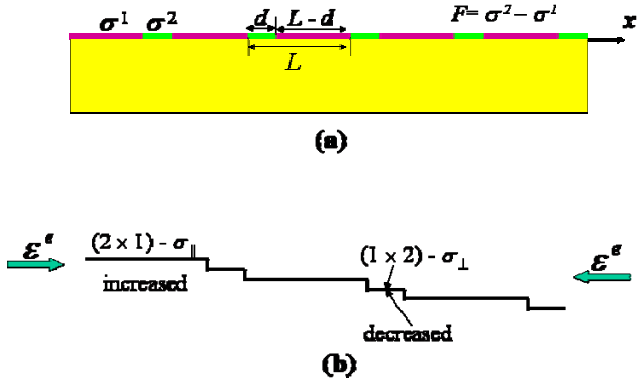

Figure 4.1.1. (a) Schematic illustration of a periodic stress-domain structure with period *L*. The majority domain size is (L-*d*) having a surface stress σ^1, and the minority domain size is *d* having a surfaces stress σ^2. The magnitude of force monopole at the domain boundary is F= $|\sigma^1 - \sigma^2|$ (b) Schematic illustration of a periodic stress-domain structure form naturally on a small miscut Si(001) surface due to its intrinsic surface stress anisotropy, consisting of alternating (2 x 1) and (1 x 2) domains. Under a compressive external strain applied along the dimer-bond direction in the (2 x 1) domains, the (2 x 1) domains are energetically favored and grow at the expense of the (1 x 2) domains.

the whole periodic stress domains can be calculated by summing up the pair interactions between infinite-long line forces (monopoles), given by equation (§.2.2.6), as

$$\begin{aligned} E_{stress} &= -C\ln\frac{d}{a_r} - C\sum_{n=1}^{\infty}(\ln\frac{nL+d}{a_r} + \ln\frac{nL-d}{a_r}) + 2C\sum_{n=1}^{\infty}\ln\frac{nL}{a_r} \\ &= -C\ln\frac{d}{a_r} - C\sum_{n=1}^{\infty}\ln[1-\frac{1}{n^2\pi^2}(\frac{d}{L}\pi)^2] \qquad , \qquad (\S.4.1.1) \\ &= -C\ln(\frac{L}{\pi a_r}\cdot\sin\frac{d}{L}\pi) \end{aligned}$$

where $C = \dfrac{(1-\nu^2)f^2}{\pi Y}$ and all the variables use the same notations as in section §.2. (Alternatively, one may derive equation §.4.1.1by Fourier transformation to reciprocal space [26], since we have a periodic structure.)

The net formation energy of the stress domains per area is then

$$E_f = \frac{E_b}{L} - \frac{C}{L}\ln(\frac{L}{\pi a_r}\cdot\sin\frac{d}{L}\pi), \qquad (\S.4.1.2)$$

where E_b is the "local" formation energy of domain boundary per unit length, due to, e.g., local bond breaking and distortion. Note that for sufficiently large period *L*, the net formation energy E_f will always become negative so that the periodic stress domains are energetically favored to form. This inidcates that thermodynamically, a single-domain large surface is intrinsically unstable and may spontaneouly break into a periodic array of energy-degenerate stress domains when there is no kinetic limitations. Minimization of E_f with respect to *L* and *d* leads to the thermodynamic equilibrium domain structure with minimum energy, which has the follwoing characteristic length scales [25,26]

$$L = \pi a_r e^{E_b/C+1}, \qquad (\S.4.1.3a)$$

$$d = L/2. \qquad (\S.4.1.3b)$$

b. Experimental Realization and Manipulation of Surface Stress-Domain Patterns

As discussed in Section §.3.3, the dimer-row formation associated with the (2 × 1) reconstruction on Si(001) introduces a highly anisotropic surface stress. (The stress $\sigma_{\parallel}$ along the dimer bond is tensile; the stress $\sigma_{\perp}$ along the dimer row is compressive.) Because the dimer rows

rotate by 90° on terraces separated by a monatomic (or odd number atomic height) step, a stress domain structure may form naturally on a vicinal Si(001) surface, consisting of a staircase of steps. Experiments [51,59] show that on a vicinal Si(001) surface with a small miscut (<1.5°), the misorientation produces only monatomic steps, which separate alternating (2 × 1) and (1 × 2) domains. 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 [51].

The first indication that the alternating the (2 × 1) and (1 × 2) domains constitute a periodic stress-domain structure was provided by an experiment, which showed the dependence of relative population of the two domains in response to an external surface strain, as illustrated in Fig. §.4.1.1b. The external strain was applied at a temperature sufficiently high so that the domain population can change via step motion [51]. Under a compressive external strain, the domain that is compressed along the dimer bond direction is favored with an increased population, indicating that the surface stress along this direction is relative tensile. This experiment was then successfully analyzed using the stress domain model [59].

Let us denote the majority and minority domain size to be respectively $L-d=(1+p)L/2$ and $d=(1-p)L/2$ under the external strain ε^e, the surface energy of the stress domain per area can be calculated as [26]

$$E_f^e=\frac{1}{4}(\sigma_{\square}+\sigma_{\perp})\varepsilon^e+\frac{1}{4}F\varepsilon^e p+\frac{E_b}{L}-\frac{C}{L}\ln(\frac{L}{\pi a_r}\cdot\cos\frac{p}{2}\pi)\,. \qquad (\S.4.1.4)$$

For a given vicinal surface with fixed miscut, the periodicity L, i.e., the average step-step separation is fixed. Minimization of E_f^e in equation §.4.1.4 with respect to p leads to the optimal relative domain population as

$$p=\frac{2}{\pi}\tan^{-1}\left(\frac{\pi\varepsilon^e}{2\varepsilon(L)}\right), \qquad (\S.4.1.5)$$

where $\varepsilon(L)=-\frac{2\pi(1-\nu^2)f}{YL}$. By fitting the measured domain population with the theoretical model, the response of relative domain population to an external strain provides a unique way to determine quantitatively the intrinsic surface stress anisotropy [59]. For Si(001), the anisotropy is fitted to be 0.07±0.01 eV/Å, as obtained from LEED and STM measurements [59], in good agreement with *ab initio* calculations [57,58].

c. Self-Organized Pattern Formation via Stress-Mediated Surface Faceting

The formation of stress-domain patterns has also important technical implications. Surface patterning is an important processing step in many device fabrication processes. The continued drive to make devices smaller and smaller has brought up the challenge to pattern surfaces in the nanometer scale where conventional lithographic techniques are no long applicable. Two different routes have been taken toward nanopatterning: one by developing new patterning techniques with nanometer resolution, such as scanning probes, the other by taking advantage of self-organization of surface patterns, such as stress-domain patterns, which occur naturally. The later approach has shown great promises because it offers an economic and parallel process for device fabrication.

Besides manipulating steps in a vicinal surface, another effective way to create and manipulate stress-domain patterns is by surface faceting process. A faceted surface, similar to a stepped surface, also constitutes naturally a stress-domain structure, as also shown first by theory [82]. The faceted surface stress domains have been demonstrated on a variety of different materials surfaces, such as the Si(001) [83-855], Si(111) [86-88], GaAs(001) [89], Au(111) [90], and Pt (001) [91] surfaces. Here, we review a recent work demonstrating self-organized nanoscale patterns on faceted vicinal Si(111) surface [92].

Ideally, one would like to create a desirable surface pattern with controllable length scales. The characteristic length scales of stress-domain patterns are determined by the competition between domain boundary energy and elastic relaxation energy. Therefore, in principle, it should be possible to control the length scales of such patterns by manipulating these energy terms. However, such kind of control is difficult to achieve in real practice. For example,

one may change the size of stress domains by applying external strain [51,59], but the domains will restore their original size upon relieving the external strain. On a vicinal surface, one may change the terrace size by tuning miscut angle, but this practice is very limited because different step structures [93,94] and facets [87,88] usually form at different miscut angles.

An example has been recently demonstrated [92] for creating the same surface pattern of controllable length scales. It is shown that when the Si(111) surfaces are miscut toward the $[\bar{2}11]$ direction, thermal annealing leads to self-organized formation of 1D surface patterns consisting of a periodic array of alternating flat terraces and high-step-density facets, with very good long-range order. Most important, all the surface patterns exhibit the same atomic structure for both the terraces [of the (7×7) reconstruction] and the facets and the same periodicity, independent of miscut angle; while the width of the facets increases linearly with increasing miscut angle. It is further proposed that such unique self-organized pattern formation is achieved by the interplay of elastic relaxation energy and surface facet energy in a surface faceting transition that involves two transition stages. The first stage, occurring at a higher temperature, originates from the stress-induced instability [25,26]. It involves mass transport over the whole surface, leading to the formation of stress-domain structures with a universal optimal periodicity and facet width (domain population) that minimize the surface elastic energy. The second stage, occurring at a lower temperature, is controlled by energy difference in surface reconstruction and facet structure. It involves mass transport over a shorter range within each period defined in the first stage, forming the lowest-energy facet.

Below, we present some main results of the experiment; the readers are referred to Ref. [92] for experimental details. Figure §.4.1.2 shows STM images of four typical ordered patterns formed on surfaces with a miscut angle of 1°, 2°, 4° and 6°, respectively, consisting of a periodic array of alternating (7×7) terraces and step-bunched facets. To quantitatively characterize the length scales of the patterns, the periodicity (terrace width + facet width) and the facet width are measured as a function of miscut angle, as shown in Figs §.4.1.3a and §.4.1.3b, respectively. The value for a given miscut angle is derived from the peak position of the Gaussian distribution

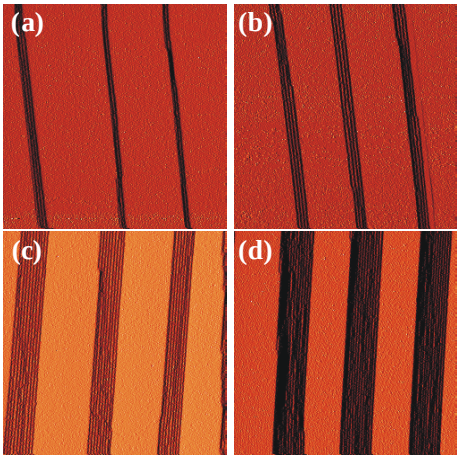

Figure 4.1.2. Ordered patterns on the Si(111) surfaces. Surface morphology of four well-annealed Si(111) samples with miscut angles of 1°, 2°, 4°, and 6° toward $[\bar{2}11]$ direction are shown. The surfaces display an ordered pattern consisting of alternating (7 x 7) reconstructed terraces and high step-density facets. All four images have the size of 2400 Å x 2400 Å. The downstairs direction runs from left to right. To emphasize the step structure in the facet, the derivative mode was used in all four images.

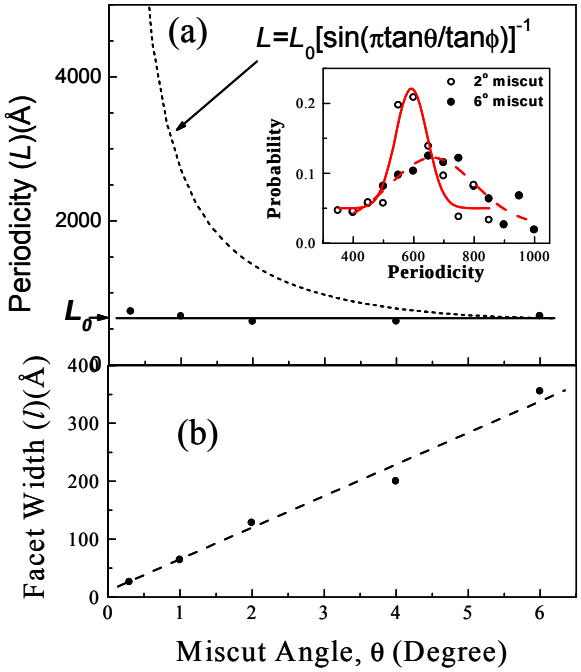

Figure 4.1.3. Dependence of faceting periodicity and facet width on miscut angle. (a) All ordered patterns adopt a universal periodicity of ~635 Å, independent of miscut angle. Values (dots) for different miscuts are obtained from the peak positions of Gaussian fits to the measured data, as shown in the inset for 2° and 6° samples. The dashed line shows the dependence of periodicity on miscut angle if a one-stage faceting transition had occurred. (b) The facet width increases linearly with increasing miscut angle. The straight line is a least squares fit to the data (dots).

fit to the experiment data, as shown in the inset of Fig. §.4.1.3a, demonstrating the preferred periodicity and good degree of long range order. The measured periodicities are 673, 593, 612, and 662 Å at miscut angles of to 1°, 2°, 4°, and 6°, respectively, with an average value of 635 Å.

Clearly, All the patterns have the same periodicity independent of the miscut angle (Fig. §.4.1.3a); while the facet width increases linearly with increasing miscut (Fig. §.4.1.3b). Such a length-scale relationship implies that the facet in all the patterns must have the same facet angle, i.e., facet structure, which is indeed confirmed by high-resolution STM images of the facet. The detailed atomic structure of the mini-terraces in a facet appears disordered and all the facets have an average step-step separation of 28 Å and a step height of 6.3 Å, amounting to a facet angle of 12.7°. The 6.3 Å equals to a double bi-layer step height in Si(111), so a transition from single bi-layer step to double bi-layer step must have occurred in addition to the bunching of steps during faceting. The same facet structure seems to be a special feature associated with azimuthal orientation of the $[\bar{2}11]$ direction. If the Si(111) surface is miscut toward other directions, different facets usually form at different miscut angles [87,88].

These unique self-organized length-scale features are believed to be originated from a faceting transition in which surface stress plays an important role in addition to surface (facet) energy [86]. Specifically, it is proposed that the periodicity and facet width are determined respectively by minimization of elastic energy and facet energy at two different stages in the transition. Figure §.4.1.4 shows schematically such a two-stage faceting transition. At a temperature above transition, the surface consists of a staircase of (1×1) terraces separated by single bi-layer steps (Fig. §.4.1.4a). As temperature is lowered below the transition temperature, the surface starts nucleation of individual domains of the reconstructed (7×7) terraces and then self-organizes into a periodic stress-domain structure [25,26,82], consisting of alternating (7×7) terraces and step-bunched (1×1) terraces (facets) (Fig. §.4.1.4b). The change of total surface free energy per unit area is

$$\Delta F(T,L,l) = \Delta f(T)(1-\frac{l}{L}) - \frac{C}{L}\ln[\frac{L}{\pi a}\sin(\frac{l}{L}\pi)] + \frac{2\gamma}{L}, \qquad (\S.4.1.6)$$

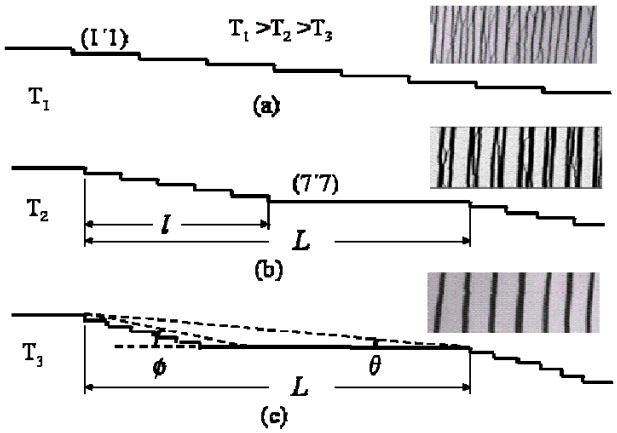

Fig. §.4.1.4. Schematic illustration and STM demonstration of the two-stage faceting transition. (a) The surface morphology above the faceting trasition temperature, consisting of equally spaced single bilayer steps. (b) A half-filling stress-domain structure formed via the first stage, consisting of alternating (7x7) reconstructed terraces and incomplete step-bunched facets. L is the faceting period and l = L/2 is the facet width.(c) The final stable morphology, formed via the second stage, consisting of the stable facet with a facet angle of ϕ = 12.7˚. θ is the surface miscut angle. The corresponding STM images show the surface morphology quenched from a 2˚ sample corresponding to the respective surfaces in Figs. (a), (b), (c), confirming the two-stage transition. They also indicate a single-to-double bilayer step transition accompanying the faceting. All the images have the size of 0.2 μm x 0.5 μm.

where T is the temperature, L is the periodicity, l is the facet width, $\Delta f(T) = f_{(7\times7)} - f_{(1\times1)}$ is the energy difference of the reconstructed (7×7) and the unreconstructed (1×1) terraces, C is a constant related to the difference of surface stress between the (7×7) terrace and the step-bunched facet and elastic constants, γ is the energy cost forming the edge of the facet, and a is a cutoff constant on the order of the (7×7) unit-cell lattice constant.

Minimization of $\Delta F(T, L, l)$ with respect to both L and l leads to

$$\Delta f(T) = -\frac{\pi c}{L}\cot(\frac{l}{L}\pi), \qquad (\S.4.1.7)$$

and $$L = \pi a_0[\sin(\frac{l}{L}\pi)]^{-1}, \qquad (\S.4.1.8)$$

where $a_0 = a\exp(1+\frac{2\gamma}{c})$. Equations (§.4.1.7) and (§.4.1.8) define the conditions and the length scales of the stress domains. $\Delta f(T)$ decreases with decreasing temperature; it is positive above the (1×1)-to-(7×7) thermodynamic phase transition temperature and negative below. The elastic relaxation drives the individual (7×7) terrace of finite size to nucleate at a temperature where $\Delta f(T) = c/(L-l)$ [86], above the (1×1)-to-(7×7) thermodynamic phase transition temperature where $\Delta f(T)$ vanishes. The ordering occurs at the lower temperatures defined by Eqn. (§.4.1.7). Interestingly, the ordering actually starts at a temperature coincides with the thermodynamic phase transition temperature, forming a domain structure with an optimal half-filling domain population, $l = L/2$ [26], as shown above in equation (§.4.1.3b). Correspondingly, from Eqn. (§.4.1.8), the domain adopts a periodicity of $L_0 = \pi a_0$, independent of miscut angle.

Taking surface stress values of 0.1855 eV/Å^2 [95] and -0.039 eV/Å^2 [96] for the (7×7) and the (1×1) terrace, respectively, and a = 47 Å, we estimate, from the observed periodicity (L_0 = 635 Å), a facet edge energy to be ~ 0.0035 eV/Å. Since the facet edge is formed by converting a (1×1) step into a (7×7) step, this value is compared with the step free

energy difference of ~ 0.004 eV/Å [97] between a (7×7) step and a (1×1) step, which is determined independently from step meandering measurements. The good agreement provides a quantitative support of the proposed model.

As the temperature is lowered further, $\Delta f(T)$ becomes negative and continues to decrease. Consequently, the (7×7) terraces in every period increase their sizes, decreasing l while maintaining the periodicity L_0, to reduce the surface energy at the expense of the elastic energy. It proceeds until the most favorable facet structure is formed with a facet angle of 12.7°, as shown in Fig. §.4.1.4c. Because the same facet structure with identical periodicity is formed in all surface patterns, it is mandatory for the facet width to increase linearly with miscut angle, as shown in Fig. §.4.1.3b.

The existence of two transition stages was further confirmed by the following experiment. The surfaces was first cooled down to a given temperature, and then rapidly quenched to low temperature, to reveal the surface morphology frozen in at different temperatures. The STM images in Fig. §.4.1.4 show three typical morphologies quenched from above and below faceting transition temperatures. The surface quenched from above transition temperature (Fig. §.4.1.4a) consists of an array of mostly double steps without facets. The surface quenched from a temperature below the first transition but above the second transition (Fig. §.4.1.4b) consists of a faceted periodic structure. It has the same universal periodicity as in the ideally faceted surface (Fig. §.4.1.4c) but imperfect facets; the overall width of the facets is about half of the periodicity. These indicate that the faceting transition has completed the first stage, defining the optimal periodicity with a half-filling domain structure, but not the second stage. The fast quenching in the second stage makes the second transition, i.e., formation of the stable facet, incomplete. These images also confirm that the faceting transition is accompanied by a single-to-double bi-layer step transition.

There are other arguments that support the proposed two-stage faceting transition. If the final ordered patterns were formed directly without the first stage, energy minimization with respect to periodicity (L) under the constraint of constant facet angle (ϕ =12.7°) would lead to the following relationship of periodicity and miscut angle (θ), $L_0 = \pi a_0[\sin(\frac{\tan\theta}{\tan\phi}\pi)]^{-1}$. Consequently, the periodicity would decrease drastically with increasing miscut angle, as shown in Fig. §.4.1.3a, rather than stay constant as observed.

In summary, the self-organized pattern formation on Si(111) surfaces with miscuts toward the $[\bar{2}11]$ direction demonstrates a unique example of stress-domain structure. The patterns consist of a periodic array of alternating (7×7) reconstructed terraces and high-step-density facets. Most interesting, the length scales of the patterns are controlled by a universal periodicity and a facet width that increases linearly with increasing miscut. Such well-ordered patterns may be used as templates for subsequent nanofabrication.

§.4.2 Self-Assembly of 2D Islands ("Quantum Platelets") at Steps

Recently there has been intense interest in employing self-assembly/self-organization processes to synthesize exotic materials, in particular, those whose structural patterns lie in the nanometer scale, as they form the building blocks for potentially novel electronic and optoelectronic devices. Among different mechanisms, stress/strain induced self-assembly/self-organization have attracted most attention for synthesizing solid-state nanoscale thin-film materials. For example, in heteroepitaxial growth (material B on A) of thin films, misfit strain due to lattice mismatch has been shown to induce self-organization of 3D islands [29,98-100], step bunch arrays [101], step bunch lattices [102], and alloy superlattices [103]. Because the focus of this chapter is on 2D structures, We review here a recent work demonstrating the strain-induced self-organized formation of monolayer-thick 2D islands, the "quantum platelets" [79]. They form at the early stage of heteroepitaxial growth on a substrate with regularly spaced steps, and align along the steps. In the direction perpendicular to substrate steps, the island position and spacing can be preselected through substrate miscut. Along the steps, the island size and density are controlled by self-organized growth.

Usually, the strain induced self-assembled and self-organized patterns and islands have a characteristic length scale, which is uniquely defined by the competition between strain relaxation energy and structural/morphological boundary energy (such as step and surface

energy), as discussed above. If more than one self-organization processes operating at different length scales can be combined, it may be possible to manipulate more than one length scale and to achieve better control of the patterns and islands that form. For example, heteroepitaxial growth can be carried out on a vicinal substrate, which often consists of a self-organized staircase of equally spaced steps that defines one length scale. Then, the self-organization of the deposited material can occur on a finer scale, superimposed on the staircase. Effectively, we have a hybrid approach to direct a self-organization process on a pre-patterned substrate.

a. 2D Analogs of 3D Equilibrium Growth Modes

On a vicinal substrate, one may expect approximate 2D analogs of all the 3D equilibrium growth modes [23,104] [Frank–van der Merwe (FV), Stranski-Krastanov (SK), and Volmer-Weber (VW)], as illustrated in Fig. §.4.2.1. They are determined by the balance between interstep energy (laterally between the film and the substrate steps, equivalent to interfacial energy in 3D) and energies of film steps and substrate steps (equivalent to surface energies in 3D). For homooepitaxy on a vicinal substrate with terrace width W [Fig. §.4.2.1(b)], one would expect "ideal step flow growth," i.e., stripes of the deposited material with width $w = \theta W$ at every step at a film coverage θ. For heteroepitaxy, i.e., if misfit strain is present, we expect, depending on the magnitude of the misfit strain, the formation of a stress domain structure, either with [Fig. §.4.2.1(d), SK growth] or without [Fig. §.4.2.1(f), VW growth] a "wetting stripe" (corresponding to the wetting layer in 3D case).

The 2D equilibrium growth modes on a vicinal substrate can be obtained by minimization of step and interstep energies. (The interstep energy may include the strain energy contribution, similar to the interface energy used in deriving 3D equilibrium growth modes.) Figure §.4.2.2 shows the resulting phase diagram for the 2D equilibrium growth modes, in terms of C_s (substrate step energy), C_f (film step energy) and C_i (interstep energy), which are denoted in Fig. §.4.2.1. For $C_i < C_s - C_f$, the grow proceeds with row-by-row step flow forming a smooth stripe along the step (analogous to layer-by-layer FM growth mode in 3D); for $C_s - C_f < C_i < C_s + C_f$, the grow proceeds with island formation along the step breaking

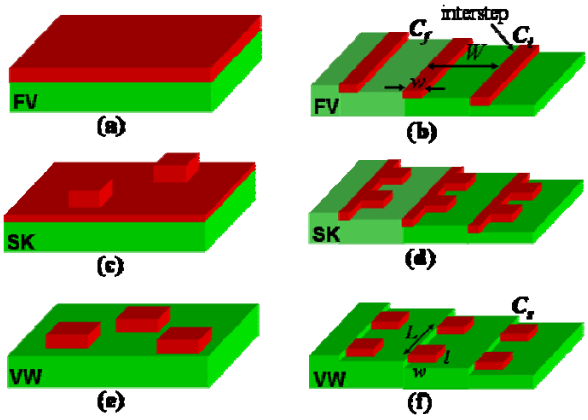

Figure 4.2.1. Two-dimensional analogs of 3D equilibrium growth mode, occurring during step flow growth on a vicinal substrate. (a) 3D Frank-van der Merwe (FV) mode, layer-by-layer growth. (b) 2D FV mode, row-by-row growth. Continuous stripes formed under "ideal" step flow. W is the terrace width of underlying substrate; w is the width of the growing stripes. (c) 3D Stranski-Krastanov (SK) mode, layer-by-layer followed by island growth. (d) 2D SK mode, row-by-row followed by island growth. (e) 3D Volmer-Weber (VW) mode, island growth. (e) 2D VM mode, island growth. To facilitate the energy calculation, the islands are assumed to be rectangular in shape and periodically arrayed. Each island has a size of length l =αL and width w =βW. L is the periodically in the direction along the substrate step.

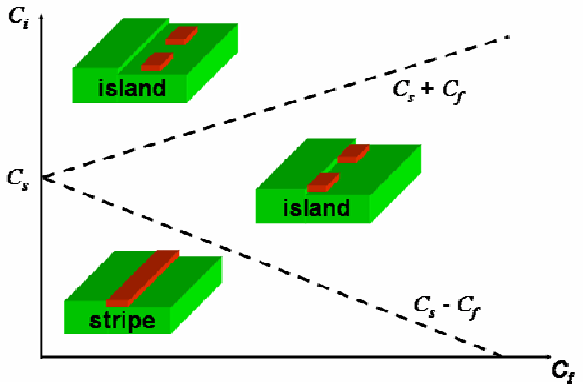

Fig. §.4.2.2 Phase diagram of 2D equilibrium growth modes, in terms of step and interstetp energies. C_s and C_f are respectively substrate and film step energy, C_8 is the interstep energy between the film and substrate.

down the step flow (analogous to islanding WV mode in 3D). If strain energy is included, the interstep energy will increase with increasing coverage so that initially $C_i < C_s - C_f$ and then $C_s - C_f < C_i < C_s + C_f$, in which the growth proceeds with row-by-row followed by island formation (analogous to SK mode in 3D). Another new interesting growth mode occurs for very large interstep energy when $C_i > C_s + C_f$, which leads to 2D island formation in the middle of terrace detaching from the substrate steps, as shown in Fig. §.4.2.2. This scenario can't exist in 3D, as it would mean that 3D islands detach from substrate surface hanging in the air! Or it would correspond to the VW growth of 3D islands having a surface contact angle larger than 180°.

SK and VW growth has been widely used for strain-induced self-organized formation of 3D nanostructures [3]. One may then ask whether a similar approach for nanofabrication can be applied in the growth of strained 2D layers. A recent work [79] demonstrated such a possibility, by both theory and experiment. It is shown that misfit strain between the deposited material and the substrate induces a generic instability against step flow at the initial stage of heteroepitaxial growth on a vicinal surface, leading to the formation of arrays of 2D (monolayer thick) islands along the substrate steps, which look like small plates and named as "quantum platelets". Furthermore, strain induces a long-range elastic interaction between islands on different terraces, as they form simultaneously along all the steps (multiple growth fronts, in contrast to only one growth front in the 3D case). This interaction plays an important role in defining the evolution of island morphology. In the direction perpendicular to steps, the quantum platelets have a perfect positional order defined by substrate miscut; along the substrate steps, the platelet size and density is controlled by the self-organized growth process.

b. Theoretical Formulation

Let's consider the heteroepitaxial growth of a layer on a vicinal substrate with terrace width W [Figs. §.4.2.1(b), (d), and (f)]. If the growth proceeded by "ideal" step flow, stripes of the deposited material with width $w = \theta W$ would form at every step at a film coverage θ [Fig. §.4.2.1(b)]. Misfit strain induces a force monopole on both edges of the stripes with a magnitude proportional to the misfit strain and the step height, as discussed in section 2, forming in effect a stress domain structure. One edge of the stripe is a free step edge with energy C_f, the other is buried, forming an "interstep" with the buried substrate step edge with energy C_i; as indicated in Fig. §.4.2.1(b). Following derivation of equation Fig. §.4.1.2, the total energy per unit area, E_{stripe} can be calculated as

$$E_{stripe} = \frac{1}{W}[C_i + C_f - C\ln(\frac{W}{\pi a}\cdot \sin\theta\pi)] \qquad (\S.4.2.1)$$

where C_i is the "interstep" energy per unit length, i.e., the vertical interface between the "buried" (in a lateral sense) stripe edge and the buried substrate step, C_f is the step energy per unit length of the free edge of the stripe, and C is a constant related to the force monopole and the elastic constants of the substrate. a is a cutoff length of the order of the surface lattice constant.

The strain contribution to E_{stripe} [the last term in Eq. (§.4.2.1)] depends on the coverage θ, which is shown in Fig. §.4.2.3 (solid line). It is very high at very low coverage because of the strong *intrastripe* repulsion between the monopoles on the opposite edges of the same stripe. It decreases with increasing film coverage and reaches a minimum at half-monolayer coverage ($\theta =$ ½). At higher coverages, the energy increases and becomes very high again as the coverage approaches one monolayer, because of the strong *interstripe* repulsion between monopoles on opposing edges of two neighboring stripes.

The initial divergence of strain energy can be greatly relaxed by roughening of the lateral growth fronts, which may occur in different forms. To illustrate the instability, we consider a particular form of relaxation, corresponding to the 2D analog of the VW mode, in which the stripe breaks up into arrays of 2D islands attached to the substrate steps, with no "wetting stripes" between the islands [Fig. §.4.2.1(f)].

We assume islands form a 2D periodic array with a periodicity L along the substrate step direction (in the orthogonal direction, the periodicity is fixed by the terrace width W), with a length $l = \alpha L$ and width $w = \beta W$ [Fig. §.4.2.1(f)]. It is easy to see that $\alpha\beta = \theta,\ \theta \le \alpha \le 1;\ \theta \le \beta \le 1$. Thus, at a given coverage, the island configuration is defined by two independent variables, L and α (or β). It is shown [79] that the optimal island configuration (i.e., optimal values of L and α) must satisfy the following equation:

$$\beta[\pi \cot \beta\pi + \frac{W}{L_0}(\frac{2}{\alpha}\sin \alpha\pi - \cos \alpha\pi] + \ln \frac{\pi L_0}{eW \sin \beta\pi} + \frac{C_i - C_s - C_f}{C} = 0 \qquad (\text{§.4.2.2})$$

where $L_0 = ae^{(C_f/C)+1}$ is the stable size of an isolated island, as derived in section §.2.3 (equation §.2.3.5), and C_s is the substrate step energy. The total energy per unit area of the corresponding island array is

$$E_{island} = \frac{C}{W}[-\beta \frac{W \sin \alpha\pi}{\pi L_0} + \alpha \ln \frac{\pi L_0}{eW \sin \beta\pi} + \frac{\alpha}{C}(C_i - C_s - C_f) + \frac{C_s}{C}]. \qquad (\text{§.4.2.3})$$

As expected, E_{island} in Eq. (§.4.2.3) reduces to E_{stripe} in Eq. (§.4.2.1) for $\alpha = 1$ and $\beta = \theta$.

For systems in which strain energy dominates step energies (i.e., $C >> |C_i - C_s - C_f|$), the island morphology is solely defined by a single parameter, W/L_0, the ratio of the substrate

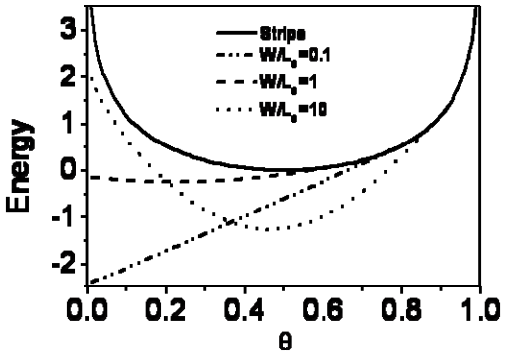

Fig. §.4.2.3 The strain energy per unit area of the stripe configuration (the solid line) and of the island configuration (the broken line) for different values of W| L_0, that is, the ratio of the substrate terrace width (W) and the size of stable isolated island (L_0). The island configuration always has lower strain energy than the stripe configuration at the beginning of the growth (low coverage) for any value of W|L_0.

terrace width, and the stable size of an isolated island. The strain energy of the optimal island array [neglecting the last two terms in Eq. (§.4.2.2)] is plotted in Fig. §.4.2.3 as a function of coverage for different values of W/L_0, in comparison to that of the continuous stripes. The islands have a lower energy than stripes at low coverages, independent of the choice of W/L_0, indicating a generic breakdown of initial step flow.

The evolution of equilibrium island morphology with increasing coverage is determined by energy minimization of stress domains [78,79]. It varies with W/L_0, under the specific boundary conditions of constrained island locations and shapes along the substrate steps. Figure §.4.2.4a shows the calculated island aspect ratio (l/w) as a function of coverage θ for different values of W/L_0, illustrating the island morphological evolution. If the terrace size is much larger than the stable size of a single island, i.e., $W >> L_0$, the interaction between islands on different terraces is initially negligible. Thus, at very low coverage, when $\theta << 1$, $\alpha << 1$, and $\beta << 1$,we have $L = L_0^2 / W\theta$ and $\alpha L = \beta W = L_0$, indicating that all islands initially form with a size of $L_0 \times L_0$, as isolated stable platelets; and the platelet separation along the substrate steps is defined by the coverage, with $L = L_0^2 / W\theta$, which decreases with increasing coverage. As the coverage increases to a point where the separation approaches the optimal value ~L_0, platelets begin to grow *perpendicular* to the substrate steps, in an attempt to form optimally spaced stripes perpendicular to the steps. As the coverage increases further, the repulsion between platelets on neighboring terraces becomes significant, and platelets stop growing perpendicular to the steps. They begin to grow *along* the steps, and gradually merge into continuous stripes, eventually converting to step flow growth. This case is shown in Fig. §.4.2.4a for $W/L_0 = 10$. The island aspect ratio starts with a value of $l/w = 1$ reflecting a square island of $L_0 \times L_0$. With increasing coverage, it first decreases, reflecting growth of island width (w) in the direction *perpendicular* to the substrate steps, then increases, reflecting growth of island length (l) *along* the substrate steps, and finally approaches rapidly toward infinity, reflecting islands merging into continuous stripes.

If the terrace size is comparable to the stable size of a single island, the repulsion between islands on different terraces (i.e., in the β direction) is strong from the very beginning. Consequently, islands initially adopt an optimal length of L_0 along the steps and a width of ~$W/2$ perpendicular to the steps. As the coverage increases, they first increase their density (reducing their separation) and then grow predominantly along the steps until coalescence, after which growth continues via step flow. This case is shown in Fig. §.4.2.4a for $W/L_0 = 1$. The island aspect ratio starts with a value of $l/w = 2$ for an island of $L_0 \times (W/2)$, and increases continuously toward infinity as the islands grow along the steps.

The equilibrium island configuration is determined by minimizing the repulsive interaction between islands. At a given growth stage (coverage), which side of the islands grows preferentially depends on the direction (α or β) in which the island-island repulsion is dominant. Eventually, the interaction between islands on different terraces can effectively suppress the roughening and always drives the islands, beyond a critical coverage below one monolayer, to merge into continuous stripes along the substrate steps. Such a morphological transition from initially rough to smooth represents a kind of inverse SK growth mode. Figure §.4.2.4a

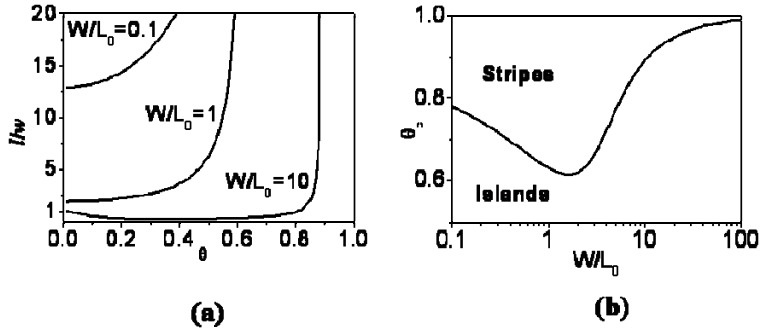

Figure 4.2.4. (a) The 2D island (platelet) shape aspect ratio, l/w, as a function of coverage θ for different values W/L_0, the ratio of the width of substrate terrace and the stable size of a single island. (b) The upper limit of critical coverage θ_c, for transition from rough growth fronts (platelets) to smooth growth fronts (stripes), as a function of W/L_0.

clearly shows that for different values of W/L_0, the island aspect ratio reaches infinity, which defines the critical point of the transition, at different coverages. The critical coverage (i.e., the stripe width) for the transition can be determined by the condition that the minimum of E_{island} is at the configuration $\alpha = 1$ and $\beta = \theta$. From Equation (§.4.2.2), we obtain

$$\theta_c[\pi\cos(\theta_c\pi)+\frac{W}{L_0}]+\ln\frac{\pi L_0}{eW\sin\theta_c\pi}+\frac{C_i-C_s-C_f}{C}=0 \qquad (§.4.2.4)$$

The first two terms on the left-hand side of Equation (§.4.2.4) make a monotonically decreasing function of θ_c between 0 and 1; the last term is negative for islands attaching to the substrate step edges, as shown by Fig. §.4.2.2 and related discussion. Consequently, solving the equation with the last term neglected (i.e., neglecting the step energy contribution) will give the upper limit for the critical coverage θ_c, which is shown in Fig. §.4.2.4b. It is solely determined by W/L_0. The effect of step energy (in particular, a large C_i) will lower the critical coverage.

b. Experimental Observation

The spontaneous island formation along the steps as shown theoretically above provides a effective pathway for growing 2D quantum dots, or quantum "platelets." It represents a natural combination of strain-induced self-organization and a patterned substrate. The islands are confined to form along the substrate steps, so we achieve a nearly perfect control of spacing in the direction perpendicular to the substrate steps. The island morphology, i.e., island size and density (spacing along the substrate steps), can also be controlled by tuning the parameter W/L_0, where W can be adjusted by changing the substrate miscut angle and L_0 can be adjusted by choosing different film and substrate combinations and by changing the film composition when this is feasible.

Below we review briefly an experimental example supporting this picture [79]. It has been shown that large areas of high-quality stepped surfaces on vicinal Si(111) can be reproducibly obtained using a special preparation procedure [105]. As an example of a strained heteroepitaxial system, CaF was deposited on such vicinal Si(111) substrates. The lattice mismatch is 0.6% at room temperature and 2.4% at typical growth temperatures of 600 ~ 650 ^{o}C [79]. Si(111) samples miscut 1^{o} towards $(\bar{1}\bar{1}2)$ direction were used as substrates. CaF_2 was deposited from a resistively heated BN cell at a growth rate of approximately 1 monolayer (ML) per minute.

Figure §.4.2.5 shows STM images of two typical growth morphologies of CaF_2 on vicinal Si(111). At low coverages, the growth results in strings of islands attached to step edges (Fig. §.4.2.5a), i.e., something resembling a 2D VW mode. The islands remain at nearly constant width of half of the substrate terrace width, as they grow along the substrate steps and merge into continuous stripes (Fig. §.4.2.5b). This experiment corresponds directly to the theory. The constant island width of half of the substrate terrace width before islands merge indicates the system is likely in the regime of $L_0 \sim W$.

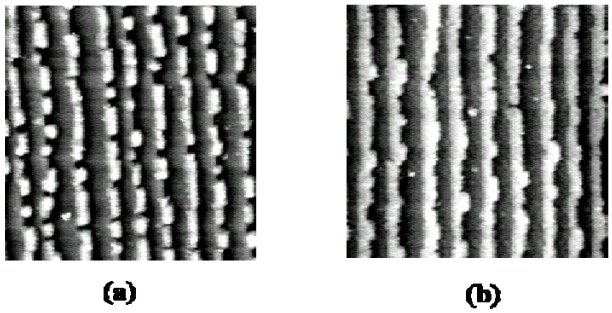

Figure 4.2.5. STM images of CaF on stepped Si (111), showing the transition from platelets to stripes. The two samples are prepared under the same conditions, except for the coverage, 0.23 ML in (a) and 0.40 ML in (b).

At high coverages, the step flow growth mode is observed for the growth conditions described above. These continuous stripes are obtained only for coverages above ~0.4 ML. This feature arises also in the calculation. Thus, the rough-to-smooth morphological transition beyond a critical coverage observed in the experiment is in good agreement with the unusual inverse SK growth mode predicted by the above theory. The measured critical coverage (~0.4 ML) is below the predicted theoretical *minimum* upper limit (~0.6 ML), indicating a large interstep energy between CaF and Si(111). The experimentally measured critical coverage is increased to ~0.5 ML when the substrate terrace width is decreased by a factor of ~2 by doubling the substrate miscut angle. This further indicates that the system now falls into the regime of $W < L_0$ (i.e., on the left side of the minimum in Fig. §.4.2.5b).

In summary, there exists a strain-induced instability at the early stage (submonolayer coverage) of heteroepitaxial growth on vicinal substrates. The growth fronts are initially rough, forming 2D platelike islands to reduce strain energy. As the growth proceeds, the strain-induced interaction between the islands formed on different terraces eventually smoothens the growth front whereupon step flow growth begins. This latter morphological transition, from rough to smooth, is a novel aspect of growth with multiple laterally interacting growth fronts. The upper limit of the critical coverage is solely determined by the ratio of the substrate terrace width and the stable size of an isolated island under strain. It is expected that this growth mode, in its initial form, may be useful for fabricating arrays of quantum platelets with potentially novel electronic properties. At the later stage, the critical smoothening phenomenon may be used for growing quantum wires with a control of minimum wire width. In particular, the smoothening mechanism tends to make the wires smoother.

§.4.3 Self-Organization of 2D Islands via Strain-Mediated Coarsening

So far, we have presented theoretical analysis of self-assembly and self-organization of surface patterns and 2D islands through total-energy minimization to determine the optimal ordering and length scales of self-assembled and self-organized periodic patterns and 2D island arrays. In this section, we will present an example of computer simulation of self-organization process of 2D islands, in particular, the strain-mediated island coarsening process. The simulation will allow us to study the dynamics of self-organized ordering process of 2D islands beyond theoretical modeling of their static properties.

Coarsening, a fundamental process in the formation of a condensed phase of islands from a supersaturated 2D vapor phase of adatoms, plays a critical role in determining the island size distribution during and/or after the growth of islands by redistributing mass between islands in accordance with their chemical potentials. Mass flows from high-chemical-potential to low-chemical-potential regions. The recent interest in studying coarsening behavior of strained 2D and 3D islands in heteroepitaxial growth [28,29,78-81,100,106-120] is driven by the desire to understand the formation of islands that have a narrow size distribution, a prerequisite condition for potential use of these islands as quantum dots. Coarsening can help to achieve island size uniformity, aided either thermodynamically by a strain-induced equilibrium stable island size [29,78-81,100] or kinetically by self-limiting effects [114-117]. Most studies have focused on the influence of strain on island size stability; while less attention has been paid to its influence on the actual island size and spatial distribution. Usually, initial configurations of islands with a uniform size and spatial distribution have been assumed. For example, thermodynamic analyses we presented above as well as in many literature, whether 2D [78-81] or 3D [29,112], have been carried out on a periodic array of equal-sized islands to investigate their stability against coarsening (Ostwald ripening) to form a sparser array of equal-sized larger islands or a single larger island.

In actual film growth, islands of course nucleate randomly at different positions and grow into different sizes. This initial randomness can have significant effects on the coarsening behavior in the presence of strain, especially when all physical manifestations of strain are included. To study a random array of islands, it is much more effective to carry out computer simulations than to perform theoretical analyses. In this section, we review one recent simulation [80], which demonstrates how strain-induced island-island interactions influence

coarsening of a random array of islands. These effects will manifest themselves *if and only if* the initial distribution of island size and spacing is random. It is shown that strain can influence not only the mass transport between islands but also the mass transport within 2D islands, thereby triggering a self-organization that narrows the island size distribution. The strain-induced island-island interaction manifests itself in two different influences on coarsened island size and spatial distribution, depending on the form of mass transport involved in the coarsening process. If only strain modified mass transport between 2D islands is added to the coarsening process, the strain-induced interaction *broadens* the island size distribution. The width of the broadening increases with decreasing island number density, in a power-law dependence. If mass transport within an island (i.e., via island edge diffusion; hence islands are effectively migrating during coarsening) is included in addition to mass transport between islands, the interaction can influence the direction of island edge diffusion and effectively lead to a strain-directed island motion. This motion, combined with a thermodynamic driving force to reach the strain-induced stable island size, can then drive a random array of 2D islands to self-organize, forming a triangular lattice of islands with greatly improved uniformity in island size and spacing.

In general, island coarsening may proceed with or without island migration. Without island migration, coarsening is governed solely by chemical potentials of individual islands, which then control the mass transport between islands. Without strain, the chemical potential depends only on island size and bond energies. Strain induces an additional term in the chemical potential. Consider an array of coherently strained 2D circular islands, as shown in Fig. §.4.3.1, formed at the early stage of heteroepitaxial growth in the submonolayer regime. The total energy of a single island (neglecting the island-island interaction for now) can be calculated as the sum of the step energy and island strain energy (see derivation of equation §.2.3.10 in section §.2.3),

$$E_i^{\sin}(R_i) = (2\pi R_i)[\gamma - C\ln\frac{R_i}{a}], \qquad (\S.4.3.1)$$

where R_i is the radius of the i-th island and γ the island edge (step) energy per unit length. $C = (1-\nu^2)F^2/(\pi Y)$ (note a minor error was made in this constant in the original publication [80]), which is related to the misfit strain-induced elastic force monopole (F) along the island edge and elastic constants ν (Poisson ratio) and Y (Young's modulus). a is a normalized cutoff length, of the order of a surface lattice constant. The chemical potential of a single island, as dervied in section §.2.3 (equation §.2.3.12), is then

$$u_i^{\sin}(R_i) = s\left[\frac{\gamma - C}{R_i} - \frac{C}{R_i}\ln\frac{R_i}{a_0}\right], \qquad (\S.4.3.2)$$

where s is the surface area per atom in the island.

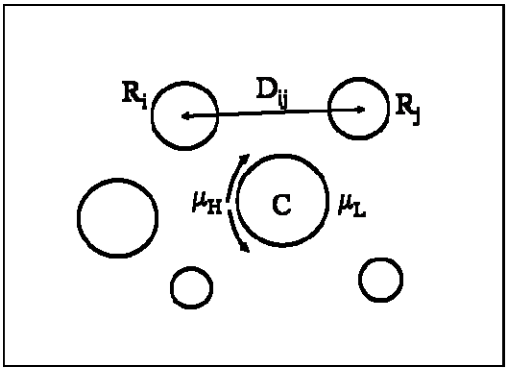

Fig. §.4.3.1. Schematic diagram of an array of circular 2D islands. R_i and R_j denote the radius of the ith and jth island, respectively, with a separation D_{ij}. The island at the center labeled "C" has three neighbors on the left and two on the right. The chemical potential on its left-side edge (u_H) is higher than on the right-side edge (u_L) (i.e., $u_H > u_L$), because of the strain-induced island-island interaction. The two arrows on its left edge indicate the direction of net mass flow via edge diffusion, causing effectively an island drifting toward the right.

Without strain (as in homoepitaxy), the island chemical potential is inversely proportional to its size, and coarsening leads to a continuous increase of average island size indefinitely. Strain introduces an island size that is thermodynamically stable against further coarsening. (While a strain-induced equilibrium stability always exists for 2D islands [78-81] as shown in section §.2.3, it may or may not exist for 3D islands, depending on whether the 3D islands have a negative surface energy in reference to the wetting layer and/or a strong island edge effect [29].) From Equation (§.4.3.2), the island chemical potential has a minimum of $u_0 = (sC)/R_0$, at $R_0 = a_0 e^{\gamma/C}$. Consequently, coarsening will tend to drive all the islands toward this stable size, forming a uniform island size distribution. We note that at finite temperatures, the contribution of entropy broadens the island sizes into a Gaussian distribution [111] and moves the mean island size away from R_0 [120].

Because strain not only contributes to the total energy of a single island, but also induces an elastic island-island interaction between islands, it will further modify the island chemical potential and hence changes the coarsening behavior of an island array. The strain energy of a single island is calculated by integrating the interaction between strain-induced force monopoles along its edge. Similarly, the island-island interaction energy can be calculated by integrating the interaction between force monopoles along the edges of different islands. In an island array, this additional interaction energy contribution to the *i*-th island can be calculated as derived in section §.2.3 (Eqn. §.2.3.14):

$$\Delta E_i^{\text{int}}(R_i) = (\pi\beta)\sum_{j\neq i}\frac{R_i^2 R_j^2}{D_{ij}^3}, \qquad (\S.4.3.3)$$

where R_j is the radius of the *j*-th island, D_{ij} is the separation between the *i*-th and *j*-th island, and $\beta = (1-\nu^2)F^2/Y$ (note the error of the sign and a factor of 2 in the original paper [80]). To arrive at Equation (§.4.3.3), we assume the size, R, of islands is much smaller than their separation, D, so that the higher-order interaction terms are neglected. Correspondingly, the correction to the chemical potential of the *i*-th island is

$$\Delta u_i^{\text{int}}(R_i) = s\beta\sum_{j\neq i}\frac{R_j^2}{D_{ij}^3}. \qquad (\S.4.3.4)$$

The interaction makes the chemical potentials of the islands dependent on their environment (i.e., the size and position of neighboring islands). Because the islands, in general, nucleate at random positions during epitaxial growth, with different neighboring island configurations, they will converge in a coarsening process to different stable sizes, R_i, rather than to the same size, R_0, for strained islands without island-island interactions.

With this strain-induced island-island interaction included, the equilibrium island size distribution is broader than that obtained without considering the interaction. As coarsening drives the system toward equilibrium, all the islands reach the same chemical potential, so

$$u_i^{\text{sin}}(R_i) + \Delta u_i^{\text{int}}(R_i) = const.. \qquad (\S.4.3.5)$$

For $R << D$, the limit we are considering, we have $\Delta u_i^{\text{int}} << u_i^{\text{sin}}$, i.e., the interaction is a small correction to the total island chemical potential. Consequently, the stable island size including the strain interaction, R_i, is not too far away from the minimum, R_0, for coarsening without the strain-induced island-island interaction (but including the strain of the single island). We can approximate u_i^{sin} using its Taylor expansion around the minimum up to the second order,

$$u_i^{\text{sin}}(R_i) = u_i^{\text{sin}}(R_0) + \frac{1}{2}\left(\frac{d^2 u_i^{\text{sin}}}{dR_i^2}\right)\Bigg|_{R_i=R_0}\cdot(R_i - R_0)^2 = u_0 + c_2(\Delta R)^2 \quad . \qquad (\S.4.3.6)$$

Furthermore, because Δu_i^{int} decays as D^{-3}, we take only the nearest-neighbor interaction,

$$\Delta u_i^{\text{int}}(R_i) = \nu\beta\sum_{j\neq i}\frac{R_{NN}^2}{D_{NN}^3} = k\frac{\overline{R}^2}{\overline{D}^3} = k'\overline{R}^2 n^{3/2}. \qquad (\S.4.3.7)$$

To derive Equation (§.4.3.7), we assume the islands are distributed on the surface with a spatially uniform average number density of n. From dimensional arguments, we have $\overline{D} \propto \sqrt{1/n}$; or more rigorously, we can calculate the probability of finding a nearest neighbor at distance D for a random distribution to be $P(D) = 2\pi nD\exp(-\pi nD^2)$, which leads also to $\overline{D} \propto \sqrt{1/n}$. Substituting Equations (§.4.3.6) and (§.4.3.7) into (§.4.3.5), we have

$$u_0 + c_2(\Delta R)^2 + k'\overline{R}^2 n^{3/2} \approx u_0, \qquad (§.4.3.8)$$

and

$$\frac{\Delta R}{\overline{R}} \sim n^{3/4}. \qquad (§.4.3.9)$$

Therefore, the width of the island size distribution (i.e., the size deviation, ΔR, from R_0) caused by the strain-induced island-island interaction scales with island number density in a power law with an exponent of $3/4$. As the island density increases, the islands get closer, and the island-island interactions get stronger. Consequently, the islands are driven farther away from the stable size without island-island interaction, R_0, and the width of their size distribution, ΔR, increases.

Computer simulations of coarsening of strained 2D islands confirm quantitatively the power law dependence of the width of the island size distribution on number density. In a standard coarsening process, in which mass transport between islands is limited by attachment/detachment of adatoms to/from island perimeters, the rate of change of island size can be expressed as

$$\frac{dR_i}{dt} = C_{AD}(\overline{u} - u_i), \qquad (§.4.3.10)$$

where C_{AD} is a coefficient related to the adatom attachment/detachment rate and the atomic area. $\overline{u}$ is the mean chemical potential averaged over all the islands. We start the simulations with a random distribution of island sizes and positions, to mimic the initial island nucleation and subsequent growth. At each time step, the chemical potential of each island is calculated, including the contribution from island-island interactions, and the island sizes are updated according to Equation (§.4.3.10). Periodic boundary conditions are employed in calculating the island-island interaction, using a cutoff distance slightly larger than half of the simulation cell size. (Because the interaction scales inversely with the third power of island separation, the convergence is guaranteed in two dimensions.)

Figures §.4.3.2a and §.4.3.2b show the initial and final island configurations, respectively, from a typical simulation run. Figure §.4.3.3a shows the time evolution of the island size distribution. At time $t = 0$, the island sizes are randomly chosen between 0 and 100. As time proceeds, small islands evaporate quickly, while medium and large islands converge into a narrow

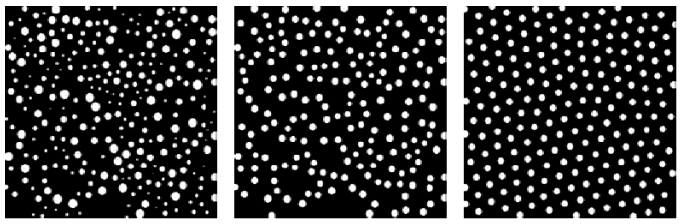

Fig. §.4.3.2 Snapshots of simulated island configurations. (a) The initial configuration with random island sizes and positions. (b) The final converged island configuration from coarsening simulations *without* island migration, using the initial configuration of (a). (c) The final converged island configuration from coarsening simulations *with* island migration, using the initial configuration of (a). The islands form a triangle lattice with much improved size and spatial uniformity relative to those in (b).

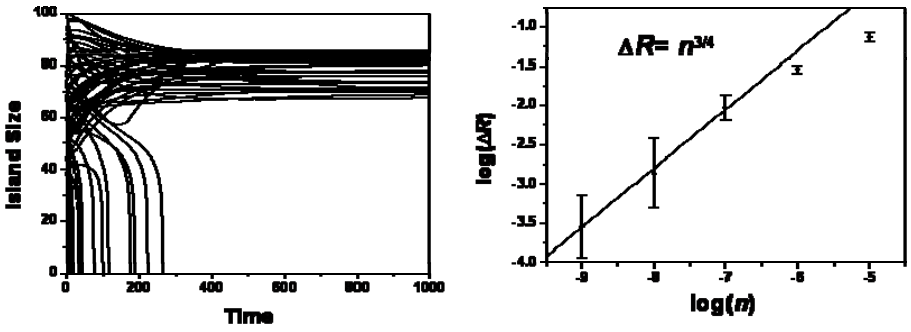

Fig. §.4.3.3. (a) Time evolution of the island size distribution from a typical simulation of coarsening without island migration. The initial size distribution is represented by the y-intercepts at $t = 0$. Islands may grow (shrink) first then shrink (grow). In general, the smallest islands disappear (dissolution), while larger islands may either grow or shrink. All converge toward a finite size distribution centered around R_0=76. (b) Log-log plot of the width of the island size distribution as a function of island number density. The straight line is a linear fit to the simulation data using the slope of ¾. The error bars on the data reflect averaging over many different runs.

size distribution around the optimum size, $R_0 = 76$. For a given island number density, the width of the converged island size distribution is determined by averaging over a great number of runs starting with different initial island configurations of the same number density. Figure §.4.3.3b shows the simulated width of the island size distribution as a function of island number density. In the low-density limit, the width scales with density in a power law with an exponent of $3/4$, in excellent agreement with the analytical solution. The power law fails at higher densities, because the assumption that island separation is much smaller than island size no longer holds and higher-order interaction terms become significant.

Next, we consider coarsening with island migration. The strain induced island-island interaction not only changes the average chemical potential of an island, but also introduces a chemical potential gradient within the island. In a random island array, the configuration of neighboring islands of a given island is, in general, anisotropic. For example, in Fig. §.4.3.1, there are three neighboring islands on the left side of the center island (labeled island C) but only two islands on the right. Consequently, the center island feels a stronger interaction from the left and the chemical potential is higher along its left edge (u_H) than along the right edge (u_L), as indicated in Fig. §.4.3.1. As the island-island interaction is repulsive, it gives rise to a positive contribution to the chemical potential. This chemical-potential gradient will then direct a net mass flow from the left edge to the right edge on the center island, if edge diffusion is activated (comparable to surface diffusion) during coarsening, effectively driving the island to move toward the right.

We have incorporated such strain-directed island motion into simulations to investigate its influence on coarsening. At each time step, in addition to mass exchange between islands that is determined by the average island chemical potential, we calculate the gradient of chemical potential at the center of each island and let the island drift along the gradient direction. The speed of island motion is set to be proportional to the magnitude of the gradient and inversely proportional to the island radius. (A larger island moves more slowly than a smaller island, because it takes longer for the atoms to travel the distance of the island perimeter for an island to move one atomic unit.) Fig. §.4.3.2c shows the final island configuration simulated from the initial configuration of Fig. §.4.3.2a. The islands form a triangular lattice, neglecting the effect of substrate symmetry as isotropic step energy and diffusion are assumed in the simulation, with much improved spatial and size uniformity, compared to the final configuration in Fig. §.4.3.2b, which was simulated from the same initial configuration (Fig. §.4.3.2a) but without island motion.

The obvious degeneracy of several domains of triangular lattice with different orientations indicates that the ordering originates locally in different regions, forming patches of different orientations; domain boundaries (or dislocations) are formed when these patches meet. The much improved island size uniformity correlates with the spatial uniformity. As the repulsive island-island interaction drives the island into a close-packed ordered array, the local

environment of an island becomes more uniform; every island has a similar hexagonal neighboring island configuration and hence converges toward an almost equal stable island size.

This self-organization, induced by strain-directed island motion during coarsening, provides a plausible mechanism and pathway for the formation of a triangular lattice of Ag vacancy islands on Ru(0001) observed in experiment [106]. It is shown experimentally that an isolated vacancy island is very mobile and the interaction between vacancy islands is elastic in nature, conclusions that are completely consistent with the physical assumptions of our model. The simulations show that the self-organized ordering is more effective at higher island number densities when the island-island interaction is stronger, also in good agreement with experimental observations [106] that islands only form a triangular lattice at sufficiently high island density with much improved size uniformity. The fast-rising island-island repulsion at short distance effectively suppresses the coalescence of islands during coarsening, as observed in both experiment and simulation. The island size distribution is broader in the experimental lattice than in the simulated lattice, primarily because thermal broadening is not included in the simulation.

In summary, in this section we illustrate how theoretical analysis and computer simulation can be combined for investigating self-organization of random 2D islands via strain-mediated island coarsening process. We show that strain-induced island-island interactions can have distinctive effects on two standard coarsening processes in arrays of 2D islands. For coarsening without island motion, the interaction gives rise to a universal power-law dependence of the island size distribution on island number density with an exponent of $3/4$ when the island density is low. So far, there seemed no experiments available to be compared to this prediction. For coarsening with island motion, via island edge diffusion, the interaction directs the island motion, leading to the self-organized formation of a regular lattice of islands with uniform island size. The simulated self-organization explains the formation of a triangular lattice of Ag vacancy islands on Ru(0001) [106]. It is also likely to be a general mechanism in other strain-driven growth systems. Practically, it provides a unique method for creating nanoscale templates with uniform size and spacing for patterning and growth of nanostructures.

§.5 Concluding Remarks

In this chapter, we have focused our discussion on 2D patterns and islands. However, the basic physical concepts and principles as well as mathematical treatments and derivations we present here can be readily applied to 3D islands [121]. For example, for a strained 2D island, the elastic force monopoles are distributed around the island boundary, as shown in Fig. §.2.3.1, while for a strained 3D island, the elastic force monopoles are distributed over the island surface, as shown in Fig. §.2.1.1d. Consequently, the elastic energy of a 2D island is calculated by integration of force monopole-monopole interactions along the 2D island boundary as discussed in section §.2, while the elastic energy of a 3D islands can be calculated by integration of force monopole-monopole interactions over the 3D island surface [121]. Another class of related problems we didn't discuss is the step-flow growth of strained films on vicinal surfaces, as illustrated in Fig. §.2.1.1c, which have important implications in growth of quantum wires and quantum-wire lattices [101-103].

So far, majority studies of strain induced self-assembly and self-organization of 2D and 3D islands have been performed in the framework of island stability by analyzing the island total energy and/or chemical potential as a function of island size, shape, and density. Therefore, these are mostly static thermodynamic studies. More future work is needed for developing theoretical and computational/simulation methods to study the dynamics and kinetics of strained 2D and 3D island growth and ordering process. One emerging approach for growth of strained islands is to combine self-assembly with surface patterning, i.e., the directed or guided self-assembly [122]. Experiments have shown great improvements in both island spatial ordering and in island size uniformity [122]. However, the physical mechanisms underlying the directed self-assembly on patterned surfaces, especially the interplay between strain relaxation and the patterned surface/interface morphology, are very much unclear. It is essentially an open area for future theoretical/computational study.

Finally, I would like to thank many of my former and current students, postdocs, and co-workers for their invaluable contribution to the writing of this chapter. Special thanks to M. Lagally and J. Tersoff, who encouraged and helped me to get into this fascinating research field, for their many years of collaboration. Thanks to A. Li, N.Y. Ma, A. Pradhan, Y. Zhao, and V. Zielasek for their contribution to some of the work described in this chapter. Special thanks to N.Y. Ma for deriving and re-deriving some mathematical formula presented in this chapter. Thanks to M. Lagally, F.K. Men and F. Himpsel for their experimental images and collaboration. Also, I would like to acknowledge US Department of Energy and National Science Foundation for their continued funding support to our research program.

References

1. M. F. Crommie, C. P. Lutz, D. M. Eigler, Science 262, 218 (1993).
2. H. C. Manoharan, C. P. Lutz, D. M. Eigler, Nature 403, 512 (2000).
3. For an earlier review on Ge island growth on Si, see Feng Liu and M. G. Lagally, Surf. Sci. 386, 169 (1997).
4. D.J. Eaglesham and M. Cerullo, Phys. Rev. Lett. 64, 1943 (1990).
5. Y.W. Mo, D.E. Savage, B.S. Swartzentruber, and M.G. Lagally, Phys. Rev. Lett. 65, 1020 (1990).
6. D. Leonard M. Krishnamurthy, C.M. Reaves, S.P. Denbaars, and P.M. Petroff, Appl. Phys. Lett., 63, 3203 (1993).
7. J.M. Moison, F. Houzay, F. Barthe, L. Leprince, E. André, O. Vatel, Appl. Phys. Lett., 64, 196 (1994).
8. R. Nötzel, T. Fukui, H. Hasegawa, J. Temmyo and T. Tamamura, Appl. Phys. Lett. 65, 2854 (1994).
9. V. Bressler-Hill, S. Varma, and W.H. Weinberg, Phys. Rev. Lett. 74, 3209 (1995).
10. Q. Xie, A. Madhukar, P. Chen, and N.P. Kobayashi, Phys. Rev. Lett. 75, 2542 (1995).
11. P.W. Deelman, T. Thundat, and L.J. Schowalter, Appl. Surf. Sci. 104/105, 510 (1996).
12. J. Tersoff, C. Teichert, and M.G. Lagally, Phys. Rev. Lett. 76, 1675 (1996).
13. C. Teichert, M.G. Lagally,L. J. Peticolas, J. C. Bean, and J. Tersoff, Phys. Rev. Lett. 76, 1675 (1996).
14. G.R. Carlow and M. Zinke-Allmang, Phys. Rev. Lett. 78, 4601 (1997).
15. S. Lee, I. Daruka, C.S. Kim, A.L. Barabási, J.L. Merz, and J.K. Furdyna, Phys. Rev. Lett. 81, 3479 (1998).
16. G. Medeiros-Ribeiro, A.M. Bratkovski, T.I. Kamins, D.A. Ohlberg, R.S. Williams, Science 279, 353 (1998).
17. F. M. Ross, J. Tersoff, and R. M. Tromp, Phys. Rev. Lett. 80, 984 (1998).
18. J.A. Floro, G.A. Lucadamo, E. Chason, L.B. Freund, M. Sinclair, R.D. Twesten, and R.Q. Huang, Phys. Rev. Lett. 80, 4717 (1998).
19. Feng Liu and M.G. Lagally, in *Precision Science and Technology for Perfect Surfaces*, Y. Furukawa, Y. Mori, and T. Kataoka (eds.) Japan Society for Precision Engineering, Tokyo, Japan (1999).
20. Van der Merwe J.H., J. Appl. Phys. 34, 117; 34, 123 (1963).
21. J.W. Matthews and A.E. Blakeslee, J. Cryst. Growth, 29, 273; 32, 265 (1975).
22. Feng Liu, F. Wu, and M. G. Lagally, Chemical Reviews, 97, 1045 (1997).
23. Feng Liu, M. Hohage, and M. G. Lagally, in *Encyclopedia of Appl. Phys*., H. Immergut and G. Trigg (eds.), Supplement Volume, 321-352 (1999).
24. L.D. Landau and E.M. Lifshitz, *Theory of Elasticity*, Pergamon Press, (1959).
25. V.I. Marchenko, Sov. Phys. JEPT Lett. 33, 381 (1981).
26. O. L. Alerland, D. Vanderbilt, R. D. Meade, and J. D. Joannopoulos, Phys. Rev. Lett. 61, 1973 (1988).
27. J. Tersoff, Phys. Rev. Lett. 80, 2018 (1998).
28. J. Tersoff and F.K. LeGoues, Phys. Rev. Lett. 72, 2570 (1994).
29. V. A. Shchukin, N. N. Ledentsov, P. S. Kop'ev, and D. Bimberg, Phys. Rev. Lett. 75, 2968 (1995).
30. V. I. Marchenko and A. Y. Parshin, Sov. Phys. JETP 52, 129 (1980).
31. K.H. Lau and W. Kohn, Surf. Sci., 65, 607 (1977).
32. A. Li, Feng Liu, and M. Lagally, Phys. Rev. Lett. 85, 1922 (2000).
33. Y.W. Mo, B.S. Swartzentruber, R. Kariotis, M.B. Webb, and M.G. Lagally, Phys. Rev. Lett., 63, 2393 (1989).
34. S. Günther, E. Kopatzki, M.C. Bartelt, J.W. Evans, R.J. Behm, Phys. Rev. Lett., 73, 553 (1994).
35. Y.W. Mo, J. Kleiner, M.B. Webb, and M.G. Lagally, Phys. Rev. Lett., 268, 275 (1992).
36. R.Q. Hwang, J. Schröder, C. Günther, and R.J. Behm, Phys. Rev. Lett., 67, 3279 (1991).
37. Z.Y. Zhang and M.G. Lagally, Science 276, 377 (1997).
38. Y.W. Mo, J. Kleiner, M.B. Webb, and M.G. Lagally, Phys. Rev. Lett. 66, 1998 (1991).
39. W. Świcęch and E. Bauer, Surf. Sci. 255, 218 (1991).
40. N.C. Bartelt, R.M. Tromp, and E.D. Williams, Phys. Rev. Lett. 73, 1656 (1994).
41. D.C. Schlösser, L.K. Verheij, G. Rosenfeld, and G. Comsa, Phys. Rev. Lett. 82, 3843 (1999).
42. V. Zielasek, Feng Liu, Y. Zhao, J.B. Maxson and M.G. Lagally, Phys. Rev. B 64, R201320 (2001).
43. A. Pradhan, N.-Y. Ma and Feng Liu, Phys. Rev. B, in press, (2004).
44. G.Z. Wulff, Kristallogr. 34, 449 (1901).
45. W.K. Burton, N. Cabrera and F.C. Frank, Phil. Trans. Roy. Soc. London A, 243, 299 (1951).
46. J.J. Eggleston, G.B. McFadden and P.W. Voorhees, Physica D, 150, 91 (2001).

47. J. Tersoff and R.M. Tromp, Phys. Rev. Lett. 70, 2782 (1993).
48. W.C. Johnson and J.W. Cahn, Acta Metall. 32, 1925 (1984).
49. M.E. Thompson, C.S. Su and P.W. Voorhees, Acta Metall. Mater. 42, 2107 (1994).
50. Feng Liu and M.G. Lagally, Phys. Rev. Lett. 76, 3156 (1996).
51. F.K. Men, W.F. Packard, and M.B. Webb, Phys. Rev. Lett. 61, 2469 (1988).
52. W. Theis, N.C. Bartelt, and R.M. Tromp, Phys. Rev. Lett. 75, 3328 (1995).
53. A. Pimpinelli and J.J. Metois, Phys. Rev. Lett. 72, 3566 (1994).
54. D.J. Chadi, Phys. Rev. Lett. 59, 1691 (1987).
55. B.S. Swartzentruber, Y.W. Mo, R. Kariotis, M.G. Lagally, and M.B. Webb, Phys. Rev. Lett. 65, 1913 (1990).
56. H.J.W. Zandvliet, H.B. Elswijk, E.J. van Leonen, and D. Dijkkamp, Phys. Rev. B 45, 5965 (1992).
57. A. Garcia and J.E. Northrup, Phys. Rev. B 48, 17350 (1993).
58. J. Dabrowski, E. Pehlke, and M. Scheffler, Phys. Rev. B 49, 4790 (1994).
59. M.B. Webb, F.K. Men, B.S. Swartzentruber, R. Kariotis, and M.G. Lagally,Surf. Sci. 242, 23 (1991).
60. J.B. Hannon, N.C. Bartelt, B.S. Swartzentruber, J.C. Hamilton, and G.L. Kellogg, Phys. Rev. Lett. 79, 4226 (1997).
61. M.T. Middel, H.J.W. Zandvliet, B. Poelsema, Phys. Rev. Lett. 88, 196105 (2002).
62. H.J.W. Zandvliet, R. van Gastel, O. Gurlu, B. Poelsema, Phys. Lett. A, 326, 457 (2004).
63. V.P. LaBella, Z. Ding, D.W. Bullock, C. Emery, P.M. Thibado, J. Vac. Sci. & Tech., B19, 1640 (2001).
64. G.E. Thayer, J.B. Hannon and R.M. Tromp, Nature Materials, 3, 95 (2004).
65. R. van Moere, H.J.W. Zandvliet, B. Poelsema, Phys. Rev. B 67, 193407 (2003).
66. D.J. Bottomley, H. Omi and T. Ogino, J. Cryst. Growth, 225, 16 (2001).
67. J. Wang, M. Li, E.I. Altman, Surf. Sci. 560, 12 (2004).
68. G.J. Xu, S.V. Khare, K.S. Nakayama, C.M. Aldao, J.H. Weaver, Phys. Rev. B 68, 235318 (2003).
69. D. Sander, S. Ouazi, V.S. Stepanyuk, D.I. Bazhanov, J. Kirschner, Surf. Sci. 512, 281 (2002).
70. H.L. Meyerheim, D. Sander, R. Popescu, J. Kirschner, O. Robach, S. Ferrer, P. Steadman, Phys. Rev. B 67, 155422 (2003).
71. D. Loretto, F.M. Ross and C.A. Lucas, Appl. Phy. Lett, 68, 2363 (1996).
72. K.L Kavanagh, M.C Reuter and R.M Tromp, J. Crys. Growth 173, 393 (1997).
73. S.H. Brongersma, M.R. Castell, D.D. Perovic, and M. Zinke-Allmang, Phys. Rev. Lett, 80, 3795 (1998).
74. C. Preinesberger, S. Vandre, T. Kalka and M. Dahne-Prietsch, J. Phys. D 31, L43 (1998).
75. Y. Chen, D.A.A. Ohlberg, G. Medeiros-Ribeiro, Y.A. Chang, and R.S. Williams, Appl. Phy. Lett, 76, 4004 (2000).
76. J. Nogami, B.Z. Liu, M.V. Katkov and C. Ohbuchi, Phys. Rev. B, 63, 233305 (2001).
77. D.J. Bottomley, H. Omi and T. Ogino, J. Cryst. Growth, 225, 16 (2001).
78. K. O. Ng and D. Vanderbilt, Phys. Rev. B 52, 2177 (1995).
79. Adam H. Li, Feng Liu, D.Y. Petrovykh, J.-L. Lin, J. Viernow, F.J. Himpsel, and M.G. Lagally, Phys. Rev. Lett. 85, 5380 (2000).
80. Feng Liu, Adam H. Li, and M.G. Lagally, Phys. Rev. Lett. 87, 126103 (2001).
81. W. Lu and Z. Suo, Phys. Rev. B, 65, 085401 (2002).
82. V.I. Marchenko, Sov. Phys. JEPT 54, 605 (1981).
83. F.K. Men, W.E. Packard, and M.B. Webb, Phys. Rev. Lett. 61, 2469 (1988).
84. X. Tong and P.A. Bennett, Phys. Rev. Lett. 67, 101 (1991).
85. B.S. Swartzentruber, N. Kitamura, M.G. Lagally, and M.B. Webb, Rev. B 47, 13432 (1992).
86. R.J. Phaneuf, N.C. Bartelt, E.D. Williams, W. Swiech, and E. Bauer, Phys. Rev. Lett. 67, 2986 (1991).
87. H. Hibino, T. Fukuda, M. Suzuki, Y. Homma, T. Sato, M. Iwatsuki, K. Miki, and H. Tokumoto, Phys. Rev. B 47, 13027 (1993).
88. J.-L. Lin, D.Y. Petrovykh, J. Viernow, F.K. Men, D.J. Seo, and F.J. Himpsel, J. Appl. Phys. 84, 255 (1998).
89. M. Kasu and N. Kobayashi, Appl. Phys. Lett. 62, 1262 (1993).
90. S. Rousset, F. Pourmir, J.M. Berroir, J. Klein, J. Lecoeur, P. Hecquet, and B. Salanon, Surf. Sci. 422, 33 (1999).
91. G.M. Watson, Doon Gibbs, D.M. Zehner, Mirang Yoon, and S.G.J. Mochrie, Phys. Rev. Lett. 71, 3166 (1993).
92. F.-K Men, Feng Liu, P.J. Wang, C.H. Chen, D.L. Cheng, J.L. Lin, and F.J. Himpsel, Phys. Rev. Lett 88, 096105 (2002).
93. P.E. Wierenga, J.A. Kubby, and J.E. Griffith, Phys. Rev. Lett. 59, 2169 (1987).
94. O.L. Alerhand, A.N. Berker, J.D. Joannopoulos, D. Vanderbilt, R.J. Hamers, and J.E. Demuth, Phys. Rev. Lett. 64, 2406 (1990).
95. R.E. Martinez,W.M. Augustyniak, and J.A. Golovchenko, Phys. Rev. Lett. 64, 1035 (1990).
96. D. Vanderbilt, Phys. Rev. Lett. 59, 1456 (1987).
97. E.D. Williams, R.J. Phaneuf, J. Wei, N.C. Bartelt, and T.L. Einstein, Surf. Sci. 294, 219 (1993).
98. J. Tersoff, C. Teichert, and M. G. Lagally, Phys. Rev. Lett.76, 1675 (1996).
99. Feng Liu, S. E. Davenport, H. M. Evans, and M. G. Lagally, Phys. Rev. Lett. 82, 2528 (1998).
100. Feng Liu, Phys. Rev. Lett 89, 246105 (2002).
101. Feng Liu, J. Tersoff, and M. G. Lagally, Phys. Rev. Lett. 80, 1268 (1998).
102. L. Bai, J. Tersoff, and Feng Liu, Phys. Rev. Lett 92, 225503 (2004).

103. P. Venezuela, J. Tersoff, J. A. Floro, E. Chason, D. M. Foll-staedt, Feng Liu, and M. G. Lagally, Nature 397, 678 (1998).
104. E. Bauer, Z. Kristallogr. 110, 372 (1958).
105. J. Viernow, J.-L. Lin, D. Y. Petrovykh, F. M. Leibsle, F. K. Men, and F. J. Himpsel, Appl. Phys. Lett. 72, 948 (1998).
106. K. Pohl, M.C. Bartelt, J. de la Flguera, N.C. Bartelt, J. Hrbek, and R.Q. Hwang, Nature 397, 238 (1999).
107. F.M. Ross, J. Tersoff, and R.M. Tromp, Phys. Rev. Lett. 80, 984 (1998).
108. S. Lee, I. Daruka, C.S. Kim, A-L. Barabási, J.L. Merz, and J.K. Furdyna, Phys. Rev. Lett. 81, 3479 (1998).
109. G.R. Carlow and M. Zinke-Allmang, Phys. Rev. Lett. 78, 4601 (1997).
110. J. Drucker, Phys. Rev. B 48, 18203 (1993).
111. C. Priester and M. Lannoo, Phys. Rev. Lett. 75, 93 (1995).
112. I. Daruka and A.-L. Barabási, Phys. Rev. Lett. 79, 3708 (1997).
113. G. Medeiros-Ribeiro, A.M. Bratkovski, T.I. Kamins, D.A. Ohlberg, and R.S. Williams, Science 279, 353 (1998).
114. Y. Chen and J. Washburn, Phys. Rev. Lett. 77, 4046 (1996).
115. A-L. Barabási, Appl. Phys. Lett. 70, 2565 (1997).
116. D.E. Jesson, G. Chen, K.M. Chen, and S.J. Pennycook, Phys. Rev. Lett. 80, 5156 (1998).
117. M. Kästner and B. Voigtländer, Phys. Rev. Lett. 82, 2745 (1999).
118. J.A. Floro, E. Chason, M.B. Sinclair, L.B. Freud, and G.A. Lucadamo, Appl. Phys. Lett. 73, 951 (1997).
119. J.A. Floro, M.B. Sinclair, E. Chason, L.B. Freud, R.D. Twesten, R.Q. Hwang, and G.A. Lucadamo, , Phys. Rev. Lett. 84, 701 (2000).
120. V.A. Shchukin et al., Phys. Stat. Sol. (b) 224, 503 (2001).
121. For a review on strained 3D island stability, see V.A. Shchukin and D. Bimberg, Rev. of Modern Phys. 71, 1125 (1999).
122. Bin Yang, Feng Liu and Max G. Lagally, Phys. Rev. Lett 92, 025502 (2004).