CHAPTER II

REVIEW ON QUANTUM NANOSTRUCTURES

By the end of 1980s the properties of quantum wells and superlattices were rather well understood and the researchers have looked further to the lower-dimensional nanostuctures including quantum wire and quantum dot. The advantages of QDs present a new chapter for novel devices, such as single electron transistors and QD lasers and allow them to be used as a basic structure in many device applications, especially, quantum computing. The review on quantum nanostructures, their properties and their predicted application are illustrated in this chapter. The theoretical explanation of the effects of strain, resulting from the lattice mismatch, on band structure is also included.

2.1 Low-dimension nanostructures

QD, a 3-dimensional confinement of carrier structure, is a semiconductor crystal whose size is on the order of just a few nanometers. Complete reduction of the remaining 'infinite' extension of a quantum well in two dimensions to atomic values leads to carrier localization in all three dimensions and breakdown of the classical band structure model of a continuous dispersion of energy as a function of momentum [20]. The density of state of bulk semiconductor and the other low-dimensional nanostructures including quantum wells, quantum wires and quantum dots are illustrated in Figure 2.1. Three-dimensional confinement leads to quantized energy levels resulting in the delta-like density of states, unlike the other quantum structures such as quantum well and quantum wire.

In order to determine the density of states of QDs, the density of states of higher-dimensional structures were considered. At first, in case of quantum wells, an effective-mass approximation is widely used for the quantized energy levels calculation. The main assumption of the effective-mass approximation is that the envelope wave function does not significant vary in the units cell with a length scale of subnanometers, therefore this assumption is applied to all low-dimensional nanostructures.

Assuming parabolic band dispersion, band-edge electron states of semiconductors can be described by Schrödinger equation as

$$\left[-\frac{\hbar^2}{2m^*}\nabla^2 + V(\mathbf{r})\right] F(\mathbf{r}) = EF(\mathbf{r}) \tag{2.1}$$

where, m^* is the effective mass, $\hbar$ is the reduced Planks constant, $\mathbf{r} = (x, y, z)$ is the

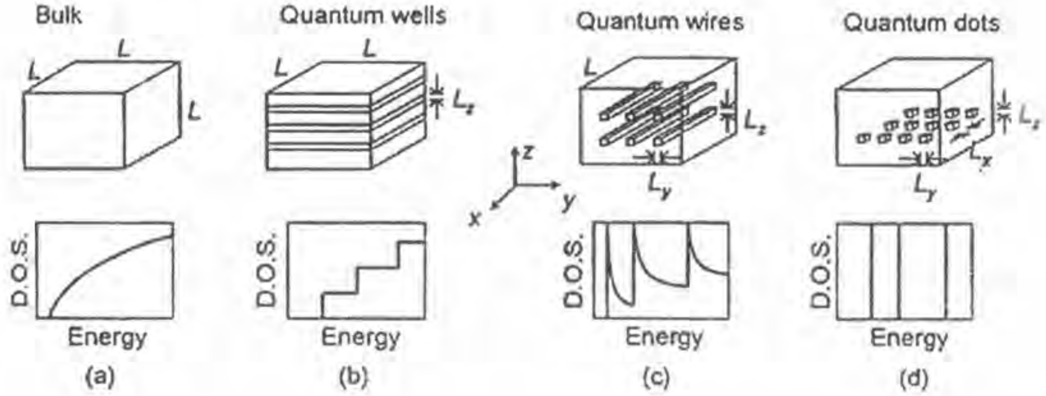

Figure 2.1: Schematic view and graphs of (a) bulk, (b) quantum wells, (c) quantum wires, and (d) quantum dots and their density of states (D.O.S.) [15].

carrier position vector, $V(\mathbf{r})$ is the confinement potential due to band offset, $F(\mathbf{r})$ is the envelope wave function, and E is the carrier energy.

From eq. (2.1), the carrier energy E in case of bulk, quantum well (QW), quantum wire (QWR), and quantum dot (QD) can be written as

$$E_{\text{bulk}} = E(\mathbf{k}) = \frac{\hbar^2 k^2}{2m^*} \tag{2.2}$$

$$E_{\text{QW}} = E(\mathbf{k}) = \frac{\hbar^2 {k_\parallel}^2}{2m^*} + E_{n,z} \tag{2.3}$$

$$E_{\text{QWR}} = E(\mathbf{k}) = \frac{\hbar^2 {k_\perp}^2}{2m^*} + E_{m,y} + E_{n,z} \tag{2.4}$$

$$E_{\text{QD}} = E_{l,x} + E_{m,y} + E_{n,z} \tag{2.5}$$

where, $\mathbf{k} = (k_x, k_y, k_z)$ is the wave vector of carriers, $k^2 = {k_x}^2 + {k_y}^2 + {k_z}^2, {k_\parallel}^2 = {k_x}^2 + {k_y}^2$ and ${k_\perp}^2 = {k_x}^2$. The energies $E_{l,x}$, $E_{m,y}$ and $E_{n,z}$, which are the function of the potential $V(\mathbf{r})$, depend on the quantum numbers l, m, and n. These energies can be determined by solving eq. (2.1) using either analytical method (with some approximations) or numerical methods. The density of states per unit volume (D.O.S), which is the number of states between the energy E and $E + dE$, of each quantum nanostructure is written as

$$D_{\text{bulk}}(E) = \frac{1}{2\pi^2}\left(\frac{2m^*}{\hbar^2}\right)^{3/2} E^{1/2} \tag{2.6}$$

$$D_{\mathrm{QW}}(E) = \frac{m^*}{\pi\hbar^2}\sum_n \Theta(E - E_{n,z}) \tag{2.7}$$

$$D_{\mathrm{QWR}}(E) = \frac{N_{wi}}{\pi}\frac{\sqrt{2m^*}}{\hbar}\sum_{m,n}\frac{1}{\sqrt{E - E_{m,y} - E_{n,z}}} \tag{2.8}$$

$$D_{\mathrm{QD}}(E) = 2N_D\sum_{l,m,n}\delta(E - E_{l,x} - E_{m,y} - E_{n,z}) \tag{2.9}$$

where, Θ is the Heaviside's unit step function, N_{wi} is the area density of the quantum wires (the number of quantum wires divided by the quantum wire region area in the y-z plane), δ is the delta function, and N_D is the volume density of QD [15]. The density of states in eqs. (2.6)-(2.9) are schematically shown with the corresponding structures in Figure 2.1.

Quantum dot has a delta-like density of states which is a sharper density of states than higher-dimensional structures. As a result, it promises as a competent candidate in not only quantum computing but also optical communication. Moreover, it plays a significant role in several frontier technologies. For instance, QD lasers incorporate the use of quantum dots as an active light-emitting area region of semiconductor lasers providing high efficiency [16]. It is also being researched for use in amplifiers, biological sensors, solar cells and other exciting novel application, including information technology. In the next section, the effect of strain, resulting from the lattice mismatch, on the band structure of QDs is presented.

2.2 Strain effects on energy gap of semiconductors

In the late 1940s, the possibility of the growth of lattice-mismatched crystalline monolayers was exhibited by Frank and van der Merwe. In lattice-matched epitaxial growth, the epitaxial film and the substrate join together with the same lattice constant as shown in Figure 2.2 (a). Conversely, in strained layer growth, the lattice constant of the epitaxial layer in the direction parallel to the interface is forced to be equal to the lattice constant of the substrate. The lattice constant of the epitaxial layer perpendicular to the substrate will be changed because of the Poisson effect, as illustrated in Figures 2.2 (b) and (c).

On the one hand, if the parallel lattice constant is forced to shrink when a compress via strain is applied, the perpendicular lattice constant will grow. On the other hand, if the parallel lattice constant is forced to expand under a tensile strain, the perpendicular lattice constant will shrink. The lattice mismatch, or misfit, is defined as

$$\frac{\triangle a}{a} = \frac{a - a_0}{a} = -\varepsilon \tag{2.10}$$

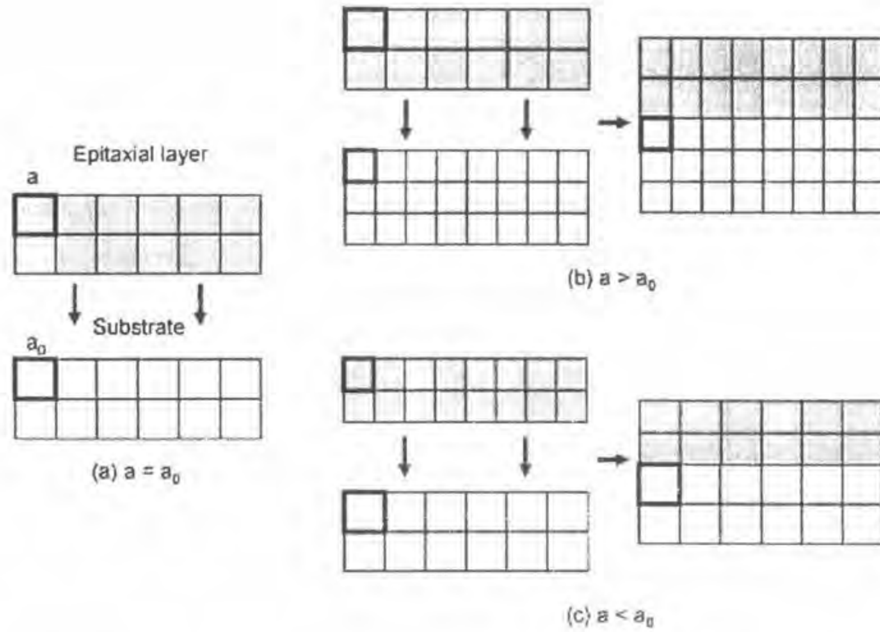

Figure 2.2: Accommodation of lattice of epitaxial layer with that of substrate for different cases: (a) lattice-matched growth ($a = a_0$), (b) biaxial compressive strain($a > a_0$), and (c) biaxial tensile strain ($a < a_0$) [18].

In case of the growth of a strained epitaxial film of thickness d_1 with lattice constant a on the substrate of thickness d_2 with lattice constant a_o, where $a > a_o$, the lattice constants of the strained film in the direction parallel to the interface can be derived as

$$a_{\parallel} = \frac{ad_1 + a_0 d_2}{d_1 + d_2} \tag{2.11}$$

In general, $d_2 >> d_1$, and therefore

$$a_{\parallel} = a_0 = (1 + \varepsilon)a \tag{2.12}$$

The lattice constant in the direction perpendicular to the interface is

$$a_{\perp} = (1 - \sigma_{ST}\varepsilon)a \tag{2.13}$$

where σ_{ST} is a material parameter given by

$$\sigma_{ST} = \frac{c_{11}}{2c_{12}} \tag{2.14}$$

Here c_{11} and c_{12} are elastic constant of the epitaxial layer. Assuming that the in-plane coordinate directions are x and y, and the growth occur along z-direction which coincides with the (001) crystallographic axis. Strained growth generates stress upon the x and y directions. The stress components in the x and y directions are equal and force the lattice constants in both of the parallel directions to be equal. Therefore,

$$\varepsilon_{xx} = \varepsilon_{yy} = \varepsilon_{\parallel} = \varepsilon \tag{2.15}$$

No compressive strain is placed upon the epitaxial layer in the z-direction, but the tensile strain in this direction will be nonzero. The strain in the z-direction can be determined by writing Hooke's law in terms of the elastic stiffness constants. Therefore, the strain in the z-direction is

$$\varepsilon_{zz} = -\frac{2c_{12}}{c_{11}}\varepsilon \tag{2.16}$$

The strain energy is given by

$$E_{st} = \varepsilon^2 \left(c_{11} + c_{12} - \frac{2{c_{12}}^2}{c_{11}} \right) d \tag{2.17}$$

which increases linearly with the thickness of the film. However, when the mismatch or thickness of the epitaxial layer is large, the strain energy stored in the crystal can be reduced by the formation of dislocations at the interface. The critical thickness h_c arises because of a composition between strain energy and chemical energy. The critical thickness is dependent upon the amount of the lattice mismatch and material parameters as well as the properties of the dislocations that form in the particular material. The critical thickness is defined as

$$h_c = \frac{a_0(1 - \frac{v_{PR}}{4}) \left[ln\left(\frac{h_c\sqrt{2}}{a_0} \right) + 1 \right]}{2\sqrt{2}\pi f(1 + v_{PR})} \tag{2.18}$$

It is evident that the larger the mismatch, the smaller the critical thickness. The critical thickness can be approximated as $h_c \cong a/2\varepsilon$ which is derived from energy minimization considerations. $v_{PR} = c_{12}/(c_{11} + c_{12})$ is the Poisson ratio. Its values are close to 1/3 for most semiconductors.

The use of strained layers was originally advanced as a technique to alter the bandgap. When the thickness of mismatched layer is larger than h_c, this layer will be

relaxed and become strain-free mismatched layer. The effects of strain on the band structure can be expressed in terms of deformation potentials. The conduction band is shifted upwards for compressive strain and downwards for tensile strain.

There are drastically changes in valence band structure because of the valence band is doubly degenerate resulting in heavy-hole and light-hole. In presence of biaxial tensile strain the light-hole is expected to be above the heavy-hole state, while for the compressive strain the reverse is expected [18]. Biaxial compressive strain increased the splitting, and tensile strain decreases it, because it moves the bands in the opposite directions. This ability of biaxial strain to split heavy-hole and light-hole states is useful optical properties in strained materials. The changes in band structure are presented in Figure 2.3.

In case of self-assembled QD, the strain distribution is more complicated resulting from the three-dimensional confinement structure. It has been widely determined by many researches based on various approaches including using Green's tensor (Faux and Pearson, 2000). Nevertheless, the lack of knowledge in exact QD shape, size and composition make the accurate prediction of this effect more complicate. The shape and composition of self-assembled QD are properties that are very difficult to determine and often only available by indirect means [19]. The strain distribution in the (x, z) plane of pyramidal QD calculated by finite difference scheme is shown in Figure 2.4 [20]. In next section, both top-down and bottom-up fabrication techniques for QDs are presented.

2.3 Fabrication techniques for quantum dots

There are many alternative techniques used in the past decade to fabricate QDs. The top-down techniques like lithographic patterning and etching of quantum well structures have been employed. However, these top-down techniques are facing the miniaturization problems. Therefore, the concept of self-assembled QDs becomes more significant.

2.3.1 Quantum dot grown by top-down techniques

By the end of 1980s the fabrication of QDs by patterning of quantum wells was considered to be the most effective way of QD fabrication. Although the well-known patterning techniques, such as optical lithography, X-ray lithography, electron or ion beam lithography (EBL or FIBL) and scanning tunnelling microscopy (STM), have several advantages, they are limited by the resolution of lithography. The schematics of various top-down approaches are illustrated in Figure 2.5.

Miniaturization of QD size is facing significant technological limitations, such as the short channel effects, the high power dissipation associated with quantum tunnelling through the gate oxide and the depletion regions, and the difficulties in doping shallow

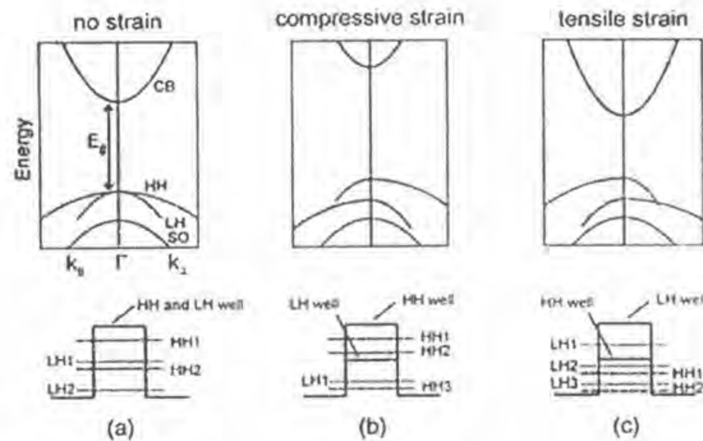

Figure 2.3: (a) Schematic representation of the band structure of an unstrained direct gap tetrahedral semiconductor. The light-hole (LH) and heavy-hole (HH) bands are degenerate at the Brillouin zone center Γ and the spin split-off bandlies lower in energy. The lowest conduction band (CB) is separated by the bandgap energy E_g from the valence bands. Note that the $k_{||}$ is perpendicular to the growth and strain direction. (b) Under biaxial compression, the hydrostatic component of the compression increase the mean band gap, while the uniaxial component splits the degeneracy of the valence band maximum and introduce an anisotropic valence structure. (c) Under biaxial tension, the mean band gap reduces and the valence band splitting is reserved. The lower panel shows the valence band diagram of the quantum well structure in case of (a) no strain, (b) compressive strained, and (c) tensile strain [26].

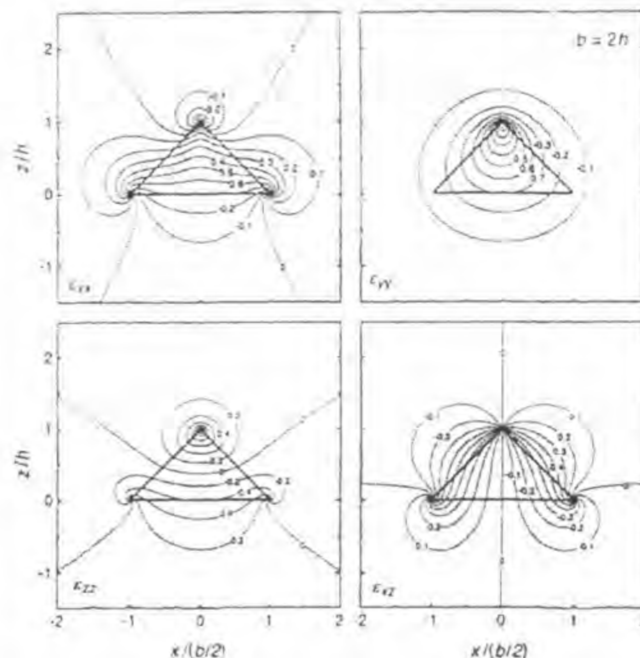

Figure 2.4: Strain distribution for a pyramidal QD with 45° facet angle in the (xz) plane through the pyramid top. Identical isotropic elastic constants and $\sigma = 1/3$ are taken throughout the structure. $\varepsilon_{xx}, \varepsilon_{yy}, \varepsilon_{zz}, \varepsilon_{xz}$ are shown; due to symmetry in this plane, ε_{xy} and ε_{yz} are zero [20].

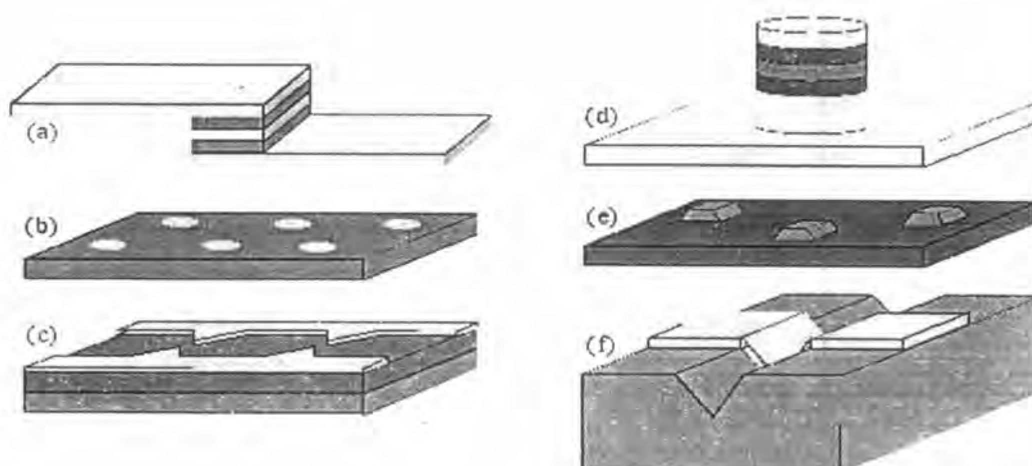

Figure 2.5: Schematic of various top-down approaches i.e., (a) metal-oxide QD using lithography technique, (b) metallic dots by chemical suspensions, (c) lateral quantum dots using heterostructures, (d) vertical quantum dots by wet-etching quantum well, (e) pyramit QDs by self-assembled growth and (f) trench quantum wire.

junctions and uniform channels [2]. Therefore, the new fabrication technique based on concept of self-organized growth with defect-free has been proposed.

2.3.2 Self-assembled quantum dots

Without the artificial patterning, the concept of self-assembled QD is based on the lattice mismatch between deposited layer and surface substrate. There are three possible modes of crystals growth on surface. In case of the island, or Volmer-Weber mode, small clusters are nucleated directly on the substrate surface and then grow into island of condensed phase, as shown in Figure 2.6 (c). This happens when the atoms, or molecules, of deposit are more strongly bond to each other than to the substrate. On the other hand, in case of the layer-by-layer, or Frank-van der Merwe mode, the atoms are more strongly bound to the substrate than each other as shown in Figure 2.6 (a). The first atoms to condense form a complete monolayer on the surface, which are covered with a somewhat less tightly bound second layer.

While the layer plus island, or Stranski-Krastanov, growth mode is an intermediate case as shown in Figure 2.6 (b). After forming the first monolayer, or a few monolayers, subsequent layer growth is unfavourable and islands are formed in top of this intermediate layer. The nanostructures formed in Stranski-Krastanov growth mode are called self-assembled QDs, which can confine carriers in three dimensions. The schematic representation of QD is illustrated in Figure 2.7. Self-assembled QD grown by SK mode play a significant role in this work. The growth steps and parameters are presented in the next chapter. The theoretical formation of self-assembled QD by SK mode is explained in following section.

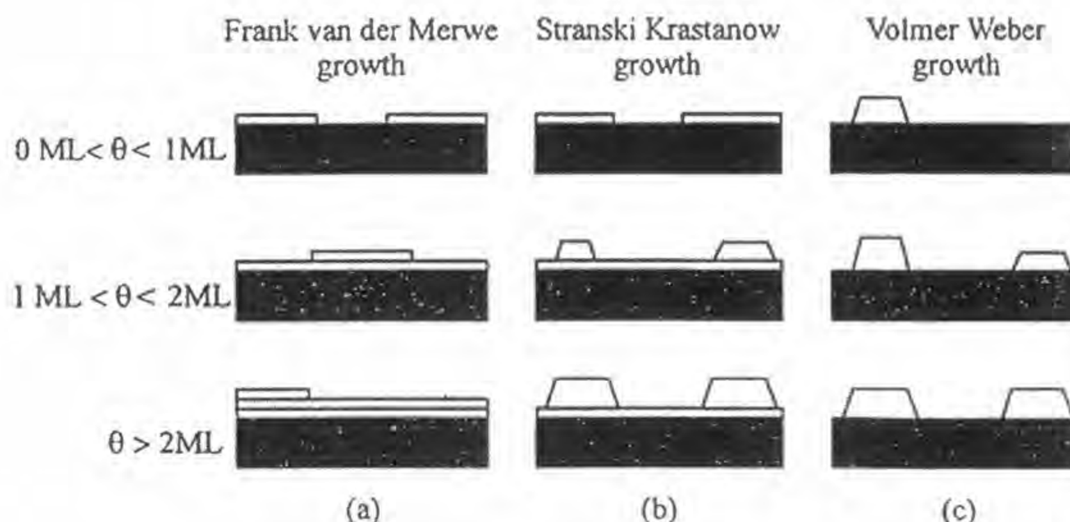

Figure 2.6: Schematic representation of the three crystal growth modes (a) layer-by-layer, or Frank-vander Merwe mode, (b) layer plus island, or Stranski-Krastanov mode and (c) the island, or Volmer-Weber mode. θ represents the coverage in monolayers [17].

2.4 Stranski-Krastanow (SK) growth mode

If the film deposited on the substrate is initially not too thick, island can develop without dislocation and subsequently alter as mass flows along the wetting layer. In SK mode the film must exceed a critical thickness before it becomes unstable resulting from the wetting layer effect, which tends to hold the film flat against the substrate. The kinetics of SK island growth rely on several factors, including the initial thickness of film; the relative size of the rate of attachment of adatoms during vapor deposition and the rate associated with surface diffusion; the composition between the strain and the surface energies, and their associated anisotropic pies; and the wetting layer thickness [21].

A two-dimensional, large deformation numerical model of the deposition process has been developed within a global, various frameworkwhich allows the relevant kinetic processes and driving forces to be incorporated consistently over time, via deposition and the diffusion of material over the surface under law of mass conservation, as shown in Figure 2.8. The elasticity problem is solved using the finite element method and the resulting elastic field is used to determine the elastic driving forces acting on each surface node. The additional driving force due to surface energy is calculated using the appropriate surface energy density function [22].

In addition, they also proposed that the effect of surface energy anisotropy can be clearly seen in the strongly facetted pyramidal structures that real dots adopt at length scales of 50-150 Å. The simulation of three types of surface energy function is shown in Figure 2.9. The effect is very noticeable (even with only 5 % fluctuation in the energy).

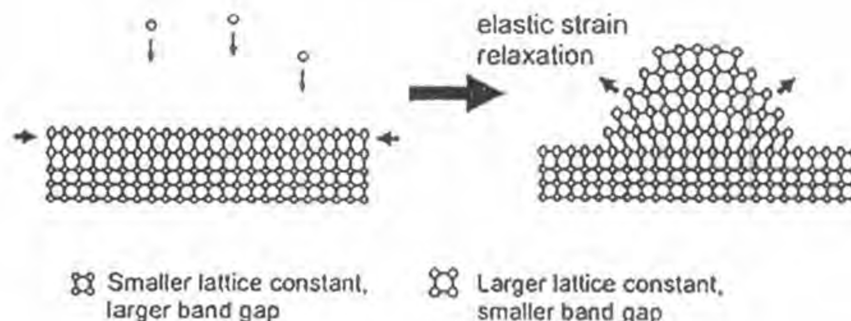

Figure 2.7: Schematic representation of island formation during epitaxial growth of a semiconductor material on top of another semiconductor with a smaller lattice constant in Stranski-Krastanov growth mode [26].

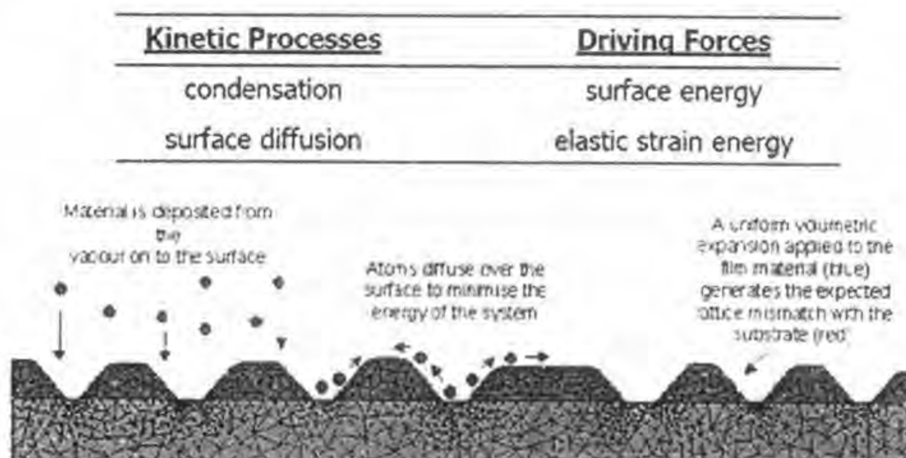

Figure 2.8: A schematic of the deposition and relaxation process [22].

The two anisotropic energy functions produce dots of the expected form although inspection of the growth path of these dots eliminates the smooth anisotropic form as a practicable candidate. The simulation of QD formation from 2D to 3D island with a cusped anisotropic surface energy is shown in Figure 2.10.

In case of experimental observations of growth kinetics of InAs on GaAs (001). Initial growth occurs by the appearance of small quasi-3D (Q3D) clusters which vary from 2-4 ML in thickness and up to 200 Å in lateral size. 3D islands start to appear after a critical thickness of 1.75 ML has been deposited. At this point the 2D layer thickness reduces to about 1.2 ML and the remaining material forms the islands. The initial half-base width is about 120 Å increasing to about 200 Å at 3 ML. The aspect ratio remains roughly constant with the height increasing with the base width.

When the 3D islands are formed, the surface strain would be modified by the strain relaxation of the islands on the surface. The surface strain can be approximated by using an analytic solution of the stress filed in 2D islands mound on a strained semi-

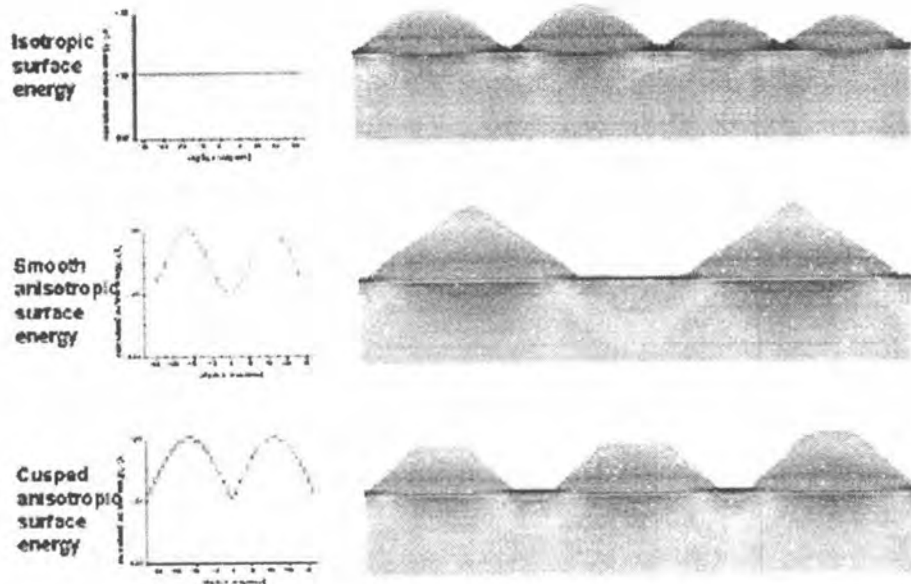

Figure 2.9: The simulations shows the effect of isotropy and anisotropy surface energy on growth kinetics [22].

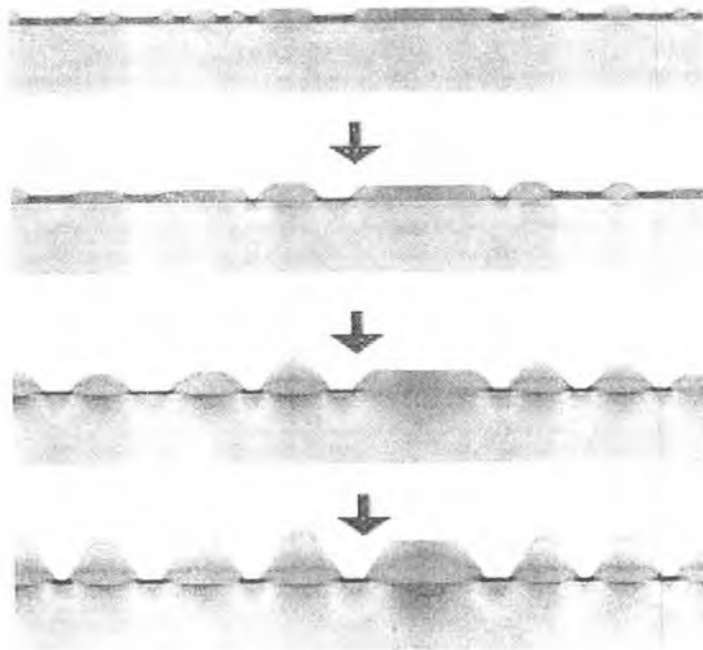

Figure 2.10: An initial perturbation allows the elastic mismatch strain to destabilise the film. The lateral stress contours show that areas of high compression (blue) are relieved by the morphological change so that the underlying substrate goes into tension (red). This simulation is with a cusped anisotropic surface energy [22].

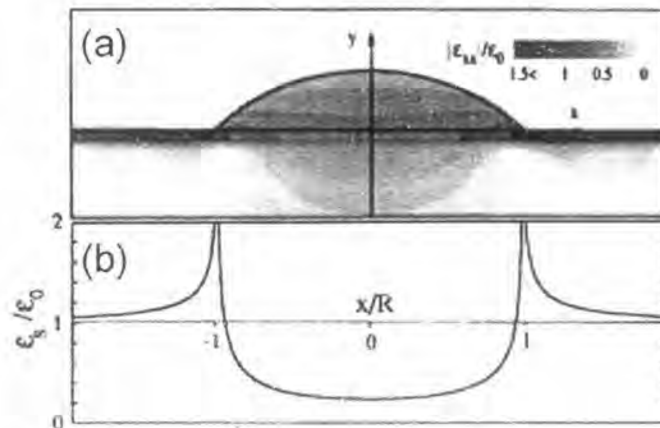

Figure 2.11: (a) A contour diagram showing strain ϵ_{xx} in the island and (b) the variation of the surface strain ϵ_s along the system surface [27].

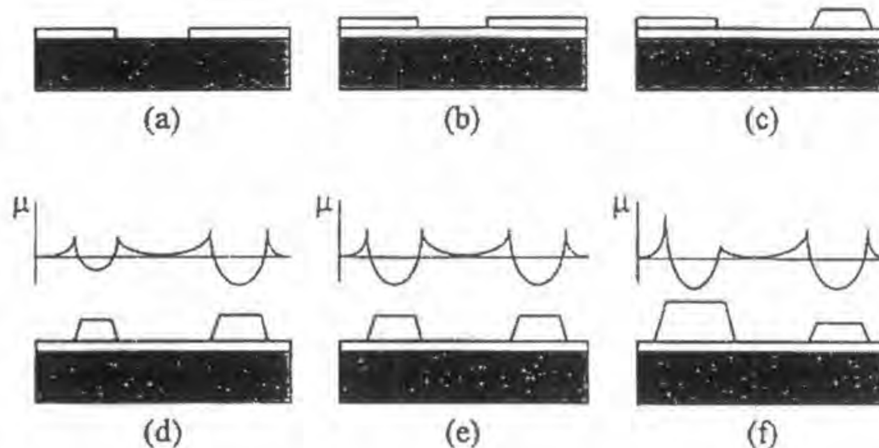

Figure 2.12: Schematic illustration of SK QDs formation process (a) initial stage of wetting formation, (b) the wetting layer formation, (c) 2D to 3D island transition, (d) non-uniform 3D islands, (e) self-regulation process and (f) misfit dislocation formation [27].

infinite substrate as a contour plot of absolute value of ϵ_{xx} and a surface tangential strain ϵ_s along the system surface shown in Figure 2.11 [27].

The partially relaxation at the top of the island induces extra strain in the substrate and increases the strain in the wetting layer. There is the high strain energy at the island edge resulting from the strain relaxation of the strain material. Cent et al (1996) proposed that the 3D islands will grow by accumulation of adatoms. These adatoms are supposed to overcome the energy barrier to attach the islands. The attached adatoms to the islands tend to increase the strain concentration at the island edge, then increase the barrier for the migrating adatoms to attach the islands. Therefore, the larger the island, the slower the growth rate is. The summarized process of the QDs formation was illustrated in Figure 2.12. When the lattice mismatched material was deposited on the substrate, the 2D islands formed (Figure 2.12 (a)), and then subsequent deposited

material formed wetting layer (Figure 2.12 (b)). Further deposition, the strain energy in the wetting layer increase until overcome the surface energy. The 2D island would transform to 3D island (Figure 2.12 (c)).

The smaller islands pushes the lower energy barrier due to the less strain energy, resulting in the directional migration of the deposited adatoms from the larger to the smaller island leading to the size uniformity (Figure 2.12 (d) and (e)). When the island size increases further, the misfit dislocation was formed to relieve the strain concentration. The strain relief in the island with dislocation decreases the energy barrier, and then permits the further growth of the island (Figure 2.12 (f)).

2.5 Quantum computing

Quantum computing is a future technology for designing computers based on quantum mechanics, the science of atomic structure and function. It uses the "qubit," or quantum bit, which can hold an infinite number of values. Quantum computers would allow a bit to store a value of 0 and 1 simultaneously. Unlike a classical computer that operates on a 3 bit register, which the bits in the register are in a definite state, such as 101, at any given time, quantum computer use the qubits which can be in a superposition of all the classically allowed states. It is believed that such a device can handle multiple operations simultaneously and can factor large numbers 10,000 times faster than today's computers.

Due to the advantages of QDs, such as delta-like function D.O.S., allow them to be used as a basic structure in many device applications,especially, quantum computing. One approach for quantum computing is based on Coulomb repulsion of like charges, named "quantum cellular automata (QCA)". The unit cell of a QCA consists of four QDs positioned at the corners of a rectangle [9, 10, 11]. Another approach is based on a quantum mechanical property, a spin, of electrons in semiconductors [12]. The two approaches are candidates to overcome the current problems of miniaturisation.

2.5.1 Quantum Cellular Automata (QCA)

The principle of quantum cellular automata was first proposed by Lent et al [9]. A basic QCA cell consists of five dots, as shown in Figure 2.13. Four dots are positioned at each corner of the square and the fifth dot is located at the center. Each cell contains two excess electrons which are able to tunnel between the dots. When the inter-dot barriers are high, the electrons will be localized on individual dots. The Coulomb repulsion between the electrons tends to make them align at antipodal sites with two possible configurations as shown in Figure 2.13. These two stable states can be represented as a cell polarization $P = +1$ and $P = -1$ which may be interpreted as binary information: $P = +1$ corresponds to a bit ´1´ and $P = -1$ corresponds to a bit ´0´ [10].

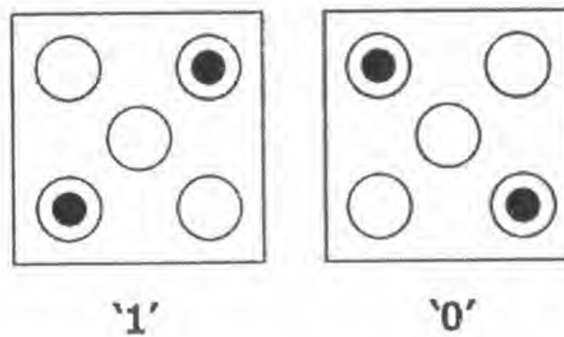

Figure 2.13: The basic QCA cell with two possible charge configurations [9]

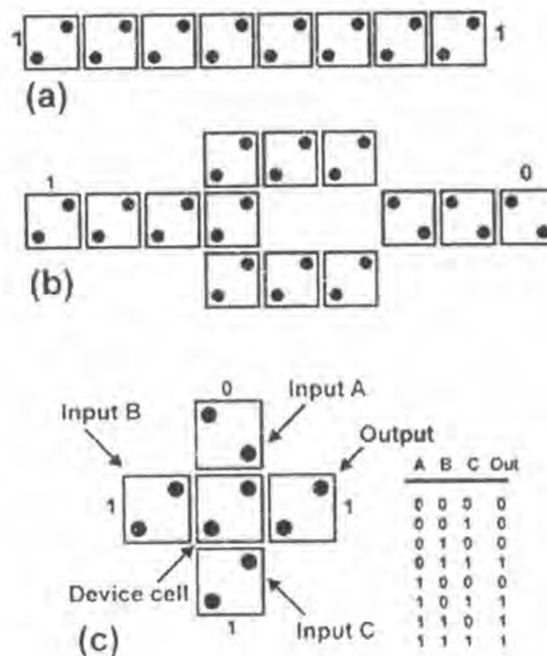

Figure 2.14: QCA arrays working as (a) a wire, (b) an inverter or (c) a majority gate [11]

The Coulomb interaction between cells causes two adjacent cells having the same polarization. Notice that only binary information, not charge, is transmitted. This means that there is no current flow between the cells. This is the reason why QCA structures consume little power. Figure 2.14 shows examples of some QCA arrays. A QCA wire is shown in figure 2.14 (a). A QCA inverter is shown in Figure 2.14 (b). And a QCA majority gate is shown in Figure 2.14 (c). The logic gate in Figure 2.14 (c) can be used as an OR gate when input A is kept fixed at ´1´, or as an AND gate when input A is fixed at ´0´.

2.6 Quantum computation using electrically controlled semiconductor spins [12]

Taylor et al. (2005), developed an architecture for quantum computation using electrically controlled semiconductor spins by extending the Loss Di Vincenzo scheme and by combining actively protected quantum memory and long-distance coupling mechanisms, and developed a universal set of quantum gates compatible with active error suppression for these encoded qubits and an effective long-range interaction between the qubits by controlled electron transport. They also described and analysed a modular solid-state architecture, based on spin qubits in double quantum dots, which can be used for scalable quantum computation. The qubits were chosen for potential long quantum memory times, and they developed a long-range, low-error transport mechanism for them. They proposed that this approach yields a scalable architecture with favourable error thresholds for fault-tolerant operation. A fully electrical architecture for controlling and coupling qubits are shown in Figure 2.15. The details of this approach were described in "Fault-tolerant architecture for quantum computation using electrically controlled semiconductor spins" authored by Taylor et al. (Nature Physics 1, 177-183 (2005)).

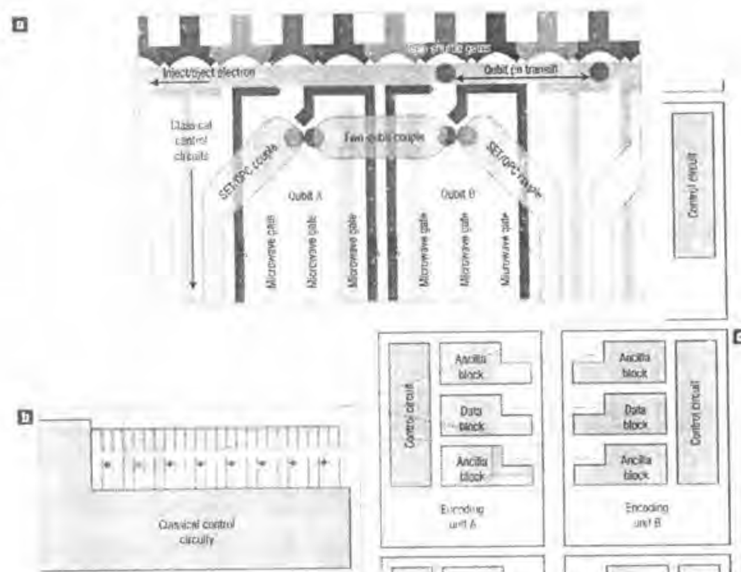

Figure 2.15: (a) Two adjacent depletion-gate-defined double quantum dots built on a semiconductor heterostructure, with tuning gates (green), high-frequency microwave gates (grey, blue) and capacitive coupling gates (orange) forms a node. Coupling to a nearby quantum point contact (QPC) or single-electron transistor (SET) provides for charge measurement. Alternating gates along the top form the spin shuttle and a Fermi sea can be selectively coupled to the spin shuttle for electron injection/ejection (preparation and erasure). The transport-channel double dot (qubit B) is only occupied during two-qubit gates. (b) Qubits (red) connected by a spin shuttle (yellow), along with semi-autonomous classical control circuitry (bluegrey), forming a single block. (c) A single data block, used for the logical computation, is adjacent to R ancilla blocks, making up an encoding unit [12].