Chapter 5
Mechanical Properties

Abstract Lattice vibrations and phonons are treated with one-dimensional models and examples for real phonon dispersions for several semiconductors including phonons in alloys and disordered materials are given. Then the theory of linear elasticity and its application to semiconductors with regard to epitaxial strain, substrate bending and sheet-scrolling is given. Finally plastic relaxation effects such as critical thickness and wafer breakage are discussed.

5.1 Introduction

The atoms making up the solid have an average position from which they can deviate since they are elastically bonded. The typical atomic interaction potential looks like the one shown in Fig. 2.1. The atoms thus perform a vibrational motion (including zero point fluctuations) and the solid is elastic. The potential is essentially asymmetric, being steeper for small distances due to quantum-mechanical overlap of orbitals. However, for small amplitudes around the minimum a harmonic oscillator can be assumed (harmonic approximation). Beyond the elastic regime, plastic deformation occurs such as generation of defects, e.g. dislocations. Eventually also the crystal can break.

5.2 Lattice Vibrations

In the following we will discuss the dispersion relations for lattice vibrations, i.e. the connection between the frequency ν (or energy $h\nu = \hbar\omega$) of the wave and its wavelength λ (or k-vector $k = 2\pi/\lambda$). Acoustic and optical vibrations are introduced in one-dimensional models. A detailed treatment of the physics of lattice vibrations is given in [328].

M. Grundmann, *The Physics of Semiconductors*, Graduate Texts in Physics,
DOI 10.1007/978-3-319-23880-7_5

5.2.1 Monoatomic Linear Chain

The essential physics of lattice vibrations can best be seen from a one-dimensional model that is called the linear chain. The mechanical vibrations will also be called *phonons*, although technically this term is reserved for the quantized lattice vibrations resulting from the quantum-mechanical treatment.

In the monoatomic linear chain the atoms of mass M are positioned along a line (x-axis) with a period (lattice constant) a at the positions $x_{n_0} = na$. This represents a one-dimensional Bravais lattice. The Brillouin zone of this system is $[-\pi/a, \pi/a]$.

The atoms will interact with a harmonic potential, i.e. the energy is proportional to the displacement $u_n = x_n - x_{n_0}$ to the second power. The total (mechanical) energy of the system is then:

$$U = \frac{1}{2} C \sum_n (u_n - u_{n+1})^2. \tag{5.1}$$

The model assumes that the mass points are connected via massless, ideal springs with a spring constant C. If $\phi(x)$ is the interaction energy between two atoms, C is given by $C = \phi''(a)$. Again, the harmonic approximation is only valid for small displacements, i.e. $u_n \ll a$. The displacement of the atoms can be along the chain (longitudinal wave) or perpendicular to the chain (transverse wave), see Fig. 5.1. We note that for these two types of waves the elastic constant C must not be the same.

When the sum in (5.1) has a finite number of terms ($n = 0, \ldots, N-1$), the boundary conditions have to be considered. There are typically two possibilities: The boundary atoms are fixed, i.e. $u_0 = u_{N-1} = 0$, the boundary conditions are periodic (Born–von Karman), i.e. $u_i = u_{N+i}$. If $N \gg 1$, the boundary conditions play no significant role anyway, thus those with the greatest ease for subsequent math are chosen. In solid-state physics typically periodic boundary conditions are

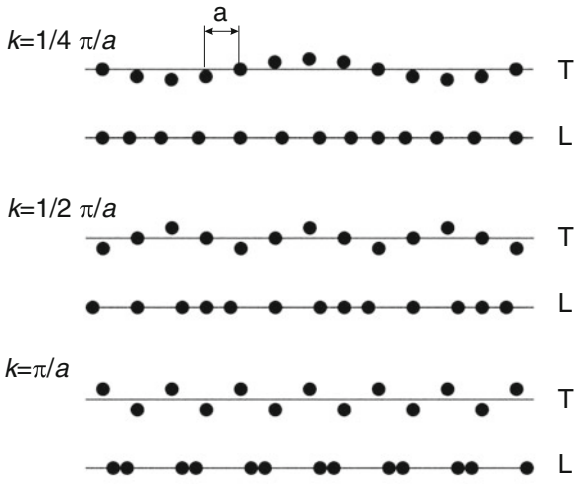

Fig. 5.1 Visualization of transverse ('T') and longitudinal ('L') waves in a linear monoatomic chain at different wavevectors

used. Boundary phenomena, such as at surfaces, are then treated separately (see Sect. 11.6.1).

The equations of motion derived from (5.1) are

$$M\ddot{u}_n = F_n = -\frac{\partial U}{\partial u_n} = -C\,(2u_n - u_{n-1} - u_{n+1}). \tag{5.2}$$

We solve for solutions that are periodic in time (harmonic waves), i.e. $u_n(x,t) = u_n \exp(-\mathrm{i}\omega t)$. Then the time derivative can be executed immediately as $\ddot{u}_n = -\omega^2 u_n$ and we obtain:

$$M\omega^2\, u_n = C\,(2u_n - u_{n-1} - u_{n+1}). \tag{5.3}$$

If, also, the solution is periodic in space, i.e. is a (one-dimensional) plane wave, i.e. $u_n(x,t) = v_0 \exp(\mathrm{i}(kx - \omega t))$ with $x = na$, we find from the periodic boundary condition $\exp(\mathrm{i}kNa) = 1$ and thus

$$k = \frac{2\pi}{a}\,\frac{n}{N},\ n \in \mathbf{N}. \tag{5.4}$$

It is important that, when k is altered by a reciprocal space vector, i.e. $k' = k + 2\pi n/a$, the displacements u_n are unaffected. This property means that there are only N values for k that generate independent solutions. These can be chosen as $k = -\pi/a, \ldots, \pi/a$, so that k lies in the Brillouin zone of the lattice. In the Brillouin zone there is a total number of N k-values, i.e. one for each lattice point. The distance between adjacent k-values is

$$\frac{2\pi}{Na} = \frac{2\pi}{L}, \tag{5.5}$$

L being the lateral extension of the system.

The displacements at the lattice points n and $n+m$ are now related to each other via

$$\begin{aligned} u_{n+m} &= v_0\, \exp(\mathrm{i}k(n+m)a) \\ &= v_0\, \exp(\mathrm{i}kna)\, \exp(\mathrm{i}kma) = \exp(\mathrm{i}kma)\, u_n. \end{aligned} \tag{5.6}$$

Thus, the equation of motion (5.3) reads

$$M\omega^2 u_n = C\,\left[2 - \exp(-\mathrm{i}ka) - \exp(\mathrm{i}ka)\right] u_n. \tag{5.7}$$

Using the identity $\exp(\mathrm{i}ka) + \exp(-\mathrm{i}ka) = 2\cos(ka)$, we find the dispersion relation of the monoatomic linear chain (Fig. 5.2):

$$\omega^2(k) = \frac{4C}{M}\,\frac{1-\cos(ka)}{2} = \frac{4C}{M}\,\sin^2\left(\frac{ka}{2}\right). \tag{5.8}$$

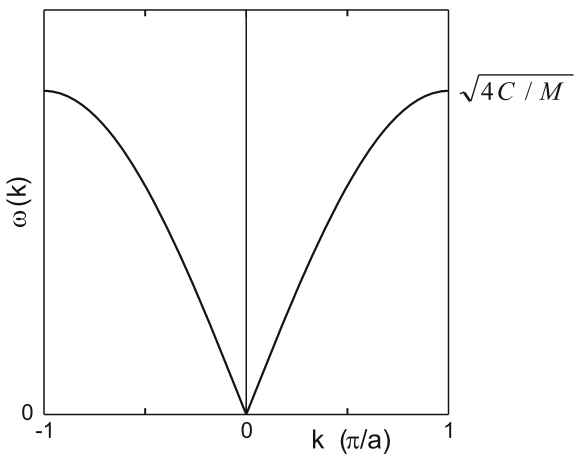

Fig. 5.2 Dispersion relation for a monoatomic linear chain

The solutions describe plane waves that propagate in the crystal with a phase velocity $c = \omega/k$ and a group velocity $v_g = \mathrm{d}\omega/\mathrm{d}k$

$$v_g = \sqrt{\frac{4C}{M}}\,\frac{a}{2}\,\cos\left(\frac{|k|a}{2}\right). \tag{5.9}$$

In the vicinity of the Γ point, i.e. $k \ll \pi/a$ the dispersion relation is linear in k

$$\omega(k) = a\sqrt{\frac{C}{M}}\,|k|. \tag{5.10}$$

We are used to such linear relations for sound (and also light) waves. The phase and group velocity are the same and do not depend on k. Thus, such solutions are called *acoustic*. The sound velocity of the medium is given by $v_\mathrm{s} = a\sqrt{C/M}$.

It is characteristic of the nonhomogeneous medium that, when k approaches the boundary of the Brillouin zone, the behavior of the wave is altered. For $k = \pi/a$ the wavelength is just $\lambda = 2\pi/k = 2a$, and thus samples the granularity of the medium. The maximum phonon frequency ω_m is

$$\omega_\mathrm{m} = \sqrt{\frac{4C}{M}}. \tag{5.11}$$

The group velocity is zero at the zone boundary, thus a standing wave is present.

Since the force constants of the longitudinal and transverse waves can be different, the dispersion relations are different. The transverse branch of the dispersion relation is 2-fold degenerate, unless the two directions that are perpendicular to x are not equivalent.

5.2.2 Diatomic Linear Chain

Now we consider the case that the system is made up from two different kinds of atoms (Fig. 5.3). This will be a model for semiconductors with a diatomic base, such as zincblende. We note that the diamond structure also needs to be modeled in this way, although both atoms in the base are the same.

The lattice will be the same and the lattice constant will be a. Alternating atoms of sort 1 and 2 with a relative distance of $a/2$ are on the chain. The displacements of the two atoms are labeled u_n^1 and u_n^2, both belonging to the lattice point n. The atoms have the masses M_1 and M_2. The force constants are C_1 (for the 1–2 bond within the base) and C_2 (for the 2–1 bond between different bases).

The total energy of the system is then given as

$$U = \frac{1}{2} C_1 \sum_n \left(u_n^1 - u_n^2\right)^2 + \frac{1}{2} C_2 \sum_n \left(u_n^2 - u_{n+1}^1\right)^2 . \tag{5.12}$$

The equations of motion are

$$M_1 \ddot{u}_n^1 = -C_1 \left(u_n^1 - u_n^2\right) - C_2 \left(u_n^1 - u_{n-1}^2\right) \tag{5.13a}$$

$$M_2 \ddot{u}_n^2 = -C_1 \left(u_n^2 - u_n^1\right) - C_2 \left(u_n^2 - u_{n+1}^1\right) . \tag{5.13b}$$

With the plane-wave ansatz $u_n^1(x,t) = v_1 \exp\left(\mathrm{i}(kna - \omega t)\right)$ and $u_n^2(x,t) = v_2 \exp\left(\mathrm{i}(kna - \omega t)\right)$ and periodic boundary conditions we find

$$0 = -M_1 \omega^2 v_1 + C_1 (v_1 - v_2) + C_2 \left[v_1 - v_2 \exp(-\mathrm{i}ka)\right] \tag{5.14a}$$

$$0 = -M_2 \omega^2 v_2 + C_1 (v_2 - v_1) + C_2 \left[v_2 - v_1 \exp(\mathrm{i}ka)\right]. \tag{5.14b}$$

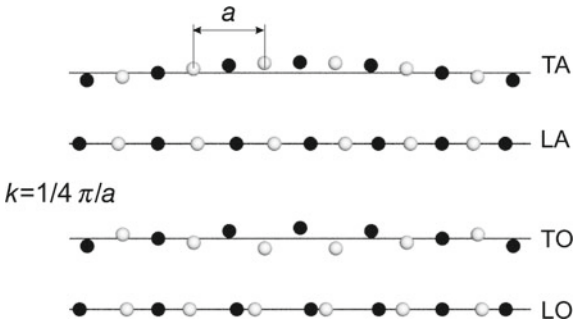

Fig. 5.3 Visualization of acoustic and optical waves in a diatomic linear chain

These equations for v_1 and v_2 can only be solved nontrivially if the determinant vanishes, i.e.

$$0 = \begin{vmatrix} M_1\,\omega^2 - (C_1 + C_2) & C_1 + \mathrm{e}^{-\mathrm{i}ka}\,C_2 \\ C_1 + \mathrm{e}^{\mathrm{i}ka}\,C_2 & M_2\,\omega^2 - (C_1 + C_2) \end{vmatrix} \tag{5.15}$$
$$= M_1 M_2\,\omega^4 - (M_1 + M_2)(C_1 + C_2)\,\omega^2 + 2C_1C_2\,[1 - \cos(ka)].$$

Using the substitutions $C_+ = (C_1 + C_2)/2$, $C_\times = \sqrt{C_1C_2}$, the arithmetic and geometrical averages, and accordingly for M_+ and $M_\times$, the solution is

$$\omega^2(k) = \frac{2C_\times}{\gamma\,M_\times}\left[1 \pm \sqrt{1 - \gamma^2\,\frac{1 - \cos(ka)}{2}}\right], \tag{5.16}$$

with

$$\gamma = \frac{C_\times\,M_\times}{C_+M_+} \le 1. \tag{5.17}$$

The dispersion relation, as shown in Fig. 5.4, now has (for each longitudinal and transverse mode) two branches. The lower branch ('−' sign in (5.16)) is related to the acoustic mode; neighboring atoms have similar phase (Fig. 5.3). For the acoustic mode $\omega = 0$ at the Γ point and the frequency increases towards the zone boundary. The maximum phonon frequency ω_m is in the upper branch ('+' sign in (5.16)) at the zone center

$$\omega_\mathrm{m} = \sqrt{\frac{4C_\times}{\gamma\,M_\times}} = 2\sqrt{\frac{C_+\,M_+}{M_\times^2}}. \tag{5.18}$$

The upper branch is called the optical mode (since it can interact strongly with light, see Sect. 9.9) and neighboring atoms have opposite phase. In the vicinity of the Γ point the dispersion of optical phonons is parabolic with negative curvature:

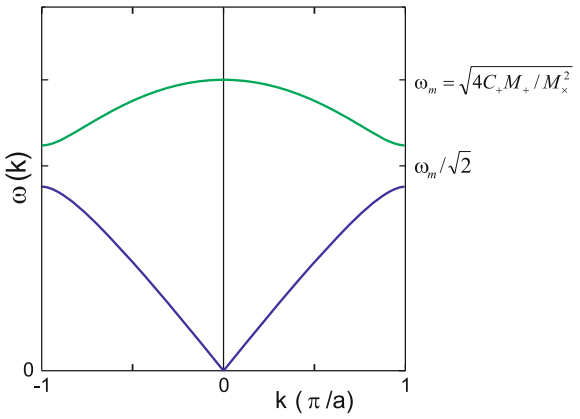

Fig. 5.4 Dispersion relation for a diatomic linear chain with acoustic (*blue*) and optical (*green*) branch

$$\omega(k) \cong \omega_{\mathrm{m}} \left[1 - \frac{1}{2} \left(\frac{\gamma a}{4} \right)^2 k^2 \right]. \tag{5.19}$$

Thus, four different vibrations exist that are labeled TA, LA, TO, and LO. Both the TA and TO branches are degenerate.

At the zone boundary (X point) a frequency gap exists. The gap center is at

$$\overline{\omega_{\mathrm{X}}} = \frac{\omega_{\mathrm{m}}}{\sqrt{2}}, \tag{5.20}$$

and the total width of the gap is

$$\Delta\omega_X = \omega_{\mathrm{m}} \sqrt{1-\gamma} = 2\sqrt{\frac{C_+ M_+ - C_\times M_\times}{M_\times^2}}. \tag{5.21}$$

The group velocity is zero for optical and acoustic phonons at $k = \pi/a$ and for optical phonons at the Γ point.

Usually two cases are treated explicitly: (i) atoms with equal mass ($M = M_1 = M_2$) and different force constants or (ii) atoms with unequal mass and identical force constants $C = C_1 = C_2$. For the case $C_1 = C_2$ and $M_1 = M_2$, $\gamma = 1$ and thus $\Delta\omega_X = 0$. Then the dispersion relation is the same as for the monoatomic chain, except that the k space has been folded since the actual lattice constant is now $a/2$.

$M_1 = M_2$

In this case, $M_+ = M_\times = M$ and the dispersion relation is

$$\omega^2 = \frac{2C_+}{M} \left[1 \pm \sqrt{1 - \frac{C_\times^2}{C_+^2} \frac{1-\cos(ka)}{2}} \right]. \tag{5.22}$$

At the zone boundary the frequencies for the acoustic and the optical branch are $\omega_{X,1} = \sqrt{2C_1/M}$ with $v_1 = v_2$ and $\omega_{X,2} = \sqrt{2C_2/M}$ with $v_1 = -v_2$, respectively (assuming $C_2 > C_1$). The motion for $k = \pi/a$ is phase shifted by 180° for adjacent bases. Additionally, for the acoustic branch the atoms of the base are in phase, while for the optical branch the atoms of the base are 180° out of phase. The vibration looks as if only one of the springs is strained.

$C_1 = C_2$

In this case, $C_+ = C_\times = C$ and the dispersion relation is

$$\omega^2 = \frac{2CM_+}{M_\times^2} \left[1 \pm \sqrt{1 - \frac{M_\times^2}{M_+^2} \frac{1-\cos(ka)}{2}} \right]. \tag{5.23}$$

At the zone boundary the frequencies for the acoustic and the optical branch are $\omega_{X,1} = \sqrt{2C/M_1}$ with $v_2 = 0$ and $\omega_{X,2} = \sqrt{2C/M_2}$ with $v_1 = 0$, respectively

(assuming $M_2 < M_1$). In the vibration for $k = \pi/a$ thus only one atom species oscillates, the other does not move. Close to the Γ point the atoms are in phase in the acoustic branch, i.e. $v_1 = v_2$. For the optical branch, the frequency at the Γ point is given by $\omega = \sqrt{2C/M_r}$ (with the reduced mass $M_r^{-1} = M_1^{-1} + M_2^{-1} = 2M_+/M_\times^2$) and the amplitude ratio is given by the mass ratio: $v_2 = -(M_1/M_2)\, v_1$, i.e. the heavier atom has the smaller amplitude.

5.2.3 *Lattice Vibrations of a Three-Dimensional Crystal*

When calculations are executed for a three-dimensional crystal with a monoatomic base, there are $3N$ equations of motion. These are transformed to normal coordinates and represent 3 acoustic branches (1 LA phonon mode and 2 TA phonon modes) of the dispersion relation. In a crystal with a base with p atoms, there are also 3 acoustic branches and $3(p-1)$ optical branches. For a diatomic base (as in the zincblende structure) there are 3 optical phonon branches (1 LO phonon mode and 2 TO phonon modes). The total number of modes is $3p$. The dispersion $\omega(\mathbf{k})$ now has to be calculated for all *directions* of $\mathbf{k}$.

In Figs. 5.5 and 5.6, the phonon dispersion in silicon, GaAs and GaP is shown along particular lines in the Brillouin zone (cf. Fig. 3.34b). The degeneracy of the LO and TO energies at the Γ point for the covalent group-IV semiconductor is lifted for the III–V semiconductors due to the ionic character of the bond and the macroscopic electric field connected with the long-wavelength LO phonon (see Sect. 5.2.9). Comparing GaAs [329] and GaP [330], the quite different mass of Ga- and P-atoms ($M_\times/M_+ \approx 0.92$) leads to the formation of a clear gap between the acoustic and optical branches, while for GaAs $M_\times/M_+ \approx 0.9994$ is close to 1 and no gap forms.

We note that the degeneracy of the TA phonon is lifted for propagation along the $\langle 110\rangle$ directions (Σ) because the two transverse directions $\langle 001\rangle$ and $\langle 1\bar{1}0\rangle$ are not equivalent.

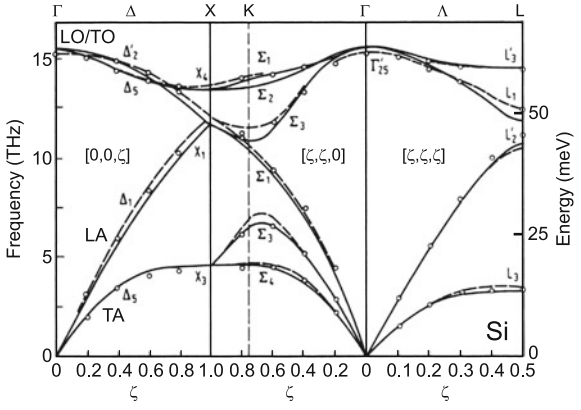

Fig. 5.5 Phonon dispersion in Si, experimental data and theory (*solid lines*: bond charge model, *dashed lines*: valence force field model). Adapted from [146]

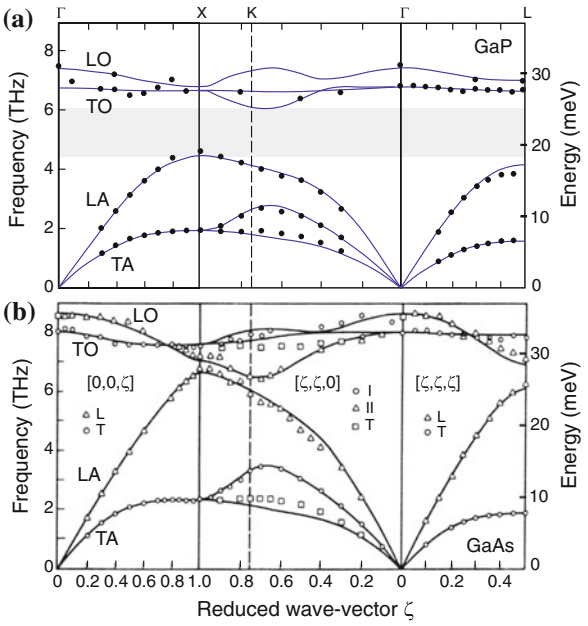

Fig. 5.6 Phonon dispersion in (**a**) GaP and (**b**) GaAs. Experimental data (*symbols*) and theory (*solid lines*, 14-parameter shell model). 'L' and 'T' refer to longitudinal and transverse modes, respectively. 'I' and 'II' (along $[\zeta, \zeta, 0]$) are modes whose polarization is in the $(1, \bar{1}, 0)$ plane. The *grey area* in (**a**) denotes the gap between acoustical and optical states. (**a**) Adapted from [331], (**b**) adapted from [329]

In boron nitride the masses of the two constituents are so similar that no gap exists between acoustical and optical branches (Fig. 5.7). Also the density of states (averaged over the entire Brillouin zone) is depicted (see next Sect.).

The displacement of atoms is shown in Fig. 5.8 for the different phonon modes present in zincblende crystals and in Fig. 5.9 for wurtzite crystals. The modes are labeled with their symmetry (in molecular notation) according to group theory (see remark in Sect. 6.2.5).

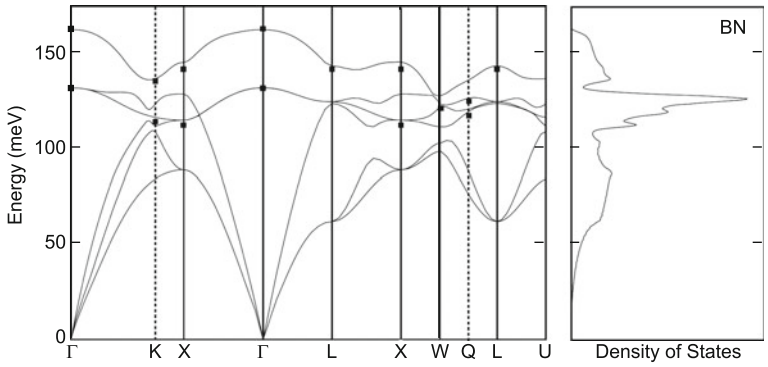

Fig. 5.7 Phonon dispersion in BN (*left panel*), experimental data (*symbols*) and theory (*solid lines*, first principles pseudopotential model). In the *right panel* the density of states is depicted. Adapted from [332]

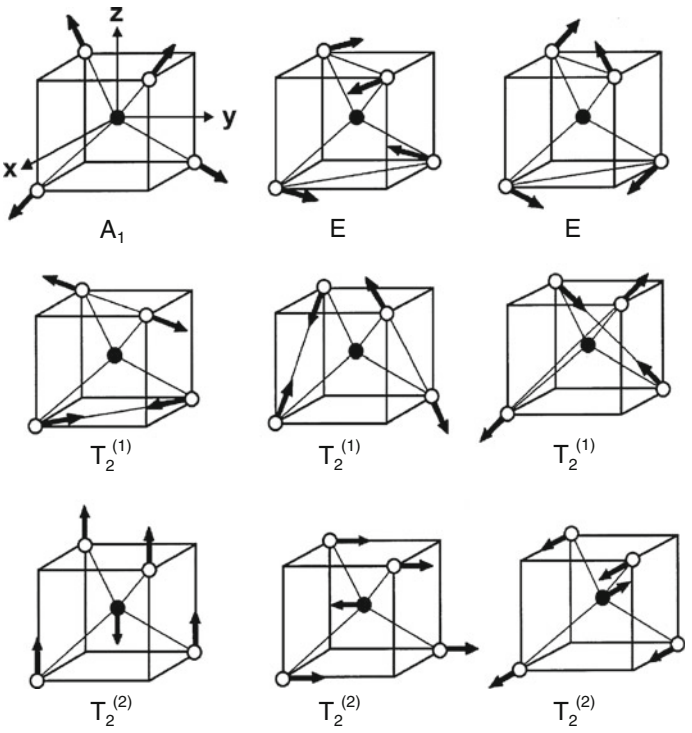

Fig. 5.8 Displacement of atoms for various phonon modes in zincblende crystals. Adapted from [333]

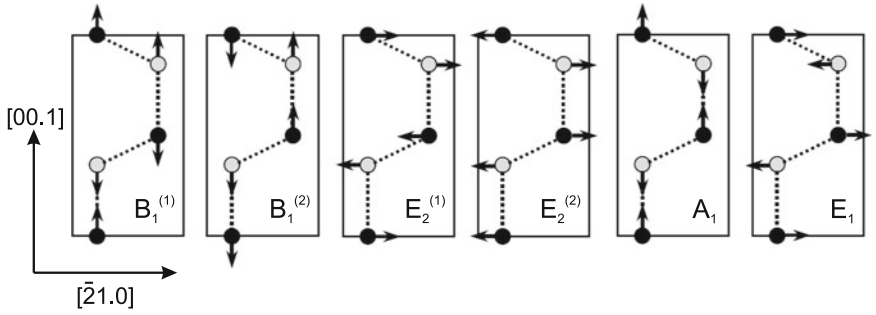

Fig. 5.9 Displacement of atoms for various phonon modes in wurtzite crystals. Adapted from [334]

The dependence of the phonon frequency on the mass of the atoms ($\propto 1/\sqrt{M}$) can be demonstrated with the isotope effect, visualized for GaAs in Fig. 5.10. The dependence of the phonon frequencies on the stiffness of the spring can be seen from Fig. 5.11; the smaller lattice constant provides the stiffer spring.

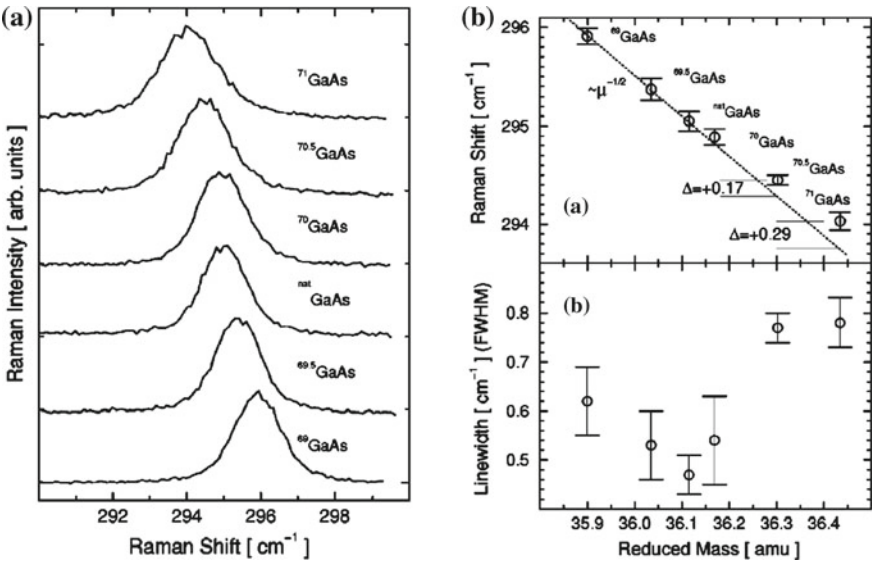

Fig. 5.10 (**a**) Raman spectra of GaAs with different isotope content as labeled. (**b**) Energy of optical phonons in GaAs with different isotope content [using the Raman spectra shown in (**a**)]. Reprinted with permission from [335], © 1999 APS

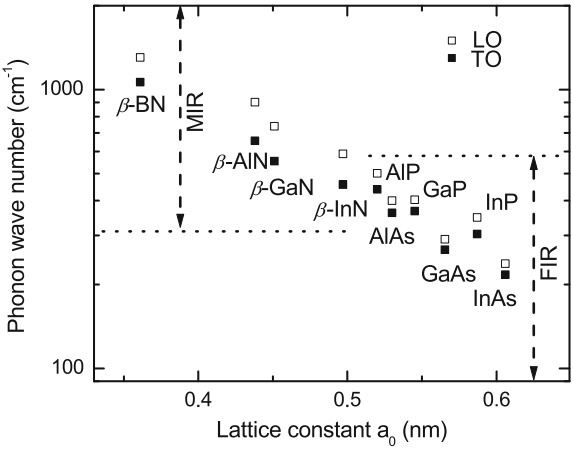

Fig. 5.11 Optical phonon frequencies (TO: *filled squares*, LO: *empty squares*) for a number of III–V compounds with different lattice constant a_0. 1 meV corresponds to 8.065 wave numbers (or cm^{-1}). Adapted from [336]

5.2.4 Density of States

The *density of states* (DOS) tells how many of the total $3pN$ modes are in a given energy interval. The states are spaced equally in **k**-space but not on the energy scale (see also Sect. 6.11).

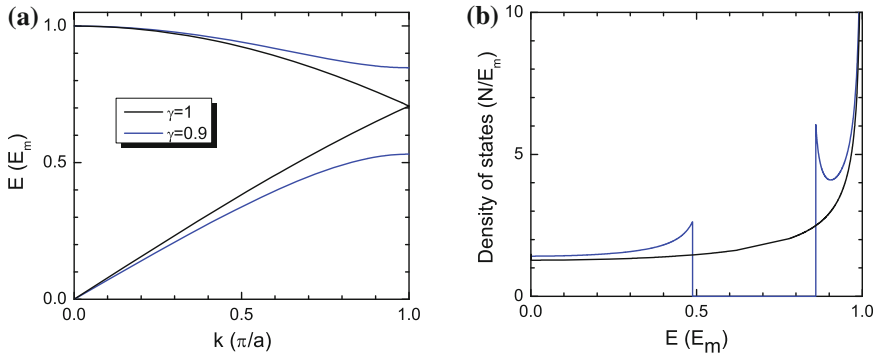

Fig. 5.12 (**a**) Phonon dispersion for the diatomic linear chain model for $\gamma = 1$ (*black line*) and $\gamma = 0.9$ (*blue lines*). (**b**) Corresponding density of states (in units of N/E_m)

For the monoatomic linear chain model, the number of states $N(E')$ from $E = 0$ up to $E = \hbar\omega = E'(k')$ for the dispersion of the acoustic phonons (5.8) is given as

$$N(E') = k' \frac{N}{\pi/a} = \frac{L}{\pi} k'. \tag{5.24}$$

Using (5.8), we find for one polarization ($E_\mathrm{m} = \hbar\omega_\mathrm{m}$)

$$N(E) = \frac{2N}{\pi} \arcsin\left(\frac{E}{E_\mathrm{m}}\right). \tag{5.25}$$

The DOS $D(E)$ is given by

$$D(E) = \frac{\mathrm{d}N(E)}{\mathrm{d}E} = \frac{2\,N}{\pi\,E_\mathrm{m}} \frac{1}{\sqrt{1-(E/E_\mathrm{m})^2}}. \tag{5.26}$$

Often the density of states is scaled by the (irrelevant) system size and given per atom (D/N) or per volume (D/L^3), per area (D/L^2) or per length (D/L) for three-, two- or one-dimensional systems, respectively.

In the diatomic linear chain model, additionally the optical phonons contribute to the density of states. In Fig. 5.12 the phonon density of states is shown for $\gamma = 0.9$ and for comparison for $\gamma = 1$ (gapless phonon dispersion). For small wavevector, the density of states is $4N/(\pi\,E_\mathrm{m})$.[1] Within the gap the density of states vanishes. At the edges of the band gap the density of states is enhanced. The total number of states for both dispersions is the same.

[1] The factor 2 compared to (5.26) stems from the folded Brillouin zone compared to the monoatomic chain model.

In a three-dimensional solid, (5.24) is modified to (for three degenerate polarizations)

$$N(E') = \frac{4\pi}{3} \frac{3}{(2\pi/L)^3} k'^3, \tag{5.27}$$

taking into account all states within a sphere in **k**-space of radius k'. Assuming a linear dispersion $\omega = v_s\, k$, we obtain

$$N(E) = \frac{V}{2\pi^2} \frac{E^3}{\hbar^3\, v_s^3}. \tag{5.28}$$

Thus the density of states is proportional to E^2,

$$D(E) = \frac{3\,V}{2\pi^2} \frac{E^2}{\hbar^3\, v_s^3}. \tag{5.29}$$

This dependence is the base for Debye's law for the T^3 temperature dependence of the heat capacity. As realistic example the phonon density of states for BN is depicted next to the dispersion in Fig. 5.7.

5.2.5 *Phonons*

Phonons are the quantized quasi-particles of the lattice vibrations (normal modes). The energy of a phonon can take the discrete values of a harmonic oscillator

$$E_{ph} = \left(n + \frac{1}{2}\right) \hbar\omega, \tag{5.30}$$

where n denotes the quantum number of the state, which corresponds to the number of energy quanta $\hbar\omega$ in the vibration. The amplitude of the vibration can be connected to n via the following discussion. For the classical oscillation $u = u_0\, \exp \mathrm{i}(kx - \omega t)$ the space and time average for the kinetic energy yields

$$E_{\mathrm{kin}} = \frac{1}{2}\, \rho\, V\, \overline{\left(\frac{\partial u}{\partial t}\right)^2} = \frac{1}{8}\, \rho\, V\, \omega^2\, u_0^2, \tag{5.31}$$

where ρ is the density and V the volume of the (homogeneous) solid. The energy of the oscillation is split in half between kinetic and potential energy. From $2E_{\mathrm{kin}} = E_{\mathrm{ph}}$ we find

$$u_0^2 = \left(n + \frac{1}{2}\right) \frac{4\hbar}{\rho\, V\, \omega}. \tag{5.32}$$

The number of phonons with which a vibrational mode is populated is thus directly related to the classical amplitude square.

Phonons act with a momentum $\hbar\mathbf{k}$, the so-called crystal momentum. When phonons are created, destroyed or scattered the crystal momentum is conserved, except for an arbitrary reciprocal-space vector $\mathbf{G}$. Scattering with $\mathbf{G} = 0$ is called a *normal* process, otherwise (for $\mathbf{G} \neq 0$) it is called an *umklapp* process.

5.2.6 Localized Vibrational Modes

A defect in the crystal can induce localized vibrational modes (LVM). The defect can be a mass defect, i.e. one of the masses M is replaced by M_d, or the force constants in the neighborhood are modified to C_d. A detailed treatment can be found in [337]. LVM are discussed, e.g., in [338–340].

First we consider the LVM for the one-dimensional, monoatomic chain. If the mass at lattice point $i = 0$ is replaced by $M_\mathrm{d} = M + \Delta M$ ($\epsilon_\mathrm{M} = \Delta M/M$), the displacements are given by $u_i = AK^{|i|}$, A being an amplitude, with

$$K = -\frac{1+\epsilon_\mathrm{M}}{1-\epsilon_\mathrm{M}}, \tag{5.33}$$

and the defect phonon frequency ω_d is

$$\omega_\mathrm{d} = \omega_\mathrm{m}\sqrt{\frac{1}{1-\epsilon_\mathrm{M}^2}}. \tag{5.34}$$

A real frequency is obtained for $|\epsilon_\mathrm{M}| < 1$. ω_d is then higher than the highest frequency of the bulk modes $\omega_\mathrm{m} = \sqrt{4C/M}$ (5.11). For $\epsilon_\mathrm{M} < 0$, i.e. the mass of the defect is smaller than the mass of the host atoms, K is negative and $|K| < 1$. Thus, the displacement can be written as

$$u_i \propto (-|K|)^{|i|} = (-1)^{|i|}\exp\left(+|i|\log|K|\right). \tag{5.35}$$

The numerical value of the exponent is negative, thus the amplitude decreases exponentially from the defect and indeed makes a localized vibrational mode. For small mass $M_\mathrm{d} \ll M$ (5.34) yields approximately $\omega_\mathrm{d} = \sqrt{2C/M_\mathrm{d}}$. This approximation is the so-called one-oscillator model. Since typically the extension of the localized mode is only a few lattice constants, the picture of LVM remains correct for impurity concentrations up to $\sim 10^{18} - 10^{20}\,\mathrm{cm}^{-3}$. For higher concentrations the concept of alloy modes has to be invoked (cf. Sect. 5.2.7).

For the case of group-III or -V substitutional impurities in group-IV semiconductors the change in force constants (treated below) can be neglected to some extent. For silicon ($M = 28$) and germanium ($M = 73$) the effect of various substitutions is shown in Fig. 5.13.

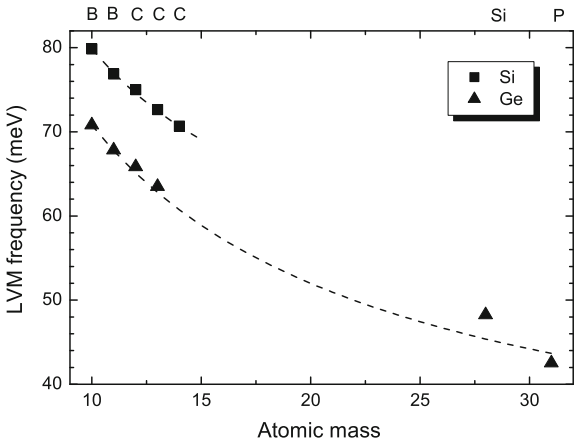

Fig. 5.13 Energy of local vibrational modes in Si and Ge. Experimental values at $T = 300\,\mathrm{K}$ (B in Ge: $T = 80\,\mathrm{K}$) taken from [337] and references therein and from [341] (C in Ge). The *dashed lines* are the mass dependence according to (5.34) scaled to the experimental frequency of the ^{10}B LVM

Now, additionally the force constants left and right of the defect are replaced by $C_\mathrm{d} = C + \Delta C$ ($\epsilon_\mathrm{C} = \Delta C/C$). The displacements are still given by $u_i = AK^{|i|}$, now with

$$K = -\frac{(1+\epsilon_\mathrm{M})\,(1+\epsilon_\mathrm{C})}{1-\epsilon_\mathrm{M}-2\epsilon_\mathrm{C}}. \tag{5.36}$$

An exponential decrease of the LVM amplitude occurs for negative K that is ensured for $\epsilon_\mathrm{M} + 2\epsilon_\mathrm{C} < 0$ (and $\epsilon_\mathrm{M} > -1$ and $\epsilon_\mathrm{C} > -1$). The defect frequency is given by

$$\omega_\mathrm{d} = \omega_\mathrm{m}\sqrt{\frac{(1+\epsilon_\mathrm{C})\,(2+\epsilon_\mathrm{C}\,(3+\epsilon_\mathrm{M}))}{2(1+\epsilon_\mathrm{M})\,(2\epsilon_\mathrm{C}+1-\epsilon_\mathrm{M})}}. \tag{5.37}$$

We note that for $\epsilon_\mathrm{C} = 0$ (5.33) and (5.34) are recovered.

For a given mass defect, the change of frequency with ΔC is (in linear order, i.e. for $\epsilon_\mathrm{C} \ll 1$)

$$\frac{\partial\omega_\mathrm{d}(\epsilon_\mathrm{M},\epsilon_\mathrm{C})}{\partial\epsilon_\mathrm{C}} = \frac{1-4\,\epsilon_\mathrm{M}-\epsilon_\mathrm{M}^2}{4\,(1-\epsilon_\mathrm{M})\sqrt{1-\epsilon_\mathrm{M}^2}}\,\epsilon_\mathrm{C}. \tag{5.38}$$

The linear coefficient diverges for $\epsilon_\mathrm{M} \to -1$. For ϵ_M between -0.968 and 0 the linear coefficient varies between 2 and 1/4. Therefore, a larger force constant ($\epsilon_\mathrm{C} > 0$) increases the LVM frequency of the defect, as expected for a stiffer spring.

In a binary compound the situation is more complicated. We assume here that the force constants remain the same and only the mass of the substitution atom M_d is different from the host. The host has the atom masses M_1 and M_2 with $M_1 < M_2$. Substitution of the heavy atom with a lighter one creates a LVM above the optical branch for $M_\mathrm{d} < M_2$. Additionally, a level in the gap between the optical and acoustic

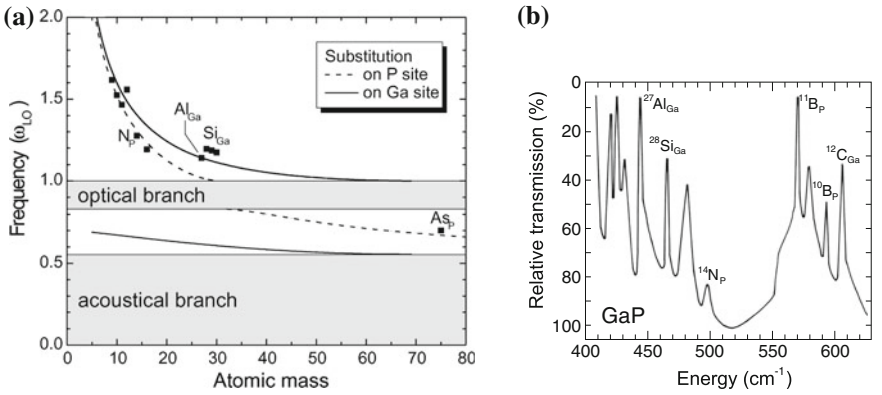

Fig. 5.14 (**a**) Numerical simulation of a linear chain model for GaP ($M_1 = 31$, $M_2 = 70$). Energy of local vibrational modes (*dashed* (*solid*) *line*): substitution on P (Ga) site in units of the optical phonon frequency at Γ ($\omega_m = 45.4$ meV). The *grey areas* indicate the acoustic and optical phonon bands. *Solid squares* are experimental data (from [337]), scaled to the theoretical curve for the $^{27}Al_{Ga}$ LVM frequency. (**b**) Differential transmission spectrum of GaP structure (nitrogen-doped layer on zinc-doped compensated substrate) against pure crystal ($T = 77$ K). Data from [342]

branch is induced. Such LVM is called a *gap mode*. Substitution of the lighter atom of the binary compound induces a LVM above the optical branch for $M_d < M_1$. A gap mode is induced for $M_d > M_1$. The situation for GaP is depicted in Fig. 5.14. LVM in GaAs have been reviewed in [338].

The energy position of a local vibrational mode is sensitive to the isotope mass of the surrounding atoms. In Fig. 5.15, a high-resolution ($0.03\,cm^{-1}$) spectrum of the $^{12}C_{As}$ LVM in GaAs is shown together with a theoretical simulation. The various theoretical peak positions are given as vertical bars, their height indicating the oscillator

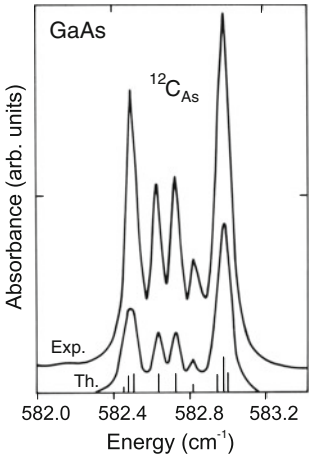

Fig. 5.15 Experimental (Exp., $T = 4.2$ K, resolution $0.03\,cm^{-1}$) and theoretical (Th., artificial Lorentzian broadening) infrared spectra of LVM of $^{12}C_{As}$ in GaAs. The positions and oscillator strengths of the theoretical transitions involving different configurations with ^{69}Ga and ^{71}Ga isotopes are shown as *vertical bars*. Data from [338]

strength. Five experimental peaks are obvious that are due to a total of nine different transitions. The C atom can experience five different surroundings (see Table 3.8) with the four neighbors being ^{69}Ga or ^{71}Ga. The natural isotope mix is an 'alloy' $^{69}Ga_x{}^{71}Ga_{1-x}As$ with $x = 0.605$. The configurations with T_d symmetry contribute one peak each, the lowest (^{71}Ga surrounding) and highest (^{69}Ga surrounding) energy transitions. The configurations with C_{3v} and C_{2v} symmetry contribute each with 2 and 3 nondegenerate modes, respectively.

The vibrations of impurity complexes have been discussed in [343].

5.2.7 *Phonons in Alloys*

In an alloy of the type $AB_{1-x}C_x$ the phonon frequencies will depend on the ternary composition. For the binary end materials AB and AC clearly TO and LO frequencies exist. The simplest behavior of the alloy is the one-mode behavior (Fig. 5.16d) where the mode frequencies vary continuously (and approximately linearly) with the composition. The oscillator strength (LO–TO splitting, (9.74)) remains approximately constant. In many cases, the two-mode behavior is observed where the LO–TO gap closes (accompanied by decreasing oscillator strength) and a localized vibrational mode and a gap mode occur for the binary end materials (Fig. 5.16a). Also, a mixed-mode behavior (Fig. 5.16b, c) can occur.

The masses of the three constituent atoms will be M_A, M_B, and M_C. Without limiting the generality of our treatment, we assume $M_B < M_C$. From the considerations in Sect. 5.2.6 on LVM and gap modes, the condition

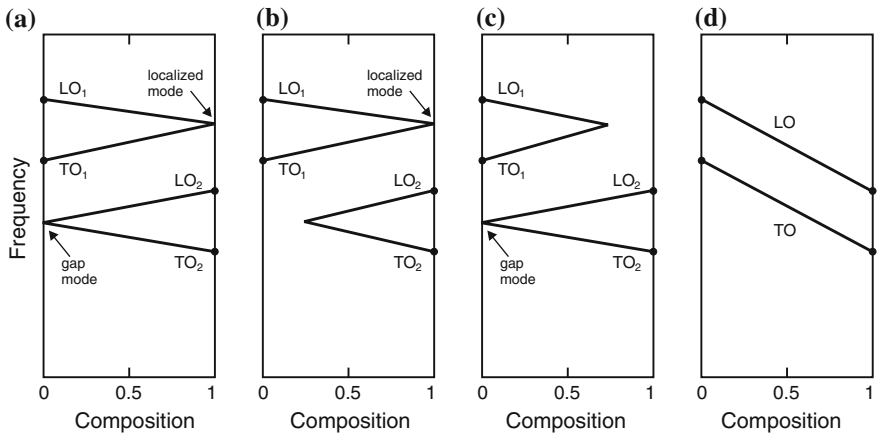

Fig. 5.16 Schematic behavior of phonon modes in an alloy. (**a**) Two-mode behavior with gap mode and localized mode, (**b**, **c**) mixed-mode behavior, (**b**) only localized mode allowed, (**c**) only gap mode allowed, (**d**) one-mode behavior with neither localized mode nor gap mode allowed

Table 5.1 Atomic masses of the constituents of various ternary compounds, reduced mass μ_{AC} (5.40), fulfillment of the relation from (5.40) ('+': fulfilled, '−': not fulfilled) and experimental mode behavior ('2': two-mode, '1': one-mode)

Alloy	A	B	C	M_A	M_B	M_C	μ_{AC}	Rel.	Modes
$GaP_{1-x}As_x$	Ga	P	As	69.7	31.0	74.9	36.1	+	2
$GaAs_{1-x}Sb_x$	Ga	As	Sb	69.7	74.9	121.8	44.3	−	1
$CdS_{1-x}Se_x$	Cd	S	Se	112.4	32.1	79.0	46.4	+	2
$Cd_xZn_{1-x}S$	S	Zn	Cd	32.1	65.4	112.4	25.0	−	1
$Mg_xZn_{1-x}O$	O	Mg	Zn	16.0	24.3	65.4	12.9	−	1

$$M_B < M_A, M_C \tag{5.39}$$

for two-mode behavior can be deduced. This ensures a LVM of atom B in the compound AC and a gap mode of atom C in the compound AB. However, it turns out that this condition is not sufficient, e.g. $Na_{1-x}K_xCl$ fulfills (5.39) but exhibits one-mode behavior. From a modified REI[2] model (for $\mathbf{k} \sim 0$ modes) it has been deduced that

$$M_B < \mu_{AC} = \frac{M_A\, M_C}{M_A + M_C} < M_A, M_C \tag{5.40}$$

is a necessary and sufficient condition (unless the force constants between A–B and A–C are significantly different) for two-mode behavior [344]. A detailed discussion is given in [345]. Equation (5.40) is a stronger condition than the previous one (5.39). If (5.40) is not fulfilled the compound exhibits one-mode behavior. As an example, we show the mass relations for $CdS_{1-x}Se_x$ and $Cd_xZn_{1-x}S$ in Table 5.1 and the experimental phonon energies in Fig. 5.17. Also in Table 5.1 the masses for $GaP_{1-x}As_x$ ($GaAs_{1-x}Sb_x$) exhibiting two- (one-) mode behavior are shown.

If the binary end components of a ternary alloy have different crystal structure, a transition between the two occurs which is reflected in the phonon structure (energies and mode symmetries). As an example, the optical phonon energies of $Mg_xZn_{1-x}O$ are depicted in Fig. 5.18 (cmp. Fig. 3.39).

5.2.8 *Disorder*

An example of 'small' disorder are the localized vibrational modes due to a single defect. Here we consider in our one-dimensional model random fluctuations of the model parameters. To that avail we set up a numerical implementation of an one-dimensional chain with masses $M_1 = M_2$ and spring constants $C_1 \neq C_2$, here $C_2 = 2\,C_1$. Now each spring constant varies randomly by $\pm\xi$ percent. The density

[2] Random element isodisplacement.

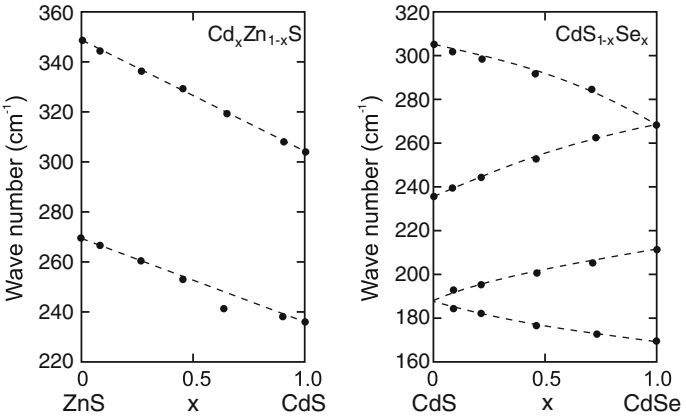

Fig. 5.17 Phonon energies of $Cd_xZn_{1-x}S$ and $CdS_{1-x}Se_x$ as a function of the ternary composition. Experimental data (*solid circles*) are from [344], *dashed lines* are guides to the eye

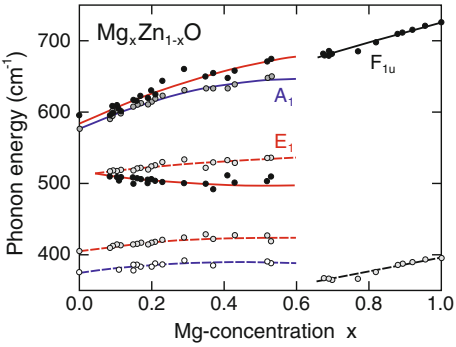

Fig. 5.18 LO (*solid lines*) and TO (*dashed lines*) phonon energies of $Mg_xZn_{1-x}O$ in the wurtzite structure (A_1 symmetry: *blue lines*, E_1 symmetry: *red lines*) and in the rocksalt phase (F_{1u} symmetry: *black lines*). Experimental data are shown as *symbols*. Adapted from data of [346]

of states is displayed for $\xi = 10, 20, 40$ and 60 %. The effects as shown in Fig. 5.19 are broadening of peaks in the DOS, broadening of the band edges, the development of band tails into the gap and eventually a closing of the gap. This is a typical behavior that also exists for electronic states (cmp. Fig. 6.46).

5.2.9 *Electric Field Created by Optical Phonons*

Adjacent atoms oscillate with opposite phase in an optical phonon. If the bond has (partial) ionic character, this leads to a time-dependent polarization and subsequently to a macroscopic electric field. This additional field will influence the phonon frequencies obtained from a purely mechanical approach. We consider in the following the case $\mathbf{k} \approx 0$. The phonon frequency for TO and LO vibrations is given by

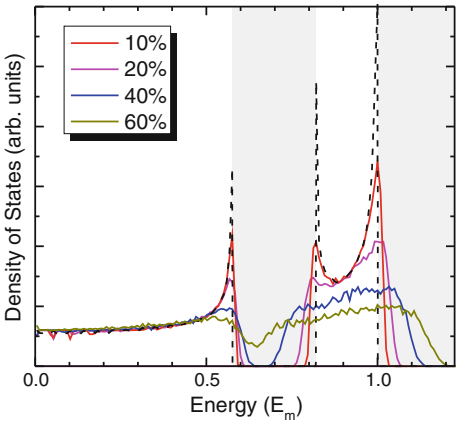

Fig. 5.19 Density of states versus energy (in units of maximum phonon energy $E_\mathrm{m} = \hbar\omega_\mathrm{m}$) for diatomic linear chain model ($M_1 = M_2$, $C_2 = 2\,C_1$, 2^9 atoms, average over 2^7 configurations) for various levels of random relative variations of the spring constants (*solid lines*). As *dashed lines* the density of states of the perfect chain is depicted (cmp. Fig. 5.12b) with the forbidden energy ranges shown in *light grey*

$$\omega_0 = \sqrt{\frac{2\,C}{M_\mathrm{r}}}, \tag{5.41}$$

where M_r is the reduced mass of the two different atoms (cf. Sect. 5.2.2). $\mathbf{u}$ is the relative displacement $\mathbf{u}_1 - \mathbf{u}_2$ of the two atoms in a diatomic base. When the interaction with the electric field $\mathbf{E}$ (which will be calculated self-consistently in the following) is considered, the Hamiltonian for the long-wavelength limit is given by [347]:

$$\hat{H}(\mathbf{p}, \mathbf{u}) = \frac{1}{2}\left(\frac{1}{M_\mathrm{r}}\,\mathbf{p}^2 + b_{11}\,\mathbf{u}^2 + 2b_{12}\,\mathbf{u}\cdot\mathbf{E} + b_{22}\,\mathbf{E}^2\right). \tag{5.42}$$

The first term is the kinetic energy ($\mathbf{p}$ stands for the momentum of the relative motion of the atoms 1 and 2 in the base, $\mathbf{p} = M_\mathrm{r}\dot{\mathbf{u}}$), the second the potential energy, the third the dipole interaction and the fourth the electric-field energy. The equation of motion for a plane wave $\mathbf{u} = \mathbf{u}_0 \exp\left[-\mathrm{i}(\omega t - \mathbf{k}\cdot\mathbf{r})\right]$ ($\ddot{\mathbf{u}} = -\omega^2\mathbf{u}$) yields

$$M_\mathrm{r}\,\omega^2\,\mathbf{u} = b_{11}\,\mathbf{u} + b_{12}\,\mathbf{E}. \tag{5.43}$$

Thus, the electric field is

$$\mathbf{E} = (\omega^2 - \omega_\mathrm{TO}^2)\,\frac{M_\mathrm{r}}{b_{12}}\,\mathbf{u}. \tag{5.44}$$

Here, the substitution $\omega_\mathrm{TO}^2 = b_{11}/M_\mathrm{r}$ was introduced that is consistent with (5.41) and $b_{11} = 2C$. ω_TO represents the mechanical oscillation frequency of the atoms

undisturbed by any electromagnetic effects. Already now the important point is visible. If ω approaches $\omega_{\rm TO}$, the system plus electric field oscillates with the frequency it has without an electric field. Therefore the electric field must be zero. Since the polarization $\mathbf{P} = (\epsilon - 1)\epsilon_0 \mathbf{E}$ is finite, the dielectric constant ϵ thus diverges.

The polarization is

$$\mathbf{P} = -\nabla_{\mathbf{E}} \hat{H} = -\left(b_{12}\,\mathbf{u} + b_{22}\,\mathbf{E}\right). \tag{5.45}$$

The displacement field is

$$\mathbf{D} = \epsilon_0\,\mathbf{E} + \mathbf{P} = \epsilon_0\,\mathbf{E} - \left(b_{22} - \frac{b_{12}^2/M_{\rm r}}{\omega_{\rm TO}^2 - \omega^2}\right)\mathbf{E} = \epsilon_0\,\epsilon(\omega)\,\mathbf{E}. \tag{5.46}$$

Therefore, the dielectric function is

$$\epsilon(\omega) = \epsilon(\infty) + \frac{\epsilon(0) - \epsilon(\infty)}{1 - (\omega/\omega_{\rm TO})^2}. \tag{5.47}$$

Here, $\epsilon(\infty) = 1 - b_{22}/\epsilon_0$ is the high-frequency dielectric constant and $\epsilon(0) = \epsilon(\infty) + b_{12}^2/(b_{11}\epsilon_0)$ the static dielectric constant. The relation (5.47) is shown in Fig. 5.20.

From the Maxwell equation $\nabla \cdot \mathbf{D} = 0$ for zero free charge we obtain the relation

$$\epsilon_0\,\epsilon(\omega)\,\nabla \cdot \mathbf{E} = 0. \tag{5.48}$$

Thus, either $\epsilon(\omega) = 0$ or $\nabla \cdot \mathbf{E} = 0$, i.e. $\mathbf{u}$ is perpendicular to $\mathbf{k}$. In the latter case we have a TO phonon and, neglecting retardation effects, using $\nabla \times \mathbf{E} = 0$ we find

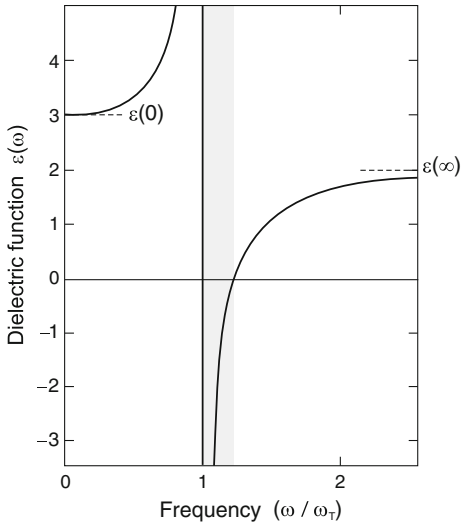

Fig. 5.20 Dielectric function according to (5.47) with $\epsilon(0) = 3$ and $\epsilon(\infty) = 2$ (without damping). *Grey area* denotes the region of negative ϵ

$\mathbf{E} = 0$ and therefore $\omega = \omega_{\mathrm{TO}}$, justifying our notation. In the case of $\epsilon(\omega) = 0$, we call the related frequency ω_{LO} and find the so-called Lyddane–Sachs–Teller (LST) relation

$$\frac{\omega_{\mathrm{LO}}^2}{\omega_{\mathrm{TO}}^2} = \frac{\epsilon(0)}{\epsilon(\infty)}. \tag{5.49}$$

This relation holds reasonably well for optically isotropic, heteropolar materials with two atoms in the basis, such as NaI and also GaAs. Since at high frequencies, i.e. $\omega \gg \omega_{\mathrm{TO}}$, only the individual atoms can be polarized, while for low frequencies the atoms can also be polarized against each other, $\epsilon(0) > \epsilon(\infty)$ and therefore also $\omega_{\mathrm{LO}} > \omega_{\mathrm{TO}}$. For GaAs, the quotient of the two phonon energies is 1.07. Using the LST relation (5.49), we can write for the dielectric function

$$\epsilon(\omega) = \epsilon(\infty) \left(\frac{\omega_{\mathrm{LO}}^2 - \omega^2}{\omega_{\mathrm{TO}}^2 - \omega^2} \right). \tag{5.50}$$

The (long-wavelength) TO-phonon does not create a long-range electric field. Using $\nabla \cdot \mathbf{D} = 0$ and (5.46) and looking at the longitudinal fields, we have

$$\epsilon_0 \, \mathbf{E} = b_{12} \, \mathbf{u} + b_{22} \, \mathbf{E}. \tag{5.51}$$

This can be rewritten as

$$\mathbf{E} = -\omega_{\mathrm{LO}} \sqrt{\frac{M_{\mathrm{r}}}{\epsilon_0}} \sqrt{\frac{1}{\epsilon(\infty)} - \frac{1}{\epsilon(0)}} \, \mathbf{u} \propto -\mathbf{u}. \tag{5.52}$$

The (long-wavelength) LO-phonon thus creates a long-range electric field acting *against* the ion displacement and represents an additional restoring force; this is consistent with the fact that $\omega_{\mathrm{LO}} > \omega_{\mathrm{TO}}$.

5.3 Elasticity

The elastic properties of the semiconductor are important if the semiconductor is subjected to external forces (pressure, temperature) or to lattice mismatch during heteroepitaxy.

5.3.1 Thermal Expansion

The lattice constant depends on temperature. The (linear) thermal expansion coefficient is defined as

$$\alpha(T_0) = \left. \frac{\partial a_0(T)}{\partial T} \right|_{T=T_0} \tag{5.53}$$

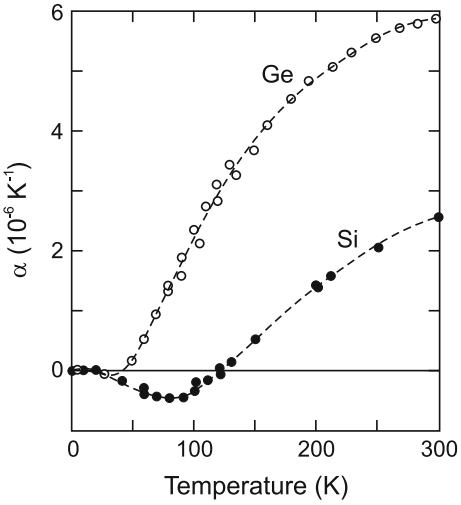

Fig. 5.21 Linear thermal expansion coefficient of silicon (*solid circles*) and germanium (*open circles*). Adapted from [348] based on experimental data from various sources. *Dashed lines* are guide to the eyes

and is temperature dependent. The temperature dependence of α for silicon and germanium is shown in Fig. 5.21. α is approximately proportional to the heat capacity (C_V) except at low temperatures. The negative values are due to negative Grüneisen parameters [348]. These anharmonicity effects are discussed in detail in [328].

5.3.2 Stress–Strain Relation

In this section, we recall the classical theory of elasticity [349]. The solid is treated as a continuous medium (piecewise homogeneous) and the displacement vector is thus a continuous function $\mathbf{u}(\mathbf{r})$ of the spatial coordinates. When the spatial variation $\nabla\mathbf{u}$ of $\mathbf{u}$ is small, the elastic energy can be written as

$$U = \frac{1}{2}\int \frac{\partial u_l}{\partial x_k}\, C_{klmn}\, \frac{\partial u_n}{\partial x_m}\, \mathrm{d}^3\mathbf{r}, \tag{5.54}$$

where $\mathbf{C}$ is the (macroscopic) tensor of the elastic coefficients. 21 components of this tensor can be independent. For crystals with cubic symmetry the number of independent constants is reduced to 3. An exchange $k \leftrightarrow l$ and $m \leftrightarrow n$ does not matter, only six indices have to be considered (xx, yy, zz, yz, xz, and xy). The strain components ϵ_{ij} are symmetrized according to

$$\epsilon_{ij} = \frac{1}{2}\left(\frac{\partial u_j}{\partial x_i} + \frac{\partial u_i}{\partial x_j}\right). \tag{5.55}$$

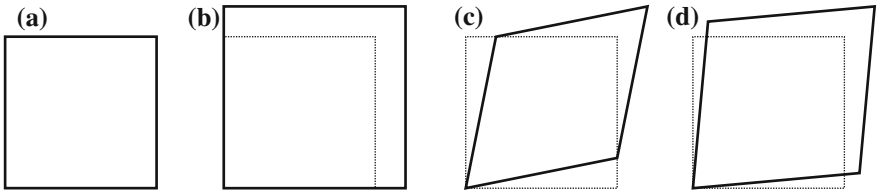

Fig. 5.22 Deformation of a square (**a**). (**b**) Pure hydrostatic deformation ($\epsilon_{xx} = \epsilon_{yy} = 0.2$, $\epsilon_{xy} = 0$), (**c**) pure shear deformation ($\epsilon_{xx} = \epsilon_{yy} = 0$, $\epsilon_{xy} = 0.2$), and (**d**) mixed deformation ($\epsilon_{xx} = \epsilon_{yy} = 0.1$, $\epsilon_{xy} = 0.1$)

The strains ϵ_{xx} are along the main axes of the crystal as visualized in Fig. 5.22. The stresses[3] σ_{kl} are then given by

$$\sigma_{kl} = C_{klmn}\,\epsilon_{mn}. \tag{5.56}$$

The inverse relation is mediated by the stiffness tensor **S**.

$$\epsilon_{kl} = S_{klmn}\,\sigma_{mn}. \tag{5.57}$$

Typically, the strain components e_{ij} or e_i are used with the convention $xx \to 1$, $yy \to 2$, $zz \to 3$, $yz \to 4$, $xz \to 5$, and $xy \to 6$ (Voigt notation):

$$e_{ij} = \epsilon_{ij}\,(2 - \delta_{ij}). \tag{5.58}$$

Then, $\sigma_m = C_{mn}e_n$ with the C_{ij} being the elastic constants. The x, y, and z directions are the main axes of the cubic solid, i.e. the $\langle 100 \rangle$ directions.

For *zincblende* material, the stress–strain relation reads[4]

$$\begin{pmatrix} \sigma_1 \\ \sigma_2 \\ \sigma_3 \\ \sigma_4 \\ \sigma_5 \\ \sigma_6 \end{pmatrix} = \begin{pmatrix} C_{11} & C_{12} & C_{12} & 0 & 0 & 0 \\ C_{12} & C_{11} & C_{12} & 0 & 0 & 0 \\ C_{12} & C_{12} & C_{11} & 0 & 0 & 0 \\ 0 & 0 & 0 & C_{44} & 0 & 0 \\ 0 & 0 & 0 & 0 & C_{44} & 0 \\ 0 & 0 & 0 & 0 & 0 & C_{44} \end{pmatrix} \begin{pmatrix} e_1 \\ e_2 \\ e_3 \\ e_4 \\ e_5 \\ e_6 \end{pmatrix}. \tag{5.59}$$

Values of the compliances for several semiconductors are given in Table. 5.2. The inverse relation is given by the matrix

[3] The stress is a force per unit area and has the dimensions of a pressure.

[4] $C_{11} = C_{1111}$, $C_{12} = C_{1122}$ and $C_{44} = C_{1212} = C_{1221} = C_{2121} = C_{2112}$.

Table 5.2 Elastic constants (in GPa) of some cubic semiconductors at room temperature. I_K refers to the Keating criterion (5.63)

Material	C_{11}	C_{12}	C_{44}	I_K
C	1076.4	125.2	577.4	1.005
Si	165.8	63.9	79.6	1.004
Ge	128.5	48.3	66.8	1.08
BN	820	190	480	1.11
GaAs	119	53.4	59.6	1.12
InAs	83.3	45.3	39.6	1.22
AlAs	120.5	46.86	59.4	1.03
ZnS	104.6	65.3	46.3	1.33
MgO	297	156	95.3	0.80

$$\begin{pmatrix} S_{11} & S_{12} & S_{12} & 0 & 0 & 0 \\ S_{12} & S_{11} & S_{12} & 0 & 0 & 0 \\ S_{12} & S_{12} & S_{11} & 0 & 0 & 0 \\ 0 & 0 & 0 & S_{44} & 0 & 0 \\ 0 & 0 & 0 & 0 & S_{44} & 0 \\ 0 & 0 & 0 & 0 & 0 & S_{44} \end{pmatrix}, \tag{5.60}$$

with the stiffness coefficients in this notation given by

$$S_{11} = \frac{C_{11} + C_{12}}{(C_{11} - C_{12})\,(C_{11} + 2C_{12})} \tag{5.61a}$$

$$S_{12} = \frac{C_{12}}{-C_{11}^2 - C_{11}C_{12} + 2C_{12}^2} \tag{5.61b}$$

$$S_{44} = \frac{1}{C_{44}}. \tag{5.61c}$$

We emphasize that in this convention (also called the *engineering* convention), e.g. $e_1 = \epsilon_{xx}$ and $e_4 = 2\,\epsilon_{yz}$. There is also another convention (the *physical* convention) without this factor of two; in this case the matrix in (5.59) contains the elements $2C_{44}$. We introduce

$$C_0 = 2\,C_{44} + C_{12} - C_{11}, \tag{5.62}$$

and note that $C_0 = 0$ for an isotropic material. The relation

$$I_K = \frac{2\,C_{44}\,(C_{11} + C_{12})}{(C_{11} - C_{12})\,(C_{11} + 3C_{12})} = 1 \tag{5.63}$$

known as the Keating criterion [350, 351], stems from the consideration of bending and stretching of the tetrahedral bonds in the valence force field (VFF) model. It is closely fulfilled (Table 5.2) for many tetrahedrally bonded semiconductors, in particular for the covalent ones. For MgO, the Keating criterion is not fulfilled because it has (six-fold coordinated) rocksalt structure and is thus not tetrahedrally bonded.

The Young's modulus Y,

$$\sigma_{\mathbf{nn}} = Y(\mathbf{n})\,\epsilon_{\mathbf{nn}}, \tag{5.64}$$

generally depends on the normal direction $\mathbf{n}$ of a strain. It is equivalent to $1/S_{11}$ of (5.61a).

For *isotropic* material Y and the Poisson ratio ν are related to the elastic constants of cubic material by

$$Y = C_{11} - \frac{2\,C_{12}^2}{C_{11} + C_{12}} \tag{5.65a}$$

$$\nu = \frac{C_{12}}{C_{11} + C_{12}}. \tag{5.65b}$$

For isotropic materials also Lamé's constants λ and μ are used. They are given by[5] $C_{11} = \lambda + 2\mu$, $C_{12} = \lambda$ and $C_{44} = \mu$ (note that C_0 according to (5.62) is zero).

The bulk modulus B (inverse of the compressibility),

$$\frac{1}{B} = -\frac{1}{V}\frac{\partial V}{\partial p}, \tag{5.66}$$

for the zincblende crystal is given as

$$B = \frac{C_{11} + 2\,C_{12}}{3}. \tag{5.67}$$

We note that Y, ν and C_{ij} of typical materials are both positive. Materials with negative Poisson ratio are called *auxetic* [352–354]. Also materials with negative compressibility are possible [355].

Beyond the dependence of the elastic constants on the bond length (as materialized in the phonon frequencies in Fig. 5.11), they depend on the ionicity. In Fig. 5.23, the elastic constants of various zincblende semiconductors are shown as a function of the ionicity f_i. The values for the elastic constants are normalized by e^2/a^4, a being the average nearest-neighbor distance.

For *wurtzite* crystals, five elastic constant are necessary for the stress–strain relation that reads[6]

$$C_{ij} = \begin{pmatrix} C_{11} & C_{12} & C_{13} & 0 & 0 & 0 \\ C_{12} & C_{11} & C_{13} & 0 & 0 & 0 \\ C_{13} & C_{13} & C_{33} & 0 & 0 & 0 \\ 0 & 0 & 0 & C_{44} & 0 & 0 \\ 0 & 0 & 0 & 0 & C_{44} & 0 \\ 0 & 0 & 0 & 0 & 0 & \frac{1}{2}(C_{11} - C_{12}) \end{pmatrix}. \tag{5.68}$$

[5] For an isotropic material, $C_{ijkl} = \lambda\,\delta_{ij}\,\delta_{kl} + \mu\,(\delta_{ik}\,\delta_{jl} + \delta_{il}\,\delta_{jk})$.

[6] $(C_{11} - C_{12})/2 = C_{1212}$, $C_{44} = C_{1313} = C_{2323}$.

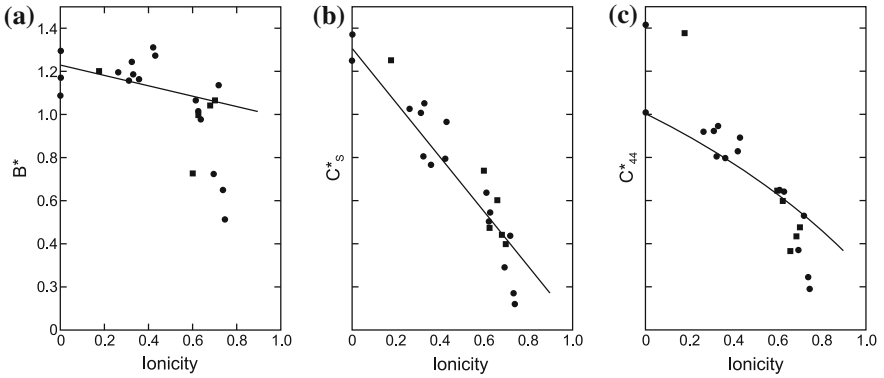

Fig. 5.23 Elastic constants as a function of ionicity for various semiconductors with diamond or zincblende (*circles*) and wurtzite (*squares*) structure. Constants are normalized by the modulus $C_0 = e^2/d^4$, d being the average nearest-neighbor distance. (**a**) Bulk modulus, $B^* = (C_{11} + 2C_{12})/(3C_0)$, (**b**, **c**) shear moduli, (**b**) $C_S^* = (C_{11} - C_{12})/C_0$, (**c**) $C_{44}^* = C_{44}/C_0$. *Solid lines* are a simple model as discussed in [356]. Adapted from [357]

Table 5.3 Elastic constants (in GPa) of some wurtzite semiconductors

Material	C_{11}	C_{12}	C_{13}	C_{33}	C_{44}	Reference
GaN	391	143	108	399	103	[359]
AlN	410	149	99	389	125	[360]
ZnS	124	60.2	45.5	140	28.6	[361]
ZnO	206	118	118	211	44	[362]

Experimental values for wurtzite materials are given in Table 5.3. The relation of the elastic tensor of wurtzite and zincblende materials, in particular viewed along the ⟨111⟩-direction has been discussed in [357, 358].

The bulk modulus of the wurtzite crystal is given by

$$B = \frac{(C_{11} + C_{12})\, C_{33} - 2\, C_{13}^2}{C_{11} + C_{12} + 2\, C_{33} - 4\, C_{13}}. \tag{5.69}$$

5.3.3 *Biaxial Strain*

In heteroepitaxy (cf. Sect. 12.2.6), a biaxial strain situation arises, i.e. layered material is compressed (or expanded in the case of tensile strain) in the interface plane and is expanded (compressed) in the perpendicular direction. Here, we assume that the substrate is infinitely thick, i.e. that the interface remains planar. Substrate bending is discussed in Sect. 5.3.5.

The simplest case is epitaxy on the (001) surface, i.e. $e_1 = e_2 = \epsilon_\parallel$. The component e_3 is found from the condition $\sigma_3 = 0$ (no forces in the z direction). All shear strains are zero. For zincblende material it follows

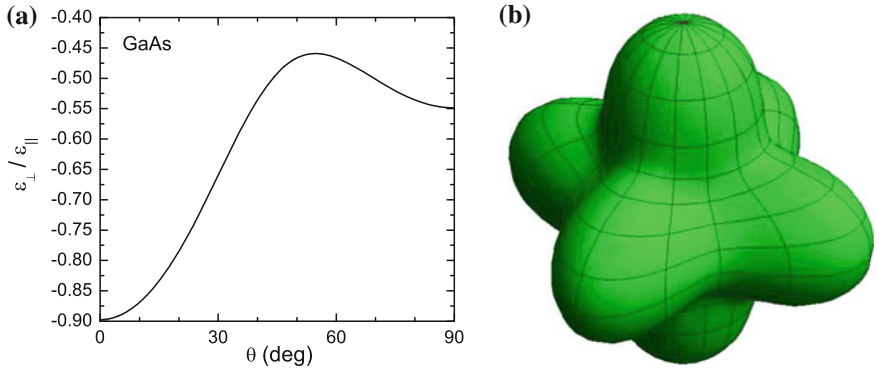

Fig. 5.24 Ratio $-\epsilon_\perp/\epsilon_\parallel$ for GaAs under symmetric biaxial strain. The angle θ denotes the surface normal in the $\langle 110\rangle$-azimuth ($\theta = 0$: [001], $\theta = 90°$: [110], the maximum of $\epsilon_\perp/\epsilon_\parallel$ is for [111]). (**b**) Is a three-dimensional visualization

$$\epsilon_\perp^{100} = e_3 = -\frac{C_{12}}{C_{11}}\,(e_1 + e_2) = -\frac{2\,C_{12}}{C_{11}}\,\epsilon_\parallel. \tag{5.70}$$

In Fig. 5.24 the ratio $\epsilon_\perp/\epsilon_\parallel$ is depicted for GaAs and various crystal orientations; the formulas for other orientations are more involved [363]:

$$\epsilon_\perp^{110} = -\frac{2C_{12} - C_0/2}{C_{11} + C_0/2}\,\epsilon_\parallel \tag{5.71}$$

$$\epsilon_\perp^{111} = -\frac{2C_{12} - 2C_0/3}{C_{11} + 2C_0/3}\,\epsilon_\parallel. \tag{5.72}$$

For wurtzite crystals and pseudomorphic growth along [00.1] the strain along the epitaxial direction (c-axis) is given by

$$\epsilon_\perp = -\frac{C_{13}}{C_{33}}\,(e_1 + e_2) = -\frac{2C_{13}}{C_{33}}\,\epsilon_\mathrm{a}, \tag{5.73}$$

where $\epsilon_\perp = \epsilon_\mathrm{c} = (c - c_0)/c_0$ and $\epsilon_\mathrm{a} = (a - a_0)/a_0$. For symmetric biaxial in-plane strain, the ratio $\epsilon_\perp/\epsilon_\parallel$ is shown in Fig. 5.25 for GaN and varying angle θ of the c-axis against the epitaxy direction. For the growth of wurtzite on wurtzite for $\theta \neq 0$, the epitaxial strain is actually *asymmetric* in the interface plane. For $\theta = 90°$, e.g. the epitaxy on m-plane substrate (cmp. Fig. 3.33) (c-axis lies in-plane), the in-plane strains are $e_1 = \epsilon_\mathrm{a}$ and $e_2 = \epsilon_\mathrm{c}$. For $\theta = 90°$, we find

$$\epsilon_\perp = -\frac{C_{12}\,\epsilon_\mathrm{a} + C_{13}\,\epsilon_\mathrm{c}}{C_{11}}. \tag{5.74}$$

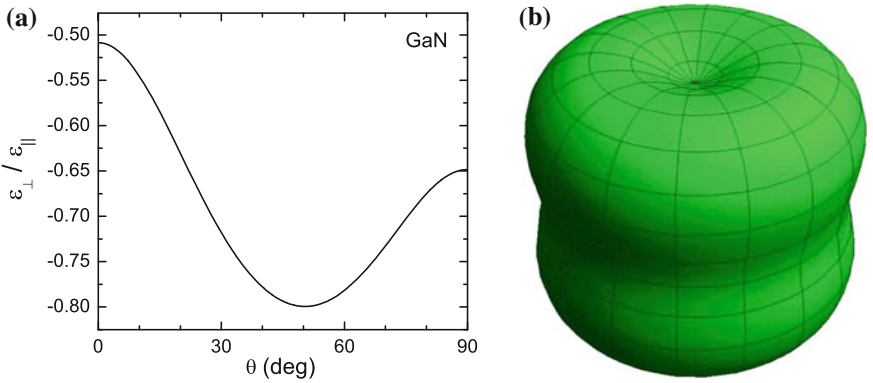

Fig. 5.25 Ratio $-\epsilon_\perp/\epsilon_\parallel$ for GaN under symmetric biaxial strain. In (**a**) θ denotes the angle of the c-axis with respect to the surface normal, (**b**) is a three-dimensional visualization, showing the in-plane isotropy

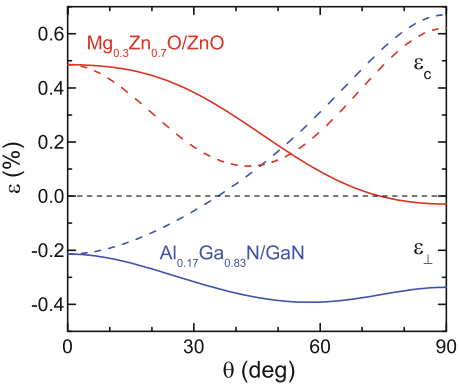

Fig. 5.26 Strains ϵ_c (*dashed lines*) and $\epsilon_\perp$ (*solid lines*) for $Al_{0.17}Ga_{0.83}N$/GaN (*blue*) and $Mg_{0.3}Ga_{0.7}O$/ZnO (*red*) as a function of the interface tilt angle θ with respect to [00.1]

The situation for pseudomorphic growth in the (Al,Ga,In)N system has been discussed for various interface orientations in [364] (cmp. also Fig. 15.14). The strains $\epsilon_\perp$ along the epitaxy direction and ϵ_c along the *c*-direction are depicted for $Al_{0.17}Ga_{0.83}N$/GaN and $Mg_{0.3}Ga_{0.7}O$/ZnO in Fig. 5.26. The different behavior of the nitride and the oxide system, e.g. regarding the sign change of ϵ_c, is due to the fact that ϵ_a is negative (positive) for $Al_xGa_{1-x}N$/GaN ($Mg_xGa_{1-x}O$/ZnO) ($\epsilon_c < 0$ for both cases) [365].

5.3.4 Three-Dimensional Strain

The strain distribution in two-dimensional or three-dimensional objects such as quantum wires and dots (see also Sect. 14) is more complicated.

A simple analytical solution for the problem of a strained inclusion is only possible for isotropic material parameters [366].

The solution for a sphere can be extended to yield the strain distribution of an inclusion of arbitrary shape. This scheme applies only for isotropic materials and identical elastic properties of the inclusion and the surrounding matrix. The solution will be given in terms of a surface integral of the boundary of the inclusion, which is fairly easy to handle. Several disconnected inclusions can be treated by a sequence of surface integrals.

The strain distribution for the inner and outer parts of a sphere with radius ρ_0 is given (in spherical coordinates) by

$$\epsilon_{\rho\rho}^{\text{in}} = \frac{2}{3}\,\epsilon_0\,\frac{1-2\nu}{1-\nu} = \epsilon_{\theta\theta}^{\text{in}} = \epsilon_{\phi\phi}^{\text{in}} \tag{5.75}$$

$$\epsilon_{\rho\rho}^{\text{out}} = \frac{2}{3}\,\epsilon_0\,\frac{1+\nu}{1-\nu}\left(\frac{\rho_0}{\rho}\right)^3 = -2\epsilon_{\theta\theta}^{\text{out}} = -2\epsilon_{\phi\phi}^{\text{out}}, \tag{5.76}$$

where ρ denotes the radius, ν the Poisson ratio, and ϵ_0 the relative lattice mismatch of the inclusion and the matrix. The radial displacements are

$$u_\rho^{\text{in}} = \frac{2}{3}\,\epsilon_0\,\frac{1-2\nu}{1-\nu}\,\rho \tag{5.77}$$

$$u_\rho^{\text{out}} = \frac{2}{3}\,\epsilon_0\,\frac{1-2\nu}{1-\nu}\,\rho_0^3\,\frac{1}{\rho^2}. \tag{5.78}$$

Dividing the displacement by the sphere's volume, we obtain the displacement per unit volume of the inclusion. From the displacement we can derive the stress σ_{ij}^0 per unit volume.

$$\sigma_{ii}^0 = \frac{1}{4\pi}\,\frac{Y\,\epsilon_0}{1-\nu}\,\frac{2x_i^2 - x_j - x_k}{\rho^5} \tag{5.79}$$

$$\sigma_{ij}^0 = \frac{3}{2}\,\frac{1}{4\pi}\,\frac{Y\,\epsilon_0}{1-\nu}\,\frac{x_i\,x_j}{\rho^5}, \tag{5.80}$$

where i, j and k are pairwise unequal indices. Due to the linear superposition of stresses, the stress distribution σ_{ij}^V for the arbitrary inclusion of volume V can be obtained by integrating over V

$$\sigma_{ij}^V = \int_V \sigma_{ij}^0\,(\mathbf{r} - \mathbf{r_0})\,\mathrm{d}^3\mathbf{r}. \tag{5.81}$$

The strains can be calculated from the stresses.

When ϵ_0 is constant within V, the volume integral can be readily transformed into an integral over the surface ∂V of V using Gauss' theorem. With the 'vector potentials' $\mathbf{A}_{ij}$ we fulfill $\mathrm{div}\mathbf{A}_{ij} = \sigma_{ij}$.

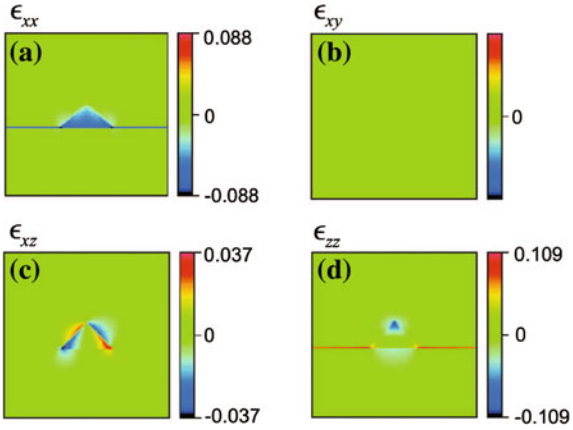

Fig. 5.27 Strain components in an InAs pyramid (quantum dot with {101} faces), embedded in GaAs. The cross section is through the center of the pyramid. The lattice mismatch between InAs and GaAs amounts to $\approx -7\,\%$. Reprinted with permission from [367], © 1995 APS

$$\mathbf{A}_{ii} = -\frac{1}{4\pi}\,\frac{Y\,\epsilon_0}{1-\nu}\,\frac{x_i \mathbf{e}_i}{\rho^3} \tag{5.82}$$

$$\mathbf{A}_{ij} = -\frac{1}{2}\,\frac{1}{4\pi}\,\frac{Y\,\epsilon_0}{1-\nu}\,\frac{x_i\,\mathbf{e}_j + x_j\,\mathbf{e}_i}{\rho^3}. \tag{5.83}$$

Equation (5.83) is valid for the case $i \neq j$. $\mathbf{e}_i$ is the unit vector in the ith direction. However, special care must be taken at the singularity $\mathbf{r} = \mathbf{r}_0$ if $\mathbf{r}_0$ lies within V because the stress within the 'δ-inclusion' is not singular (in contrast to the electrostatic analog of a δ-charge). Thus, we find

$$\sigma_{ij}^V(\mathbf{r}_0) = \oint_{\partial V} \mathbf{A}_{ij}\,\mathrm{d}\mathbf{S} + \delta_{ij}\,\frac{Y\,\epsilon_0}{1-\nu}\int_V \delta(\mathbf{r}-\mathbf{r}_0)\,\mathrm{d}^3\mathbf{r}. \tag{5.84}$$

As an example, we show in Fig. 5.27 the numerically calculated strain components [367] (taking into account the different elastic properties of the dot and matrix materials) in the cross section of a pyramidal InAs quantum dot in a GaAs matrix on top of a two-dimensional InAs layer. The strain component ϵ_{zz} is positive in the 2D layer, as expected from (5.70). However, in the pyramid ϵ_{zz} exhibits a complicated dependence and even takes negative values at the apex.

5.3.5 *Substrate Bending*

If a lattice-mismatched layer is pseudomorphically grown on top of a substrate it suffers biaxial strain. For finite substrate thickness part of the strain will relax via substrate bending. If the substrate is circular, a spherical cap is formed. If the lattice constant of the film is larger (smaller) than that of the substrate, the film is under compressive (tensile) strain and the curvature is convex (concave) with respect to

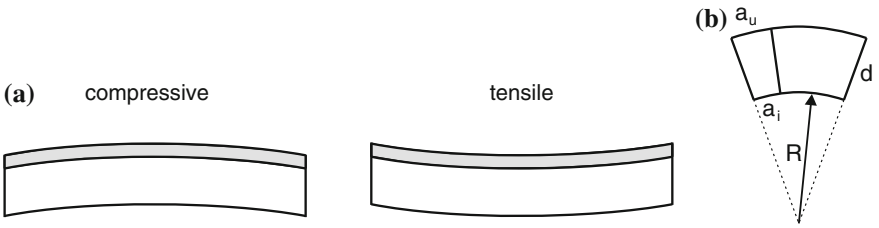

Fig. 5.28 (**a**) Schematic bending of a film/substrate system for compressive (*left*) and tensile (*right*) film strain. (**b**) Schematic deformation of curved film of thickness d. The lattice constants at the inner and outer surface are a_i and a_u, respectively

the outward normal given by the growth direction (Fig. 5.28a). Substrate bending can also be induced by a mismatch of the thermal expansion coefficients $\alpha_\mathrm{th}^\mathrm{f}$ and $\alpha_\mathrm{th}^\mathrm{s}$ of the film and substrate, respectively. If a film/substrate system is flat at a given temperature, e.g. growth temperature, a decrease of temperature, e.g. during cooling, will lead to compressive (tensile) strain if $\alpha_\mathrm{th}^\mathrm{f}$ is smaller (larger) than $\alpha_\mathrm{th}^\mathrm{s}$.

In a curved structure, the lattice constant in the tangential direction increases from a_i^t at the inner surface ($r = R = \kappa^{-1}$) to a_u^t at the outer surface ($r = R + d$). Thus, the tangential lattice constant varies with the radial position

$$a^\mathrm{t}(r) = a_\mathrm{i}^\mathrm{t}\,(1 + r\,\kappa)\,, \tag{5.85}$$

where d is the layer thickness (Fig. 5.28b). Therefore $a_\mathrm{u} = a_\mathrm{i}(1 + d/R)$. We note that (5.85) holds in all layers of a heterostructure, i.e. the film *and* the substrate.

The lattice constant in the radial direction a^r, however, depends on the lattice constant a_0 of the local material and is calculated from the biaxial strain condition, such as (5.70). The in-plane strain is $\epsilon_\parallel = (a^\mathrm{t} - a_0)/a_0$ (we assume a spherical cap with $\epsilon_\parallel = \epsilon_{\theta\theta} = \epsilon_{\phi\phi}$). For an isotropic material we find $a^\mathrm{r} = a_0\,(1 + \epsilon_\perp)$ with $\epsilon_\perp = -2\nu\epsilon_\parallel/(1 - \nu)$. The local strain energy density U is given by

$$U = \frac{Y}{1 - \nu}\,\epsilon_\parallel^2. \tag{5.86}$$

The total strain energy per unit area U' of a system of two layers with lattice constants a_1, a_2, Young's moduli Y_1, Y_2 and thickness d_1, d_2 (we assume the same Poisson constant ν in both layers) is

$$U' = \int_0^{d_1} U_1\,\mathrm{d}r + \int_{d_1}^{d_2} U_2\,\mathrm{d}r. \tag{5.87}$$

The total strain energy needs to be minimized with respect to a_i and R in order to find the equilibrium curvature κ. We find

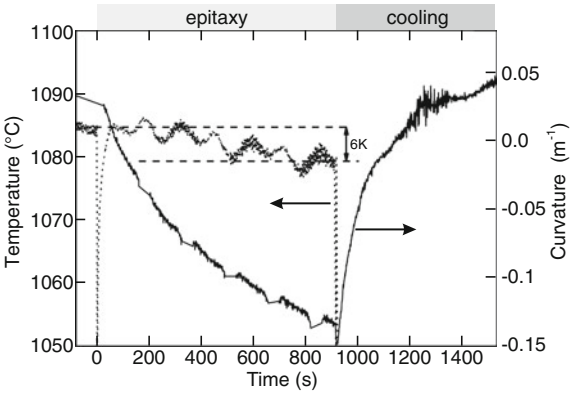

Fig. 5.29 Curvature of the middle of a Si wafer during GaN growth on an AlN interlayer grown at low temperatures on GaN and subsequent cooling. During growth the decrease in curvature indicates convex bowing due to compressive stress; during cooling the wafer flattens and becomes concave due to thermally induced tensile stress. Adapted from [372]

$$\kappa = \frac{6a_1a_2\,(a_2-a_1)\,d_1d_2\,(d_1+d_2)\,Y_1Y_2}{a_2^3\,d_1^4\,Y_1^2+\alpha\,Y_1Y_2+a_1^3\,d_2^4\,Y_2^4} \tag{5.88}$$
$$\alpha = a_1a_2\,d_1d_2\,\left[-a_2d_1(2d_1+3d_2)+a_1(6d_1^2+9d_1d_2+4d_2^2)\right].$$

For $a_2 = a_1\,(1+\epsilon)$ we develop κ to first order of ϵ and find ($\chi = Y_2/Y_1$) [368, 369]

$$\kappa = \frac{6\chi\,d_1\,d_2\,(d_1+d_2)}{d_1^4+4\chi\,d_1^3\,d_2+6\chi\,d_1^2\,d_2^2+4\chi\,d_1\,d_2^3+\chi^2\,d_2^4}\,\epsilon. \tag{5.89}$$

In the case of a substrate (d_s) with a thin epitaxial layer ($d_f \ll d_s$), the radius of curvature is approximately (Stoney's formula [370])

$$\kappa = 6\,\epsilon\,\frac{d_f}{d_s^2}\,\frac{Y_f}{Y_s}. \tag{5.90}$$

Conversely, if the radius of curvature is measured [371], e.g. optically, the film curvature (and through models also the film strain) can be determined during epitaxy as depicted in Fig. 5.29.

5.3.6 Scrolling

In some cases cylindrically scrolled structures are important, e.g. for thin-film flexible electronics, nanotubes, nanoscrolls or nanohelixes. The scrolling of thin layers must be avoided by suitable strain management for thin layers that are lifted off from their substrate for transfer to another flat substrate. If the film remains attached to its substrate, a scroll can be fabricated as schematically shown in Fig. 5.30. Such structures were first reported in [373], a review can be found in [374]. The shape of such scroll is investigated in [375] without a priori assumptions on its shape.

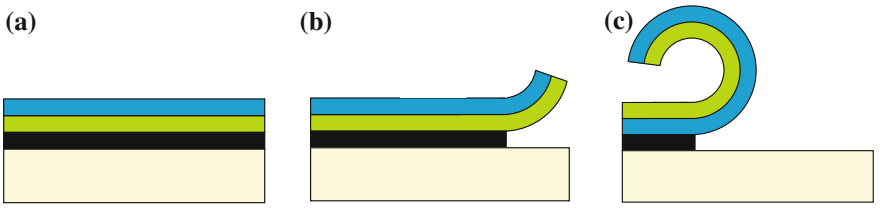

Fig. 5.30 Schematic representation of nanoscroll formation. (**a**) Strained heterostructure (*blue/green*) that is planar due to large substrate thickness, (**b**) starting removal of sacrificial layer (*black*), (**c**) release of thin film into nanoscroll geometry

If bending strain occurs only in *one* of the tangential directions, the energy density is given by

$$U = \frac{Y}{2\,(1-\nu^2)}\,(\epsilon_\mathrm{t}^2 + \epsilon_y^2 + 2\nu\,\epsilon_\mathrm{t}\,\epsilon_y), \tag{5.91}$$

where ϵ_y is the strain in the unbent direction (cylinder axis) as shown in Fig. 5.31a. For a strained heterostructure made up from two layers the curvature is given by (calculated analogous to (5.89), $\chi = Y_2/Y_1$ [369])

$$\kappa = \frac{6(1+\nu)\,\chi\,d_1\,d_2\,(d_1+d_2)}{d_1^4 + 4\chi\,d_1^3\,d_2 + 6\chi\,d_1^2\,d_2^2 + 4\chi\,d_1\,d_2^3 + \chi^2\,d_2^4}\,\epsilon, \tag{5.92}$$

which differs from (5.89) only by the factor $1+\nu$ in the nominator.

For cubic material and a (001) surface the energy is given as

$$U_{100} = \frac{C_{11}-C_{12}}{2C_{11}}\left[C_{11}\,(\epsilon_\mathrm{t}^2 + \epsilon_y^2 + C_{12}\,(\epsilon_\mathrm{t}+\epsilon_y)^2)\right] \tag{5.93}$$

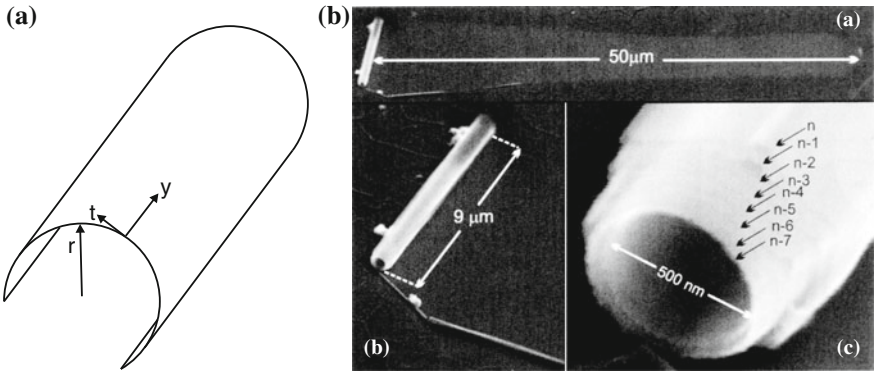

Fig. 5.31 (**a**) Schematic representation of a cylindrically rolled sheet with radial direction r, tangential direction t and direction along the cylinder axis y. (**b**) SEM images of multiwall InGaAs/GaAs nanoscroll rolled up over about 50 µm. Part (**b**) from [376]

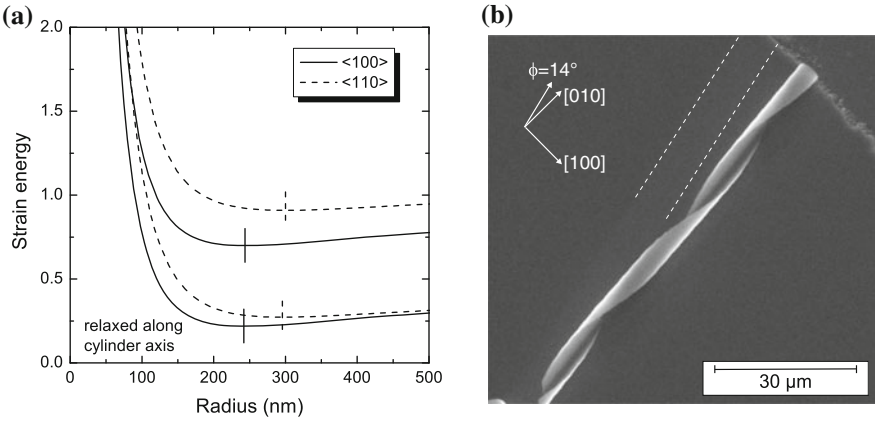

Fig. 5.32 (**a**) Strain energy (in units of the strain energy of the flat pseudomorphic layers) of a scroll of a 4-layer SiGe structure ($Si_{0.3}Ge_{0.7}$, $Si_{0.6}Ge_{0.4}$ and $Si_{0.8}Ge_{0.2}$, each 3 nm thick and a 1-nm Si cap) as a function of radius for winding directions along ⟨100⟩ and ⟨110⟩. *Top* (*bottom*) curves without (with complete) strain relaxation along the cylinder axis. *Vertical lines* indicate the positions of the respective energy minima [369]. (**b**) SEM image of curled InGaAs/GaAs nanoscroll rolled $\phi = 14°$ off ⟨100⟩. The stripe from which the film was rolled off is indicated by *white dashed lines*. Part (**b**) from [379]

for a scrolling direction along ⟨100⟩. When the (001)-oriented film winds up along a direction ⟨$hk0$⟩ having an angle ϕ with the [100] direction ($\phi = 45°$ for ⟨110⟩), the strain energy is given by (C_0 is given by (5.62))

$$U_\phi = U_{100} + C_0 \left(\frac{\epsilon_t - \epsilon_y}{2} \right)^2 \sin^2(2\phi). \tag{5.94}$$

The strain energy versus bending radius ($= \kappa^{-1}$) is shown for a SiGe nanoscroll in Fig. 5.32. First, the relaxation along the cylinder axis plays a minor role. The smallest strain energy is reached for scrolling along ⟨100⟩, also yielding the smaller bending radius (larger curvature). Therefore, the film preferentially scrolls along ⟨100⟩. This explains the observed 'curl' behavior of scrolls winding up for $\phi \neq 0$ [373, 377] (Fig. 5.32b). The effect of surface strain needs to be included to yield improved quantitative agreement with experimental values of $\kappa(\epsilon, d)$ [378].

5.4 Plasticity

5.4.1 Critical Thickness

Strained epitaxial films are called pseudomorphic when they do not contain defects and the strain relaxes elastically, e.g. by tetragonal distortion. When the layer thick-

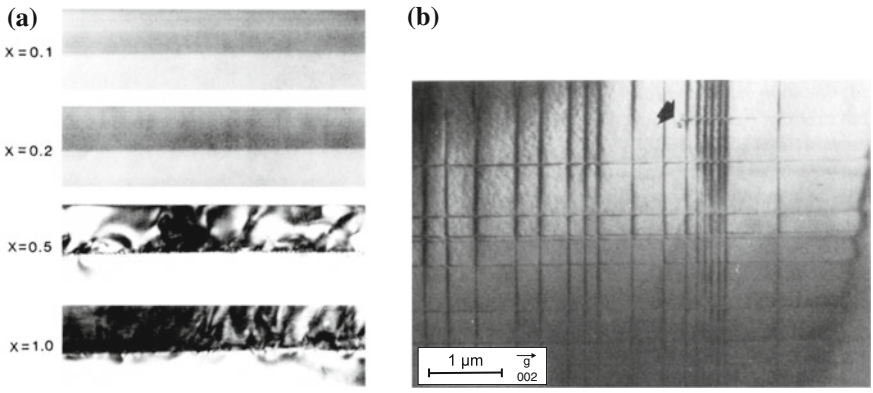

Fig. 5.33 (**a**) Series of cross-sectional TEM images of 100-nm thick Ge_xSi_{1-x} layers on Si(001) with different ternary compositions $x = 0.1, 0.2, 0.5$, and 1.0. The growth temperature was 550°C. The transition from commensurate to incommensurate growth is obvious. Adapted from [380]. (**b**) Plan view ⟨022⟩ TEM bright field image of a 250-nm $Ge_{0.15}Si_{0.85}$ layer on Si (001), annealed at about 700°C. The arrow denotes the position of a dislocation loop. Reprinted with permission from [381], © 1989 AVS

ness increases, however, strain energy is accumulated that will lead at some point to plastic relaxation via the formation of defects. In many cases, a grid of misfit dislocations forms at the interface (Figs. 4.18 and 5.33).

In Fig. 5.34 the strain around misfit dislocations at a GaAs/CdTe heterointerface, as calculated from a TEM image (Fig. 4.14), is shown.

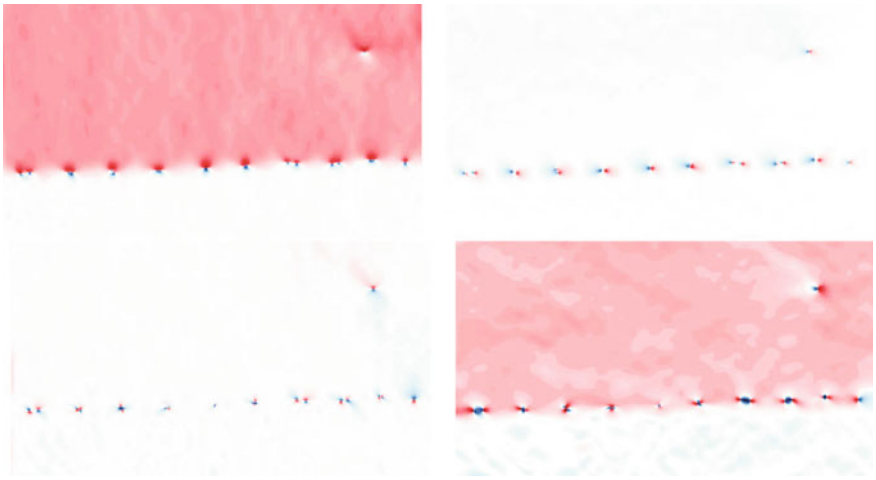

Fig. 5.34 Components $\begin{pmatrix} \epsilon_{xx} & \epsilon_{xz} \\ \epsilon_{zx} & \epsilon_{zz} \end{pmatrix}$ of the strain tensor (with respect to the GaAs lattice constant) of the dislocation array shown in Fig. 4.14, *red*/*blue*: positive/negative value, *white*: zero. From [286]

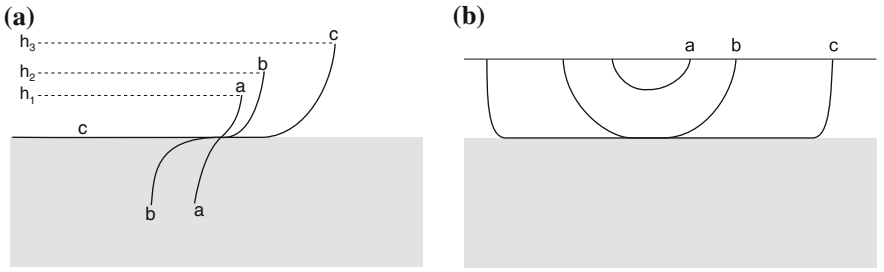

Fig. 5.35 Schematic formation of misfit dislocations by (**a**) elongation of a grown-in threading dislocation and (**b**) by the nucleation and growth of dislocation half-loops. (**a**) depicts a threading dislocation. Initially, for thickness h_1 the interface is coherent 'a', for larger thickness h_2 the interface is critical and the force of the interface on the dislocation is equal to the tension in the dislocation line, 'b'. For larger thickness, e.g. h_3, the dislocation line is elongated in the plane of the interface, 'c'. In (**b**) 'a' denotes a subcritical dislocation half-loop, 'b' depicts a half-loop being stable under the misfit stress and for 'c' the loop has grown under the misfit stress into a misfit dislocation line along the interface

The average distance p of the dislocations is related to the misfit $f = (a_1 - a_2)/a_2$ and the edge component $b_\perp$ of the Burger's vector (for a 60° dislocation $b_\perp = a_0/\sqrt{8}$)

$$p = \frac{b_\perp}{f}. \tag{5.95}$$

Two mechanisms have been proposed for the formation of misfit dislocations (Fig. 5.35), the elongation of a grown-in threading dislocation [382] and the nucleation and growth of dislocation half-loops [383]. For the modeling of such systems a mechanical approach based on the forces on dislocations [382] or an energy consideration based on the minimum strain energy necessary for defect formation [383–386] can be followed. Both approaches have been shown to be equivalent [387] (if a periodic array of dislocations is considered). In [388] it was pointed out that the finite speed of plastic flow also has to be considered to explain experimental data. Temperature affects the observed critical thickness and a kinetic model is needed. Another way of introducing dislocations is the plastic relaxation at the edge of coherent strained islands (cmp. Fig. 14.26).

In the following, isotropic materials and identical elastic constants of substrate and thin film are assumed, following [387]. The interface plane is the (x,y)-plane, the growth direction is z. The energy E_d of a periodic dislocation array with period p and Burgers vector $\mathbf{b} = (b_1, b_2, b_3)$ is

$$E_\mathrm{d} = \frac{Y}{8\pi\,(1-\nu^2)}\,\beta^2$$

$$\beta^2 = \left[b_1^2 + (1-\nu)\,b_2^2 + b_3^2\right] \ln\left(\frac{p\,\left[1-\exp(-4\pi h/p)\right]}{2\pi\,q}\right)$$

$$+\left(b_1^2-b_3^2\right)\frac{4\pi h}{p}\,\frac{\exp(-4\pi h/p)}{1-\exp(-4\pi h/p)}$$
$$-\frac{1}{2}\left(b_1^2+b_3^2\right)\left(\frac{4\pi h}{p}\right)^2\frac{\exp(-4\pi h/p)}{\left[1-\exp(-4\pi h/p)\right]^2}$$
$$+b_3^2\,\frac{2\pi h}{p}\,\frac{\exp(-2\pi h/p)}{1-\exp(-2\pi h/p)}, \tag{5.96}$$

where h is the film thickness and q denotes the cutoff length for the dislocation core, taken as $q = b$. The misfit strain including the relaxation due to dislocations with Burger's vectors $\mathbf{b}$ and $\hat{\mathbf{b}}$ in the two orthogonal interface $\langle 110\rangle$ directions $\mathbf{n}$ and $\hat{\mathbf{n}}$. We chose the coordinate system such that $\mathbf{n} = (1, 0, 0)$ and $\hat{\mathbf{n}} = (0, 1, 0)$ (the z direction remains). With respect to these axes the Burger's vectors are $\left(\pm\frac{1}{2}, \frac{1}{2}, \frac{1}{\sqrt{2}}\right) a_0/\sqrt{2}$. The misfit strain $\epsilon_{ij}^{\mathrm{m}}$ is reduced due to the dislocation formation to the 'relaxed' misfit strain $\epsilon_{ij}^{\mathrm{r}}$ with

$$\epsilon_{ij}^{\mathrm{r}} = \epsilon_{ij}^{\mathrm{m}} + \frac{b_i\, n_j + b_j\, n_i}{2p} + \frac{\hat{b}_i\, \hat{n}_j + \hat{b}_j\, \hat{n}_i}{2p}, \tag{5.97}$$

with an associated stress σ_{ij}. The strain energy E_{s} of the layer due to the relaxed misfit is then

$$E_{\mathrm{s}} = \frac{1}{2}\, h\, \sigma_{ij}\, \epsilon_{ij}^{\mathrm{r}} \tag{5.98}$$
$$\lim_{p\to\infty} E_{\mathrm{s}} = 2h\,\frac{Y(1+\nu)}{1-\nu}\, f^2. \tag{5.99}$$

The total strain energy E is given by

$$pE = 2E_{\mathrm{d}} + 2E_{\mathrm{c}} + p\, E_{\mathrm{s}} \tag{5.100}$$
$$E_\infty = \lim_{p\to\infty} E, \tag{5.101}$$

with the core energy E_{c} of the dislocation that needs to be calculated with an atomistic model (not considered further here). This energy is shown in Fig. 5.36a for the material parameters of $Ge_{0.1}Si_{0.9}$/Si(001) (misfit -0.4%) for various layer thicknesses as a function of $1/p$. This plot looks similar to that for a first-order phase transition (with $1/p$ as the order parameter). For a certain critical thickness h_{c1} the energy of the layer without any dislocation and the layer with a particular dislocation density p_1 are identical ($E - E_\infty = 0$) and additionally $\partial E/\partial p|_{p=p_1} = 0$. However, between $p \to \infty$ and $p = p_0$ there is an energy barrier. The critical thickness h_{c2} is reached when

$$\partial E/\partial p\,|_{p\to\infty} = 0, \tag{5.102}$$

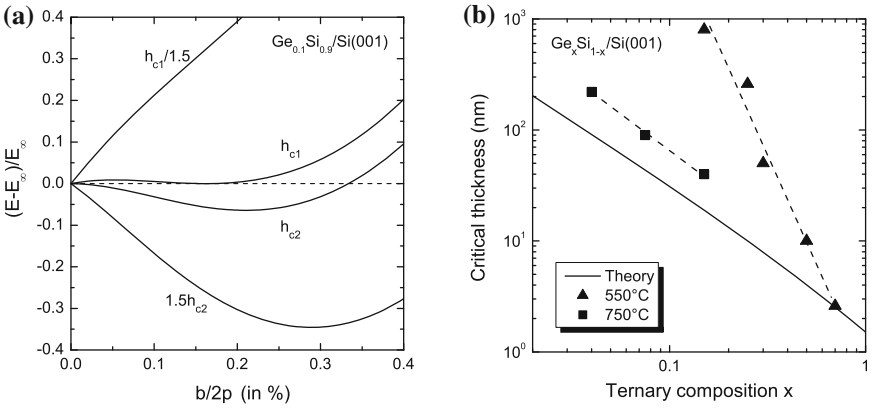

Fig. 5.36 (**a**) Theoretical calculation for the strain energy versus inverse dislocation density for various thicknesses of $Ge_{0.1}Si_{0.9}$ layers on Si (001). The ordinate is $b/2p$, $b/2$ being the edge component of the Burgers vector and p being the dislocation spacing. The abscissa is the strain energy E scaled with E_∞ (5.101). (**b**) Critical thickness for Ge_xSi_{1-x} layers on Si (001). The *solid line* is theory (h_{c2}) according to (5.103). Data points are from [389] (*squares*, growth temperature of 750 °C) and from [380] (*triangles* for growth temperature of 550 °C)

i.e. the energy decreases monotonically for increasing dislocation density up to the global energy minimum at a certain equilibrium dislocation spacing p_2. Equation (5.102) leads to the following implicit equation for the determination of h_{c2}:

$$h_{c2} = \frac{b\left[-16 + 3b^2 + 8\,(-4+\nu)\,\ln\,(2h_{c2}/q)\right]}{128\,f\,\pi\,(1+\nu)}, \tag{5.103}$$

with the length of the Burgers vector $b = a_0/\sqrt{2}$.

The theoretical dependence of h_{c2} for Ge_xSi_{1-x}/Si(001) with varying composition is shown in Fig. 5.36b together with experimental data. The critical thickness for a fairly high growth temperature is much closer to the energetic equilibrium than that deposited at lower temperature. This shows that there are kinetic limitations for the system to reach the mechanical equilibrium state. Also, the experimental determination of the critical thickness is affected by finite resolution for large dislocation spacing, leading generally to an overestimate of h_c.

In zincblende materials two types of dislocations are possible, α and β, with Ga- and As-based cores, respectively. They have $[\bar{1}10]$ and $[110]$ line directions for a compressively strained interface. The α dislocation has the larger glide velocity. Therefore, strain relaxation can be anisotropic with regard to the $\langle 110\rangle$ directions for zincblende material, e.g. InGaAs/GaAs [390, 391].

5.4.2 *Cleaving*

The cleavage planes of the diamond structure are {111} planes (Fig. 5.37a). It is easiest to break the bonds connecting the double layers in the ⟨111⟩ directions.

The cleavage planes of the zincblende structure are {110} planes (Fig. 5.37b). Due to the ionic character, breaking the bonds connecting the double layers in the ⟨111⟩ directions would leave charged surfaces, which is energetically unfavorable. The {100} planes contain only one sort of atom and would also leave highly charged surfaces. The {110} planes contain equal amounts of A and B atoms and are neutral. Ideally, the cleaving plane is atomically flat (Fig. 5.38a) or exhibits large mono-atomically flat terraces. However, certain dopants in high concentrations, e.g. GaAs:Te, can induce a rough surface due to lattice distortion [392].

The natural cleavage planes of wurtzite (GaN) are $\{1\bar{1}.0\}$ (m-type) planes [393].

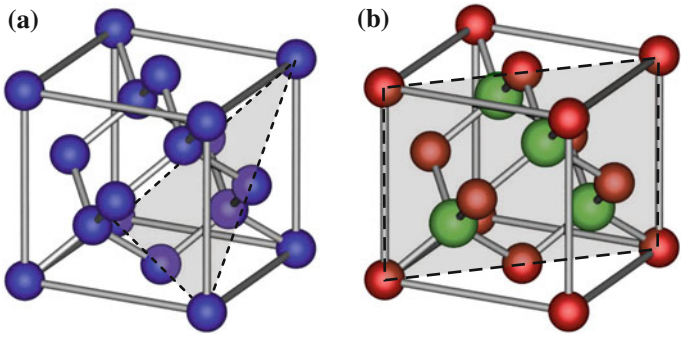

Fig. 5.37 Cleaving planes of the (**a**) diamond and (**b**) zincblende lattice

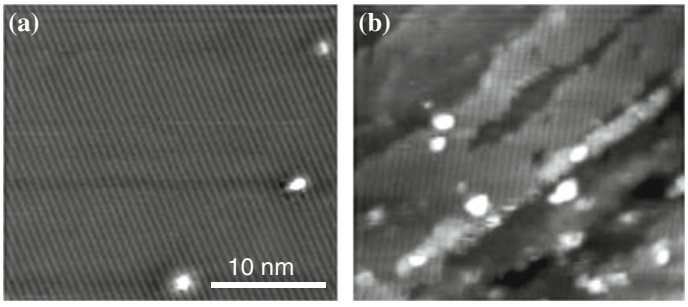

Fig. 5.38 Scanning tunneling microscopy images of a cleaved GaAs (110) surface with (**a**) good cleave and (**b**) bad cleave with defects dominating. Adapted from [394]

5.4.3 Wafer Breakage

The thickness and thus strength of wafers for semiconductor production (cmp. Sect. 12.2.2) is an important issue. The wafer should be as thin as possible for saving expensive materials but thick enough to avoid loss due to stress during handling, in particular during the later steps in a process since the value of a wafer increases with number of process steps it has undergone.

Reasons for wafer breakage is the mechanical handling (pick, place, transport) [395], stress loads due to processing (dielectrics, metals, asymmetric structures) and stress during processing for example due to thermal loads in annealing or deposition steps and cutting/dicing. The problem is less important in microelectronics but especially severe in photovoltaics (PV) industry, handling large total areas; on the other hand the profit loss per broken wafer is much higher in microelectronics industry. An additional problem poses the grain structure of multicrystalline silicon wafers used for PV [396] and the effect of surface cracks and irregularities at wafer edges and corners. Just going from a wafer thickness of 270 to 250 μm can more than double the breakage rate at certain process steps [397]. The minimum strength of a wafer with surface cracks is about 100 MPa, while the strength of wafers with cracks at the edge can reach rather low values around 20 MPa. Also the careful shaping of the wafer edge is important to avoid breakage [398].