

Chapter 7

Crystals, Lattices, Lattice Vibrations and Phonons

In this chapter we start to discuss topics that are specific to crystalline solids and, starting with Chap. 8, to semiconductors. We shall inspect the elementary excitations and quasi particles in semiconductors in Chaps. 7–10. These will be needed to describe and understand the linear optical properties in Chaps. 11–18. More details about these elementary excitations are found in textbooks on solid state physics; see for examples the references given in Chap. 1 like [81A1]a, h or [75Z1, 81M1, 89K1, 93K1, 95C1, 95I1, 97S1] and many others.

7.1 Adiabatic Approximation

If we want to describe a semiconductor, all we have to do in principle, is to solve the Schrödinger equation for the problem. It depends on the coordinates of the ion cores, consisting of the nucleus and the tightly bound electrons in the inner shells and the outer or valence electrons with coordinates \mathbf{R}_j and \mathbf{r}_i , and masses M_j and m_0 , respectively. The Hamiltonian reads (see e.g. Ref. [72B1] of Chap. 8):

$$H = -\frac{\hbar^2}{2} \sum_{j=1}^M \frac{1}{M_j} \Delta_{\mathbf{R}_j} - \frac{\hbar^2}{2m_0} \sum_{i=1}^N \Delta_{\mathbf{r}_i} + \frac{1}{4\pi\epsilon_0} \times \left(\sum_{j>j'} \frac{e^2 Z_j Z_{j'}}{|\mathbf{R}_j - \mathbf{R}_{j'}|} + \sum_{i>i'} \frac{e^2}{|\mathbf{r}_i - \mathbf{r}_{i'}|} + \sum_{i,j} \frac{e^2 Z_j}{|\mathbf{R}_j - \mathbf{r}_i|} \right). \quad (7.1)$$

Z_j is the effective charge of the ion core j and the indices j and i run over all M ion cores and N electrons, respectively.

We want to stress here that out of the four fundamental interactions so far known, namely strong, electromagnetic, weak, and gravitational interaction, only the electromagnetic one is of importance for all chemical properties including binding and thus also for the typical properties of semiconductors such as the transport and

the optical properties discussed in this book. Within the electromagnetic interaction we restrict ourselves here mainly to the electric ones (including, however, the Zeeman effect, diamagnetic shift (Sect. 16.1) magnetic polarons (Sects. 10.2 and 16.1.2), magnons (Sects. 10.2 and 12.4) and exchange interaction) since electric interactions are usually much stronger than magnetic ones, basically since electric interactions begin with monopole-monopole (i.e., Coulomb) interaction, whereas magnetic interactions start only with dipole-dipole interactions, due to the absence of magnetic monopoles (see (2.1a)). Magnetic interactions, however, do have a certain subtle importance e.g. in (diluted) magnetic semiconductors (Sect. 16.1.2) or, beyond the scope of this book in electron paramagnetic – or nuclear magnetic resonances.

The wavefunction solving (7.1) depends on all coordinates \mathbf{R}_j and \mathbf{r}_i including spins not noted explicitly here.

$$H\phi(\mathbf{r}_i, \mathbf{R}_j) = E\phi(\mathbf{r}_i, \mathbf{R}_j) \quad (7.2)$$

Since the indices j and i running from one to M and N , respectively, both count of the order of 10^{23} particles per cm^3 of semiconductor, it is obvious that there is at present no realistic chance of solving (7.1) and (7.2) though a proper solution would, in principle, contain all information about a given semiconductor. If we do not want to get stuck at this point we must use some approximations to simplify (7.1). The most important one is the so-called adiabatic or Born–Oppenheimer approximation. It starts from the fact that the mass of an ion core is three to five orders of magnitude heavier than a free electron, i.e.,

$$M_j \simeq 1836 \cdot A_j m_0 \quad (7.3)$$

where A_j is the mass number of ion j . Since the electric forces that bind the outer electrons to the atom are comparable to the ones which bind neighboring atoms or ions, and which can be described for small elongations by parabolic potentials i.e. by force constants β , we can easily see, even from classical arguments, that the highest resonance frequencies Ω with which ions can oscillate are much lower than the corresponding values ω for electrons

$$\Omega \simeq (\beta M_j^{-1})^{1/2} \ll \omega = (\beta m_0^{-1})^{1/2} . \quad (7.4)$$

Consequently, the electrons can practically instantaneously follow the motion of the ion cores, but not vice versa. This is the essence of the adiabatic approximation. On this basis we can separate $\phi(\mathbf{r}_i, \mathbf{R}_j)$ into a product of a wavefunction which depends only on the \mathbf{R}_j and describes the motion of the ion cores, and another one which gives the wavefunction of the electron system depending on the momentary values of the \mathbf{R}_j . In a next step we will further assume that all ions are fixed at their equilibrium positions \mathbf{R}_{j0} resulting finally in

$$\phi(\mathbf{r}_i, \mathbf{R}_j) = \phi(\mathbf{r}_i)\phi(\mathbf{R}_{j0}) , \quad (7.5)$$

and treating both the interaction between electrons and the deviation of the ions from their equilibrium positions in perturbation theory.

Before we start to inspect both factors of (7.5), we shall briefly outline how we describe a periodic lattice.

7.2 Lattices and Crystal Structures in Real and Reciprocal Space

In most cases we shall consider crystalline semiconductors. Disordered or amorphous systems will be mentioned explicitly. Crystalline solids have a periodic spatial arrangement of atoms, i.e., they show long-range order. We can define in such a case three non-coplanar elementary translation vectors \mathbf{a}_i ($i = 1, 2, 3$) with the property that if we start at a special atom, e.g., a Ga atom in a GaAs crystal, we reach an identical atom if we move by a vector \mathbf{R} given by

$$\mathbf{R} = n_1 \mathbf{a}_1 + n_2 \mathbf{a}_2 + n_3 \mathbf{a}_3 \quad (7.6)$$

with $n_i = 0, \pm 1, \pm 2, \dots$

The vector \mathbf{R} is called a translation vector of the lattice. If we shift the (infinitely large) lattice by \mathbf{R} it comes to a position which is identical to the starting one.

The vectors \mathbf{a}_i define a parallelepiped which is called the unit cell (see also Figs. 7.1 and 7.2). The whole volume of a crystal is completely filled with identical unit cells. The unit cell and the vectors \mathbf{a}_i are called primitive if the unit cell has the minimum possible volume. This definition is not unique as we explain for a two-dimensional cubic lattice in Fig. 7.1, where we show a non-primitive unit cell and two primitive ones. By convention, a special primitive unit cell is agreed upon. In our case the one defined by \mathbf{a}_1 and \mathbf{a}_2 .

The vectors \mathbf{R} evidently form for an infinite crystal an Abelian group which is called the translational group (Chap. 26). The positions of the atoms in the unit cell are given by the so-called basis. In Fig. 7.1 the basis consists of two atoms, one atom A at $(0, 0)$ and one atom B at $(1/2\mathbf{a}_1, 1/2\mathbf{a}_2)$. The translation vectors \mathbf{a}_i and the basis is all that we need to describe a crystal structure.

The translation vectors \mathbf{a}_i define an abstract, translation invariant point lattice, the basis gives the information where the atoms are really located in the primitive unit cell. The lattice and the basis define together the crystal structure.

Apart from the translational group there is another type of symmetry operation which transforms the lattice into itself, but for which at least one point is kept fixed. These symmetry operations also form a group which is called the point group. The elements of this group are for example reflections at mirror planes, rotations around axes with two-, three-, four- or six-fold symmetry or the inversion through the origin.

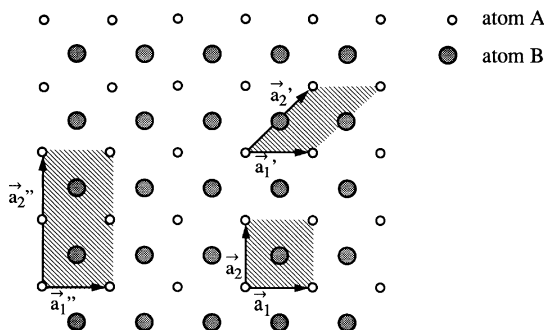


Fig. 7.1 Two primitive (*r.h.s.*) and a non-primitive (*l.h.s.*) unit cells in a two-dimensional cubic lattice with a basis consisting of two different atoms per primitive unit cell

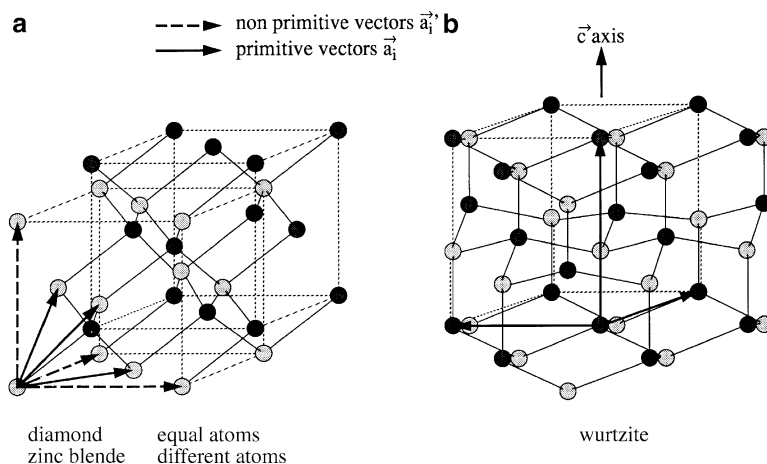


Fig. 7.2 The unit cells for the diamond and zinc-blende-type crystal structures (a) and of the wurtzite crystal structures (b) (see [82L1] of Chap. 1)

Furthermore there may be screw axes or glide planes, which combine either a rotation axis or a mirror plane with a translation by a rational fraction of the \mathbf{a}_i . The abstract translation lattices can be grouped into 14 Bravais lattices namely one triclinic, two monoclinic, four orthorhombic, two tetragonal, three cubic, one trigonal (rhombohedral) and one hexagonal lattice.

If we include the positions of the atoms and the translational invariance we find from all possible combinations of the symmetry operations which transform an infinite crystal into itself a total of 230 so-called space groups, out of which 73 can be written as a product of the translation group and the point group. For details see Chap. 26 and the references therein. The most important point groups for semiconductors are O_h (realized e.g. in the crystal structures of diamond, Si, Ge, Cu_2O or NaCl), T_d (realized e.g. in the zincblende type crystal structure of ZnS,

ZnSe, GaAs, InP, CuCl or AgBr) and C_{6v} (realized e.g. in the wurtzite type crystal structure of ZnS, ZnO, CdS or GaN). We give in Fig. 7.2 the crystal structures of diamond, zinblende and wurtzite. The diamond crystal structure consist of C atoms occupying the lattice points of two face-centered cubic (fcc) lattices shifted by $1/4$ of the space diagonal of the cubic unit cell. For zinblende one has the same principle, however one of the two sublattices is occupied by atoms A, the other by B. This situation is depicted in Fig. 7.2a. If the grey and the black “atoms” are identical, one has the diamond crystal structure. In Fig. 7.2a we show three non-primitive translation vectors \mathbf{a}'_i by the dashed arrows, which just define one of the fcc cubes. By solid arrows we show the primitive translation vectors \mathbf{a}_i , which start at a corner of the cube and point to the three nearest identical atoms, which are situated in the middle of the three faces intersecting in the corner. The wurtzite crystal structure is hexagonal with a polar crystallographic \mathbf{c} -axis. In this case, one has two primitive translation vectors frequently called \mathbf{a} and \mathbf{b} of equal length which include an angle of 120° and a third one (called \mathbf{c}) of different length and perpendicular to the plane spanned by \mathbf{a} and \mathbf{b} . Usually the z -axis is chosen parallel to \mathbf{c} . In all three cases, one atom is surrounded tetrahedrally by its four nearest neighbours. The difference between zinblende and wurtzit structures is in the positions of the next nearest neighbours only. Therefore several of the above mentioned compound semiconductors can crystallize in both structures like ZnS (which is notorious for these two polytypes), CdS or GaN. It is recommended that the reader visualizes these differences using some crystal models.

The chemical binding of the semiconductors is covalent for the elements (C, Si, Ge) with sp^3 hybridization, and acquires an increasing and finally dominant ionic admixture when going to the III–V, II^b–VI and I^b–VII compounds.

Now we want to introduce the so-called reciprocal lattice. It is defined by its elementary translation vectors \mathbf{b}_i in the same way as the lattice in real space. The \mathbf{b}_i are given by:

$$\mathbf{b}_1 = \frac{2\pi}{V_{uc}} \mathbf{a}_2 \times \mathbf{a}_3 \quad (7.7)$$

and cyclic permutations of the indices, V_{uc} is the volume of the unit cell given by

$$V_{uc} = \mathbf{a}_1(\mathbf{a}_2 \times \mathbf{a}_3). \quad (7.8)$$

A general translation vector of the reciprocal lattice is usually called \mathbf{G}

$$\mathbf{G} = l_1 \mathbf{b}_1 + l_2 \mathbf{b}_2 + l_3 \mathbf{b}_3 \quad l_i = 0, \pm 1, \pm 2, \dots \quad i = 1, 2, 3. \quad (7.9)$$

Without trying to be complete, we give some properties of the reciprocal lattice and its connections with the real one.

Every periodic function in real space which is sufficiently smooth and has a periodicity given by $f(\mathbf{r} + \mathbf{R}) = f(\mathbf{r})$ and \mathbf{R} defined by (7.6) can be written as a Fourier series summing over all vectors of the reciprocal lattice

$$f(\mathbf{r}) = \sum_{\mathbf{G}} f_{\mathbf{G}} e^{i\mathbf{G}\mathbf{r}} \quad (7.10)$$

with

$$f_{\mathbf{G}} = V_{\text{uc}}^{-1} \int_{\text{uc}} f(\mathbf{r}) e^{-i\mathbf{G}\mathbf{r}} d\tau .$$

The scalar product of \mathbf{R} and \mathbf{G} always fulfills

$$\mathbf{R} \cdot \mathbf{G} = 2\pi m; \quad m = 0, \pm 1, \pm 2, \dots \quad (7.11)$$

As a consequence, we can choose to describe effects occurring in periodic lattices in real space or in reciprocal space. The latter is the appropriate space for wave vectors \mathbf{k} or (quasi-)momenta $\hbar\mathbf{k}$. The “translation” from one space into the other is given by the three-dimensional Fourier series of (7.10).

In a crystal lattice we no longer have invariance with respect to infinitesimal translations in space (Sect. 3.1.3) but only invariance with respect to translations by integer multiples of \mathbf{a}_i . The conservation law for the momentum $\hbar\mathbf{k}$ which follows from an invariance with respect to infinitesimal translations according to Noether’s theorem (3.14b) is modified for a periodic lattice so that $\hbar\mathbf{k}$ is conserved only to within integer multiples of the \mathbf{b}_i , i.e., we can add to a given \mathbf{k} -vector a vector from the reciprocal lattice \mathbf{G} :

$$\mathbf{k} \lesseqgtr \mathbf{k} + \mathbf{G} . \quad (7.12)$$

This is a very important statement which forms, together with energy conservation, the basis, for example, of Ewald’s construction for the diffraction of X-rays or neutrons from a periodic lattice.

From (7.12) it is clear that we do not have to consider the whole \mathbf{k} -space, but can restrict ourselves to a “unit-cell” defined by the vectors \mathbf{b}_i . Every \mathbf{k} -vector that is outside the unit cell can be shifted inside the unit cell by adding an appropriate \mathbf{G} . Usually one does not work in reciprocal space with unit cells defined as in Figs. 7.1 or 7.2, but uses another construction explained in Fig. 7.3 for the two-dimensional case. One constructs the planes perpendicularly bisecting the lines connecting one point of the reciprocal lattice, which is chosen as the origin, with all others. The figure thereby enclosed around the origin is called the first Brillouin zone; the equivalent pieces which are next together form the second Brillouin zone, and so on. All Brillouin zones have equal area or volume in two or three dimensions, respectively. All higher Brillouin zones can be shifted into the first one by adding appropriate \mathbf{G} vectors. The Brillouin zones also form a type of elementary cells, but constructed according to Fig. 7.3 and not according to Fig. 7.1. The first cell constructed in real space according to Fig. 7.3 is known as a Wigner–Seitz cell. The names of points and lines of high symmetry in the first Brillouin zone are indicated in Fig. 7.4.

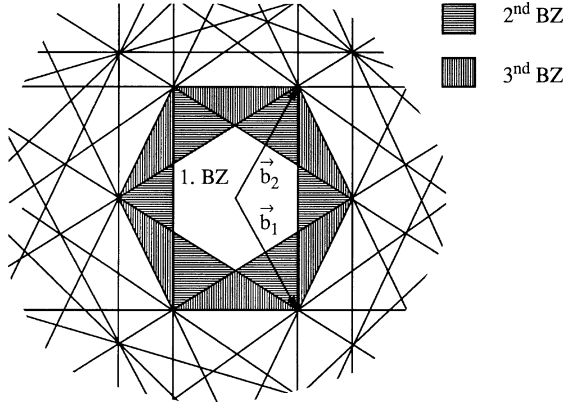


Fig. 7.3 The first Brillouin zones of a two-dimensional, hexagonal lattice

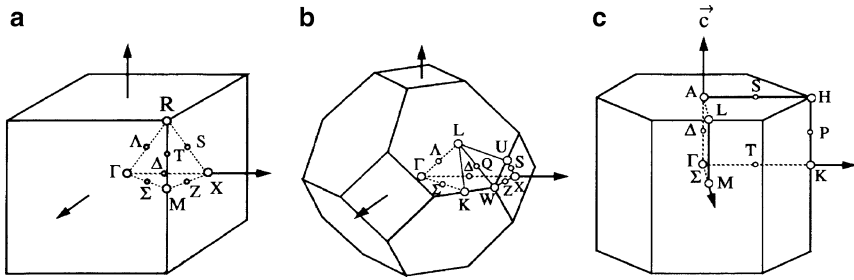


Fig. 7.4 The first Brillouin zones for a simple cubic lattice (a), of the diamond and zinc-blende-type structures (point groups O_h and T_d , respectively) (b) and of the hexagonal wurtzite-type structure (C_{6v}) (c). The names of points and of directions of high symmetry are indicated (According e.g. to Ref. [82L1] of Chap. 1)

For a simple cubic lattice with

$$\mathbf{a}_1 = (a, 0, 0), \quad \mathbf{a}_2 = (0, a, 0), \quad \mathbf{a}_3 = (0, 0, a), \quad (7.13)$$

the \mathbf{b}_i are also orthogonal with

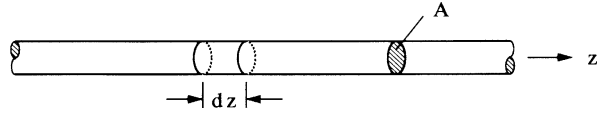
$$\mathbf{b}_1 = \left[\frac{2\pi}{a}, 0, 0 \right], \quad \mathbf{b}_2 = \left[0, \frac{2\pi}{a}, 0 \right], \quad \mathbf{b}_3 = \left[0, 0, \frac{2\pi}{a} \right], \quad (7.14)$$

and the first Brillouin zone is a cube which extends in all three directions from

$$-\frac{\pi}{a} \leq k_i \leq +\frac{\pi}{a}, \quad i = x, y, z. \quad (7.15)$$

In many of the following figures and equations, we give for simplicity the values of (7.15) for the borders of the first Brillouin zone, though most solids including

Fig. 7.5 A piece of a string as a model to explain the derivation of (7.19)



semiconductors are not crystallizing in a simple cubic lattice but frequently in the structures discussed above and below. See e.g. Figs. 7.9, 7.15 and 7.17.

In Fig. 7.4 we give the first Brillouin zones for this simple cubic lattice and for the point groups T_d , O_h and C_{6v} using the primitive unit cell including the notation for some special points and directions. The center of the first Brillouin zone $\mathbf{k} = (0, 0, 0)$ is always called the Γ -point, other points of high symmetry are labelled by capital Latin letters and directions of high symmetry by capital Greek letters. As an example: in T_d symmetry, when leaving the Γ -point in the Σ -direction one arrives in at the point K at the border of the first Brillouin zone.

The quantity $\hbar\mathbf{k}$ of an excitation in a periodic lattice is usually called quasi-momentum if one wants to stress the difference compared to the momentum $\mathbf{p} = \hbar\mathbf{k}$ of a free particle in vacuum, e.g., a photon or an electron, where, in contrast to (7.12) no reciprocal lattice vector may be added. Actually it is possible to make a transition from one case to the other: if the lattice constant a goes to zero, the system regains translational invariance with respect to infinitesimally small shifts in real space. On the other hand, the \mathbf{b}_i go to infinity in this limit (7.7) and the first Brillouin zone fills the whole \mathbf{k} -space, so that reciprocal lattice vectors become physically meaningless. For a more detailed discussion of the term “quasi”-momentum see e.g. Sect. 7.6 or [98B1] of Chap. 2.

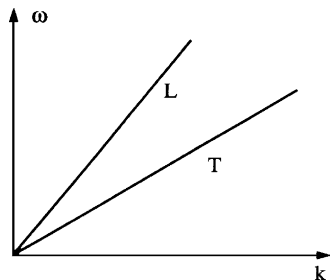
7.3 Vibrations of a String

In Sects. 7.3–7.6 we treat the lattice vibrations and the resulting quanta, the phonons, in the way introduced in [31F1] and adopted in many textbooks, i.e., we start with a homogeneous string, proceed to monatomic and diatomic chains and finally arrive at the three-dimensional solid.

Let us first consider a quasi one-dimensional string, as shown schematically in Fig. 7.5. Two types of waves can propagate along it, transverse and longitudinal ones. The direction of the elongation is either perpendicular to the direction of propagation, i.e., in the x - y plane, or parallel to it, i.e., in the z -direction, respectively. We start with the latter case. The mass density of the string is ρ its cross-section A , and the elongation of an infinitesimally small piece dz of the string at z from its equilibrium position is $u(z)$. Then, Newton’s equation of motion reads

$$dm \frac{\partial^2 u}{\partial t^2} = \rho A \cdot dz \cdot \frac{\partial^2 u}{\partial z^2} = F. \quad (7.16)$$

Fig. 7.6 The dispersion relation of waves on a homogeneous string



The force F is connected to the elasticity modulus E via

$$F = A \cdot E \frac{\partial^2 u}{\partial z^2} . \quad (7.17)$$

The appearance of the second derivative in (7.17) is for some students surprising, bearing in mind Hooke's law. However, we must consider that the stress σ is indeed given by

$$\sigma(z) = E \frac{\partial u}{\partial z} . \quad (7.18)$$

If the stress is the same on both sides of the infinitesimal element of length dz , the resulting forces at z and $z + dz$ compensate each other to zero. The restoring force F is therefore given by $d\sigma/dz$ leading to (7.17).

Putting (7.16) and (7.17) together leads to the standard harmonic wave equation

$$\rho \frac{\partial^2 u}{\partial t^2} = E \frac{\partial^2 u}{\partial z^2} . \quad (7.19)$$

With the ansatz

$$u = u_0 \exp [i(kz - \omega t)] \quad (7.20)$$

for a plane wave we find the dispersion relation for longitudinal waves

$$\omega_L = (E/\rho)^{1/2} k . \quad (7.21)$$

This is a linear relation as shown in Fig. 7.6. Consequently phase and group velocity are constant and equal, namely, with (2.13):

$$v_{ph}^L = v_g^L = (E/\rho)^{1/2} . \quad (7.22)$$

For the two degenerate, transverse waves we find in a similar way

$$\omega_T = (G/\rho)^{1/2} k \quad (7.23)$$

or

$$v_{\text{ph}}^{\text{T}} = v_{\text{g}}^{\text{T}} = (G/\rho)^{1/2} . \quad (7.24)$$

where G is the shear or torsion modulus.

Since it is known from the theory of elasticity that

$$G \leq E , \quad (7.25)$$

we find

$$v_{\text{ph}}^{\text{T}} \leq v_{\text{ph}}^{\text{L}} , \quad (7.26)$$

a result comparable to (4.28).

7.4 Linear Chains

We now should consider the regime of validity of the above calculation. We assumed a homogeneous string, neglecting the fact that a solid is made up from atoms. Therefore the above approximation can only be valid for wavelengths much longer than the lattice constant or for wave vectors close to the center of the first Brillouin zone, i.e.,

$$\lambda \gg a \quad \text{or} \quad k \ll \frac{\pi}{a} . \quad (7.27)$$

For shorter wavelengths we have to consider the atomic structure of solids. The interaction potential between neighboring atoms as a function of the lattice constant a looks qualitatively approximately like Fig. 7.7 for all types of binding, e.g., covalent, ionic, metallic or binding by van de Waals forces. For sufficiently large lattice constants there is no interaction between the atoms, i.e., $V = 0$; then comes an attractive regime (without which there would be no solids); and this is finally followed by a steep repulsive increase due to Pauli's exclusion principle when the filled inner shells of neighboring atoms start to overlap. Different analytic approximations to $V(a)$ are known for example the Born-Mayer or Lennard-Jones potentials. These details have at present no relevance for us. We note that a crystal left to itself will come to a state close to the energetic minimum, i.e., to the equilibrium position a_0 . In the vicinity of a_0 , $V(a)$ can be approximated by a parabola, that is by a harmonic potential, shown by the dashed line. It is at least qualitatively clear that this harmonic approximation is valid only very close to a_0 . For larger deviations from a_0 significant anharmonicities (i.e., deviations from the harmonic potential) have to be expected. The anharmonicities are characteristic for lattice vibrations and manifest themselves, among other things, in the thermal expansion of solids and in phonon-phonon interaction.

For the moment, however, the harmonic approximation is good enough and we consider a linear model solid in which every atom with mass M is connected to its neighbors by a "spring" with a force constant D , representing the harmonic potential, leading to the linear-chain model of Fig. 7.8a in which we indicate also