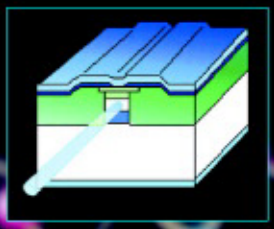
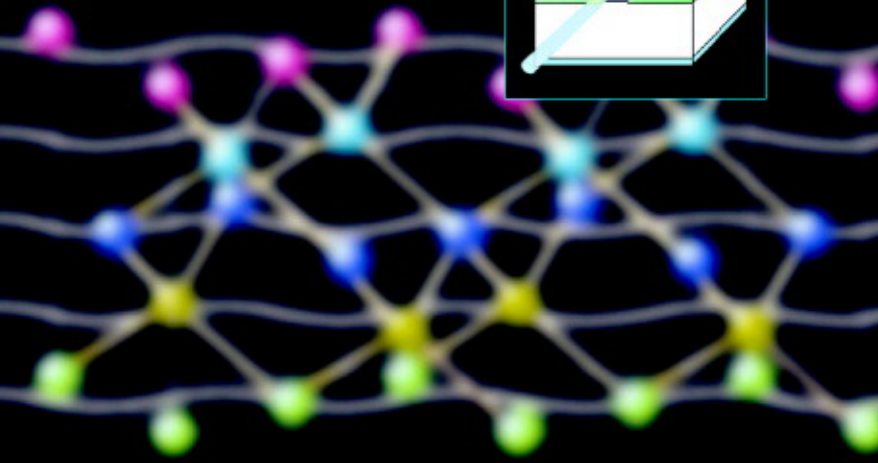
Electronic and Optoelectronic Properties of Semiconductor Structures





Jasprit Singh

CAMBRIDGE

CAMBRIDGE more information - www.cambridge.org/9780521823791

This page intentionally left blank

Electronic and Optoelectronic Properties of Semiconductor Structures presents the underlying physics behind devices that drive today's technologies. The book covers important details of structural properties, bandstructure, transport, optical and magnetic properties of semiconductor structures. Effects of low-dimensional physics and strain – two important driving forces in modern device technology – are also discussed. In addition to conventional semiconductor physics the book discusses self-assembled structures, mesoscopic structures and the developing field of spintronics.

The book utilizes carefully chosen solved examples to convey important concepts and has over 250 figures and 200 homework exercises. Real-world applications are highlighted throughout the book, stressing the links between physical principles and actual devices.

Electronic and Optoelectronic Properties of Semiconductor Structures provides engineering and physics students and practitioners with complete and coherent coverage of key modern semiconductor concepts. A solutions manual and set of viewgraphs for use in lectures is available for instructors.

JASPRIT SINGH received his Ph.D. from the University of Chicago and is Professor of Electrical Engineering and Computer Science at the University of Michigan, Ann Arbor. He has held visiting positions at the University of California, Santa Barbara and the University of Tokyo. He is the author of over 250 technical papers and of seven previous textbooks on semiconductor technology and applied physics.

Electronic and Optoelectronic Properties of Semiconductor Structures

Jasprit Singh

University of Michigan, Ann Arbor

CAMBRIDGE
UNIVERSITY PRESS

CAMBRIDGE UNIVERSITY PRESS
Cambridge, New York, Melbourne, Madrid, Cape Town, Singapore, São Paulo

Cambridge University Press
The Edinburgh Building, Cambridge CB2 2RU, United Kingdom

Published in the United States of America by Cambridge University Press, New York

www.cambridge.org

Information on this title: www.cambridge.org/9780521823791

First published in print format 2003

ISBN-13 978-0-511-07552-0 eBook (Adobe Reader)
ISBN-10 0-511-07552-9 eBook (Adobe Reader)

ISBN-13 978-0-521-82379-1 hardback
ISBN-10 0-521-82379-X hardback

CONTENTS

PREFACE

Semiconductor-based technologies continue to evolve and astound us. New materials, new structures, and new manufacturing tools have allowed novel high performance electronic and optoelectronic devices. To understand modern semiconductor devices and to design future devices, it is important that one know the underlying physical phenomena that are exploited for devices. This includes the properties of electrons in semiconductors and their heterostructures and how these electrons respond to the outside world. This book is written for a reader who is interested in not only the physics of semiconductors, but also in how this physics can be exploited for devices.

The text addresses the following areas of semiconductor physics: i) electronic properties of semiconductors including bandstructures, effective mass concept, donors, acceptors, excitons, etc.; ii) techniques that allow modifications of electronic properties; use of alloys, quantum wells, strain and polar charge are discussed; iii) electron (hole) transport and optical properties of semiconductors and their heterostructures; and iv) behavior of electrons in small and disordered structures. As much as possible I have attempted to relate semiconductor physics to modern device developments.

There are a number of books on solid state and semiconductor physics that can be used as textbooks. There are also a number of good monographs that discuss special topics, such as mesoscopic transport, Coulomb blockade, resonant tunneling effects, etc. However, there are few single-source texts containing "old" and "new" semiconductor physics topics. In this book well-established "old" topics such as crystal structure, band theory, etc., are covered, along with "new" topics, such as lower dimensional systems, strained heterostructures, self-assembled structures, etc. All of these topics are presented in a textbook format, not a special topics format. The book contains solved examples, end-of-chapter problems, and a discussion of how physics relates to devices. With this approach I hope this book fulfills an important need.

I would like to thank my wife, Teresa M. Singh, who is responsible for the artwork and design of this book. I also want to thank my editor, Phil Meyler, who provided me excellent and timely feedback from a number of reviewers.

Jasprit Singh

INTRODUCTION

Semiconductors and devices based on them are ubiquitous in every aspect of modern life. From "gameboys" to personal computers, from the brains behind "nintendo" to world wide satellite phones—semiconductors contribute to life perhaps like no other manmade material. Silicon and semiconductor have entered the vocabulary of newscasters and stockbrokers. Parents driving their kids cross-country are grudgingly grateful to the "baby-sitting service" provided by ever more complex "gameboys." Cell phones and pagers have suddenly brought modernity to remote villages. "How exciting," some say. "When will it all end?" say others.

The ever expanding world of semiconductors brings new challenges and opportunities to the student of semiconductor physics and devices. Every year brings new materials and structures into the fold of what we call semiconductors. New physical phenomena need to be grasped as structures become ever smaller.

I.1 SURVEY OF ADVANCES IN SEMICONDUCTOR PHYSICS

In Fig. I.1 we show an overview of progress in semiconductor physics and devices, since the initial understanding of the band theory in the 1930s. In this text we explore the physics behind all of the features listed in this figure. Let us take a brief look at the topics illustrated.

- **Band theory:** The discovery of quantum mechanics and its application to understand the properties of electrons in crystalline solids has been one of the most important scientific theories. This is especially so when one considers the impact of band theory on technologies such as microelectronics and optoelectronics. Band theory and its outcome—effective mass theory—has allowed us to understand the difference between metals, insulators, and semiconductors and how electrons respond to external forces in solids. An understanding of electrons, holes, and carrier transport eventually led to semiconductor devices such as the transistor and the demonstration of lasing in semiconductors.

- **Semiconductor Heterostructures:** Initial work on semiconductors was carried out in single material systems based on Si, Ge, GaAs, etc. It was then realized that if semiconductors could be combined, the resulting structure would yield very interesting properties. Semiconductors heterostructures are now widely used in electronics and optoelectronics. Heterostructures are primarily used to confine electrons and holes and to produce low dimensional electronic systems. These low dimensional systems, including quantum wells, quantum wires and quantum dots have density of states and other electronic properties that make them attractive for many applications.

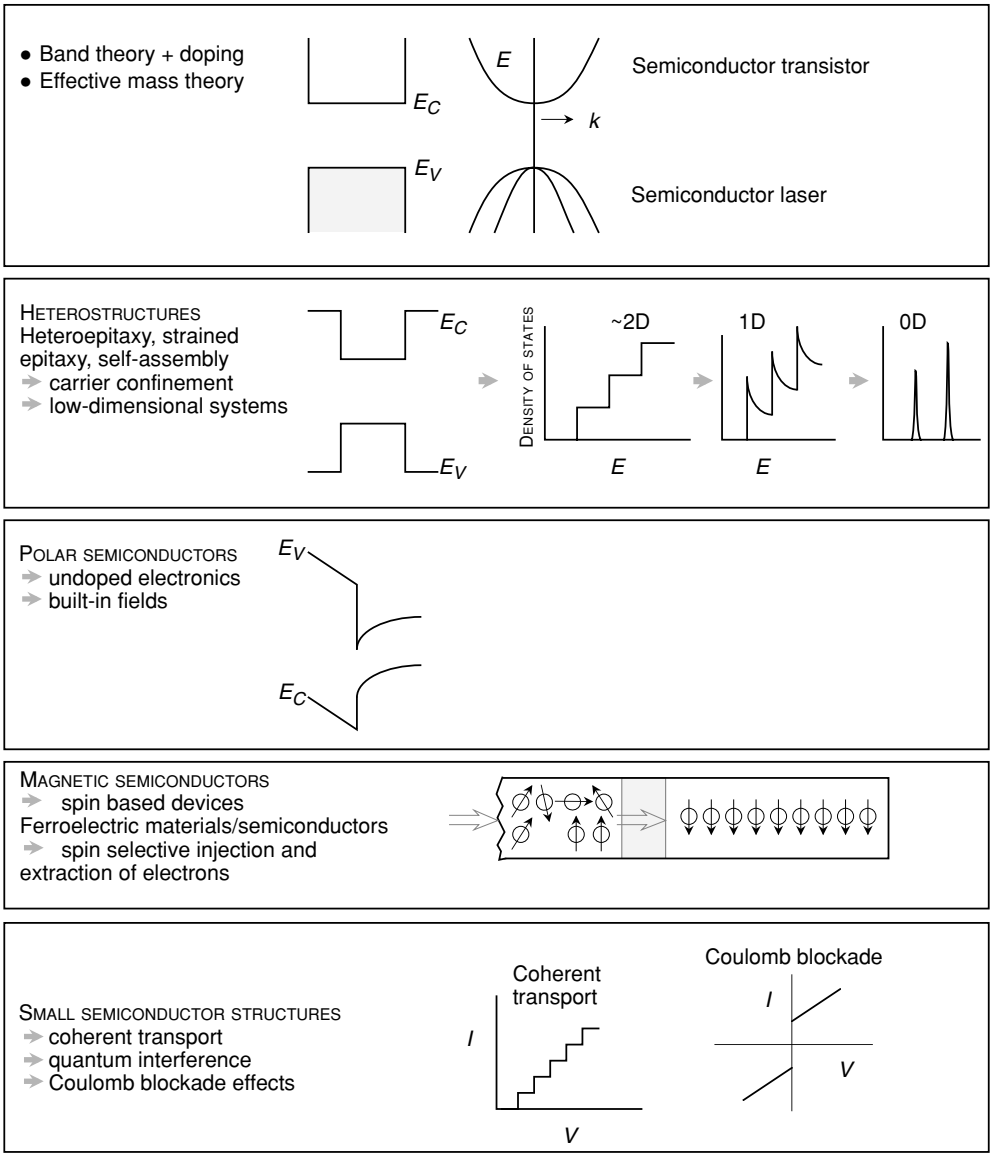


Figure I.1: Evolution of semiconductor physics and phenomena. These topics are discussed in this book.

Advances in heterostructures include strain epitaxy and self-assembled structures. In strained epitaxy it is possible to incorporate a high degree of strain in a thin layer. This can be exploited to alter the electronic structure of heterostructures. In self-assembled structures lateral structures are produced by using the island growth mode or other features in growth processes. This can produce low-dimensional systems without the need of etching and lithography.

- **Polar and Magnetic Heterostructures:** Since the late 1990s there has been a strong push to fabricate heterostructures using the nitride semiconductors (InN, GaN, and AlN). These materials have large bandgaps that can be used for blue light emission and high power electronics. It is now known that these materials have spontaneous polarization and a very strong piezoelectric effect. These features can be exploited to design transistors that have high free charge densities without doping and quantum wells with large built-in electric fields.

  In addition to materials with fixed polar charge there is now an increased interest in materials like ferroelectrics where polarization can be controlled. Some of these materials have a large dielectric constant, a property that can be exploited for design of gate dielectrics for very small MOSFETs. There is also interest in semiconductors with ferromagnetic effects for applications in spin selective devices.

- **Small Structures:** When semiconductor structures become very small two interesting effects occur: electron waves can propagate without losing phase coherence due to scattering and charging effects become significant. When electron waves travel coherently a number of interesting characteristics are observed in the current-voltage relations of devices. These characteristics are qualitatively different from what is observed during incoherent transport.

  An interesting effect that occurs in very small capacitors is the Coulomb blockade effect in which the charging energy of a single electron is comparable or larger than k_BT. This effect can lead to highly nonlinear current-voltage characteristics which can, in principle, be exploited for electronic devices.

I.2 PHYSICS BEHIND SEMICONDUCTORS

Semiconductors are mostly used for information processing applications. To understand the physical properties of semiconductors we need to understand how electrons behave inside semiconductors and how they respond to external stimuli. Considering the complexity of the problem—up to 10^{22} electrons cm^{-3} in a complex lattice of ions —it is remarkable that semiconductors are so well understood. Semiconductor physics is based on a remarkably intuitive set of simplifying assumptions which often seem hard to justify rigorously. Nevertheless, they work quite well.

The key to semiconductor physics is the band theory and its outcome—the effective mass theory. As illustrated in Fig. I.2, one starts with a perfectly periodic structure as an ideal representation of a semiconductor. It is assumed that the material can be represented by a perfectly periodic arrangement of atoms. This assumption although not correct, allows one to develop a band theory description according to which electrons act as if they are in free space except their effective energy momentum

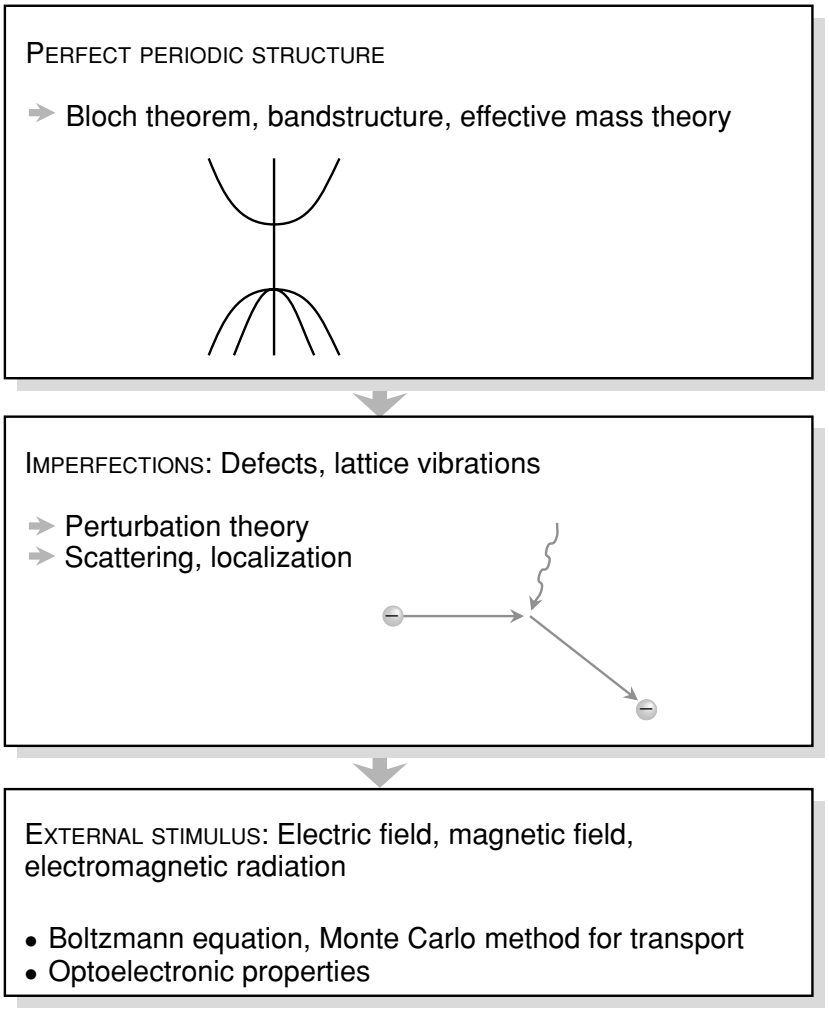


Figure I.2: A schematic of how our understanding of semiconductor physics proceeds.

relation is modified. This picture allows one to represent electrons near the bandedges of semiconductors by an "effective mass."

In real semiconductors atoms are not arranged in perfect periodic structures. The effects of imperfections are treated perturbatively—as a correction to band theory. Defects can localize electronic states and cause scattering between states. A semiclassical picture is then developed where an electron travels in the material, every now and then suffering a scattering which alters its momentum and/or energy. The scattering rate is calculated using the Fermi golden rule (or Born approximation) if the perturbation is small.

The final step in semiconductor physics is an understanding of how electrons

respond to external stimuli such as electric field, magnetic field, electromagnetic field, etc. A variety of techniques, such as Boltzmann transport equations and Monte Carlo computer simulations are developed to understand the response of electrons to external stimulus.

I.3 ROLE OF THIS BOOK

This book provides the underlying physics for the topics listed in Fig. I.1. It covers "old" topics such as crystal structure and band theory in bulk semiconductors and "new" topics such as bandstructure of stained heterostructures, self-assembled quantum dots, and spin transistors. All these topics have been covered in a coherent manner so that the reader gets a good sense of the current state of semiconductor physics.

In order to provide the reader a better feel for the theoretical derivations a number of solved examples are sprinkled in the text. Additionally, there are end-of-chapter problems. This book can be used to teach a course on semiconductors physics. A rough course outline for a two semester course is shown in Table I.1. In a one semester course some section of this text can be skipped (e.g., magnetic field effects from Chapter 11) and others can be covered in less detail (e.g., Chapter 8). If a two semester course is taught, all of the material in the book can be used. It is important to note that this book can also be used for special topic courses on heterostructures or optoelectronics.

Chapter 1	• Crystal growth; crystal structure	*1 lecture*
	• Strained heterostructures	*1 lecture*
	• Polar heterostructures	*1 lecture*
Chapter 2	• Bloch theorem, metals, semiconductors, insulators	*1 lecture*
	• Tight binding method	*1-2 lectures*
	• Spin-orbit effects, symmetry of states	*1 lecture*
	• $k \cdot p$ method	*1 lecture*
	• Intrinsic and extrinsic carrier densities	*1 lecture*
Chapter 3	• Bandstructure of alloys	*1 lecture*
	• Bandstructure in quantum wells	*2 lectures*
	• Strain effects in heterostructures	*2 lectures*
Chapter 4	• Boltzmann transport equation	*1-1/2 lectures*
	• Averaging procedures	1/2 *lecture*
	• Hall effect, Hall mobility	*1 lecture*

Table I.1: Suggested set of topics for a one semester course on semiconductor physics.

Chapter 5	• Ionized impurity scattering	*1 lecture*
	• Alloy, neutral impurity scattering	*1 lecture*
	• Carrier-carrier scattering	*1 lecture*
Chapter 6	• Phonon dispersion and statistics	*2 lectures*
	• Phonon scatteringã general	*1 lecture*
	• Acoustic phonon scattering, optical phonon scattering	*2 lectures*
Chapter 7	• Low field mobility	*1 lecture*
	• Monte Carlo techniques	*2 lectures*
	• Velocity-field result discussion	*1 lecture*
	• Transport in lower dimensions	*1 lecture*
Chapter 8	Optional Chapter	
	• Bloch oscillations	*1 lecture*
	• Resonant tunneling	*1 lecture*
	• Localization issues and disorder	*1 lecture*
	• Mesoscopic systems	*2 lectures*

Table I.2: Suggested set of topics for a one semester course on semiconductor physics (con't.).

Chapter 9	• Interband transitions: Bulk and 2D	*2 lectures*
	• Intraband transitions in quantum wells	*1 lecture*
	• Charge injection and light emission	*1 lecture*
	• Nonradiative processes	*1 lecture*
Chapter 10	• Excitonic states in 3D and lower dimensions	*2 lectures*
	• Modulation of optical properties	*2 lectures*
Chapter 11	Optional Chapter	
	• Semiclassical theory of magnetotransport	*1 lecture*
	• Landau levels	*1 lecture*
	• Aharonov Bohm effect	*1/2 lecture*
	• Magnetooptic effect	*1/2 lecture*
	• "Spintronics"	

Appendix **B**: Reading assignments

Table I.3: Suggested set of topics for a one semester course on semiconductor physics (con't.).

Chapter 1

STRUCTURAL PROPERTIES OF SEMICONDUCTORS

1.1 INTRODUCTION

Semiconductors form the basis of most modern information processing devices. Electronic devices such as diodes, bipolar junction transistors, and field effect transistors drive modern electronic technology. Optoelectronic devices such as laser diodes, modulators, and detectors drive the optical networks. In addition to devices, semiconductor structures have provided the stages for exploring questions of fundamental physics. Quantum Hall effect and other phenomena associated with many-body effects and low dimensions have been studied in semiconductor structures.

It is important to recognize that the ability to examine fundamental physics issues and to use semiconductors in state of the art device technologies depends critically on the purity and perfection of the semiconductor crystal. Semiconductors are often associated with clean rooms and workers clad in "bunny suits" lest the tiniest stray particle get loose and latch onto the wafer being processed. Indeed, semiconductor structures can operate at their potential only if they can be grown with a high degree of crystallinity and if impurities and defects can be controlled. For high structural quality it is essential that a high quality substrate be available. This requires growth of bulk crystals which are then sliced and polished to allow epitaxial growth of thin semiconductor regions including heterostructures.

In this chapter we start with a brief discussion of the important bulk and epitaxial crystal growth techniques. We then discuss the important semiconductor crystal structures. We also discuss strained lattice structures and the strain tensor for such crystals. Strained epitaxy and its resultant consequences are now widely exploited in

semiconductor physics and it is important to examine how epitaxial growth causes distortions in the crystal lattice.

1.2 CRYSTAL GROWTH

1.2.1 Bulk Crystal Growth

Semiconductor technology depends critically upon the availability of high quality substrates with as large a diameter as possible. Bulk crystal growth techniques are used mainly to produce substrates on which devices are eventually fabricated. While for some semiconductors like Si and GaAs (to some extent for InP) the bulk crystal growth techniques are highly matured; for most other semiconductors it is difficult to obtain high quality, large area substrates. Several semiconductor technologies are dependent on substrates that are not ideal. For example, the nitrides GaN, AlN, InN are grown on SiC or sapphire substrates, since there is no reliable GaN substrate. The aim of the bulk crystal growth techniques is to produce single crystal boules with as large a diameter as possible and with as few defects as possible. In Si the boule diameters have reached 30 cm with boule lengths approaching 100 cm. Large size substrates ensure low cost device production.

For the growth of boules from which substrates are obtained, one starts out with a purified form of the elements that are to make up the crystal. One important technique that is used is the Czochralski (CZ) technique. In the Czochralski technique shown in Fig. 1.1, the melt of the charge (i.e., the high quality polycrystalline material) is held in a vertical crucible. The top surface of the melt is just barely above the melting temperature. A seed crystal is then lowered into the melt and slowly withdrawn. As the heat from the melt flows up the seed, the melt surface cools and the crystal begins to grow. The seed is rotated about its axis to produce a roughly circular cross-section crystal. The rotation inhibits the natural tendency of the crystal to grow along certain orientations to produce a faceted crystal.

The CZ technique is widely employed for Si, GaAs, and InP and produces long ingots (boules) with very good circular cross-section. For Si up to 100 kg ingots can be obtained. In the case of GaAs and InP the CZ technique has to face problems arising from the very high pressures of As and P at the melting temperature of the compounds. Not only does the chamber have to withstand such pressures, also the As and P leave the melt and condense on the sidewalls. To avoid the second problem one seals the melt by covering it with a molten layer of a second material (e.g., boron oxide) which floats on the surface. The technique is then referred to as liquid encapsulated Czochralski, or the LEC technique.

A second bulk crystal growth technique involves a charge of material loaded in a quartz container. The charge may be composed of either high quality polycrystalline material or carefully measured quantities of elements which make up a compound crystal. The container called a “boat” is heated till the charge melts and wets the seed crystal. The seed is then used to crystallize the melt by slowly lowering the boat temperature starting from the seed end. In the gradient-freeze approach the boat is pushed into a furnace (to melt the charge) and slowly pulled out. In the Bridgeman approach, the boat is kept stationary while the furnace temperature is temporally varied to form

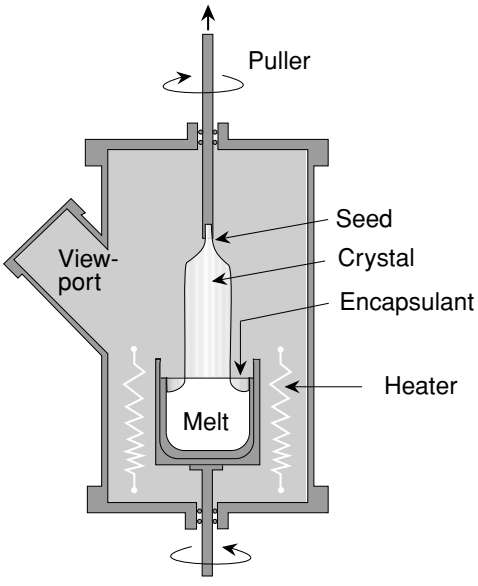


Figure 1.1: Schematic of Czochralski-style crystal grower used to produce substrate ingots. The approach is widely used for Si, GaAs and InP.

the crystal. The approaches are schematically shown in Fig. 1.2.

The easiest approach for the boat technique is to use a horizontal boat. However, the shape of the boule that is produced has a D-shaped form. To produce circular cross-sections vertical configurations have now been developed for GaAs and InP.

In addition to producing high purity bulk crystals, the techniques discussed above are also responsible for producing crystals with specified electrical properties. This may involve high resistivity materials along with n- or p-type materials. In Si it is difficult to produce high resistivity substrated by bulk crystal growth and resistivities are usually $<10^4$ Ω-cm. However, in compound semiconductors carrier trapping impurities such as chromium and iron can be used to produce material with resistivities of $\sim 10^8$ Ω cm. The high resistivity or semi-insulating (SI) substrates are extremely useful in device isolation and for high speed devices. For n- or p-type doping carefully measured dopants are added in the melt.

1.2.2 Epitaxial Crystal Growth

Once bulk crystals are grown, they are sliced into substrates or wafers about 250 μm thick. These are polished and used for growth of epitaxial layers a few micrometers thick. All active devices are produced on these epitaxial layers. As a result the epitaxial growth techniques are very important. The epitaxial growth techniques have a very slow growth rate (as low as a monolayer per second for some techniques) which allow one to control very accurately the dimensions in the growth direction. In fact, in techniques like molecular beam epitaxy (MBE) and metal organic chemical vapor

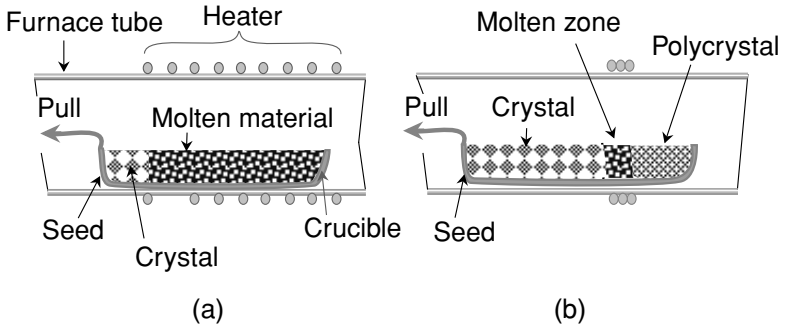


Figure 1.2: Crystal growing from the melt in a crucible: (a) solidification from one end of the melt (horizontal Bridgeman method); (b) melting and solidification in a moving zone.

deposition (MOCVD), one can achieve monolayer ($\sim$ 3 Å) control in the growth direction. This level of control is essential for the variety of heterostructure devices that are being used in optoelectronics. The epitaxial techniques are also very useful for precise doping profiles that can be achieved. In fact, it may be argued that without the advances in epitaxial techniques that have occurred over the last two decades, most of the developments in semiconductor physics would not have occurred. Table 1.1 gives a brief view of the various epitaxial techniques used along with some of the advantages and disadvantages.

Liquid Phase Epitaxy (LPE)
LPE is a relatively simple epitaxial growth technique which was widely used until 1970s when it gradually gave way to approaches such as MBE and MOCVD. It is a less expensive technique (compared to MBE or MOCVD), but it offers less control in interface abruptness when growing heterostructures. LPE is still used for growth of crystals such as HgCdTe for long wavelength detectors and AlGaAs for double heterostructure lasers. As shown in Table 1.1, LPE is a close to equilibrium technique in which the substrate is placed in a quartz or a graphite boat and covered by a liquid of the crystal to be grown (see Fig. 1.3). The liquid may also contain dopants that are to be introduced into the crystal. LPE is often used for alloy growth where the growth follows the equilibrium solid-liquid phase diagram. By precise control of the liquid composition and temperature, the alloy composition can be controlled. Because LPE is a very close to equilibrium growth technique, it is difficult to grow alloy systems which are not miscible or even grow heterostructures with atomically abrupt interfaces. Nevertheless heterostructures where interface is graded over 10-20 Å can be grown by LPE by sliding the boat over successive "puddles" of different semiconductors. For many applications such interfaces are adequate and since LPE is a relatively inexpensive growth technique, it is used in many commercial applications.

Vapor Phase Epitaxy (VPE)
A large class of epitaxial techniques rely on delivering the components that form the crystal from a gaseous environment. If one has molecular species in a gaseous form with

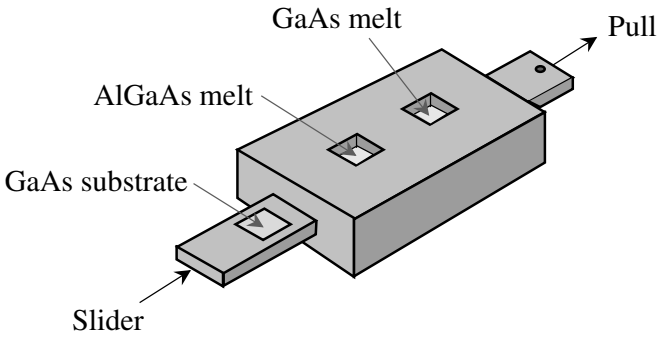


Figure 1.3: A schematic of the LPE growth of AlGaAs and GaAs. The slider moves the substrate, thus positioning itself to achieve contact with the different melts to grow heterostructures.

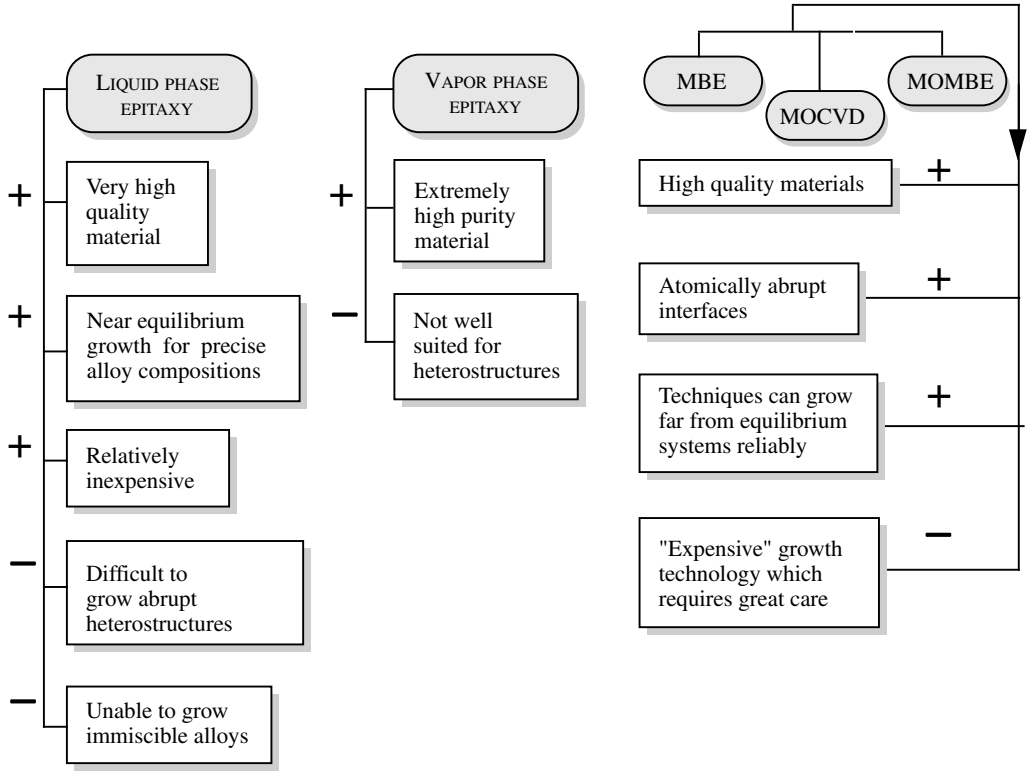


Table 1.1: A schematic of the various epitaxial crystal growth techniques and some of their positive and negative aspects.

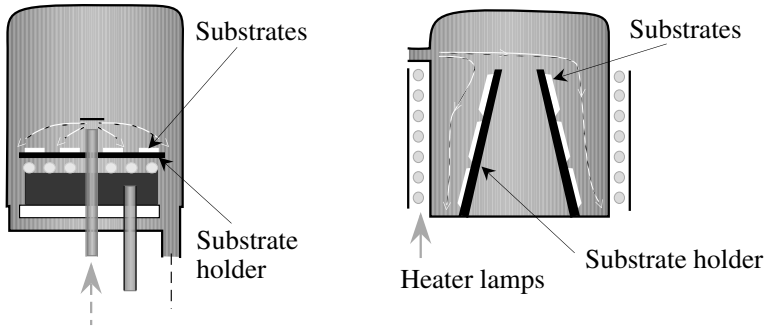


Figure 1.4: Reactors for VPE growth. The substrate temperature must be maintained uniformly over the area. This is achieved better by lamp heating. A pyrometer is used for temperature measurement.

partial pressure P, the rate at which molecules impinge upon a substrate is given by

$$F = \frac{P}{\sqrt{2\pi m k_B T}} \sim \frac{3.5 \times 10^{22} P(\text{torr})}{\sqrt{m(g)T(K)}} \text{mol./cm}^2\text{s} \tag{1.1}$$

where m is the molecular weight and T the cell temperature. For most crystals the surface density of atoms is $\sim 7 \times 10^{14}$ cm^{-2}. If the atoms or molecules impinging from the vapor can be deposited on the substrate in an ordered manner, epitaxial crystal growth can take place.

The VPE technique is used mainly for homoepitaxy and does not have the additional apparatus present in techniques such as MOCVD for precise heteroepitaxy. As an example of the technique, consider the VPE of Si. The Si containing reactant silane (SiH_4) or dichlorosilane (SiH_2Cl_2) or trichlorosilane ($SiHCl_3$) or silicon tetrachloride ($SiCl_4$) is diluted in hydrogen and introduced into a reactor in which heated substrates are placed as shown in Fig. 1.4. The silane pyrolysis to yield silicon while the chlorine containing gases react to give $SiCl_2$, HCl and various other silicon-hydrogen-chlorine compounds. The reaction

$$2SiCl_2 \rightleftharpoons Si + SiCl_4 \tag{1.2}$$

then yields Si. Since HCl is also produced in the reaction, conditions must be tailored so that no etching of Si occurs by the HCl. Doping can be carried out by adding appropriate hydrides (phosphine, arsine, etc.,) to the reactants.

VPE can be used for other semiconductors as well by choosing different appropriate reactant gases. The reactants used are quite similar to those employed in the MOCVD technique discussed later.

Molecular Beam Epitaxy (MBE)

MBE is capable of controlling deposition of submonolayer coverage on a substrate and has become one of the most important epitaxial techniques. Almost every semiconductor

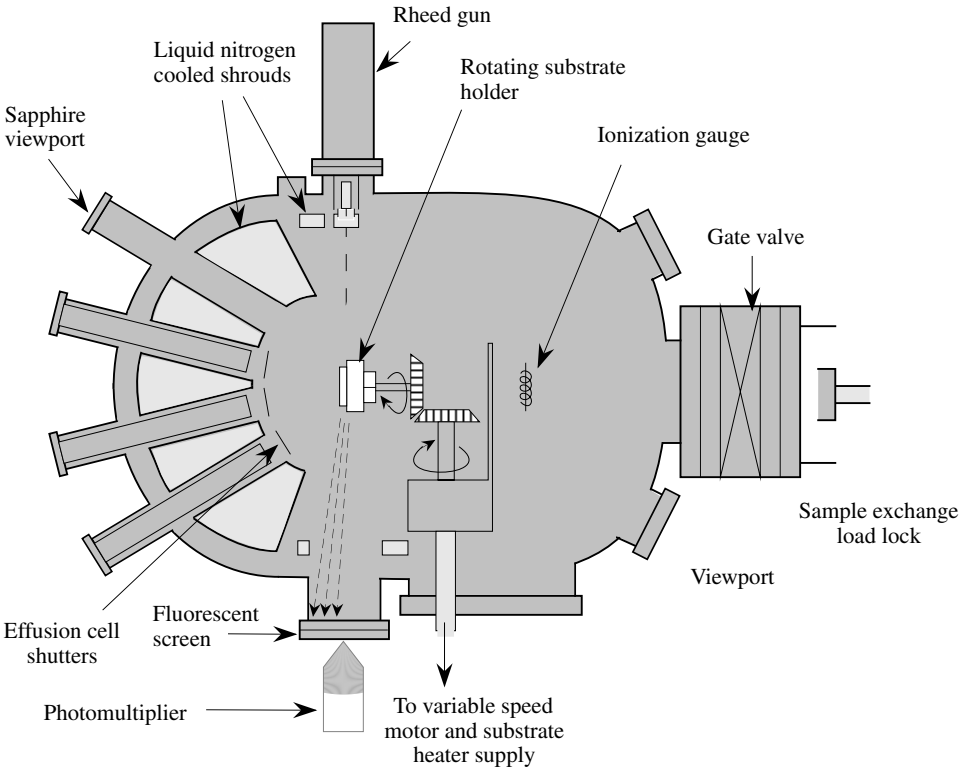


Figure 1.5: A schematic of the MBE growth system.

has been grown by this technique. MBE is a high vacuum technique ($\sim 10^{-11}$ torr vacuum when fully pumped down) in which crucibles containing a variety of elemental charges are placed in the growth chamber (Fig. 1.5). The elements contained in the crucibles make up the components of the crystal to be grown as well as the dopants that may be used. When a crucible is heated, atoms or molecules of the charge are evaporated and these travel in straight lines to impinge on a heated substrate.

The growth rate in MBE is $\sim$1.0 monolayer per second and this slow rate coupled with shutters placed in front of the crucibles allow one to switch the composition of the growing crystal with monolayer control. Since no chemical reactions occur in MBE, the growth is the simplest of all epitaxial techniques and is quite controllable. However, since the growth involves high vacuum, leaks can be a major problem. The growth chamber walls are usually cooled by liquid N_2 to ensure high vacuum and to prevent atoms/molecules to come off from the chamber walls.

The low background pressure in MBE allows one to use electron beams to monitor the growing crystal. The reflection high energy electron diffraction (RHEED) techniques relies on electron diffraction to monitor both the quality of the growing substrate and the layer by layer growth mode.

Metal Organic Chemical Vapor Deposition (MOCVD)

Metal organic chemical vapor deposition (MOCVD) is another important growth technique widely used for heteroepitaxy. Like MBE, it is also capable of producing monolayer abrupt interfaces between semiconductors. A typical MOCVD system is shown in Fig. 1.6. Unlike in MBE, the gases that are used in MOCVD are not made of single elements, but are complex molecules which contain elements like Ga or As to form the crystal. Thus the growth depends upon the chemical reactions occurring at the heated substrate surface. For example, in the growth of GaAs one often uses triethyl gallium and arsine and the crystal growth depends upon the following reaction:

$$\mathrm{Ga(CH_3)_3 + AsH_3 \rightleftharpoons GaAs + 3CH_4} \tag{1.3}$$

One advantage of the growth occurring via a chemical reaction is that one can use lateral temperature control to carry out local area growth. Laser assisted local area growth is also possible for some materials and can be used to produce new kinds of device structures. Such local area growth is difficult in MBE.

There are several varieties of MOCVD reactors. In the atmospheric MOCVD the growth chamber is essentially at atmospheric pressure. One needs a large amount of gases for growth in this case, although one does not have the problems associated with vacuum generation. In the low pressure MOCVD the growth chamber pressure is kept low. The growth rate is then slower as in the MBE case.

The use of the MOCVD equipment requires very serious safety precautions. The gases used are highly toxic and a great many safety features have to be incorporated to avoid any deadly accidents. Safety and environmental concerns are important issues in almost all semiconductor manufacturing since quite often one has to deal with toxic and hazardous materials.

In addition to MBE and MOCVD one has hybrid epitaxial techniques often called MOMBE (metal organic MBE) which try to combine the best of MBE and MOCVD. In MBE one has to open the chamber to load the charge for the materials to be grown while this is avoided in MOCVD where gas bottles can be easily replaced from outside. Additionally, in MBE one has occasional spitting of material in which small clumps of atoms are evaporated off on to the substrate. This is avoided in MOCVD and MOMBE.

EXAMPLE 1.1 Consider the growth of GaAs by MBE. The Ga partial pressure in the growth chamber is 10^{-5} Torr, and the Ga cell temperature is 900 K. Calculate the flux of Ga atoms on the substrate. The surface density of Ga atoms on GaAs grown along (001) direction is 6.3×10^{14} cm^{-2}. Calculate the growth rate if all of the impinging atoms stick to the substrate.

The mass of Ga atoms is 70 g/mole. The flux is (from Eqn. 1.1)

$$F = \frac{3.5 \times 10^{22} \times 10^{-5}}{\sqrt{70 \times 900}} = 5.27 \times 10^{14} \text{atoms/cm}^2$$

Note that the surface density of Ga atoms on GaAs is $\sim 6.3 \times 10^{14}$ cm^{-2}. Thus, if all of the Ga atoms were to stick, the growth rate would be ~0.8 monolayer per second. This assumes that there is sufficient arsenic to provide As in the crystal. This is a typical growth rate for epitaxial films. It would take nearly 10 hours to grow a 10 μm film.

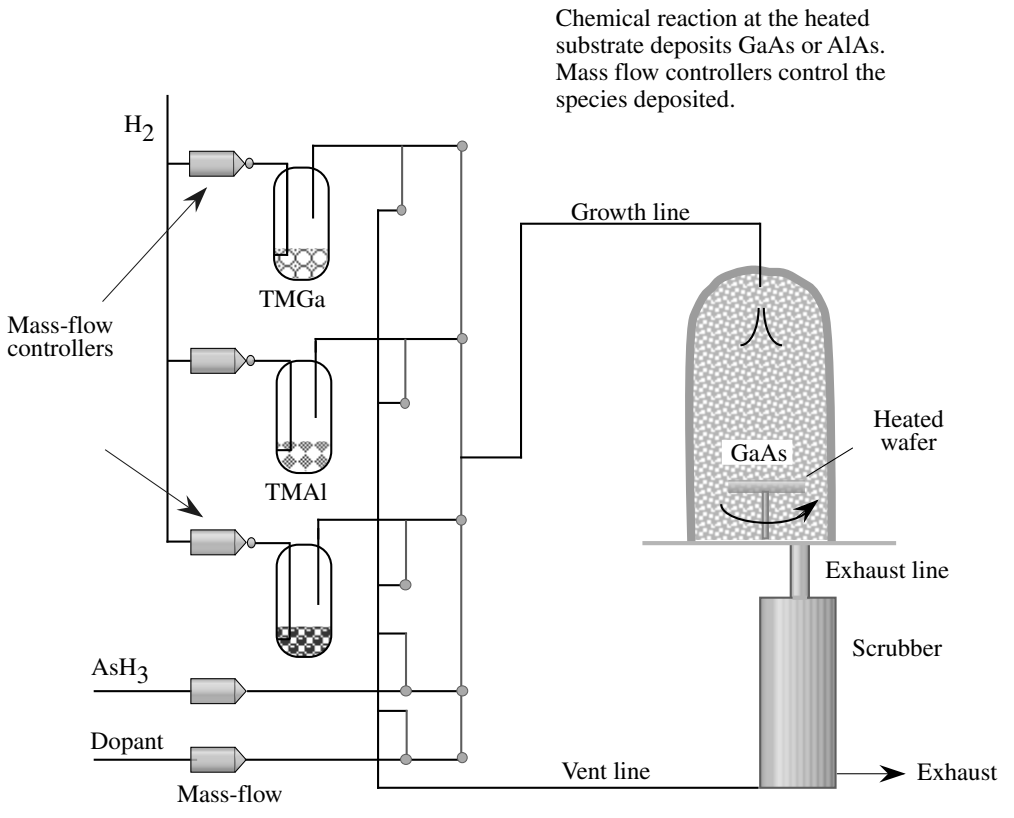


Figure 1.6: Schematic diagram of an MOCVD system employing alkyds (trimethyl gallium (TMGa) and trimethyl aluminum (TMAl) and metal hydride (arsine) material sources, with hydrogen as a carrier gas.

1.2.3 Epitaxial Regrowth

The spectacular growth of semiconductor microelectronics owes a great deal to the concept of the integrated circuit. The ability to fabricate transistors, resistors, inductors and capacitors on the same wafer is critical to the low cost and high reliability we have come to expect from microelectronics. It is natural to expect similar dividents from the concept of the optoelectronic integrated circuit (OEIC). In the OEIC, the optoelectronic device (the laser or detector or modulator) would be integrated on the same wafer with an amplifier or logic gates.

One of the key issues in OEICs involves etching and regrowth. As we will see

later, the optoelectronic devices have a structure that is usually not compatible with the structure of an electronic device. The optimum layout then involves growing one of the device structures epitaxially and then masking the region to be used as, say, the optoelectronic device and etching away the epitaxial region. Next a regrowth is done to grow the electronic device with a different structure. The process is shown schematically in Fig. 1.7. While this process looks simple conceptually, there are serious problems associated with etching and regrowth.

A critical issue in the epitaxial growth of a semiconductor layer is the quality of the semiconductor-vacuum interface. This semiconductor surface must be "clean," i.e., there should be no impurity layers (e.g., an oxide layer) on the surface. Even if a fraction of a monolayer of the surface atoms have impurities bonded to them, the quality of the epitaxial layer suffers drastically. The growth may occur to produce microcrystalline regions separated by grain boundaries or may be amorphous in nature. In either case, the special properties arising from the crystalline nature of the material (to be discussed in the next chapter) are then lost.

The issue of surface cleanliness and surface reconstruction can be addressed when one is doing a single epitaxial growth. For example, a clean wafer can be loaded into the growth chamber and the remaining impurities on the surface can be removed by heating the substrate. The proper reconstruction (which can be monitored by RHEED) can be ensured by adjusting the substrate temperature and specy overpressure. Now consider the problems associated with etching after the first epitaxial growth has occurred. As the etching starts, foreign atoms or molecules are introduced on the wafer as the semiconductor is etched. The etching process is quite damaging and as it ends, the surface of the etched wafer is quite rough and damaged. In addition, in most growth techniques the wafer has to be physically moved from the high purity growth chamber to the etching system. During this transportation, the surface of the wafer may collect some "dirt." During the etching process this "dirt" may not be etched off and may remain on the wafer. As a result of impurities and surface damage, when the second epitaxial layer is grown after etching, the quality of the layer suffers.

A great deal of processing research in OEICs focusses on improving the etching/regrowth process. So far the OEICs fabricated in various laboratories have performances barely approaching the performance of hybrid circuits. Clearly the problem of etching/regrowth is hampering the progress in OEIC technology.

It may be noted that the etching regrowth technology is also important in creating quantum wires and quantum dots which require lateral patterning of epitaxial layers.

1.3 CRYSTAL STRUCTURE

Essentially all high performance semiconductor devices are based on crystalline materials. there are some devices that use low cast amorphous or polycrystalline semiconductors, but their performance is quite poor. Crystals are made up of identical building blocks, the block being an atom or a group of atoms. While in "natural" crystals the crystalline symmetry is fixed by nature, new advances in crystal growth techniques are allowing scientists to produce artificial crystals with modified crystalline structure.

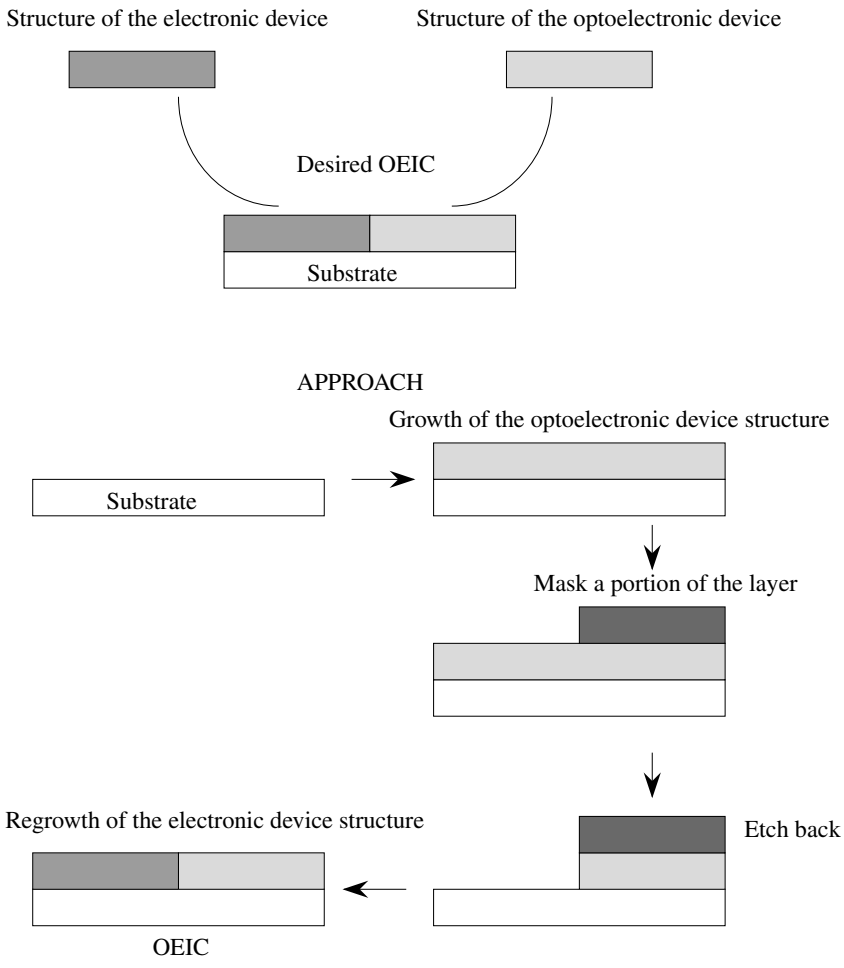


Figure 1.7: The importance of regrowth is clear when one examines the difference in the structure of electronic and optoelectronic devices. Etching and regrowth is essential for fabrication of optoelectronic integrated circuits (OEIC).

These advances depend upon being able to place atomic layers with exact precision and control during growth, leading to "superlattices". To define the crystal structure, two important concepts are introduced. The *lattice* represents a set of points in space which form a periodic structure. Each point sees an exact similar environment. The lattice is by itself a mathematical abstraction. A building block of atoms called the *basis* is then attached to each lattice point yielding the crystal structure.

An important property of a lattice is the ability to define three vectors $\mathbf{a}_1$, $\mathbf{a}_2$, $\mathbf{a}_3$, such that any lattice point $\mathbf{R}'$ can be obtained from any other lattice point $\mathbf{R}$ by a translation

$$\mathbf{R}' = \mathbf{R} + m_1\mathbf{a}_1 + m_2\mathbf{a}_2 + m_3\mathbf{a}_3 \tag{1.4}$$

where m_1, m_2, m_3 are integers. Such a lattice is called Bravais lattice. The entire lattice can be generated by choosing all possible combinations of the integers m_1, m_2, m_3 . The crystalline structure is now produced by attaching the basis to each of these lattice points.

$$\boxed{\text{lattice} + \text{basis} = \text{crystal structure}} \tag{1.5}$$

The translation vectors $\mathbf{a}_1$, $\mathbf{a}_2$, and $\mathbf{a}_3$ are called primitive if the volume of the cell formed by them is the smallest possible. There is no unique way to choose the primitive vectors. One choice is to pick

$\mathbf{a}_1$ to be the shortest period of the lattice
$\mathbf{a}_2$ to be the shortest period not parallel to $\mathbf{a}_1$
$\mathbf{a}_3$ to be the shortest period not coplanar with $\mathbf{a}_1$ and $\mathbf{a}_2$

It is possible to define more than one set of primitive vectors for a given lattice, and often the choice depends upon convenience. The volume cell enclosed by the primitive vectors is called the *primitive unit cell.*

Because of the periodicity of a lattice, it is useful to define the symmetry of the structure. The symmetry is defined via a set of point group operations which involve a set of operations applied around a point. The operations involve rotation, reflection and inversion. The symmetry plays a very important role in the electronic properties of the crystals. For example, the inversion symmetry is extremely important and many physical properties of semiconductors are tied to the absence of this symmetry. As will be clear later, in the diamond structure (Si, Ge, C, etc.), inversion symmetry is present, while in the Zinc Blende structure (GaAs, AlAs, InAs, etc.), it is absent. Because of this lack of inversion symmetry, these semiconductors are piezoelectric, i.e., when they are strained an electric potential is developed across the opposite faces of the crystal. In crystals with inversion symmetry, where the two faces are identical, this is not possible.

1.3.1 Basic Lattice Types

The various kinds of lattice structures possible in nature are described by the symmetry group that describes their properties. Rotation is one of the important symmetry groups. Lattices can be found which have a rotation symmetry of 2π, $\frac{2\pi}{2}$, $\frac{2\pi}{3}$, $\frac{2\pi}{4}$, $\frac{2\pi}{6}$. The rotation symmetries are denoted by 1, 2, 3, 4, and 6. No other rotation axes exist; e.g., $\frac{2\pi}{5}$ or $\frac{2\pi}{7}$ are not allowed because such a structure could not fill up an infinite space.

There are 14 types of lattices in 3D. These lattice classes are defined by the relationships between the primitive vectors a_1, a_2, and a_3, and the angles α, β, and γ between them. The general lattice is triclinic ($\alpha \neq \beta \neq \gamma, a_1 \neq a_2 \neq a_3$) and there are 13 special lattices. Table 1.2 provides the basic properties of these three dimensional lattices. We will focus on the cubic lattice which is the structure taken by all semiconductors.

There are 3 kinds of cubic lattices: simple cubic, body centered cubic, and face centered cubic.

System	Number of lattices	Restrictions on conventional cell axes and singles
Triclinic	1	$a_1 \neq a_2 \neq a_3$ $\alpha \neq \beta \neq \gamma$
Monoclinic	2	$a_1 \neq a_2 \neq a_3$ $\alpha = \gamma = 90^o \neq \beta$
Orthorhombic	4	$a_1 \neq a_2 \neq a_3$ $\alpha = \beta = \gamma = 90^o$
Tetragonal	2	$a_1 = a_2 \neq a_3$ $\alpha = \beta = \gamma = 90^o$
Cubic	3	$a_1 = a_2 = a_3$ $\alpha = \beta = \gamma = 90^o$
Trigonal	1	$a_1 = a_2 = a_3$ $\alpha = \beta = \gamma < 120^o, \neq 90^o$
Hexagonal	1	$a_1 = a_2 \neq a_3$ $\alpha = \beta = 90^o$ $\gamma = 120^o$

Table 1.2: The 14 Bravais lattices in 3-dimensional systems and their properties.

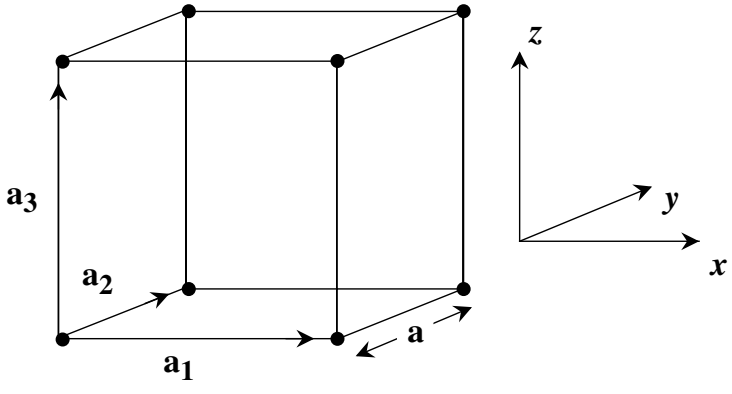


Figure 1.8: A simple cubic lattice showing the primitive vectors. The crystal is produced by repeating the cubic cell through space.

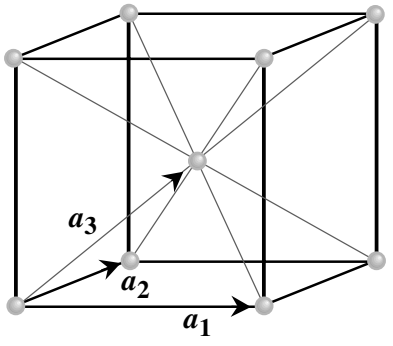


Figure 1.9: The body centered cubic lattice along with a choice of primitive vectors.

Simple cubic: The simple cubic lattice shown in Fig. 1.8 is generated by the primitive vectors

$$a\mathbf{x}, a\mathbf{y}, a\mathbf{z} \tag{1.6}$$

where the $\mathbf{x}$, $\mathbf{y}$, $\mathbf{z}$ are unit vectors.

Body-centered cubic: The bcc lattice shown in Fig. 1.9 can be generated from the simple cubic structure by placing a lattice point at the center of the cube. If $\hat{\mathbf{x}}, \hat{\mathbf{y}}$, and $\hat{\mathbf{z}}$ are three orthogonal unit vectors, then a set of primitive vectors for the body-centered cubic lattice could be

$$a_1 = a\hat{\mathbf{x}}, a_2 = a\hat{\mathbf{y}}, a_3 = \frac{a}{2}(\hat{\mathbf{x}} + \hat{\mathbf{y}} + \hat{\mathbf{z}}) \tag{1.7}$$

A more symmetric set for the bcc lattice is

$$a_1 = \frac{a}{2}(\hat{\mathbf{y}} + \hat{\mathbf{z}} - \hat{\mathbf{x}}), a_2 = \frac{a}{2}(\hat{\mathbf{z}} + \hat{\mathbf{x}} - \hat{\mathbf{y}}), a_3 = \frac{a}{2}(\hat{\mathbf{x}} + \hat{\mathbf{y}} - \hat{\mathbf{z}}) \tag{1.8}$$

Face Centered Cubic: Another equally important lattice for semiconductors is the *face-centered cubic* (fcc) Bravais lattice. To construct the face-centered cubic Bravais lattice add to the simple cubic lattice an additional point in the center of each square face (Fig. 1.10).

A symmetric set of primitive vectors for the face-centered cubic lattice (see Fig. 1.10) is

$$a_1 = \frac{a}{2}(\hat{y} + \hat{z}), a_2 = \frac{a}{2}(\hat{z} + \hat{x}), a_3 = \frac{a}{2}(\hat{x} + \hat{y}) \tag{1.9}$$

The face-centered cubic and body-centered cubic Bravais lattices are of great importance, since an enormous variety of solids crystallize in these forms with an atom (or ion) at each lattice site. Essentially all semiconductors of interest for electronics and optoelectronics have fcc structure.

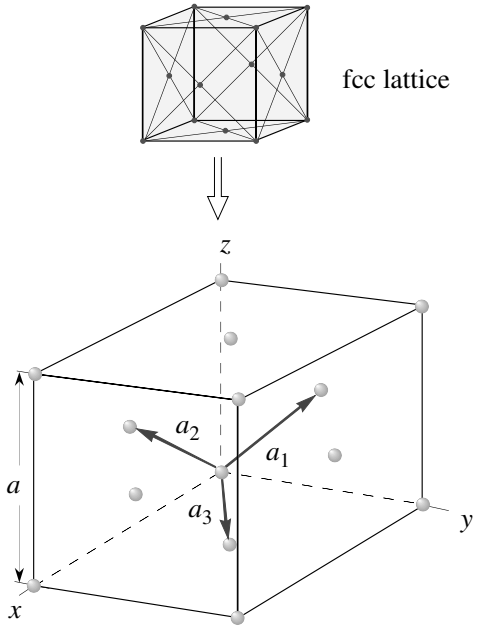


Figure 1.10: Primitive basis vectors for the face centered cubic lattice.

1.3.2 Basic Crystal Structures

Diamond and Zinc Blende Structures

Most semiconductors of interest for electronics and optoelectronics have an underlying fcc lattice. However, they have two atoms per basis. The coordinates of the two basis atoms are

$$(000) \text{ and } (\frac{a}{4}, \frac{a}{4}, \frac{a}{4}) \tag{1.10}$$

Since each atom lies on its own fcc lattice, such a two atom basis structure may be thought of as two inter-penetrating fcc lattices, one displaced from the other by a translation along a body diagonal direction $(\frac{a}{4}\frac{a}{4}\frac{a}{4})$.

Figure 1.11 gives details of this important structure. If the two atoms of the basis are identical, the structure is called diamond. Semiconductors such as Si, Ge, C, etc., fall in this category. If the two atoms are different, the structure is called the Zinc Blende structure. Semiconductors such as GaAs, AlAs, CdS, etc., fall in this category. Semiconductors with diamond structure are often called elemental semiconductors, while the Zinc Blende semiconductors are called compound semiconductors. The compound semiconductors are also denoted by the position of the atoms in the periodic chart, e.g., GaAs, AlAs, InP are called III-V (three-five) semiconductors while CdS, HgTe, CdTe, etc., are called II-VI (two-six) semiconductors.

Hexagonal Close Pack Structure The hexagonal close pack (hcp) structure is an important lattice structure and many metals have this underlying lattice. Some

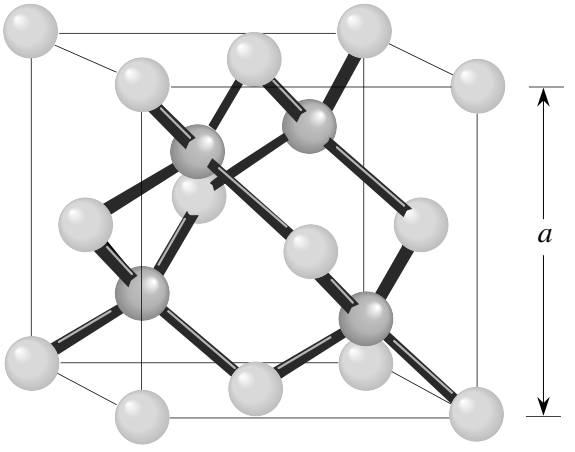


Figure 1.11: The zinc blende crystal structure. The structure consists of the interpenetrating fcc lattices, one displaced from the other by a distance $(\frac{a}{4}\frac{a}{4}\frac{a}{4})$ along the body diagonal. The underlying Bravais lattice is fcc with a two atom basis. The positions of the two atoms is (000) and $(\frac{a}{4}\frac{a}{4}\frac{a}{4})$.

semiconductors such as BN, AlN, GaN, SiC, etc., also have this underlying lattice (with a two-atom basis). The hcp structure is formed as shown in Fig. 1.12a. Imagine that a close-packed layer of spheres is formed. Each sphere touches six other spheres, leaving cavities, as shown. A second close-packed layer of spheres is placed on top of the first one so that the second layer sphere centers are in the cavities formed by the first layer. The third layer of close-packed spheres can now be placed so that center of the spheres do not fall on the center of the starting spheres (left side of Fig. 1.12a) or coincide with the centers of the starting spheres (right side of Fig. 1.12b). These two sequences, when repeated, produce the fcc and hcp lattices.

In Fig. 1.12b we show the detailed positions of the lattice points in the hcp lattice. The three lattice vectors are $a_1,\ a_2\ a_3$, as shown. The vector a_3 is denoted by c and the term c-axis refers to the orientation of a_3. In an ideal structure, if $\mid a \mid = \mid a_1 \mid = \mid a_2 \mid$,

$$\frac{c}{a} = \sqrt{\frac{8}{3}} \tag{1.11}$$

In Table 1.3 we show the structural properties of some important materials. If two or more semiconductors are randomly mixed to produce an alloy, the lattice constant of the alloy is given by Vegard's law according to which the alloy lattice constant is the weighted mean of the lattice constants of the individual components.

1.3.3 Notation to Denote Planes and Points in a Lattice: Miller Indices

A simple scheme is used to describe lattice planes, directions and points. For a plane, we use the following procedure:

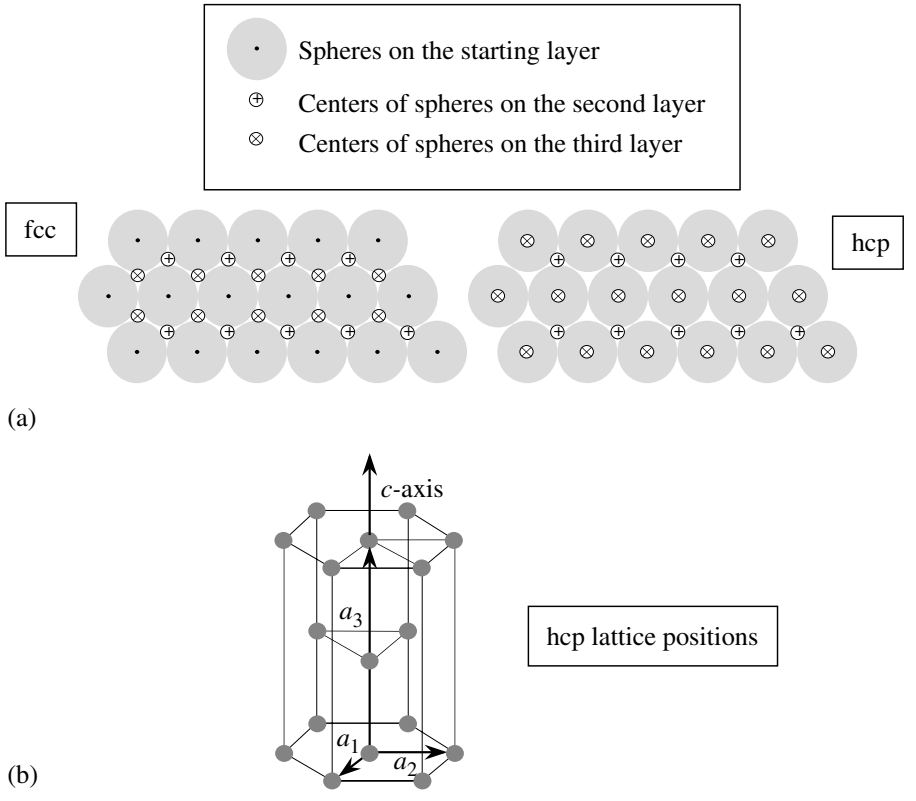


Figure 1.12: (a) A schematic of how the fcc and hcp lattices are formed by close packing of spheres. (b) Arrangement of lattice points on an hcp lattice.

(1) Define the x, y, z axes (primitive vectors).
(2) Take the intercepts of the plane along the axes in units of lattice constants.
(3) Take the reciprocal of the intercepts and reduce them to the smallest integers.

The notation (hkl) denotes a family of parallel planes.
The notation (hkl) denotes a family of equivalent planes.

To denote directions, we use the smallest set of integers having the same ratio as the direction cosines of the direction.

In a cubic system the Miller indices of a plane are the same as the direction perpendicular to the plane. The notation [ ] is for a set of parallel directions; $< >$ is for a set of equivalent direction. Fig. 1.13 shows some examples of the use of the Miller indices to define planes.

EXAMPLE 1.2 The lattice constant of silicon is 5.43 Å. Calculate the number of silicon atoms in a cubic centimeter. Also calculate the number density of Ga atoms in GaAs which

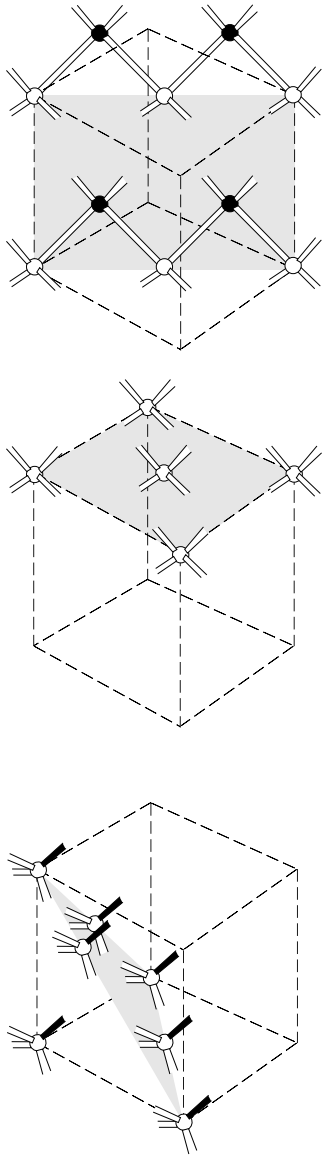
Material	Structure	Lattice Constant (Å)	Density (gm/cm^3)
C	Diamond	3.5668	3.5153
Si	Diamond	5.431	2.329
Ge	Diamond	5.658	5.323
GaAs	Zinc Blende	5.653	5.318
AlAs	Zinc Blende	5.660	3.760
InAs	Zinc Blende	6.058	5.667
GaN	Wurtzite	$a = 3.175$; $c = 5.158$	6.095
AlN	Wurtzite	$a = 3.111$; $c = 4.981$	3.255
SiC	Zinc Blende	4.360	3.166
Cd	hcp	$a = 2.98$; $c = 5.620$	8.65
Cr	bcc	2.88	7.19
Co	hcp	$a = 2.51$; $c = 4.07$	8.9
Au	fcc	4.08	19.3
Fe	bcc	2.87	7.86
Ag	fcc	4.09	10.5
Al	fcc	4.05	2.7
Cu	fcc	3.61	8.96

Table 1.3: Structure, lattice constant, and density of some materials at room temperature.

has a lattice constant of 5.65 Å.

Silicon has a diamond structure which is made up of the fcc lattice with two atoms on each lattice point. The fcc unit cube has a volume a^3. The cube has eight lattice sites at the cube edges. However, each of these points is shared with eight other cubes. In addition, there are six lattice points on the cube face centers. Each of these points is shared by two adjacent cubes. Thus the number of lattice points per cube of volume a^3 are

$$N(a^3) = \frac{8}{8} + \frac{6}{2} = 4$$



ATOMS ON THE (110) PLANE

Each atom has 4 bonds:

- 2 bonds in the (110) plane
- 1 bond connects each atom to adjacent (110) planes

$\Longrightarrow$ Cleaving adjacent planes requires breaking 1 bond per atom

ATOMS ON THE (001) PLANE

2 bonds connect each atom to adjacent (001) plane

Atoms are either Ga or As in a GaAs crystal

$\Longrightarrow$ Cleaving adjacent planes requires breaking 2 bonds per atom

ATOMS ON THE (111) PLANE

Could be either Ga or As

1 bond connecting an adjacent plane on one side

3 bonds connecting an adjacent plane on the other side

Figure 1.13: Some important planes in the cubic system along with their Miller indices. This figure also shows how many bonds connect adjacent planes. This number determines how easy or difficult it is to cleave the crystal along these planes.

In silicon there are two silicon atoms per lattice point. The number density is, therefore,

$$N_{Si} = \frac{4 \times 2}{a^3} = \frac{4 \times 2}{(5.43 \times 10^{-8})^3} = 4.997 \times 10^{22} \text{ atoms/cm}^3$$

In GaAs, there is one Ga atom and one As atom per lattice point. The Ga atom density is, therefore,

$$N_{Ga} = \frac{4}{a^3} = \frac{4}{(5.65 \times 10^{-8})^3} = 2.22 \times 10^{22} \text{ atoms/cm}^3$$

There are an equal number of As atoms.

EXAMPLE 1.3 In semiconductor technology, a Si device on a VLSI chip represents one of the smallest devices while a GaAs laser represents one of the larger devices. Consider a Si device with dimensions $(5 \times 2 \times 1)$ μm^3 and a GaAs semiconductor laser with dimensions $(200 \times 10 \times 5)$ μm^3. Calculate the number of atoms in each device.

From Example 1.1 the number of Si atoms in the Si transistor are

$$N_{Si} = (5 \times 10^{22} \text{ atoms/cm}^3)(10 \times 10^{-12} \text{ cm}^3) = 5 \times 10^{11} \text{ atoms}$$

The number of Ga atoms in the GaAs laser are

$$N_{Ga} = (2.22 \times 10^{22})(10^4 \times 10^{-12}) = 2.22 \times 10^{14} \text{ atoms}$$

An equal number of As atoms are also present in the laser.

EXAMPLE 1.4 Calculate the surface density of Ga atoms on a Ga terminated (001) GaAs surface.

In the (001) surfaces, the top atoms are either Ga or As leading to the terminology Ga terminated (or Ga stabilized) and As terminated (or As stabilized), respectively. A square of area a^2 has four atoms on the edges of the square and one atom at the center of the square. The atoms on the square edges are shared by a total of four squares. The total number of atoms per square is

$$N(a^2) = \frac{4}{4} + 1 = 2$$

The surface density is then

$$N_{Ga} = \frac{2}{a^2} = \frac{2}{(5.65 \times 10^{-8})^2} = 6.26 \times 10^{14} \text{ cm}^{-2}$$

EXAMPLE 1.5 Calculate the height of a GaAs monolayer in the (001) direction.

In the case of GaAs, a monolayer is defined as the combination of a Ga and As atomic layer. The monolayer distance in the (001) direction is simply

$$A_{m\ell} = \frac{a}{2} = \frac{5.65}{2} = 2.825 \text{ Å}$$

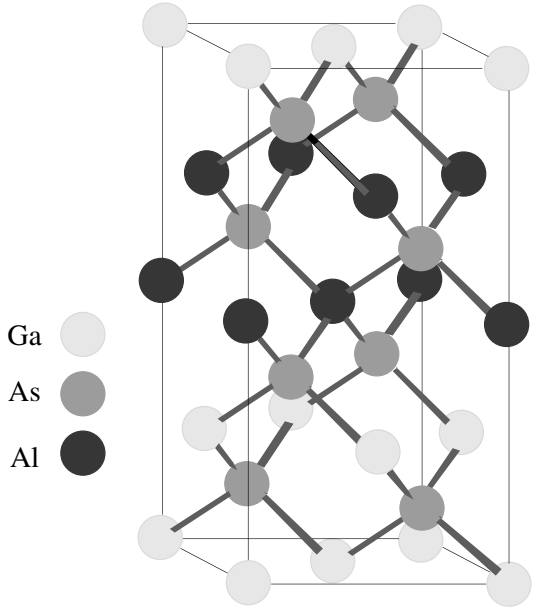


Figure 1.14: Arrangement of atoms in a $(GaAs)_2(AlAs)_2$ superlattice grown along (001) direction.

1.3.4 Artificial Structures: Superlattices and Quantum Wells

It is known that electrons and optical properties can be altered by using heterostructures, i.e., combinations of more that one semiconductor. MBE or MOCVD are techniques which allow monolayer ($\sim$3 Å) control in the chemical composition of the growing crystal. Nearly every semiconductor extending from zero bandgap (α-Sn,HgCdTe) to large bandgap materials such as ZnSe,CdS, etc., has been grown by epitaxial techniques such as MBE and MOCVD. Heteroepitaxial techniques allow one to grow heterostructures with atomic control, one can change the periodicity of the crystal in the growth direction. This leads to the concept of superlattices where two (or more) semiconductors A and B are grown alternately with thicknesses d_A and d_B respectively. The periodicity of the lattice in the growth direction is then $d_A + d_B$. A $(GaAs)_2$ $(AlAs)_2$ superlattice is illustrated in Fig. 1.14. It is a great testimony to the precision of the new growth techniques that values of d_A and d_B as low as monolayer have been grown.

It is important to point out that the most widely used heterostructures are not superlattices but quantum wells, in which a single layer of one semiconductor is sandwiched between two layers of a larger bandgap material. Such structures allow one to exploit special quantum effects that have become very useful in electronic and optoelectronic devices.

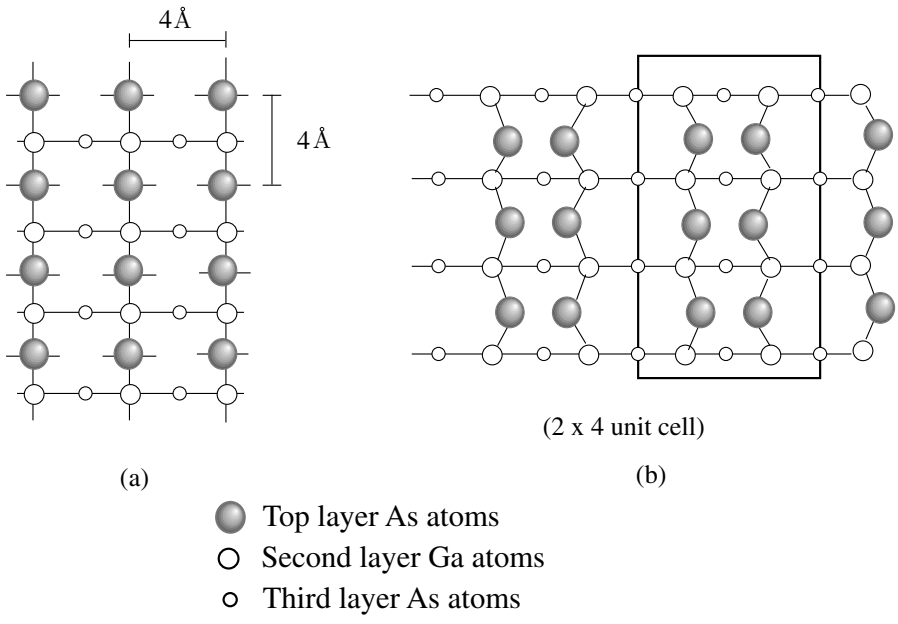


Figure 1.15: The structure (a) of the unreconstructed GaAs (001) arsenic-rich surface. The missing dimer model (b) for the GaAs (001) (2×4) surface. The As dimers are missing to create a 4 unit periodicity along one direction and a two unit periodicity along the perpendicular direction.

1.3.5 Surfaces: Ideal Versus Real

The crystalline and electronic properties are quite different from the properties of the bulk material. The bulk crystal structure is decided by the internal chemical energy of the atoms forming the crystal with a certain number of nearest neighbors, second nearest neighbors, etc. At the surface, the number of neighbors is suddenly altered. Thus the spatial geometries which were providing the lowest energy configuration in the bulk may not provide the lowest energy configuration at the surface. Thus, there is a readjustment or "reconstruction" of the surface bonds towards an energy minimizing configuration.

An example of such a reconstruction is shown for the GaAs surface in Fig. 1.15. The figure (a) shows an ideal (001) surface where the topmost atoms form a square lattice. The surface atoms have two nearest neighbor bonds (Ga-As) with the layer below, four second neighbor bonds (e.g., Ga-Ga or As-As) with the next lower layer, and four second neighbor bonds within the same layer. In a "real" surface, the arrangement of atoms is far more complex. We could denote the ideal surface by the symbol C(1×1), representing the fact that the surface periodicity is one unit by one unit along the square lattice along [110] and [$\bar{1}$10]. The reconstructed surfaces that occur in nature are generally classified as C(2×8) or C(2×4) etc., representing the increased periodicity along the [$\bar{1}$10] and [110] respectively. The C(2×4) case is shown schematically in Fig.

1.15b, for an arsenic stabilized surface (i.e., the top monolayer is As). The As atoms on the surface form dimers (along $[\bar{1}10]$ on the surface to strengthen their bonds. In addition, rows of missing dimers cause a longer range ordering as shown to increase the periodicity along the [110] direction to cause a C(2×4) unit cell. The surface periodicity is directly reflected in the x-ray diffraction pattern.

A similar effect occurs for the (110) surface of GaAs. This surface has both Ga and As atoms (the cations and anions) on the surface. A strong driving force exists to move the surface atoms and minimize the surface energy. Reconstruction effects also occur in silicon surfaces, where depending upon surface conditions a variety of reconstructions are observed. Surface reconstructions are very important since often the quality of the epitaxial crystal growth depends critically on the surface reconstruction.

EXAMPLE 1.6 Calculate the planar density of atoms on the (111) surface of Ge.

As can be seen from Fig. 1.13, we can form a triangle on the (111) surface. There are three atoms on the tips of the triangle. These atoms are shared by six other similar triangles. There are also 3 atoms along the edges of the triangle which are shared by two adjacent triangles. Thus the number of atoms in the triangle are

$$\frac{3}{6} + \frac{3}{2} = 2$$

The area of the triangle is $\sqrt{3}a^2/2$. The density of Ge atoms on the surface is then 7.29×10^{14} cm^{-2}.

1.3.6 Interfaces

Like surfaces, interfaces are an integral part of semiconductor devices. We have already discussed the concept of heterostructures and superlattices which involve interfaces between two semiconductors. These interfaces are usually of high quality with essentially no broken bonds, except for dislocations in strained structures (to be discussed later). There is, nevertheless, an *interface roughness* of one or two monolayers which is produced because of either non-ideal growth conditions or imprecise shutter control in the switching of the semiconductor species. The general picture of such a rough interface is as shown in Fig. 1.16 for epitaxially grown interfaces. The crystallinity and periodicity in the underlying lattice is maintained, but the chemical species have some disorder on interfacial planes. Such a disorder is quite important in many electronic and optoelectronic devices.

One of the most important interfaces in electronics is the Si/SiO_2 interface. This interface and its quality is responsible for essentially all of the modern consumer electronic revolution. This interface represents a situation where two materials with very different lattice constants and crystal structures are brought together. However, in spite of these large differences the interface quality is quite good. In Fig. 1.17 we show a TEM cross-section of a Si/SiO_2 interface. It appears that the interface has a region of a few monolayers of amorphous or disordered Si/SiO_2 region creating fluctuations in the chemical species (and consequently in potential energy) across the interface. This interface roughness is responsible for reducing mobility of electrons and holes in MOS devices. It can also lead to "trap" states, which can seriously deteriorate device performance if the interface quality is poor.

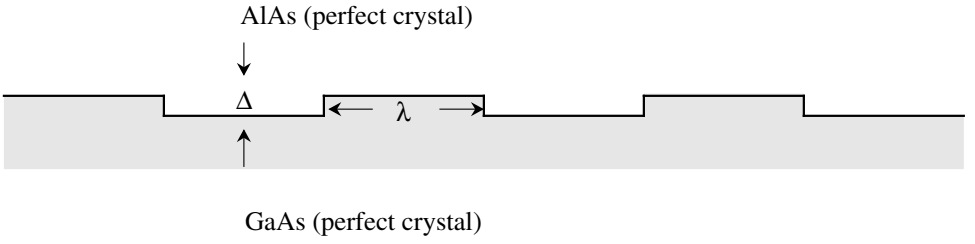


Figure 1.16: A schematic picture of the interfaces between materials with similar lattice constants such as GaAs/AlAs. No loss of crystalline lattice and long range order is suffered in such interfaces. The interface is characterized by islands of height Δ and lateral extent λ.

Finally, we have the interfaces formed between metals and semiconductors. Structurally, these important interfaces are hardest to characterize. These interfaces are usually produced in presence of high temperatures and involve diffusion of metal elements along with complex chemical reactions. The "interfacial region" usually extends over several hundred Angstroms and is a complex non-crystalline region.

1.3.7 Defects in Semiconductors

In the previous section we have discussed the properties of the perfect crystalline structure. In real semiconductors, one invariably has some defects that are introduced due to either thermodynamic considerations or the presence of impurities during the crystal growth process. In general, defects in crystalline semiconductors can be characterized as i) point defects; ii) line defects; iii) planar defects and iv) volume defects. These defects are detrimental to the performance of electronic and optoelectronic devices and are to be avoided as much as possible. We will give a brief overview of the important defects.

Point Defects

A point defect is a highly localized defect that affects the periodicity of the crystal only in one or a few unit cells. There are a variety of point defects, as shown in Fig. 1.18. Defects are present in any crystal and their concentration is given roughly by the thermodynamics relation

$$\frac{N_d}{N_{Tot}} = k_d \exp\left(-\frac{E_d}{k_B T}\right) \tag{1.12}$$

where N_d is the vacancy density, N_{Tot} the total site density in the crystal, E_d the defect formation energy, k_d is a dimensionless parameter with values ranging from 1 to 10 in semiconductors, and T, the crystal growth temperature. The vacancy formation energy is in the range of an eV for most semiconductors.

An important point defect in compound semiconductors such as GaAs is the anti-site defect in which one of the atoms, say Ga, sits on the arsenic sublattice instead of the Ga sublattice. Such defects (denoted by Ga_{As}) can be a source of reduced device performance.

Other point defects are interstitials in which an atom is sitting in a site that is in between the lattice points as shown in Fig. 1.18, and impurity atoms which involve a

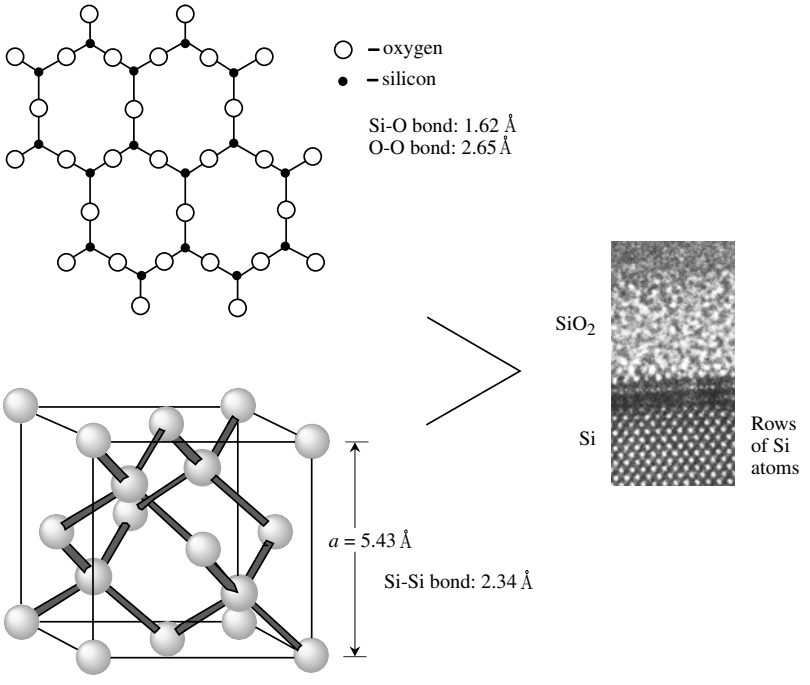


Figure 1.17: The tremendous success of Si technology is due to the Si/SiO_2 interface. In spite of the very different crystal structure of Si and SiO_2, the interface is extremely sharp, as shown in the TEM picture in this figure.

wrong chemical species in the lattice. In some cases the defect may involve several sites forming a defect complex.

Line Defects or Dislocations

In contrast to point defects, line defects (called dislocations) involve a large number of atomic sites that can be connected by a line. Dislocations are produced if, for example, an extra half plane of atoms are inserted (or taken out) of the crystal as shown in Fig. 1.19. Such dislocations are called edge dislocations. Dislocations can also be created if there is a slip in the crystal so that part of the crystal bonds are broken and reconnected with atoms after the slip.

Dislocations can be a serious problem, especially in the growth of strained heterostructures (to be discussed later). In optoelectronic devices, dislocations can ruin the device performance and render the device useless. Thus the control of dislocations is of great importance.

Planar Defects and Volume Defects

Planar defects and volume defects are not important in single crystalline materials, but can be of importance in polycrystalline materials. If, for example, silicon is grown on a glass substrate, it is likely that polycrystalline silicon will be produced. In the polycrystalline material, small regions of Si ($\sim$ a few microns in diameter) are perfectly

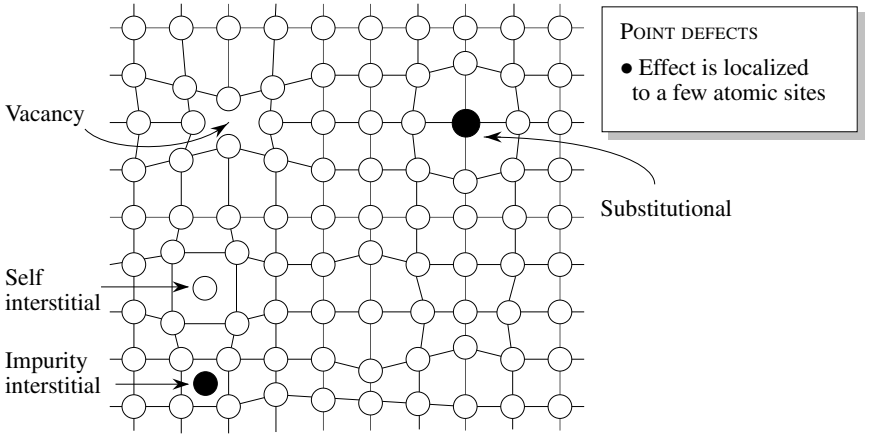


Figure 1.18: A schematic showing some important point defects in a crystal.

crystalline, but are next to microcrystallites with different orientations. The interface between these microcrystallites are called grain boundaries. Grain boundaries may be viewed as an array of dislocations.

Volume defects can be produced if the crystal growth process is poor. The crystal may contain regions that are amorphous or may contain voids. In most epitaxial techniques used in modern optoelectronics, these defects are not a problem. However, the developments of new material systems such as diamond (C) or SiC are hampered by such defects.

EXAMPLE 1.7 Consider an equilibrium growth of a semiconductor at a temperature of 1000 K. The vacancy formation energy is 2.0 eV. Calculate the vacancy density produced if the site density for the semiconductor is 2.5×10^{22} cm^{-3}. Assume that $k_d = 1$.

The vacancy density is

$$\begin{aligned} N_{vac} &= N_{Tot} \exp\left(-\frac{E_{vac}}{k_B T}\right) \\ &= (2.5 \times 10^{22}\ \text{cm}^{-3})\ \exp\left(-\frac{2.0\ \text{eV}}{0.0867\ \text{eV}}\right) \\ &= 2.37 \times 10^{12}\ \text{cm}^{-3} \end{aligned}$$

This is an extremely low density and will have little effect on the properties of the semiconductor. The defect density would be in mid 10^{15} cm^{-3} range if the growth temperature was 1500 K. At such values, the defects can significantly affect device performance.

1.4 STRAINED HETEROSTRUCTURES

In an epitaxial process, the overlayer that is grown on the substrate could have a lattice constant that may differ from that of the substrate. Such epitaxy is called strained epitaxy and is one of the important emerging areas of crystal growth studies. The motivation for strained epitaxy is two fold:

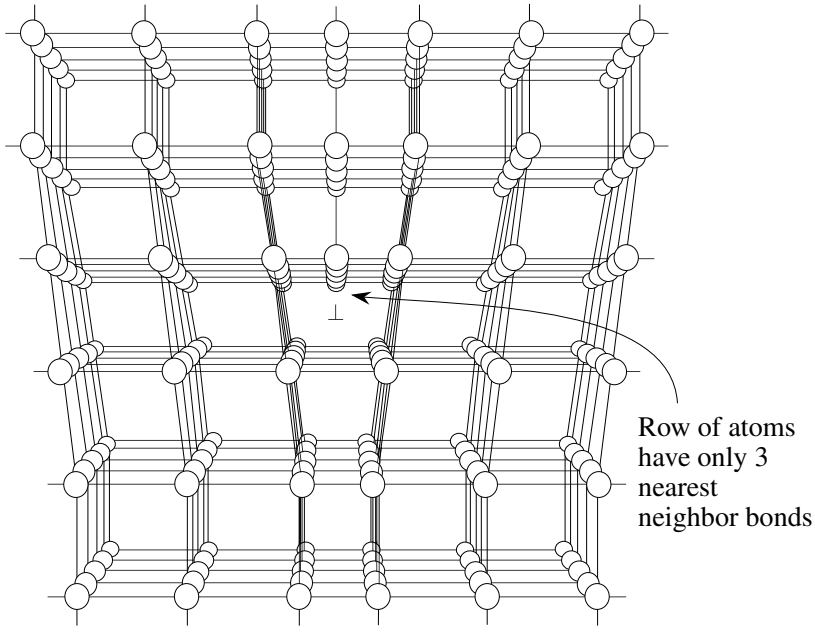


Figure 1.19: A schematic showing the presence of a dislocation. This line defect is produced by adding an extra half plane of atoms. At the edge of the extra plane, the atoms have a missing bond.

i) Incorporation of built-in strain: When a lattice mismatched semiconductor is grown on a substrate and the thickness of the overlayer is very thin (this will be discussed in detail later), the overlayer has a built-in strain. This built-in strain has important effects on the electronic and optoelectronic properties of the material and can be exploited for high performance devices.

ii) Generation of a new effective substrate: We have noted that in semiconductor technology, high quality substrates are only available for Si, GaAs and InP (sapphire and quartz substrates are also available and used for some applications). Most semiconductors are not lattice-matched to these substrates. How can one grow these semiconductors epitaxially? One solution that has emerged is to grow the overlayer on a mismatched substrate. If the conditions are right, a lot of dislocations are generated and eventually the overlayer forms its own substrate. This process allows a tremendous flexibility in semiconductor technology. Not only can it, in principle, resolve the substrate availability problem, it also allows the possibility of growing GaAs on Si, CdTe on GaAs, etc. Thus different semiconductor technologies can be integrated on the same wafer.

Coherent and Incoherent Structures

Consider a case where an overlayer with lattice constant a_L is grown on a substrate with lattice constant a_S. This situation is shown schematically in Fig. 1.20. The strain

between the two materials is defined as

$$\epsilon = \frac{a_S - a_L}{a_L} \tag{1.13}$$

Consider a conceptual exercise where we deposit a monolayer of the overlayer on the substrate. If the lattice constant of the overlayer is maintained to be a_L, it is easy to see that after every $1/\epsilon$ bonds between the overlayer and the substrate, either a bond is missing or an extra bond appears as shown in Fig. 1.20b. In fact, there would be a row of missing or extra bonds since we have a 2-dimensional plane. These defects are the dislocations. The presence of these dislocations costs energy to the system since a number of atoms do not have proper chemical bonding at the interface.

An alternative to the incoherent case is shown in Fig. 1.20c. Here all the atoms at the interface of the substrate and the overlayer are properly bonded by adjusting the in-plane lattice constant of the overlayer to that of the substrate. This causes the overlayer to be under strain and the system has a certain amount of strain energy. *This strain energy grows as the overlayer thickness increases.* In the strained epitaxy, the choice between the state of the structure shown in Fig. 1.20b and the state shown in Fig. 1.20c is decided by free energy minimization considerations. Theoretical and experimental studies have focussed on these considerations for over six decades, and the importance of these studies has grown since the advent of heteroepitaxy. The general observations can be summarized as follows:

For small lattice mismatch ($\epsilon < 0.1$), the overlayer initially grows in perfect registry with the substrate, as shown in Fig. 1.20c. However, as noted before, the strain energy will grow as the overlayer thickness increases. As a result, it will eventually be favorable for the overlayer to generate dislocations. In simplistic theories this occurs at an overlayer thickness called the critical thickness, d_c, which is approximately given by

$$d_c \cong \frac{a_S}{2|\epsilon|} \tag{1.14}$$

In reality, the point in growth where dislocations are generated is not so clear cut and depends upon growth conditions, surface conditions, dislocation kinetics, etc. However, one may use the criteria given by Eqn. 1.14 for loosely characterizing two regions of overlayer thickness for a given lattice mismatch. Below critical thickness, the overlayer grows without dislocations and the film is under strain. Under ideal conditions above critical thickness, the film has a dislocation array, and after the dislocation arrays are generated, the overlayer grows without strain with its free lattice constant.

While strained epitaxy below critical thickness is an extremely powerful tool for tailoring the optoelectronic properties of semiconductors, epitaxy beyond the critical thickness is important to provide new effective substrates for new material growth. For these applications the key issues center around ensuring that the dislocations generated stay near the overlayer-substrate interface and do not propagate into the overlayer as shown in Fig. 1.21. A great deal of work has been done to study this problem. Often thin superlattices in which the individual layers have alternate signs of strain are grown to "trap" or "bend" the dislocations. It is also useful to build the strain up gradually.

EXAMPLE 1.8 Estimate the critical thickness for $In_{0.3}Ga_{0.7}As$ grown on a GaAs substrate.

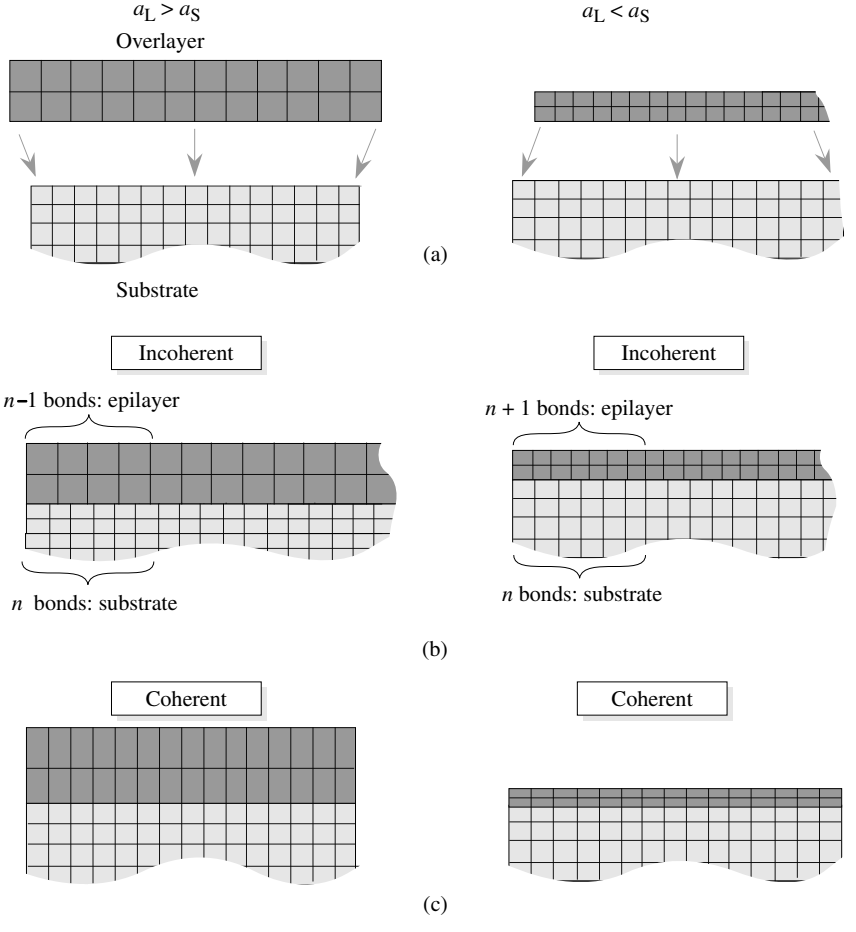


Figure 1.20: (a) The conceptual exercise in which an overlayer with one lattice constant is placed without distortion on a substrate with a different lattice constant. (b) Dislocations are generated at positions where the interface bonding is lost. (c) The case is shown where the overlayer is distorted so that no dislocation is generated.

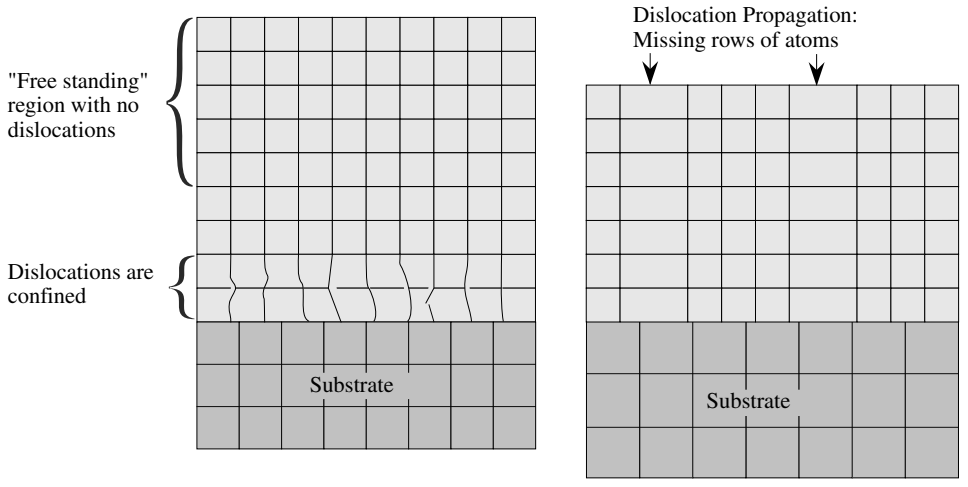


Figure 1.21: Strained epitaxy above critical thickness. On the left hand side is shown a structure in which the dislocations are confined near the overlayer-substrate interface. This is a desirable mode of epitaxy. On the right hand side, the dislocations are penetrating the overlayer, rendering it useless for most optoelectronic applications.

The lattice constant of an alloy is given by the Vegard's law:

$$\begin{aligned} a(In_{0.3}Ga_{0.7}As) &= 0.3a_{InAs} + 0.7a_{GaAs} \\ &= 5.775\ \mathring{A} \end{aligned}$$

The strain is

$$\epsilon = \frac{5.653 - 5.775}{5.653} = -0.022$$

The critical thickness is approximately

$$d_c = \frac{5.653\ \mathring{A}}{2(0.022)} = 128\ \mathring{A}$$

This thickness is quite adequate for most devices and can be used to make useful quantum well devices. If, on the other hand, the strain is, say, 5%, the critical thickness is ~ 50 Å, which is too thin for most useful device applications.

Self-Assembed Structures

When a lattice mismatched structure is grown on a substrate (which for most cases can be regarded as semi-infinite) a number of energetic and kinetic demands come into play. There is strain energy that is created in the system if the overlayer is under strain. This has to compete against the chemical bonding energy created by bond formation. Additionally, in real growth surface effects and the ability of the system to reach the free energy minimum state play an important role.

In Fig. 1.22 we show three kinds of growth mechanisms that occur when a strained overlayer is grown on a substrate under near equlibrium conditions. In Fig. 1.22

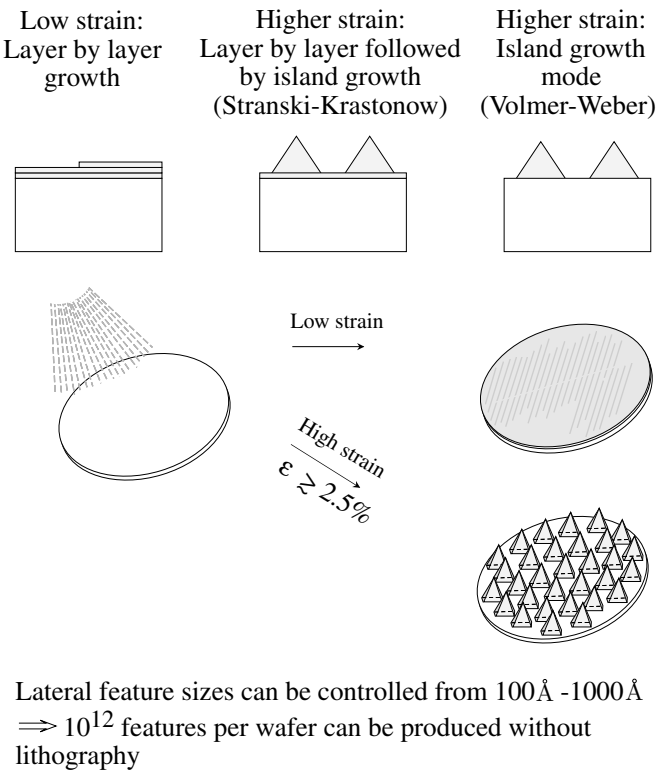


Figure 1.22: Growth modes in strained epitaxy. The island mode growth can be exploited to make "self-assembled" quantum dot structures.

we have a case where the lattice mismatch is very small ($\epsilon \leq 2\%$). The overlayer grows in the monolayer by monolayer mode since this allows the maximum chemical bonding to occur. If the lattice mismatch is increased, the growth occurs by a mode known as the Stranski-Krastanow mode where the initial growth starts out in the monolayer by monolayer growth, but then the overlayer grows in an island mode. The island growth provides fewer chemical per atom for the growing layer (since the surface area is larger), but the strain energy is minimized, since the bond lengths do not have to adjust as much to fit the substrate. Finally, at higher lattice mismatch the growth initiates directly in the island mode (the Volmer-Weber mode).

If heterostructures are to be grown with atomically abrupt interfaces between two semiconductors, one should be in the layer-by-layer growth mode described schematically in Fig. 1.22. There are, however, some advantages of growing in the island mode. By growing islands and then imbedding them by another material it is possible to grow quasi-zero dimensional systems in which electron (holes) are confined in all threee directions. Since such quantum dots are self-organized and regime no lithography/etching/ regrowth they are very attractive for many applications. Such self-organized quantum dots have been grown with InGaAs/GaAs, SiGe/Si, etc.

1.5 STRAIN TENSOR IN LATTICE MISMATCHED EPITAXY

In order to study the effect of strain on electronic properties of semiconductors, it is first essential to establish the strain tensor produced by epitaxy. In Appendix A we discuss important issues in strain and stress in materials. The reader who is unfamiliar with these issues should go over this appendix which is the basis for the results given in this section. As noted above, careful growth of an epitaxial layer whose lattice constant is close, but not equal, to the lattice constant of the substrate can result in a coherent strain. If the strain is small one can have layer-by-layer growth as shown in Fig. 1.22. In this case the lattice constant of the epitaxial layer in the directions parallel to the interface is forced to be equal to the lattice constant of the substrate. The lattice constant of the epitaxial perpendicular to the substrate will be changed by the Poisson effect. If the parallel lattice constant is forced to shrink, or a compressive strain is applied, the perpendicular lattice constant will grow. Conversely, if the parallel lattice constant of the epitaxial layer is forced to expand under tensile strain, the perpendicular lattice constant will shrink. These two cases are depicted in Fig. 1.20c. This type of coherently strained crystal is called pseudomorphic.

For layer-by-layer growth, the epitaxial semiconductor layer is biaxially strained in the plane of the substrate, by an amount $\epsilon_{\parallel}$, and uniaxially strained in the perpendicular direction, by an amount $\epsilon_{\perp}$. For a thick substrate, the in-plane strain of the layer is determined from the bulk lattice constants of the substrate material, a_S, and the layer material, a_L:

$$\begin{aligned} e_{\parallel} &= \frac{a_S}{a_L} - 1 \\ &= \epsilon \end{aligned} \tag{1.15}$$

Since the layer is subjected to no stress in the perpendicular direction, the perpendicular strain, $\epsilon_{\perp}$, is simply proportional to $\epsilon_{\parallel}$:

$$\epsilon_{\perp} = \frac{-\epsilon_{\parallel}}{\sigma} \tag{1.16}$$

where the constant σ is known as Poisson's ratio.

Noting that there is *no stress* in the direction of growth it can be simply shown that for the strained layer grown on a (001) substrate (for an fcc lattice)

$$\begin{aligned} \sigma &= \frac{c_{11}}{2c_{12}} \\ \epsilon_{xx} &= \epsilon_{\parallel} \\ \epsilon_{yy} &= \epsilon_{xx} \\ \epsilon_{zz} &= \frac{-2c_{12}}{c_{11}}\epsilon_{\parallel} \\ \epsilon_{xy} &= 0 \\ \epsilon_{yz} &= 0 \\ \epsilon_{zx} &= 0 \end{aligned} \tag{1.17}$$