Chapter 4
Band Structure in Silicon and Compound Semiconductors

In this chapter we will overview Si and Compound Semiconductor (CS) band structures and become familiar with the concept of semiconductor alloys. We will study bandgap engineering as well as strain engineering in lifting band degeneracy in various hetero-structures.

Compound Semiconductors

Nature has provided us with many semiconductor materials. Unfortunately, the quality of these semiconductors is often not good for making electronic devices and circuits. Over the past 50+ years, we have learned to synthesis semiconductors through various processes. For instance, to make Si wafers, we start by melting clean sand! and forming a molten slurry of Si in a crucible. A small crystalline seed of Si is then immersed in the crucible and withdrawn slowly as it is rotated. The result is a pure Si cylinder called ingot. Once Si ingot is cooled it is sawed into thin wafers of Si which then can be used for wafer processing. Currently 12" Si wafers are available for processing. There are other techniques to make Si and compound semiconductor ingots. Today, 8" GaAs wafers and 6" InP wafers are readily available.

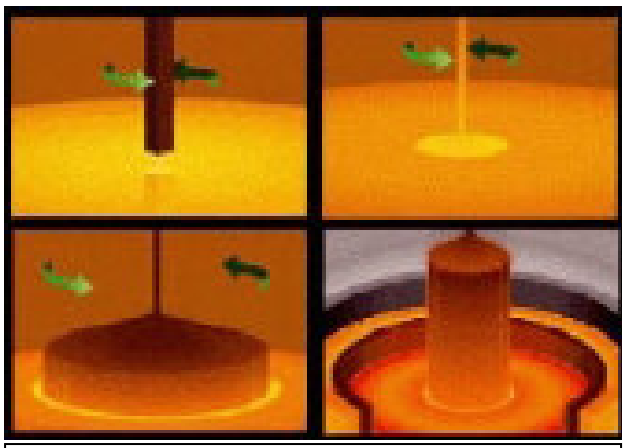

Fig. 4.1: Czochralski method to make Si ingot

A number of compound semiconductor materials are formed by making an alloy using an epitaxial growth technique. Epitaxial films of such materials are formed on top of a compound semiconductor substrate. There are two widely-used techniques to do epitaxy, namely molecular beam epitaxy (MBE) and metalorganic chemical vapor deposition (MOCVD). While the two techniques are different by nature, the resulting films qualities are often comparable. Certain alloys are grown more easily using one technique compared to the other. For instance, an alloy of

InAlAs can be easily grown by MBE but is hard to grow with MOCVD, while InP can be easily grown by MOCVD but is harder to grow using MBE.

Semiconductor substrates that are often used for epitaxy are Si GaAs, InP, SiC and GaP. Si substrate is used to grow epitaxial Si as well as epitaxial SiGe material. GaAs substrate is often used for growing AlGaAs alloy as well as InGaP alloy. InP substrate is used for growing InGaAs alloy and InAlAs alloy. SiC is used for SiC related materials as well as GaN related materials. Epitaxy is done using one of the following procedures:

1. Lattice Matched
2. Pseudomorphic
3. Metamorphic

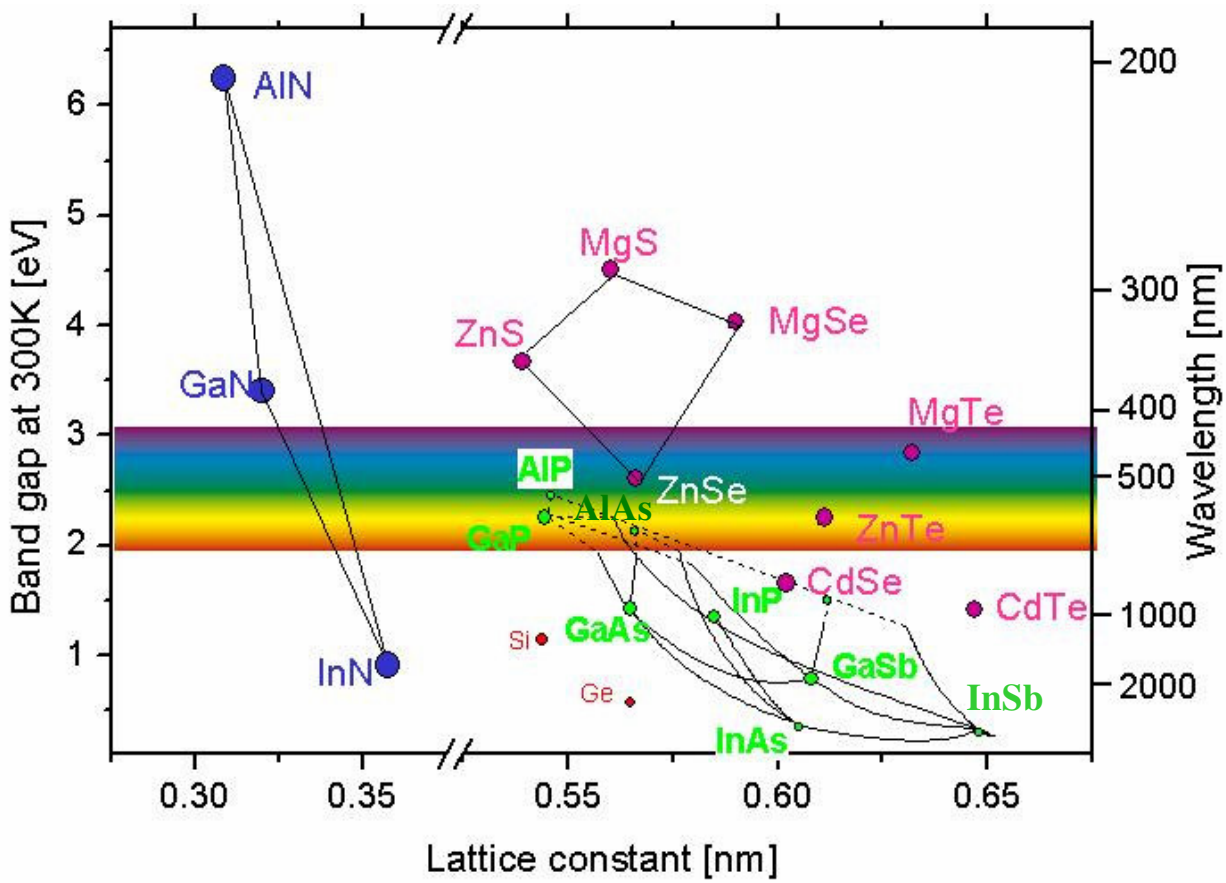

Fig. 4.2. Relationship between bandgap, absorbed wavelength and Lattice Constant of several CS

In lattice-matched epitaxy, the epitaxial layer has the same lattice constant as the substrate. An example is AlGaAs alloy grown on GaAs substrate. An example is AlGaAs alloy on GaAs substrate. For Al mole fractions of smaller than 85% AlGaAs is a direct material and has almost the same lattice constant as GaAs. AlGaAs bandgap is higher than GaAs. Another example of lattice matched epitaxy is $In_{0.48}Ga_{0.52}As$ alloy grown on InP substrate. The following figure (next page) shows the bandgap as a function of lattice constant for some III-V and II-VI compound semiconductors with a better resolution.

In pseudomorphic epitaxy an alloy on a semiconductor substrate with slightly different lattice constant is grown. The difference in lattice constant induces strain in the alloy film. The thickness of the alloy film and the mismatch between the two lattice constants sets the maximum

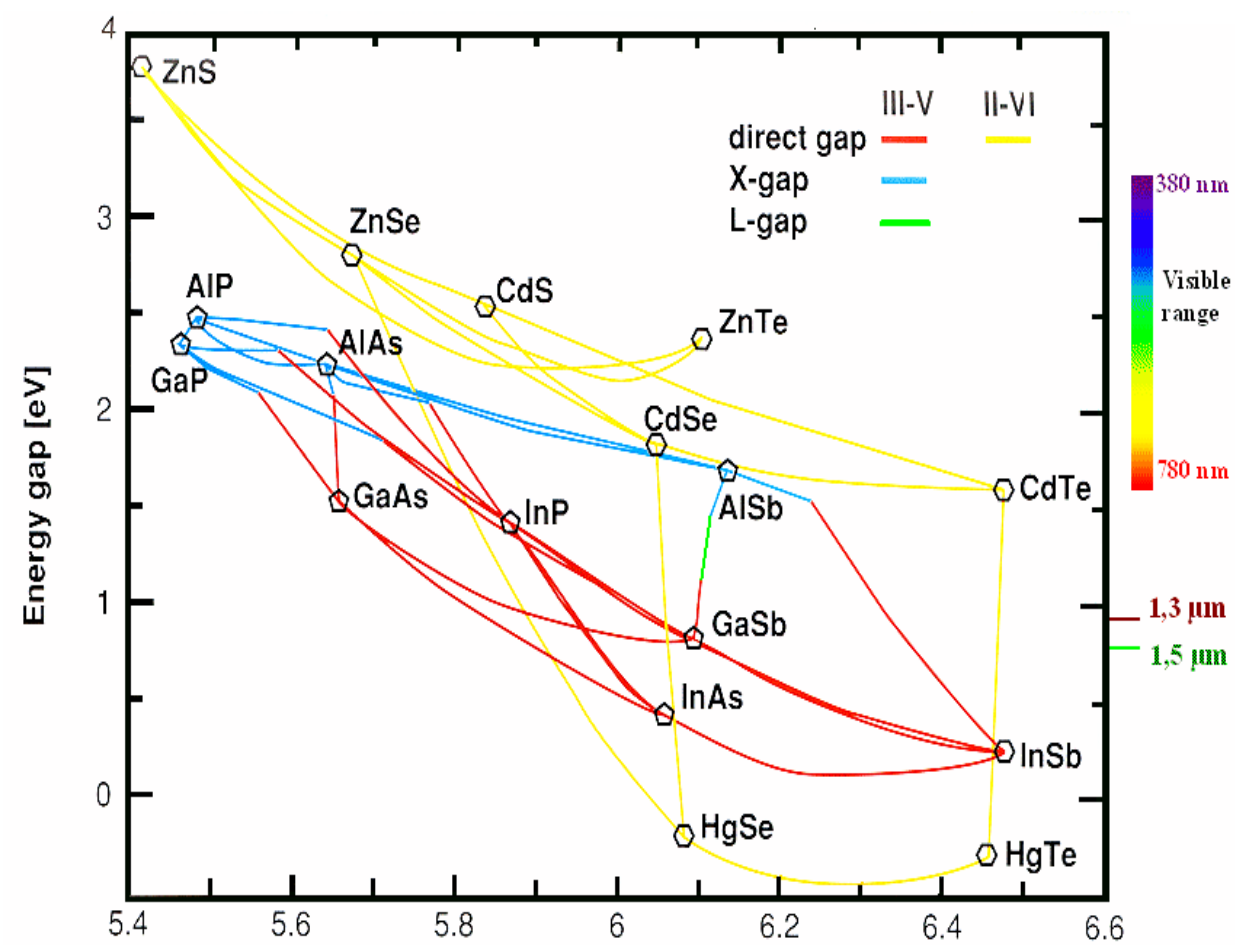

Fig. 4.3. Relationship between bandgap and Lattice Constant of several CS alloys

film thickness that can be grown without producing dangling bonds (so called critical thickness). Fig. 4.4 schematically shows how pseudomorphic epitaxy results in compressive strain in one direction and tensile strain in the perpendicular direction.

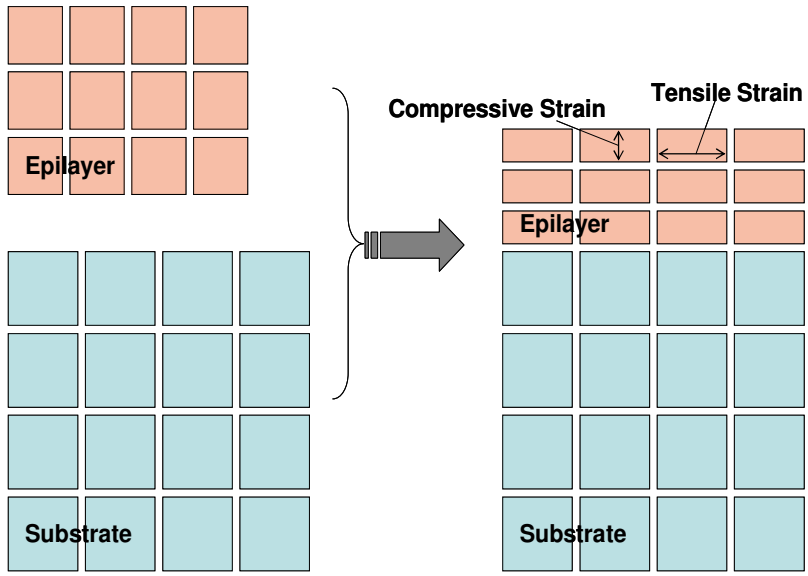

Fig. 4.4. Heteroepitaxy of pseudomorphic semiconductor materials and resulting compressive and tensile strains.

The third epitaxial method is called metamorphic growth. Metamorphic technique was developed to address the need to have high performance transistors that are often used on InP substrate as well as optoelectronic heterostructures based on InGaAs and InAlAs alloys on a more robust substrate such as GaAs. Advantages of GaAs substrate over InP substrate are lower cost, larger size substrate, easier to handle (not as brittle) and more advanced processing technology. In metamorphic growth a buffer layer with a lot of defects or a super-lattice is used to change the properties of GaAs to one of InP-based alloys. Obviously, in this technique there is no built-up strain in the device, so no boost in mobility due to stress occurs. The technology allows one to use InP-

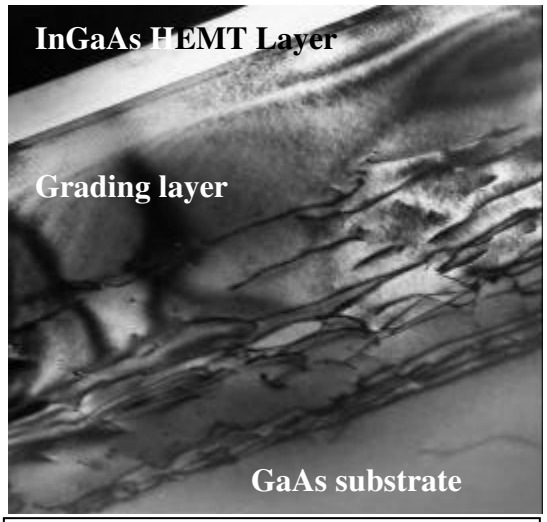

Fig. 4.5. Metamorphic Heteroepitaxy of InGaAs HEMT on GaAs substrate. A high defect grading layer is used.

based material on GaAs substrate as shown in Fig. 4.5.

Compound Semiconductor Bandstructures

There are several computational techniques that can be used to calculate the band structure of a certain semiconductor. The most universal technique is to calculate the central part of the band structure based on tight-binding approximation (applying periodicity through Bloch theorem to single atom model). For calculating the band structure near the bandedges, it is better to use a perturbation theory away from symmetry points. This technique is called Kronig-Penney method. The resulting band structure shows the calculated conduction bands and valence bands as a function of wavevector (or momentum). The following observations are made for different materials that are studied.

Si

Si has an fcc lattice (face center cubic) and is an indirect bandgap material. That means the bottom of the conduction band does not align with the top of the valence band (see E-K diagram of Fig. 4.6). Therefore, to go from valence band to conduction band, electrons need to receive energy (as in the form of photons) and momentum (phonons). The bottom of the conduction band in Si is close to the X-point of the brillouin zone ($\frac{2\pi}{a}(0.85,0,0)$). Due to symmetry of fcc lattice, there are six degenerate X-points and consequently sic conduction edge valleys. The near bandedge band structure can be represented by simple E-K function of ellipsoid nature as follows.

$$E(k) = \frac{\hbar^2 k_x^{\,2}}{2m_l^*} + \frac{\hbar^2 (k_y^{\,2} + k_z^{\,2})}{2m_t^*} \tag{4.1}$$

In the above equation, there are two different masses of the electrons. The longitudinal mass m_l^* is approximately $0.98m_0$ (m_0 being mass of a free electron), while the transverse mass m_t^* is approximately $0.19m_0$. Note that for (100) direction, the longitudinal direction is x-axis while the transverse is y-z axis. But for (010) direction, the longitudinal direction is y-axis while the transverse is x-z axis. Although there is strong anisotropy in each valley, there is not much of an anisotropy in properties such as transport of electrons which involve the moving of electrons in all of the valleys. The next valley in the conduction band is L-point valley, which is about 0.5eV

above the bandedge. Above this is the Γ-point edge which gives a direct bandgap for Si of ~2.2eV. This direct bandgap is quite important for optical properties of Si as the absorption coefficient for photons above this energy is very strong. Due to 6-fold degeneracy of the conduction bandedge, the electron transport in Si is quite poor. This is because of very large density of states near the bandedge leading to a high scattering rate.

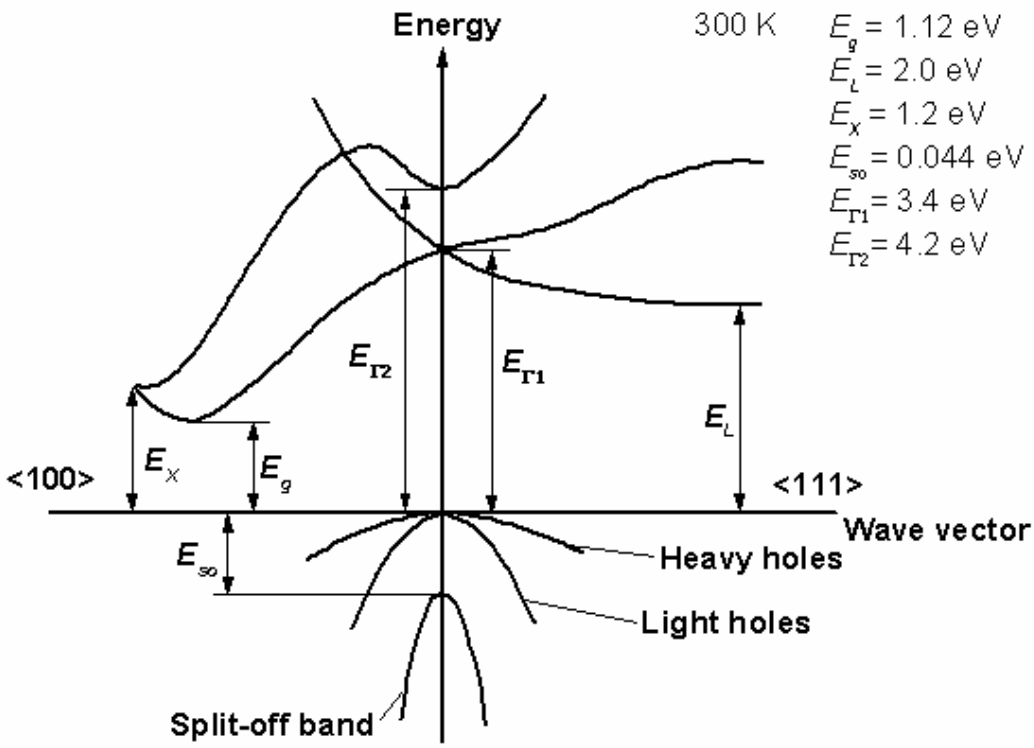

Fig. 4.6. Si band structure.

The top of valence band has the typical features seen in all semiconductor valence bands. One has the light hole (LH) and heavy hole (HH) degeneracy and the split-off band (SO) arising from spin-orbit interaction. In the case of Si, SO energy is only 44meV below the top of valence band. The large hole mass (both LH and HH) and closeness of SO band results in very poor hole transport properties in Si. Hole masses are $m_{HH}^* = 0.49m_0$ and $m_{LH}^* = 0.16m_0$.

As you can see both electron and hole transport properties in Si are very poor. Despite this deficiency, Si is one of the most utilized semiconductor for electronic devices (almost %98 of the market belongs to Si). There are two reasons: One that Si has natural oxide SiO_2 that can be grown on Si with excellent quality, as opposed to other semiconductors. The second reason is that most high speed electronic devices work under high electric field (E_{field}>20kV/cm). At such high fields the response of electrons to field characterized by electron saturation velocity in almost all semiconductors (with a few exceptions) is nearly the same.

Ge

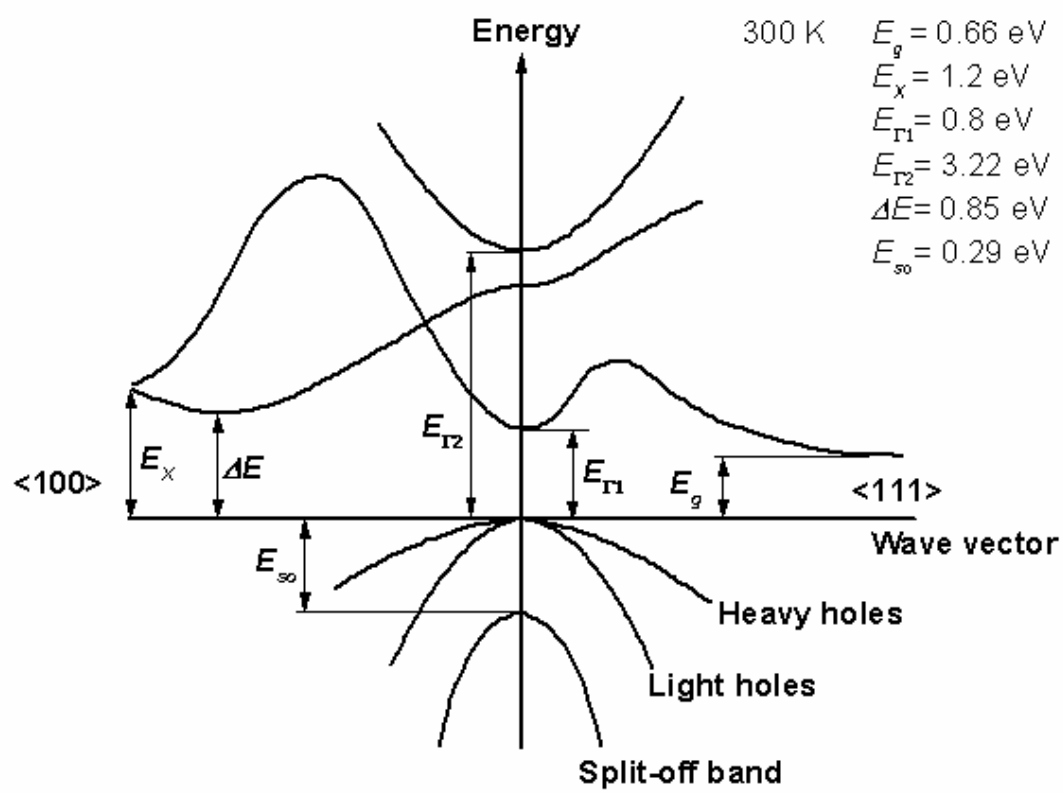

Fig. 4.7. Ge band structure.

Ge is an indirect bandgap material with its conduction band minima at L-valley with a low bandgap of 0.66eV (see Fig. 4.7). Ge was the first semiconductor to be used for making a transistor but was soon bypassed by overal better qualities of Si. The longitudinal mass m_l^* for

Ge is approximately $0.159m_0$, while the transverse mass m_t^* is approximately $0.08m_0$. The top of valence band has the typical features seen in all semiconductor valence bands. The SO energy for Ge is 0.29eV below the top of valence band. Hole masses are much lower than Si with $m_{HH}^* = 0.33m_0$ and $m_{LH}^* = 0.043m_0$.

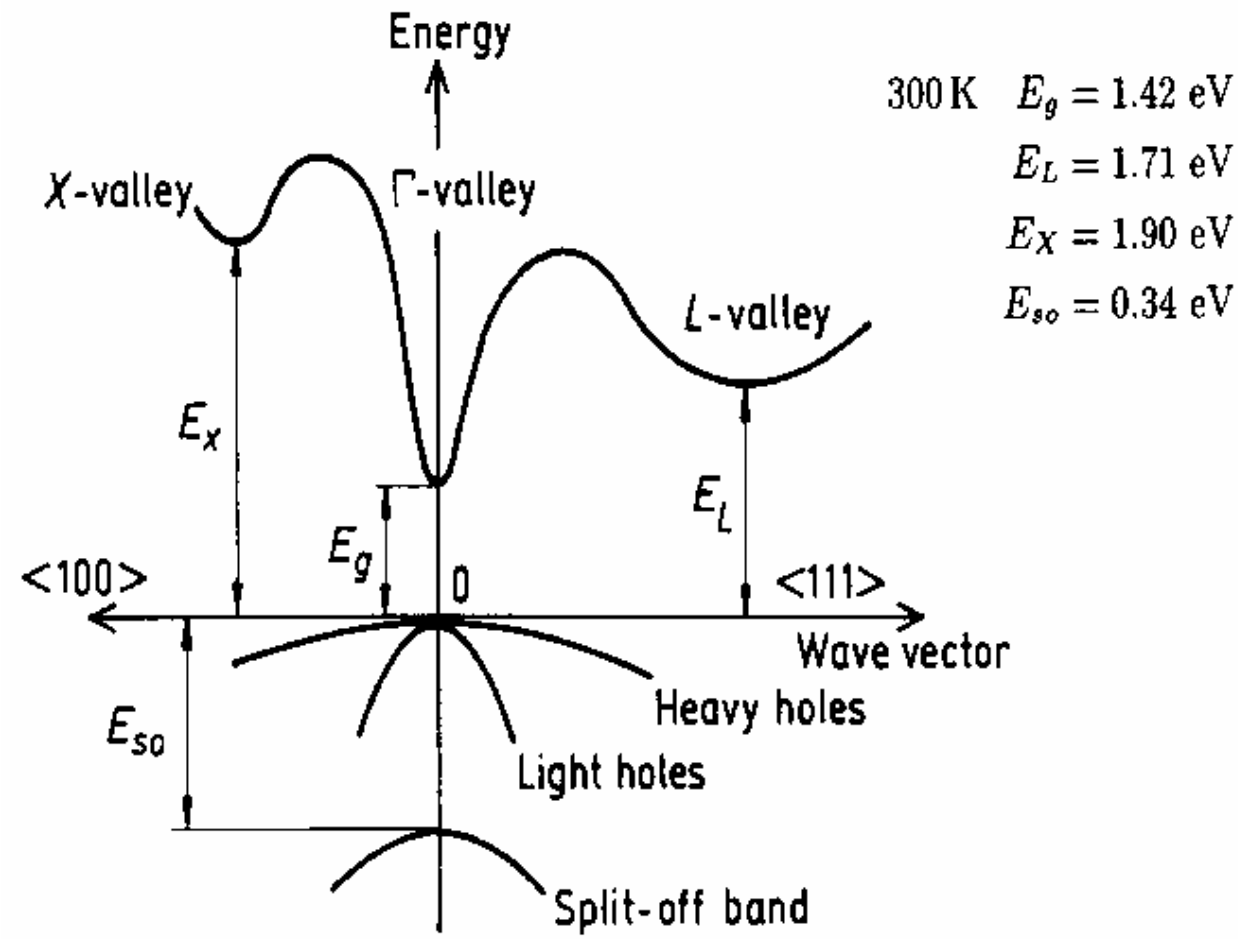

Fig. 4.8. GaAs band structure.

GaAs

Fig. 4.8 shows the E-k diagram of GaAs which is a direct bandgap (Γ-valley has the minimum energy) material with fcc lattice. GaAs has superior band structure when compared to Si. However it suffers from several drawbacks related to its material properties. GaAs does not have high quality oxide that can be used as gate dielectric or passivation layer. It is much more difficult to process GaAs which is an inherent problem of compound semiconductors. The basis

consists of two atoms which makes it easier to introduce defects (e.g. a Ga atom on the As site ~ this is called antiside defect).

The advantages of GaAs over Si make it a viable choice for many high speed electronics, high power and optoelectronics applications. The conduction bandstructure can be represented by a non parabolic relationship:

$$E(k)\left(1+\alpha E(k)\right)=\frac{\hbar^2 k^2}{2m^*} \tag{4.2}$$

where m^* is approximately $0.067m_0$ and α is 0.67eV^{-1}. For high electric field transport it is important to note that the valleys above the Γ-valley are L-valleys. There are four L-valleys with the separation of 0.29eV. GaAs effective mass in L-valley is much higher than Γ-valley ($m_L^* = 0.25m_0$). This difference leads to a negative differential resistance as well as low saturation velocity of electrons under high electric fields. X-valleys (3 of them) have about 0.48eV higher energy with a large mass of $m_X^* = 0.6m_0$.

The valence band of GaAs is again similar to all semiconductors with LH, HH and SO bands. Spin-orbit splitting (SO) energy is large and about 0.34eV. Hole masses are $m_{HH}^* = 0.45m_0$ and $m_{LH}^* = 0.08m_0$.

Larger bandgap of GaAs ensures better high temperature properties, which translates into better high power properties. Despite this advantage, GaAs has three times higher thermal resistance compared to Si, which means it is more difficult to cool GaAs devices through the substrate.

<u>BandStructure of Semiconductor Alloys:</u>

Here we consider the bandstructure of ternary alloys (for instance AlGaAs). With the advent of advanced epitaxial techniques it is also possible to grow quarternary alloys (such as GaInAsP) which will give you more degree of freedom in terms of the bandgap, electron and hole effective masses (mobility) and lattice constant (lattice matched or pseudomorphic). The discussion of

quarternary alloys is beyond the scope of this chapter, but can be further examined through a research project in this course.

In compound semiconductors, the concept of randomly mixing two or more semiconductors is motivated by two main objectives:

1. alter the bandgap of the material to a pre-selected value.
2. to create a material with a proper lattice constant. For instance $In_{0.53}Ga_{0.47}As$ alloy is lattice matched to InP substrate

For the same lattice structure materials, the lattice constant follows Vegard's law:

$$\alpha_{alloy} = x\alpha_A + (1-x)\alpha_B \tag{4.3}$$

where x is the mole fraction for alloy atom A and (1-x) is the mole fraction for alloy atom B. An alloy is called <u>phase separated</u> if in the formation of $A_xB_{1-x}C$ alloy, A atoms are localized in one region while the B atoms are localized in another region. If the separation of A atoms and B atoms follows a well order periodic structure in the direction of the epitaxy, such alloy is called <u>superlattice</u>. Random positioning of A and B atoms yield <u>random alloys</u>.

The band structure calculations for alloys are very complex and Vegard's law is just an approximation. In most alloys, there is a bowing effect arising from the increasing disorder due to alloying. The bandgap of an alloy due to this bowing phenomenon follows

$$E_{g-alloy} = a + bx + cx^2 \tag{4.4}$$

where c is the bowing parameter. Bandgap of some of the important III-V alloys are given in the following table.

Alloy	Direct Energy Gap (E_g)at 300K (eV)
$Al_xGa_{1-x}As$ (on GaAs)	1.424+1.247x
$Al_xIn_{1-x}As$ (on InP)	$0.36+2.01x+0.70x^2$
$Ga_xIn_{1-x}P$ (on GaAs)	$1.35+0.64x+0.77x^2$
$Ga_xIn_{1-x}As$ (on InP)	0.36+1.06x
Ge_xSi_{1-x} (on Si)	$1.12-0.41x + 0.008x^2$ (x<0.85) $1.86 - 1.2x$ (x>0.85)

Since

$$E_{alloy}(k) = \frac{\hbar^2 k^2}{2m^*_{alloy}} = \frac{x\hbar^2 k^2}{2m^*_A} + \frac{(1-x)\hbar^2 k^2}{2m^*_B} \tag{4.5}$$

we conclude that both electron and hole masses of an alloy can be found according to

$$\frac{1}{m^*_{alloy}} = \frac{x}{m^*_A} + \frac{(1-x)}{m^*_B} \tag{4.6}$$

In the following, we will study some important alloys that often used in conjunction with Si and compound semiconductors.

AlGaAs

AlGaAs due to its excellent match to GaAs substrates used to be the alloy of choice for many high speed transistors (both HEMTs and HBTs), optoelectronic devices and even some exotic quantum well structures. Its importance has degraded due to the advent of InGaP.

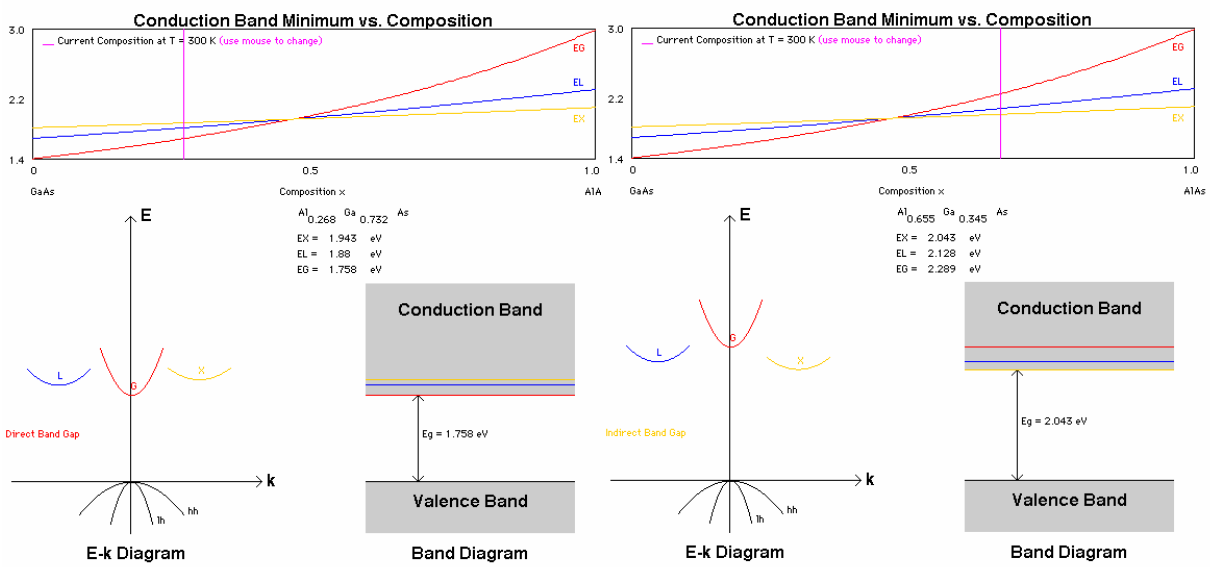

Fig. 4.9. Bandgap of AlGaAs vs. Al mole fraction. Both direct and indirect AlGaAs are shown.

Fig. 4.9 shows the conduction band minimum of different valleys (Γ, L and X) as a function of Al mole fraction. For Al mole fractions above %45, AlGaAs becomes an indirect material. For lower than 45% Al mole fraction the bandgap of AlGaAs increases linearly with no bowing effect. For Al mole fractions close to 35%, the alloy develops problems related to deep levels

known as D-X centers. These deep levels seem to be connected to the merging of Γ, X and L valleys. AlGaAs with Al mole fraction of higher than 45% is seldom used as it has poor electron transport properties.

InGaAs

InGaAs has a bandgap range of 0.39eV to 1.5eV with an overall lattice mismatch of %7. It is always a direct bandgap material. Its bandgap follows Vegard's law with minimum bowing. Its properties fall in the range of very high speed transistors as well as optical communication lasers. $In_{0.53}Ga_{0.47}As$ has an energy gap of 0.8eV and $m_e^* = 0.04m_0$ and large Γ-L and Γ-X separations. $In_{0.53}Ga_{0.47}As$ is lattice matched to InP with an E_g of 1.35eV and is also lattice matched to $In_{0.52}Al_{0.48}As$ with an energy gap of 1.45eV. There has been a growing interest in the epitaxy of InGaAs with In mole fractions of higher than 0.53 to achieve strained structures as will be discussed later.

SiGe

Si and Ge have a lattice mismatch of about %4. Therefore SiGe alloys are always strained with regards to the substrate in used (whether it is Si substrate or Ge substrate). The conduction band minima of Si is at X-point while that of Ge is at the L-point. The alloy retains an X-point character for up to %85 Ge mole fraction and then changes its character to L-like as shown in Fig. 4.10. Strained SiGe/Si system which is mostly used in electronic devices has different properties than relaxed SiGe. We will discuss the effect of strain on SiGe in a later section.

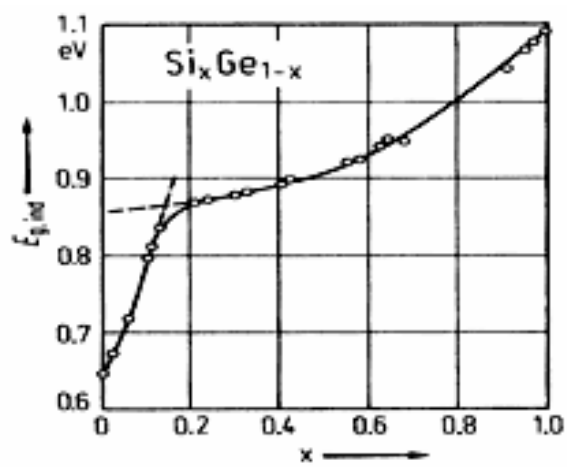

Fig. 4.10. Bandgap of SiGe alloy and transition from X-valley to L-valley.

Bandstructure Modification

When two dissimilar materials with different band gap are grown on top of each other a heterostructure is made. Band structure can be modified through using such heterostructures.

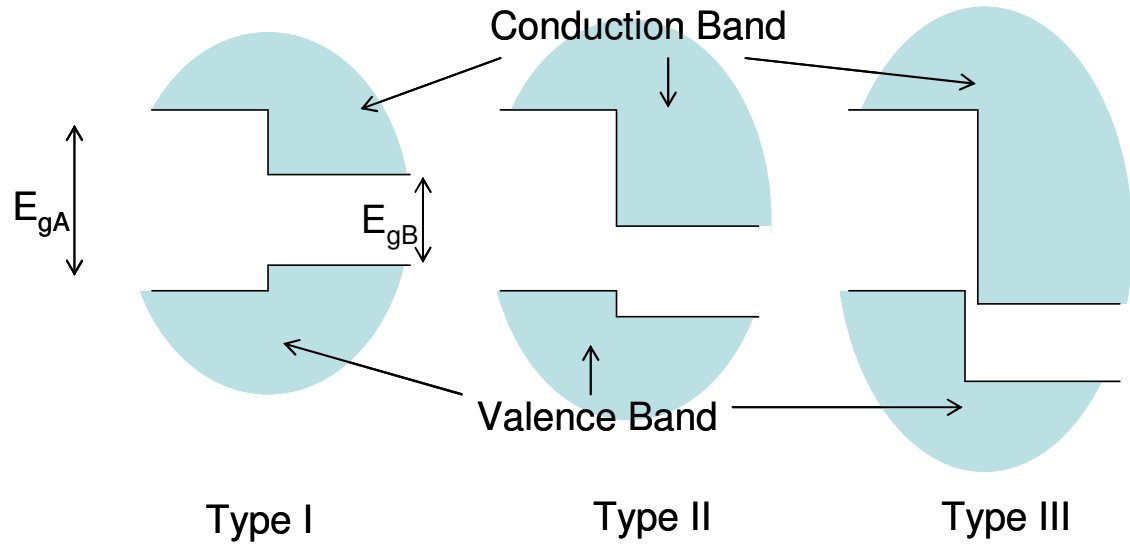

Fig. 4.11. Possible bandedge lineups in heterostructures

Generally in electronic and optoelectronic devices, heterostructures are used for band tailoring, charge confinement and built-in stress. If the interface roughness of a hetereostructure is more than a couple of monolayers, many of the advantages of heterostructures are lost. There are three possible bandedge lineups in heterostructures as shown in Fig. 4.11. The type of the lineup of heterostructure cannot be easily determined. If the electron affinity rule applies, the conduction band offset at the heterostructure is equal to the difference of electron affinities between the two semiconductors. Therefore the semiconductor with smaller electron affinity has the higher conduction band at the interface. Unfortunately, electron affinity rule does not work very well due to the fact that there is always a charge transfer across the interface due to the dissimilar nature of the chemical bonds.

It is also possible to manipulate the band structure through applying strain. This is done by pseudomorphical epitaxy of an alloy on the substrate. As long as the thickness of the epitaxial alloy is less than a certain critical thickness, the epitaxial film adjusts its lattice constant to that of the substrate (and therefore becomes strained) and maintains good interface quality. Critical thickness is calculated from the following equation

$$h_c = \frac{b}{9.9\frac{\Delta a}{a}} \tag{4.7}$$

where h_c is the critical thickness, b is a material parameter related the semiconductor substrate (e.g. b=0.38nm for Si) and $\Delta a/a$ is the mismatch between the substrate and the film. For a 1% mismatch between Si and $Si_{0.75}Ge_{0.25}$, the critical thickness is about 4nm. For SiGe device applications, typically lower Ge mole fractions are used (typically between 5 to 15%), therefore extending the critical thickness to 7 to 20nm.

Pseudomorphic alloys with internal strain often take advantage of lifting of the degeneracy in both conduction band and valence band. Provided that the right strain direction is applied (compressive or tensile), degeneracy of both conduction sub-valleys and valence sub-bands can occur leading to a decrease in the overall effective mass in the direction of the current flow in the device. Thus speed and transconductance of the device can be improved through strain manipulation. Unfortunately, to fully investigate the effect of strain on bandgap, effective mass and other parameters of the epitaxial semiconductor requires the knowledge of tensor-based equations. This is beyond the objectives of this course. Interested students are encouraged to study articles on the physics of heterostructures.