Band diagrams of heterostructures

Band diagram lineups

In a semiconductor heterostructure, two different semiconductors are brought into physical contact. How is this done in practice?

The properties of such heterostructures are of critical importance for many heterostructure devices including field-effect transistors, bipolar transistors, light-emitting diodes and lasers.

Three different alignments of the conduction and valence bands:

Straddled alignment (e. g. GaAs / $Al_xGa_{1-x}As$ heterostructure)

Staggered lineup (e. g. $Ga_xIn_{1-x}As$ / $GaAs_ySb_{1-y}$)

Broken-gap alignment (e.g. InAs / GaSb material system)

© E. F. Schubert, Rensselaer Polytechnic Institute, 2003

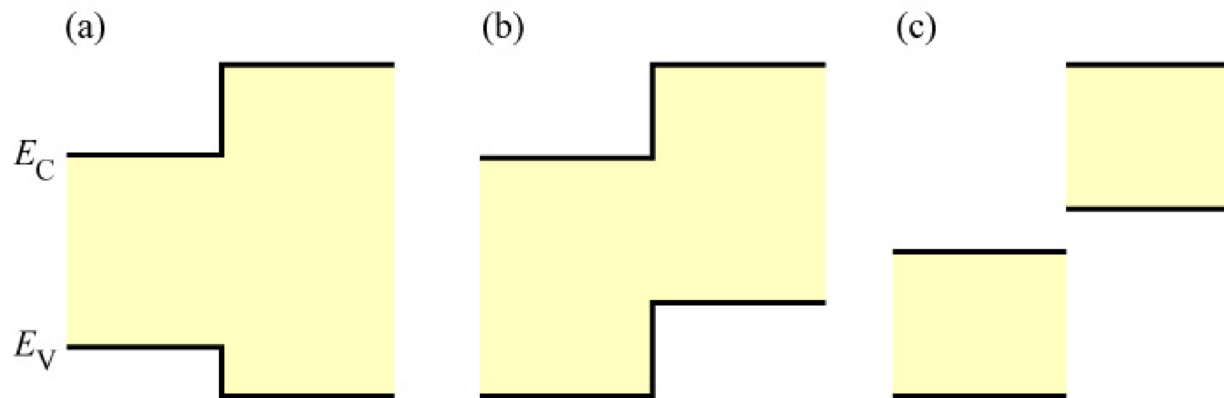


Fig. 17.1. Types of energy band lineups: (a) straddled or "Type I" lineup, (b) staggered or "Type II" lineup, and (c) broken or "Type III" lineup.

There have been numerous attempts and models to predict and calculate the energy ***band offsets*** in semiconductor heterostructures.

© E. F. Schubert, Rensselaer Polytechnic Institute, 2003

Linear superposition of atomic-like potentials

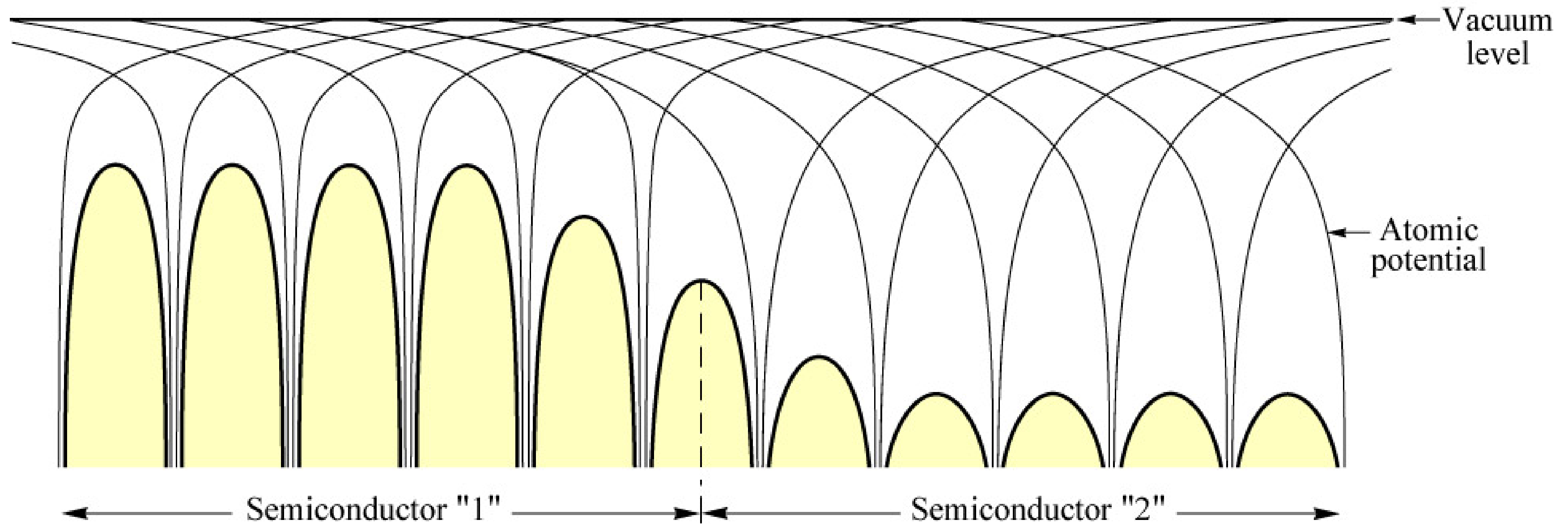


Fig. 17.2. Atomic potentials in the vicinity of two semiconductors "1" and "2". Within each semiconductor, all atomic potentials are identical. The resulting crystal potential is obtained by the superposition of all atomic potentials.

We first discuss the model of the linear superposition of atomic-like potentials (Kroemer 1975, 1985). He pointed out that the problem of theoretically understanding the relative alignment of bands is the problem of determining the relative alignment of the two periodic potentials of the

© E. F. Schubert, Rensselaer Polytechnic Institute, 2003

two semiconductors forming the heterostructure. Once the periodic potential of a semiconductor or of a heterostructure is known, the energy bands can be calculated.

The periodic potential of a semiconductor can be viewed as a linear superposition of the overlapping atomic-like potentials.

Although the model of the superposition of atomic-like potentials is very instructive, the ability of this model to *predict* offsets between semiconductors is very limited (Kroemer, 1985).

© E. F. Schubert, Rensselaer Polytechnic Institute, 2003

The electron affinity model

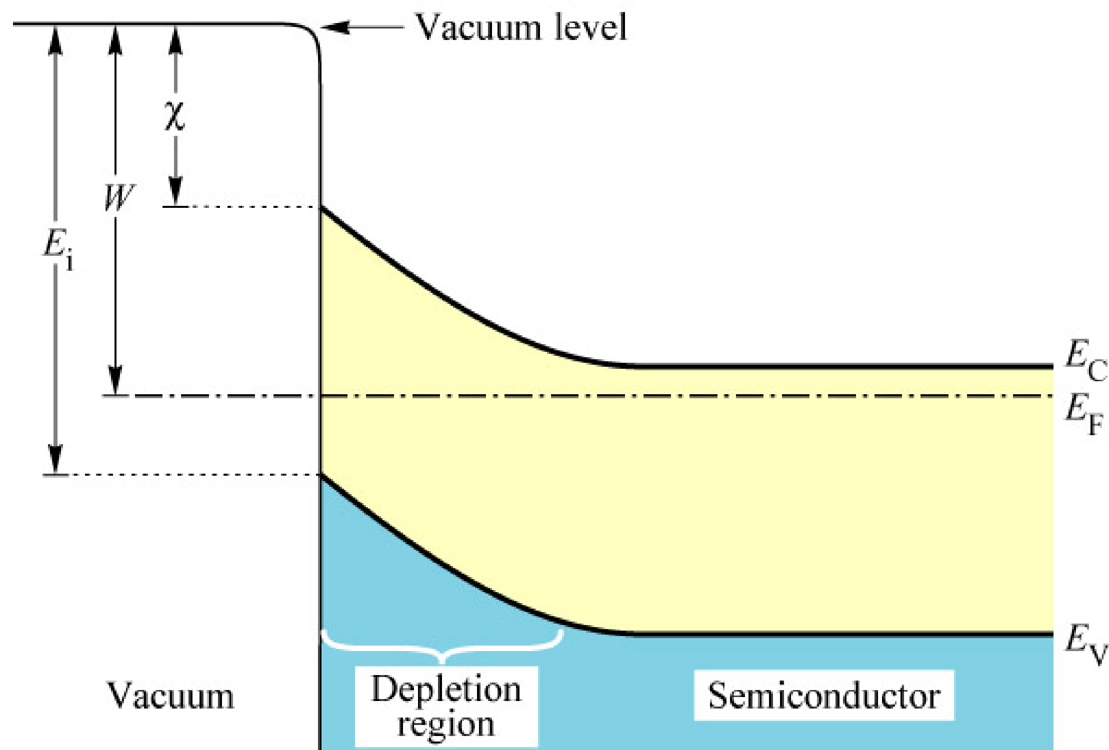


Fig. 17.3. Electron affinity χ, work function W, and ionization energy E_i of a semiconductor. The electron affinity is measured from the bottom of the conduction band at the semiconductor surface, the work function from the Fermi level, and the ionization energy from the top of the valence band at the surface.

The electron affinity model is the oldest model invoked to calculate the band offsets in semiconductor heterostructures (Anderson, 1962). This model has proven to give accurate predictions for the band offsets in

several semiconductor heterostructures, whereas the model fails for others.

Electron affinity χ

(i. e. promoting an electron from the bottom of the conduction band to the vacuum beyond the reach of image forces)

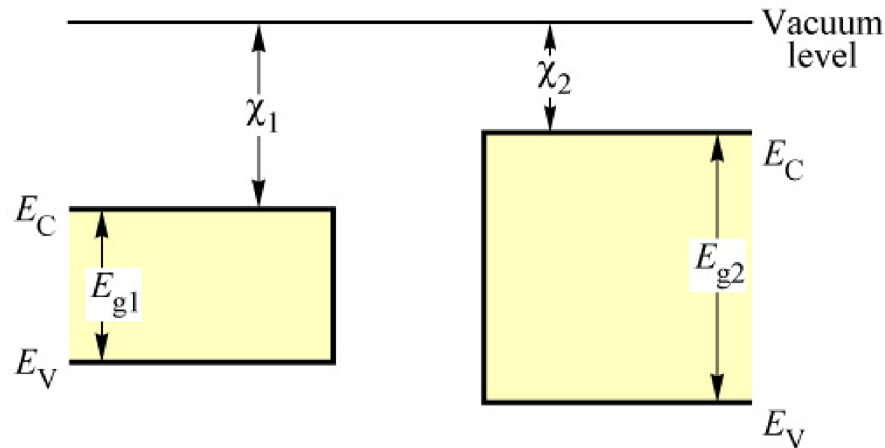
Work function W

(i. e. promoting an electron from the Fermi level to the vacuum beyond the reach of image forces)

Ionization energy E_i

The ionization energy is measured by photo-ionization experiments, in which semiconductors are illuminated by monochromatic light with a variable wavelength. The longest wavelength at which photo-ionization occurs defines the ionization energy.

© E. F. Schubert, Rensselaer Polytechnic Institute, 2003



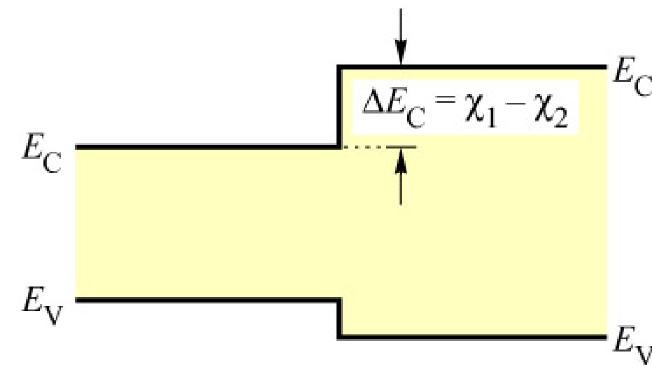


Fig. 17.4. Band diagrams of (a) two separated semiconductors and (b) two semiconductors in contact. The semiconductors have a band gap energy of E_{g1} and E_{g2} and an electron affinity of χ_1 and χ_2.

Next consider that two semiconductors are brought into physical contact.

The electron affinity model is based on the fact that the energy balance of an electron moved from the vacuum level to semiconductor "1", from there to semiconductor "2", and from there again to the vacuum level must be zero, that is $\chi_1 - \Delta E_c - \chi_2 = 0$ or

© E. F. Schubert, Rensselaer Polytechnic Institute, 2003

$$\Delta E_c = \chi_1 - \chi_2 \quad (1)$$

The valence band discontinuity then follows automatically as

$$\Delta E_v = E_{g2} - E_{g1} - \Delta E_c \quad (2)$$

Note that Eqs. (17.1) and (17.2) are valid only if the potential steps caused by atomic dipoles at the semiconductor surfaces and the heterostructure interfaces can be neglected.

The electron affinity model has successfully explained the band discontinuities of several semiconductor heterostructures. In the InAs / GaSb material system, the electron affinity rule correctly predicts a broken-gap alignment (Gobeli and Allen, 1966; Kroemer, 1985). The highly asymmetric lineup of InAs / GaAs heterostructures is also predicted well (Kroemer, 1985). In the Si / Ge heterostructure system, the electron affinity model predicts ΔE_c = 0.12 eV and ΔE_v = 0.33 eV in reasonable agreement with experimental data (Kroemer, 1985). Shay *et al*. (1976) and Phillips (1981) used the electron affinity rule to calculate

© E. F. Schubert, Rensselaer Polytechnic Institute, 2003

ΔE_c in CdS / InP heterostructures and found excellent agreement with their experimental data.

Despite the reasonable agreement between theory experiment, the electron affinity model suffers from several conceptual problems.

First, surface dipole layers affect the measurement of the electron affinity.

Second, electron correlation effects also influence the measured values of the electron affinity. Generally, the magnitude of correlation effects is small.

Due to the dipole and correlation effects, the applicability of the electron affinity rule is limited to semiconductors in which these effects are small.

© E. F. Schubert, Rensselaer Polytechnic Institute, 2003

Common anion rule

Many compound semiconductor heterostructures consist of two compounds which share a common anion element. For example in AlGaAs / GaAs heterostructures, As is the anion element on both sides of the heterostructure. It is a well established fact that the valence band wave functions evolve mainly from the atomic wave function of anions and the conduction band wave functions evolve mainly from the atomic wave functions of cations (see, for example, Harrison, 1980). Hence, the valence band structure of different semiconductors with the same anion element will be similar. Furthermore, *the valence band offsets of compound semiconductors with the same anion element are generally smaller than the conduction band offset*. This rule is clearly confirmed in the material system $Al_xGa_{1-x}As$ / GaAs where $\Delta E_c / \Delta E_v \cong 2 / 1$ for direct-gap range of $Al_xGa_{1-x}As$ ($x \leq 0.45$). The common anion rule also works well for GaAs / InAs heterostructures in which $\Delta E_c / \Delta E_v \cong 5 / 1$ (Kowalczyk *et al.*, 1982).

© E. F. Schubert, Rensselaer Polytechnic Institute, 2003

Harrison atomic orbital model

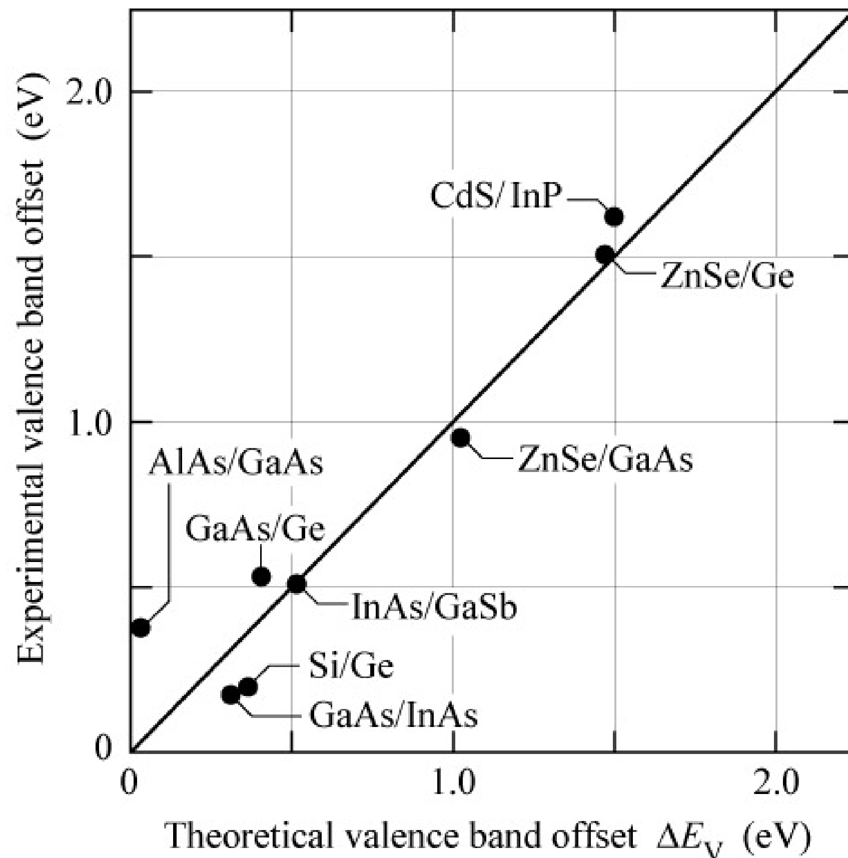


Fig. 17.5. Comparison of experimental valence band offsets with theoretical valence band offsets calculated by the Harrison atomic orbital theory. The AlAs / GaAs value is extrapolated from $Al_{0.30}Ga_{0.70}As$ / GaAs assuming $\Delta E_V / \Delta E_g = 1/3$.

Harrison (1977, 1980, and 1985) developed a theory based on atomic orbitals to predict band offsets in semiconductor heterostructures.

© E. F. Schubert, Rensselaer Polytechnic Institute, 2003

Kroemer (1985) compared the Harrison atomic orbital model and other models with experiments and he arrived at the conclusion that the Harrison model gives very good overall agreement with experimental band offsets.

The basis of the Harrison model is the *linear combination of atomic orbitals* of a very small group of atoms which is then used to calculate the band structure.

© E. F. Schubert, Rensselaer Polytechnic Institute, 2003

The effective dipole model

As we have already stated above, any dipole charges at the heterointerface will change the heterostructure band discontinuity. These dipole charges are due to the locally different atomic and electronic structure at the heterointerface as compared to the bulk atomic structure of either semiconductor. As a result of the different atomic environment at the heterointerface, valence electrons of atoms at the interface will move from their bulk equilibrium positions to new equilibrium positions. Hence, atomic dipoles are formed due to the new charge distribution at the heterointerface.

Ruan and Ching (1987) calculated heterostructure band offsets based on (i) the electron affinity model and (ii) by taking into account atomic dipoles at the interface which cause an additional shift of the band discontinuity.

© E. F. Schubert, Rensselaer Polytechnic Institute, 2003

Material system A / B	E_g^A (eV)	E_g^B (eV)	ΔE_V (eV)	$\Delta E_V / \Delta E_g$ (absolute value)	Remarks
Si / Ge	1.12	0.67	–0.16 to –0.40	0.35 to 0.89	(a)
Si / GaP	1.12	2.25	+0.80	0.71	(b)
Si / GaAs	1.12	1.42	+0.05	0.17	(b)
Si / GaSb	1.12	0.72	–0.05	0.12	(b)
Si / ZnSe	1.12	2.70	+1.25	0.79	(b)
Si / CdTe	1.12	1.52	+0.75	1.87	(b)
Ge / AlAs	0.67	2.15	+0.92	0.62	(b)
Ge / GaAs	0.67	1.42	+0.25 to +0.65	0.33 to 0.87	(b)
Ge / InP	0.67	1.34	+0.64	0.95	(b)
AlAs / GaAs	2.15	1.42	–0.40	0.55	(c)
$Al_{0.3}Ga_{0.7}As$ / GaAs	1.79	1.42	–0.12	0.32	(d)
AlSb / GaSb	1.61	0.72	–0.4	0.45	(b)
GaAs / InAs	1.42	0.36	–0.17	0.16	(b)
GaAs / ZnSe	1.42	2.70	+0.96 to +1.10	0.75 to 0.86	(b)
GaSb / InAs	0.72	0.36	+0.46	1.28	(b)
InP / CdS	1.34	2.42	+1.63	1.51	(b)
$Al_{0.48}In_{0.52}As$ / $Ga_{0.47}In_{0.53}As$	1.45	0.75	–0.21	0.30	(e)
$Ga_{0.52}In_{0.48}P$ / GaAs	1.88	1.42	–0.23	0.50	(f)
$Al_{0.48}In_{0.52}As$ / InP	1.45	1.34	+1.19	1.73	(g)
$Ga_{0.47}In_{0.53}As$ / InP	0.75	1.34	+0.40	0.68	(g)

© E. F. Schubert, Rensselaer Polytechnic Institute, 2003

(a) after Ruan and Ching (1987). Van de Walle and Martin (1986) showed that ΔE_V depends strongly on strain.
(b) after Ruan and Ching (1987)
(c) indirect gap AlAs, after Ruan and Ching (1987)
(d) direct gap $Al_xGa_{1-x}As$, after Pfeiffer *et al*. (1991) and after Menendez *et al*. (1986)
(e) after Peng *et al*. (1986) and after Sugiyama *et al*. (1986)
(f) Rao *et al*. (1987)
(g) after Tiwari and Frank (1992)

Table. Bandgap energies and valence band offsets of semiconductor heterostructures "A / B". The valence band offset ΔE_V is positive, if the top of the valence band of semiconductor "A" is higher than that of semiconductor "B".

© E. F. Schubert, Rensselaer Polytechnic Institute, 2003

Linearity of band offsets

Tiwari and Frank (1992) used experimental data of band offsets in order to plot the band edges of semiconductors as a function of the lattice constant. The Tiwari-and-Frank plot, relies on the experimental observation that the band offset from material "A" to material "B" plus the offset from material "B" to material "C" is equal to the band offset from material "A" to material "C". This linearity of offsets is consistent with the electron affinity model, and this property allows one to predict band alignments of any semiconductor heterostructure.

© E. F. Schubert, Rensselaer Polytechnic Institute, 2003

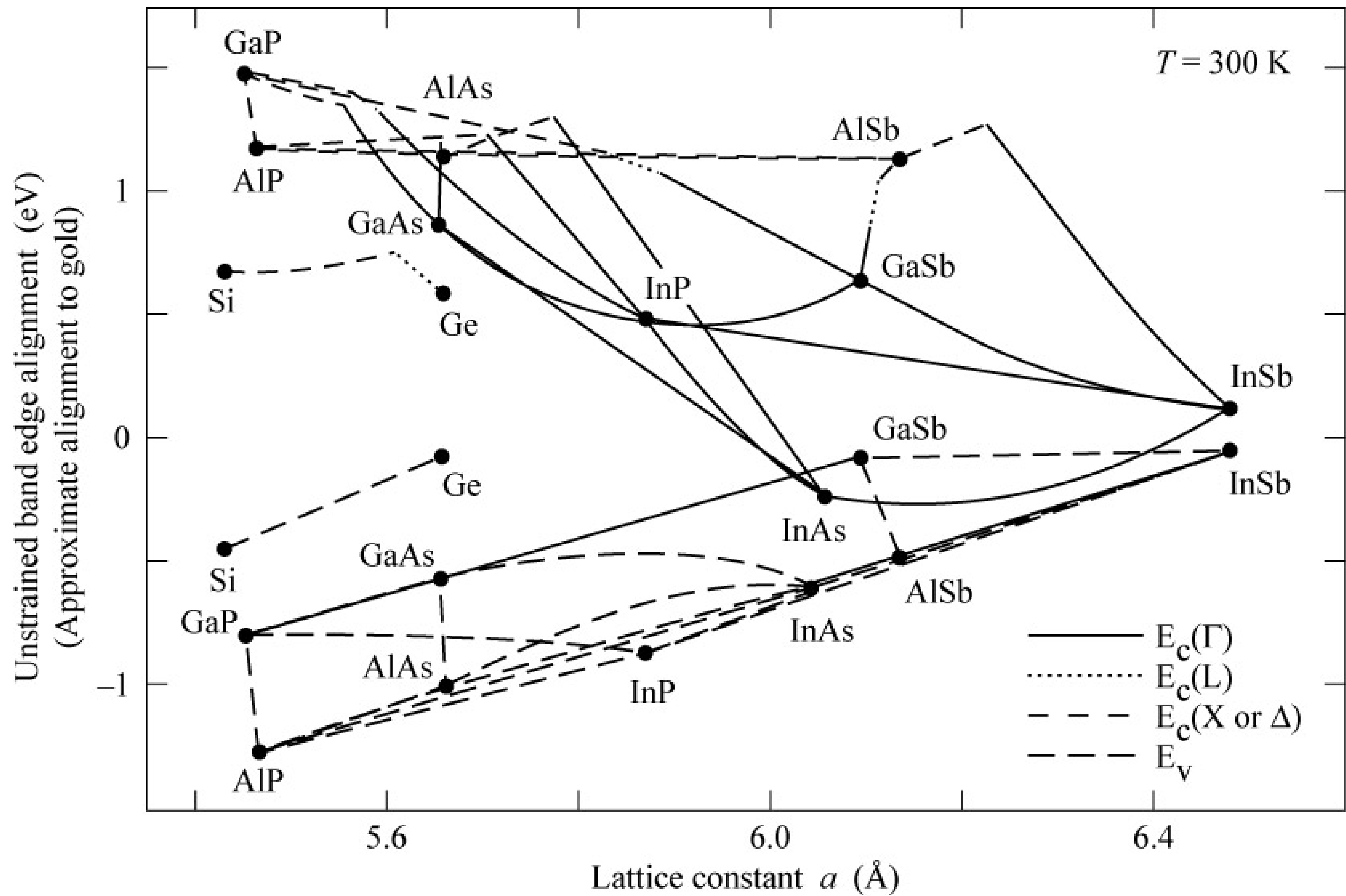


Fig. 17.6. Band edges as a function of the lattice constant. The zero energy point represents the approximate Fermi level (located mostly in the forbidden gap) of the gold / semiconductor Schottky contact (after Tiwari and Frank, 1992).

© E. F. Schubert, Rensselaer Polytechnic Institute, 2003

Boundary conditions at heterointerface

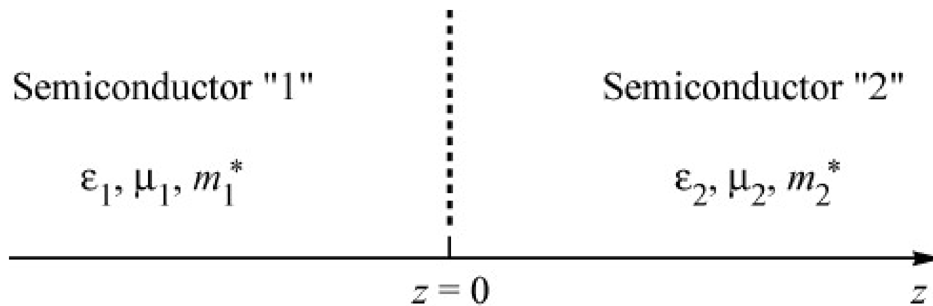


Fig. 17.7. Boundary between two semiconductors "1" and "2" which have a permittivity of ε_1, and ε_2, and an effective mass m_1^* and m_2^*, respectively.

The interface is located in the plane $z = 0$ of a cartesian coordinate system. The semiconductors "1" and "2" have a dielectric permittivity, magnetic permeability, and effective mass of ε_l, μ_1, m_1^*, and ε_2, μ_2, m_2^*, respectively.

The first boundary condition considered here concerns the Fermi level. *The* ***Fermi level*** *is constant across a heterointerface under thermal equilibrium conditions*. **Why?**

© E. F. Schubert, Rensselaer Polytechnic Institute, 2003

Four electrodynamics boundary conditions must be satisfied at heterointerfaces:

The ***magnetic boundary condition*** states that the tangential component of the magnetic field (H_t) and the normal component of the magnetic induction (B_n) are constant across interfaces.

The ***electric boundary condition*** states that the tangential component of the electric field (E_t) and the normal component of the dielectric displacement (D_n) are constant across interfaces.

© E. F. Schubert, Rensselaer Polytechnic Institute, 2003

The latter boundary condition, D_n = const, is now used to derive another "boundary condition", namely the ***charge neutrality condition***. We denote the normal component of the dielectric displacement at the interface as D_{1n} and D_{2n} in semiconductor "1" and "2", respectively.

Furthermore we assume that the dielectric displacement vanishes for sufficiently large distances from the interface. Then, using Gauss's equation, the boundary condition $D_{1n} = D_{2n}$ can be written as:

$$\mathcal{D}_{1n} = \int_{z=-\infty}^{z=0} \rho(z)\,\mathrm{d}z = -\int_{z=0}^{z=\infty} \rho(z)\,\mathrm{d}z = \mathcal{D}_{2n} \qquad (3)$$

© E. F. Schubert, Rensselaer Polytechnic Institute, 2003

We finally discuss the ***boundary conditions for the quantum-mechanical wave function*** at heterointerfaces. The interface is located in the plane $z = 0$ and we are only interested in the z dependence of the wave function $\psi(z)$. In the chapter entitled “Resume of quantum mechanical principles”, the boundary conditions for $\psi(z)$ is given by

$$\psi_1(z \to -0) \quad = \quad \psi_2(z \to +0) \tag{4}$$

That is, the wave function is continuous at the interface.

The boundary condition for the derivative of the wave function is given by:

$$\left. \frac{1}{m_1^*} \frac{d\psi_1}{dz} \right|_{z \to -0} = \left. \frac{1}{m_2^*} \frac{d\psi_2}{dz} \right|_{z \to +0} \tag{5}$$

© E. F. Schubert, Rensselaer Polytechnic Institute, 2003

Graded gap structures

Assume two semiconductors "A" and "B" that are chemically miscible. The mixed compound, also called ***semiconductor alloy***, is designated by the chemical formula A_xB_{1-x}, where x is the ***mole fraction*** of semiconductor "A" in the mixed compound. The mole fraction x is also designated as the chemical ***composition*** of the compound A_xB_{1-x}. Most semiconductors of practical relevance are completely miscible. Assume further that the gap energy of semiconductor "A" and "B" are different and that the bandgap energy depends on the composition. The dependence of the forbidden-gap energy on the composition x is usually expressed in terms of a parabolic (linear plus quadratic) dependence. The gap energy of the alloy A_xB_{1-x} is then given by:

$$E_{\mathrm{g}}^{\mathrm{AB}} = x\,E_{\mathrm{g}}^{\mathrm{A}} + (1-x)\,E_{\mathrm{g}}^{\mathrm{B}} + x\,(1-x)\,E_{\mathrm{b}} \qquad (6)$$

where the first two summands describe the linear dependence of the gap and the summand $x\,(1-x)E_{\mathrm{b}}$ describes the quadratic dependence of the gap. The parameter E_{b} is called the ***bowing parameter***. For some

© E. F. Schubert, Rensselaer Polytechnic Institute, 2003

semiconductor alloys, *e. g.* $(AlAs)_x(GaAs)_{1-x}$, the bowing parameter is vanishingly small. The bandgap of the alloy is then given by

$$E_g^{AB} = x\,E_g^{A} + (1-x)\,E_g^{B} \tag{7}$$

Equations (17.6) and (17.7) are valid for homogeneous bulk semiconductors.

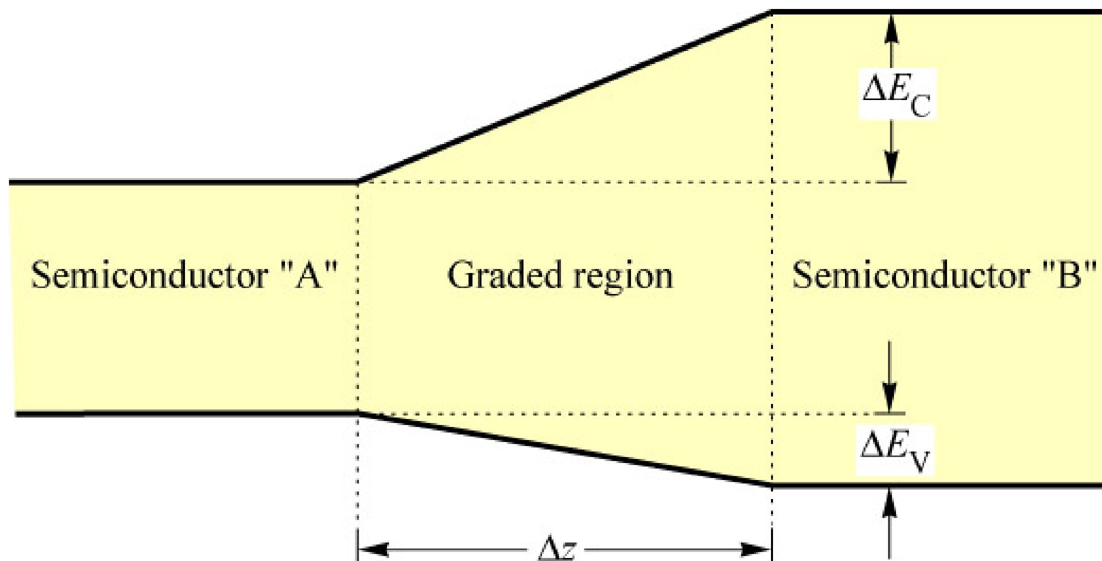


Fig. 17.8. Linearly graded region between a semiconductor "A" and "B" creating a quasi-electric field in the graded transition region.

© E. F. Schubert, Rensselaer Polytechnic Institute, 2003

Graded gap semiconductor structures were first considered by Kroemer (1957). He showed that the changes of the band-edge energies with position can be understood as ***quasi-electric fields***. The quasi-electric field of the band diagram shown in ***Fig***. 17.8 is, in the conduction band, given by

$$\left| \mathcal{E}_{\mathrm{C}} \right| = \Delta E_{\mathrm{C}} \, / \, \left(e \, \Delta z \right) \tag{8}$$

In the valence band it is given by

$$\left| \mathcal{E}_{\mathrm{V}} \right| = \Delta E_{\mathrm{V}} \, / \, \left(e \, \Delta z \right) \tag{9}$$

Figure 17.8 reveals that the electric fields in the conduction band and in the valence band have opposite polarization. Therefore, electrons and holes are driven in the *same* direction. This cannot be achieved by real electric fields in which electrons and holes are *always* driven in opposite directions. Due to this difference, Kroemer (1957) designated the fields occurring in graded semiconductor structures as ***quasi-electric fields***.

© E. F. Schubert, Rensselaer Polytechnic Institute, 2003

In some graded gap structures, electrons and holes are pulled in the same direction, as discussed for the band diagram.

In other graded gap structures, one of the bands could, *e. g*. the valence band, may be flat, while the other band could have a quasi-electric field.

Kroemer also envisioned graded-gap heterobipolar transistors which enhance the minority carrier transport through the base.

© E. F. Schubert, Rensselaer Polytechnic Institute, 2003

Example
The energy gap of the unstrained Si_xGe_{1-x} was analyzed as a function of composition by Weber and Alonso (1989). Analytical expressions were given for the lowest energy gap as a function of the composition. For $x \le 0.85$, the X band is the lowest conduction band minimum. The energy gap of unstrained Si_xGe_{1-x} is given by

$$E_g(x) = \left(1.155 - 0.43\,x + 0.206\,x^2\right)\text{eV} \quad \text{for } x \le 0.85 \qquad (10a)$$

For $x > 0.85$, the L band is the lowest conduction band minimum. The energy gap is then given by

$$E_g(x) = \left(2.010 - 1.270\,x\right)\text{eV} \qquad \text{for } x > 0.85 \qquad (10b)$$

For strained Si_xGe_{1-x} grown on Si substrates, Lang *et al*. (1985) showed that the degenerate valence band splits into two bands. The energy gap

© E. F. Schubert, Rensselaer Polytechnic Institute, 2003

between the lowest conduction band and the highest valence band is then given by

$$E_g(x) = \left(1.155 - 0.65\,x + 0.22\,x^2\right)\text{eV} \qquad \text{for} \quad x \le 0.70 \qquad \text{(10c)}$$

This equation is an analytical expression of the low-temperature (90 K) photoluminescence data of Lang *et al*. (1985).

© E. F. Schubert, Rensselaer Polytechnic Institute, 2003

Vegard's law

In graded semiconductor structures, the composition of the semiconductor is varied. This variation in chemical composition is not only accompanied by a change of the bandgap energy, but also by a change in the lattice constant.

The change in lattice constant is, for all semiconductor alloys, governed by Vegard's law. Consider a semiconductor "A" with a lattice constant a_0^A and a semiconductor "B" with the lattice constant a_0^B.

Then the lattice constant of the alloy A_xB_{1-x} is given by Vegard's law which states:

$$a_0^{AB} = a_0^A x + a_0^B (1 - x) \quad (11)$$

For most graded semiconductor structures, it is imperative that the lattice constant does not change as the composition of the alloy is varied. Such

© E. F. Schubert, Rensselaer Polytechnic Institute, 2003

structures are called ***lattice-matched*** graded semiconductors. If semiconductors are not lattice matched, graded semiconductors. If semiconductors are not lattice matched, microscopic defects occur when the composition is varied. These defects degrade the quality, *e. g*. the radiative efficiency, of the semiconductor.

© E. F. Schubert, Rensselaer Polytechnic Institute, 2003

Bandgap data

Relationship between the gap energy, the corresponding wavelength, and the lattice constant of group-IV and group III-V semiconductors:

© E. F. Schubert, Rensselaer Polytechnic Institute, 2003

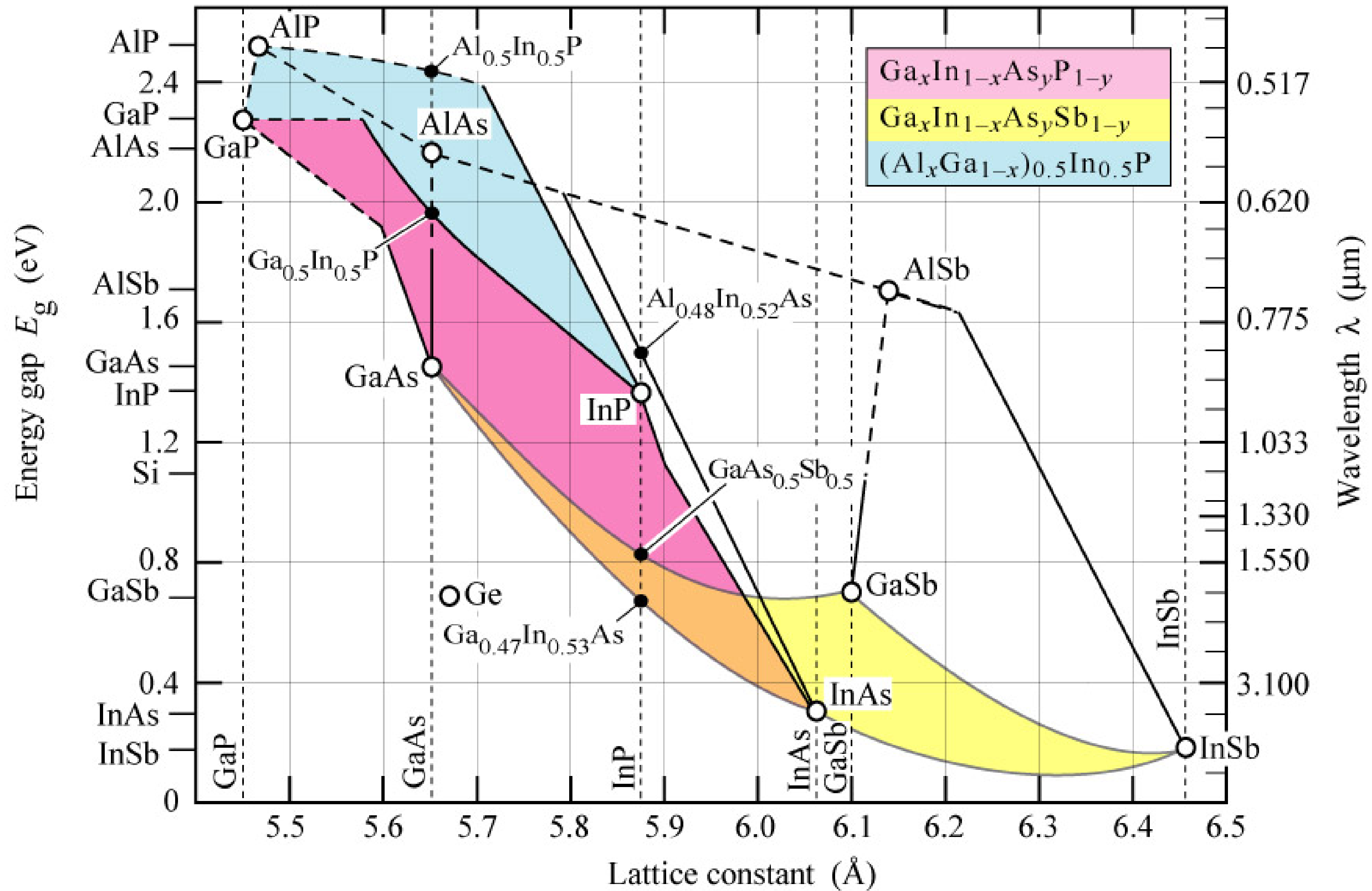


Fig. 17.9 Lattice constant versus energy gap at room temperature for various III–V semiconductors and their alloys (after Tien, 1985).

© E. F. Schubert, Rensselaer Polytechnic Institute, 2003

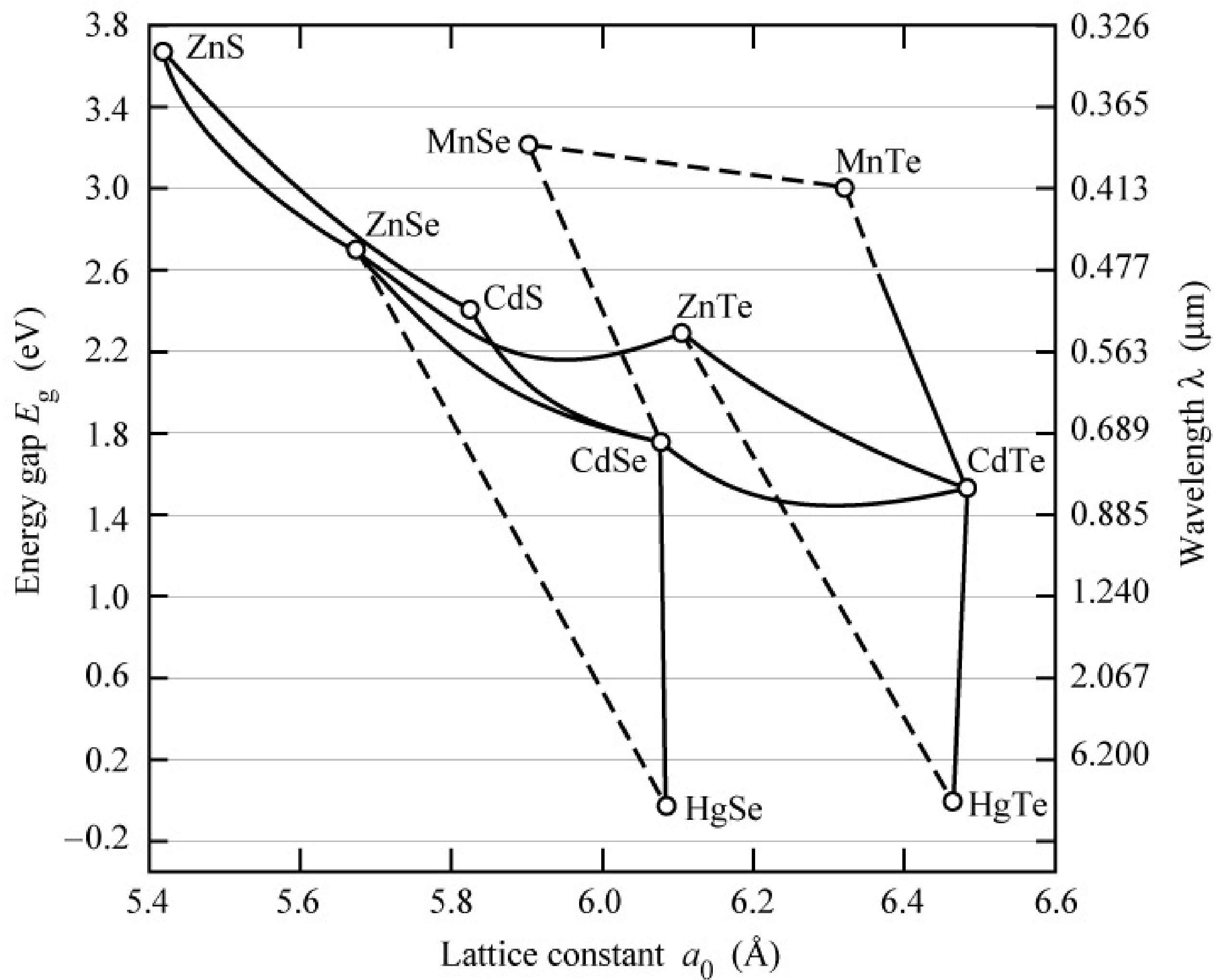
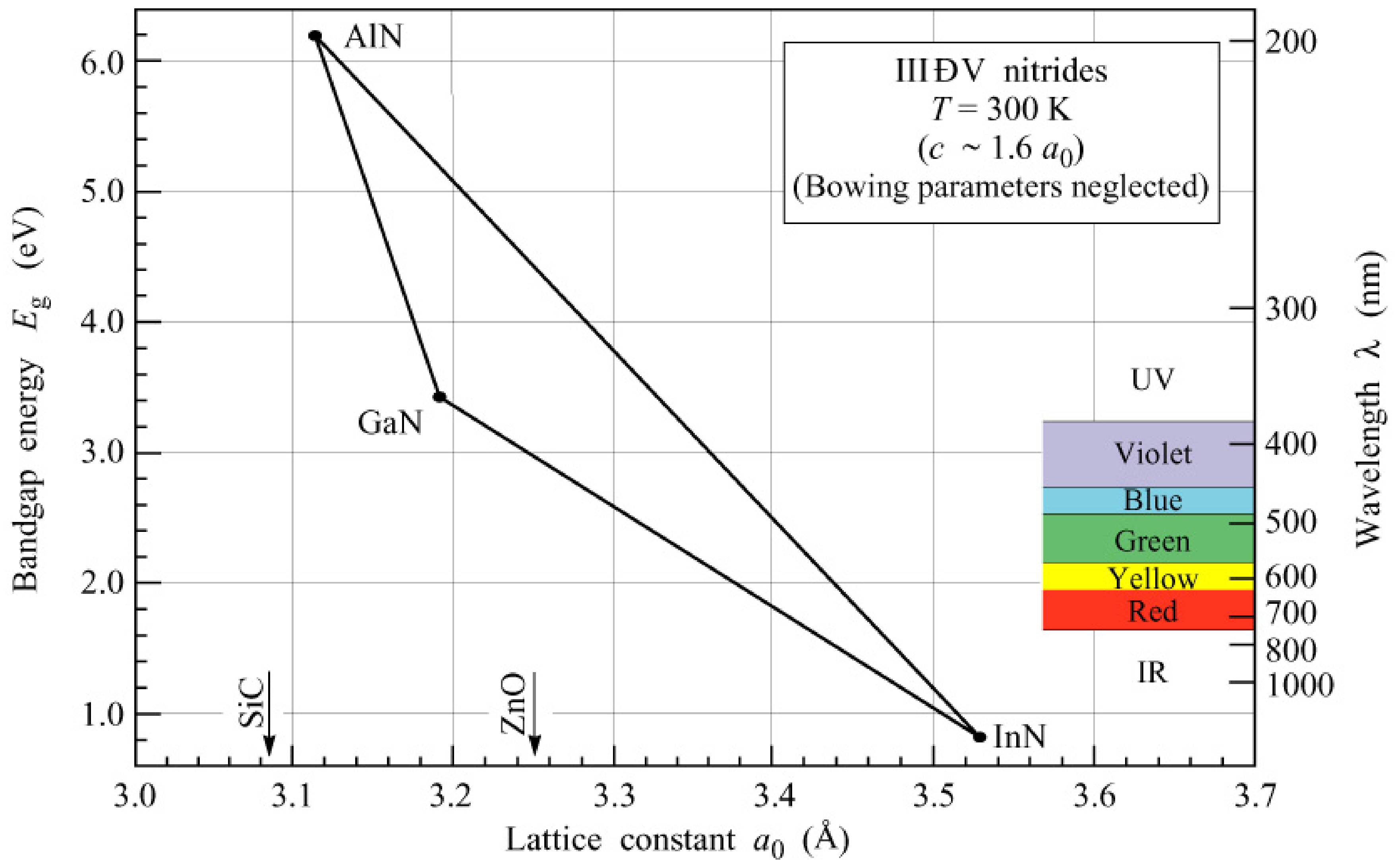


Fig. 17.11 Lattice constant versus energy gap at room temperature for various II–VI semiconductors and their alloys (after Feldman *et al*., 1992).

© E. F. Schubert, Rensselaer Polytechnic Institute, 2003



© E. F. Schubert, Rensselaer Polytechnic Institute, 2003