These operators create or annihilate excitons with quasi-momentum K and lectron-hole separation β. The transformation means the transition from ıdividual excitations with defined k or R to collective excitations in which only ˙ and β are given. Corresponding to (3.76), we then construct general *exciton perators* from linear combinations of these operators

$$b_{\nu K}^{(+)} = \sum_{\beta} U_{\nu K}(\beta) b_{\beta K}^{(+)} . \tag{3.92}$$

By considering the commutation relations of these $b_{\nu K}^{(+)}$, one can show that ıe excitons have boson character. The proof is simpler if carried out for the niting case of Frenkel excitons. We shall restrict ourselves to this. As in the evious section we put $\beta = 0$. We have then just one term left in (3.92). If we rther put $m = 0$ and $n = 1$ then, since the operators for $K \neq K'$ clearly com-ute,

$$\begin{aligned}[b_K, b_{K'}^{+}] &= \delta_{KK'}[b_K, b_K^{+}] \\ &= \frac{1}{N}\sum_{R} [c_{0R}^{+} c_{0R}(1 - c_{1R}^{+} c_{1R}) - c_{1R}^{+} c_{1R}(1 - c_{0R}^{+} c_{0R})]\delta_{KK'} .\end{aligned} \tag{3.93}$$

The product operators appearing in the last line are the particle number erators N_{1R} and N_{0R}. They have the eigenvalues 1 or 0 depending on whether electron at R is in an excited state or the ground state. The sum of the opera-s $N_{1R} + N_{0R}$ is correspondingly equal to the unit operator. (In our model re is an electron in either excited or ground state of the atom at every lattice ition). On the other hand, the product of the two operators is clearly zero. (3.93) then becomes

$$[b_K, b_K^{+}] = \delta_{KK'}\frac{1}{N}\sum_{R}(1 - 2N_{1R}) = \delta_{KK'}\left(1 - \sum_{R}\frac{2N_{1R}}{N}\right) . \tag{3.94}$$

The last term is of the order of "the number of excitons divided by the num-of unexcited lattice atoms". When one is concerned with the lowest excita-ıs of the insulator, this term can be neglected as small by comparison with 1. his approximation excitons are therefore bosons.

High exciton excitations has recently become an area of particular interest se-Einstein condensation of excitons, exciton molecules, electron-hole drops). further reading on this we refer the reader to the review article by *Haken Nikitine* [111d.73].

3.3 Ion-Ion Interaction: Phonons

3.3.1 Introduction

Having so far taken an interest only in electron motion, we want to turn now to the motion of the ion lattice itself.

Lattice dynamics is important in many areas of solid-state physics. As a result of their thermal motion, the lattice ions vibrate about their equilibrium positions. The forces which oppose this motion are those of the chemical bond. All the elastic properties, compressibility, the propagation of acoustic waves are related to it. These features are mostly described in the context of a continuum theory by disregarding the atomic structure of the lattice.

The continuum approximation is a limiting case of the microscopic theory which looks at the dynamics of the lattice ions themselves. To discuss this theory we shall first of all set up equations of motion for the lattice ions within the framework of classical mechanics and derive from these the energy and frequency of the "normal modes of oscillations" of the lattice. In describing the dispersion relations of these normal modes, we shall again meet up with the reciprocal lattice, the Brillouin zone concept, and other mathematical aids already introduced in the previous chapter. We shall be able to draw many parallels with earlier results which will allow us to limit the discussion necessary in this chapter. Section 3.3.2 is dedicated to a classical description of lattice vibrations.

If we supply thermal energy to a lattice ion, it will rapidly be distributed throughout the entire lattice by the mutual interaction between the ions. Local excitations will therefore lead to *collective vibrations* of the whole ion system. It is therefore appropriate to use collective coordinates (normal coordinates) for the mathematical description. The lattice vibrations can be readily quantized in this new representation. The associated quanta are elementary excitations called *phonons*. Phonons are bosons and thus call for the application of a different sort of statistics to electrons. Section 3.3.3 is devoted to the phonon concept.

We come to grips with our first application of this concept in Section 3.3.4 when we consider the energy content of lattice vibrations and specific heat. Section 3.3.5 then presents an overview of the calculation of phonon dispersion relations.

The phonon dispersion spectrum allows us to determine the associated density of states in a way similar to that in which the density of one-electron states followed from the band model. The close similarity means that the treatment of this question in Section 3.3.6 can be kept short.

In the last two sections of this chapter we examine the limiting case of lattice vibrations of very long wavelength. If the wavelength is large compared with the atomic separation, the microscopic structure of the solid can be neglected. This gives us, therefore, a link to the classical continuum theory.

In the approximation used throughout, the potential energy of a lattice ion xpanded in powers of the instantaneous deviation from its equilibrium posi- ı and only the first nonvanishing (harmonic) teı m is taken. This is the *monic approximation.* With it the Hamiltonian can be resolved into a sum of ependent terms with the form of Hamiltonians of harmonic oscillators. s is the basis of the quantization and with it the description of lattice vibra- ıs as a noninteracting phonon gas. The inclusion of higher anharmonic terms he expansion means an *interaction between the phonons.* This forms the con- : of a later chapter (Chap. 7).

The topic of lattice vibrations and the harmonic approximation is treated in ıy publications. Special attention is given to lattice dynamics in the books by *radudin, Montroll,* and *Weiss* [102.3]; *Wallis* [120]; *Stevenson* [121]; *Enns Haering* [133.II]; *Bak* [66]; and *Born* and *Huang* [67]. An excellent overview lso given by *Cochran* and *Cowley* [106, XXV/20], by *Leibfried* in [113a] and *rott* in [117]. The book series [101] contains numerous special articles. Phonon ›ersion relations are collected in a book by *Bilz* and *Kress* [107.10].

2 The Classical Equations of Motion

shall now consider ion-ion interaction and in doing so we shall ignore elec- ı motion. The model is therefore that of (1.8): the lattice ions vibrate about r equilibrium positions. Between them forces are present which correlate the vidual motions. The electron system is replaced by a spatially uniform ıtive space charge (inverse jellium model). Again we consider the ion lattice ıin a volume V_g with cyclic boundary conditions. Let the number of Wigner- z cells in this volume be N. Let the equilibrium positions of the ions be $= \boldsymbol{R}_n + \boldsymbol{R}_\alpha$. Here $\boldsymbol{R}_n$ is a suitable reference point inside the Wigner-Seitz and the $\boldsymbol{R}_\alpha$ are the vectors from this point to the αth basis atom. The index ll run from 1 to r for a basis made up of r ions. Let the instantaneous dis- ement of the $n\alpha$th ion from its equilibrium position be given by the time- ›ndent vector $\boldsymbol{s}_{n\alpha}(t)$.

n this section we consider the classical problem. The classical Hamiltonian ade up of the kinetic energy of all the ions and of their interaction energy. *kinetic energy* is

$$T = \sum_{n\alpha i} \frac{M_\alpha}{2} \dot{s}^2_{n\alpha i} \quad n = 1, \ldots, N, \quad \alpha = 1, \ldots, r, \quad i = 1, 2, 3. \tag{3.95}$$

here is the mass of the αth basis atom. The index i distinguishes the three ›sian coordinates of the vector $s_{n\alpha}$.

We expand the *potential energy* in increasing powers of the displacement The first (constant) term in this expansion is the potential energy of the ion ›e in equilibrium. We shall omit this component along with the negative ›ground since it does not contribute to the dynamics of lattice vibrations ›h is our sole interest here.

The second term in the expansion is linear in the $s_{n\alpha i}$. Since we are expanding about an equilibrium position, this term must disappear. The third term is quadratic in the displacement and has the form

$$\frac{1}{2} \sum_{\substack{n\alpha i \\ n'\alpha' i'}} \frac{\partial^2 V}{\partial R_{n\alpha i} \partial R_{n'\alpha' i'}} s_{n\alpha i} s_{n'\alpha' i'} = \frac{1}{2} \sum_{\substack{n\alpha i \\ n'\alpha' i'}} \Phi^{n'\alpha' i'}_{n\alpha i} s_{n\alpha i} s_{n'\alpha' i'} . \tag{3.96}$$

The matrix $\Phi^{n'\alpha' i'}_{n\alpha i}$ appearing here has $3rN$ rows and columns. We terminate the expansion at this first nonvanishing term (*harmonic approximation*). We shall not look at corrections to this approximation until Chapter 7. We recognize the meaning of the $\Phi^{n'\alpha' i'}_{n\alpha i}$ from the equations of motion

$$M_\alpha \ddot{s}_{n\alpha i} = -\frac{\partial V}{\partial s_{n\alpha i}} = -\sum_{n'\alpha' i'} \Phi^{n'\alpha' i'}_{n\alpha i} s_{n'\alpha' i'} . \tag{3.97}$$

$\Phi^{n'\alpha' i'}_{n\alpha i}$ is hereafter the force in the i-direction on the αth ion in the nth elementary cell when the α'th ion in the n'th cell is displaced by unit distance in the i'-direction.

The $\Phi^{n'\alpha' i'}_{n\alpha i}$ are called atomic force constants. They are linked by a large number of symmetry relations. First we recognize from (3.96) that the force constants are symmetric

$$\Phi^{n'\alpha' i'}_{n\alpha i} = \Phi^{n\alpha i}_{n'\alpha' i'} . \tag{3.98}$$

Furthermore they are clearly real. We obtain further relations when we use the fact that the potential energy must be invariant to an (infinitesimal) translation or rotation of the crystal. Let the translation be given by $s_{n\alpha i} = \delta s_i$ for all n, α, i and the rotation by $s_{n\alpha i} = \sum_k \delta\omega_{ik} R_{n\alpha k} (\delta\omega_{ik} = -\delta\omega_{ki})$. These operations are not allowed to result in any forces on the lattice ions, so the right-hand side in (3.97) must disappear. This leads to

$$\sum_{i'} \delta s_{i'} \sum_{n'\alpha'} \Phi^{n'\alpha' i'}_{n\alpha i} = 0 \quad \text{and} \quad \sum_{i'k'} \delta\omega_{i'k'} \sum_{n'\alpha'} \Phi^{n'\alpha' i'}_{n\alpha i} R_{n'\alpha' k'} = 0 \tag{3.99}$$

for translation and rotation, respectively, and hence to the symmetry relations

$$\sum_{n\alpha} \Phi^{n'\alpha' i'}_{n\alpha i} = 0 \tag{3.100}$$

and

$$\sum_{n\alpha} \Phi^{n\alpha i}_{n'\alpha' i'} R_{n\alpha k} = \sum_{n\alpha} \Phi^{n\alpha k}_{n'\alpha' i'} R_{n\alpha i} . \tag{3.101}$$

Along with these general relations, many others can be derived by use of lattice symmetry. We want to go into these later, but for the moment we shall

leave lattice symmetry out of our treatment. We shall then be in a better position later to distinguish between general statements and those which spring from lattice symmetry.

We are looking for solutions of the equations of motion which are periodic in time. To do this we put

$$s_{n\alpha i}(t) = \frac{1}{\sqrt{M_\alpha}} u_{n\alpha i} \exp(-i\omega t) \quad \text{with time-independent } u_{n\alpha i}. \tag{3.102}$$

Here we have drawn out a factor $M_\alpha^{-1/2}$. Using this expression, the equations of motion become

$$\omega^2 u_{n\alpha i} = \sum_{n'\alpha' i'} D_{n\alpha i}^{n'\alpha' i'} u_{n'\alpha' i'} \qquad \left(D \equiv \frac{\Phi}{\sqrt{M_\alpha M_{\alpha'}}}\right). \tag{3.103}$$

This is an eigenvalue equation for the real symmetric matrix $D_{n\alpha i}^{n'\alpha' i'}$ with $3rN$ real eigenvalues ω_j^2. The ω_j can only be either real or purely imaginary. The latter possibility can be eliminated since (3.102) would then lead to $s_{n\alpha i}$ which increase or decrease continuously with time.

The eigenvectors $u_{n\alpha i}$ of (3.103) are correspondingly characterized by the index j: $u_{n\alpha i}^{(j)}$, i.e., for each ω_j there are $3rN$ $u_{n\alpha i}^{(j)}$. They are called *normal modes*.

We now take account of the translation symmetry of the lattice. This requires that the $\Phi_{n\alpha i}^{n'\alpha' i'}$ (or $D_{n\alpha i}^{n'\alpha' i'}$) can depend not on the cell indices n' and n separately, but only on their difference $n' - n$: $\Phi_{n\alpha i}^{n'\alpha' i'} = \Phi_{\alpha i}^{\alpha' i'}(n' - n)$.

With this condition and the expression

$$u_{n\alpha i} = c_{\alpha i} \exp(i\boldsymbol{q}\cdot\boldsymbol{R}_n) \tag{3.104}$$

(3.103) becomes

$$\omega^2 c_{\alpha i} = \sum_{\alpha' i'} \left\{ \sum_{n'} \frac{1}{\sqrt{M_\alpha M_{\alpha'}}} \Phi_{\alpha i}^{\alpha' i'}(n' - n) \exp[i\boldsymbol{q}\cdot(\boldsymbol{R}_n - \boldsymbol{R}_{n'})] \right\} c_{\alpha' i'} \tag{3.105}$$

or since the summation over n' can be changed into a summation over $n' - n$

$$\omega^2 c_{\alpha i} = \sum_{\alpha' i'} D_{\alpha i}^{\alpha' i'}(\boldsymbol{q}) c_{\alpha' i'}, \tag{3.106}$$

with $D_{\alpha i}^{\alpha' i'}(\boldsymbol{q})$ defined by (3.105).

Lattice periodicity has thus reduced the system of $3rN$ equations (3.103) to a system of $3r$ equations. This system then has only $3r$ eigenvalues, that is, $3r$ ω_j. However these are functions of the vector $\boldsymbol{q}$

$$\omega = \omega_j(\boldsymbol{q}) \qquad j = 1, \ldots, 3r. \tag{3.107}$$

For each ω_j (3.106) has a solution $c_{\alpha i} = e_{\alpha i}^{(j)}(\boldsymbol{q})$. These solutions can be combined to vectors. They are defined except for a common factor which can be chosen such that the $e_\alpha^{(j)}(\boldsymbol{q})$ are normalized (and orthogonal to each other).

For the displacements $s_{n\alpha}(t)$ we then have, as special solutions of the equations of motion (3.97),

$$s_{n\alpha}^{(j)}(\boldsymbol{q}, t) = \frac{1}{\sqrt{M_\alpha}} e_\alpha^{(j)}(\boldsymbol{q}) \exp(i[\boldsymbol{q}\cdot\boldsymbol{R}_n - \omega_j(\boldsymbol{q})t]), \tag{3.108}$$

from which we can construct the general solutions.

Before looking more closely at (3.108) we want to examine the dispersion relations (3.107). $\hbar\omega_j$ is an energy, $\boldsymbol{q}$ is a vector in reciprocal space. The function $\omega_j(\boldsymbol{q})$ therefore has the same significance for lattice vibrations as the function $E_n(\boldsymbol{k})$ has for the motion of the electrons in the periodic lattice. We can adopt all the important results concerning qualitative properties from Section 2.2.

1) The function $\omega_j(\boldsymbol{q})$ is periodic in $\boldsymbol{q}$-space. So we need only consider one Brillouin zone, the form of which is given by the point group of the crystal.
2) The set of $\boldsymbol{q}$-values is finite as a result of the cyclic boundary conditions imposed on the crystal. If V_g contains N elementary cells, then there are N values of $\boldsymbol{q}$ in the Brillouin zone. Since j can take $3r$ values, there are $3rN$ different $\omega_j(\boldsymbol{q})$, that is, as many as the crystal has internal degrees of freedom.
3) $\omega_j(\boldsymbol{q})$ is an analytic function of $\boldsymbol{q}$ in the Brillouin zone in the same sense as $E_n(\boldsymbol{k})$ is analytic. However, whereas the index n in $E_n(\boldsymbol{k})$ can take any number of integer values, j has only $3r$ different values; $\omega_j(\boldsymbol{q})$ has $3r$ *branches*.
4) In the Brillouin zone $\omega_j(\boldsymbol{q})$ has the same symmetries as the band structure $E_n(\boldsymbol{k})$. In addition to the symmetries resulting from the space group of the crystal, we have by time reversal symmetry $\omega_j(\boldsymbol{q}) = \omega_j(-\boldsymbol{q})$.

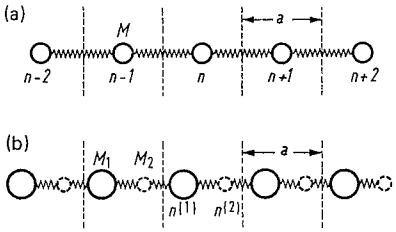


Fig. 3.6a and b. The linear chain (a) without and (b) with a basis.

The behaviour of $\omega_j(\boldsymbol{q})$ as $\boldsymbol{q} \to 0$ is important. We shall take the simple example of an oscillating linear chain to look at the most important aspects. Let us assume a chain of identical spheres linked together by springs of force constant f (Fig. 3.6a). Let a further be the separation between the spheres in equilibrium and let s_n be the displacement of the nth sphere from its equilibrium position. Eq. (3.97) then becomes

$$M\ddot{s}_n = -f(s_n - s_{n+1}) + f(s_{n-1} - s_n). \tag{3.109}$$

In accordance with (3.102) and (3.104) we put

$$s_n = \frac{1}{\sqrt{M}} c \exp\left[\mathrm{i}(qan - \omega t)\right]. \tag{3.110}$$

We then find

$$\omega^2 M = f[2 - \exp(-\mathrm{i}qa) - \exp(\mathrm{i}qa)] \tag{3.111}$$

and

$$\omega = 2\sqrt{\frac{f}{M}}\left|\sin\frac{qa}{2}\right| \tag{3.112}$$

so that ω is a periodic function of q. The first period (Brillouin zone) lies between $-\pi/a$ and $+\pi/a$ (Fig. 3.7).

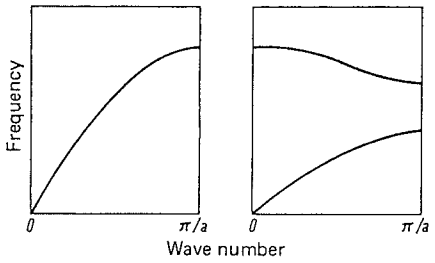


Fig. 3.7. Dispersion relations $\omega(q)$ for the linear chain without and with a basis (left and right, respectively).

If there are two atoms in the elementary cell, instead of (3.109) we have the equations (see Fig. 3.6b for the notation)

$$\begin{aligned} M_1 \ddot{s}_n^{(1)} &= -f(2s_n^{(1)} - s_n^{(2)} - s_{n-1}^{(2)}), \\ M_2 \ddot{s}_n^{(2)} &= -f(2s_n^{(2)} - s_{n+1}^{(1)} - s_n^{(1)}). \end{aligned} \tag{3.113}$$

Using the following expressions

$$\begin{aligned} s_n^{(1)} &= \frac{1}{\sqrt{M_1}} c_1 \exp\left(\mathrm{i}\left[q\left(n - \frac{1}{4}\right)a - \omega t\right]\right) \\ s_n^{(2)} &= \frac{1}{\sqrt{M_2}} c_2 \exp\left(\mathrm{i}\left[q\left(n + \frac{1}{4}\right)a - \omega t\right]\right) = \exp\left(\mathrm{i}\frac{qa}{2}\right)\frac{c_2}{c_1} s_n^{(1)} \sqrt{\frac{M_1}{M_2}} \end{aligned} \tag{3.114}$$

we find

$$\begin{aligned} -\omega^2\sqrt{M_1}\, c_1 &= -\frac{2f}{\sqrt{M_1}} c_1 + \frac{2f}{\sqrt{M_2}} c_2 \cos\frac{qa}{2}, \\ -\omega^2\sqrt{M_2}\, c_2 &= -\frac{2f}{\sqrt{M_2}} c_2 + \frac{2f}{\sqrt{M_1}} c_1 \cos\frac{qa}{2}. \end{aligned} \tag{3.115}$$

and as the solutions of the determinant

$$\begin{vmatrix} \frac{2f}{\sqrt{M_1}} - \omega^2\sqrt{M_1} & -\frac{2f}{\sqrt{M_2}}\cos\frac{qa}{2} \\ -\frac{2f}{\sqrt{M_1}}\cos\frac{qa}{2} & \frac{2f}{\sqrt{M_2}} - \omega^2\sqrt{M_2} \end{vmatrix} = 0, \tag{3.116}$$

we finally arrive at

$$\omega_\pm^2 = f\left(\frac{1}{M_1} + \frac{1}{M_2}\right) \pm f\sqrt{\left(\frac{1}{M_1} + \frac{1}{M_2}\right)^2 - \frac{4}{M_1 M_2}\sin^2\frac{qa}{2}}. \tag{3.117}$$

ω has therefore two branches, $\omega_+(q)$ and $\omega_-(q)$, which take the values $[2f(1/M_1 + 1/M_2)]^{1/2}$ and 0, respectively, at $q = 0$, and the values $(2f/M_1)^{1/2}$ and $(2f/M_2)^{1/2}$, respectively, at $q = \pm\pi/a$ (Fig. 3.7).

For the two limits $q = 0$ and $\pm\pi/a$, the ratio of the amplitudes c_2/c_1 follows from (3.115) as

$$\begin{aligned} c_2/c_1 &= +(M_2/M_1)^{1/2} \quad \text{for} \quad q = 0 \quad \text{and} \quad \omega = \omega_- \\ &= -(M_1/M_2)^{1/2} \quad \text{for} \quad q = 0 \quad \text{and} \quad \omega = \omega_+ \end{aligned} \tag{3.118}$$

$$\begin{aligned} &= \infty \quad \text{for} \quad q = \pm\pi/a \quad \text{and} \quad \omega = \omega_- \\ &= 0 \quad \text{for} \quad q = \pm\pi/a \quad \text{and} \quad \omega = \omega_+. \end{aligned} \tag{3.119}$$

These limiting cases correspond to typical modes of oscillation. Since $|q| = 2\pi/\lambda$, $q = 0$ means oscillations of infinite wavelength. All elementary cells move in the same sense. For $\omega = \omega_-$ the displacements of the two basis atoms in the cell are in the same direction, for $\omega = \omega_+$ they are in opposite directions.

The first case is the limiting case of an acoustic wave. The branch disappearing at $q = 0$ is correspondingly called the *acoustic branch*. The second type of vibration is readily excited optically in ionic crystals. The associated branch is therefore called the *optical branch*.

For $q = \pm\pi/a$ the basis atoms of one sort (M_1 or M_2) are at the nodes of the vibrations of wavelength $2a$. If each elementary cell has r basis atoms, then in addition to the acoustic branch there are $r - 1$ optical branches.

We have said nothing in (3.109) about whether these equations are meant to describe *transverse* vibrations (displacements perpendicular to the chain) or *longitudinal* vibrations (displacements along the chain). Eq. (3.109) applies to both situations as long as the vibration amplitudes are small. The meaning of the force constant f is different though for the two cases. For small amplitudes, each three-dimensional vibration of the chain can be resolved into three independent components, one longitudinal and two transverse. The two transverse

vibrations take place in two perpendicular planes whose line of intersection is the equilibrium position of the chain.

Thus we see that in the general case the function $\omega_j(q)$ for the oscillating chain has three acoustic and $3(r-1)$ optical branches.

If we now go from this example to a crystal with r atoms per elementary cell, we find that qualitatively the picture is the same. $\omega_j(\boldsymbol{q})$ is made up of three acoustic branches (degenerate at $\boldsymbol{q}=0$) and $3(r-1)$ optical branches. These branches are now functions of a vector $\boldsymbol{q}$. The degeneracy of the pair of transverse branches in the one-dimensional case only remains at points or along lines of high symmetry in the Brillouin zone. At a general point $\boldsymbol{q}$ all $3r$ branches are separated.

The expression "optical branch" cannot always be taken literally. Optical vibrations can in some cases not be excited optically, e.g., if the two basis atoms have the same effective charge. Again, away from $\boldsymbol{q}=0$ the vibrations of an optical branch are not always out of phase and those of an acoustic branch not always in phase. Complicated mixed modes of the two limiting cases (3.118) can arise here. Likewise, the lattice vibrations are only strictly longitudinal or transverse at points or along lines of high symmetry.

3.3.3 Normal Coordinates, Phonons

According to (3.95) and (3.96), the Hamiltonian function for lattice vibrations has the form

$$H=\sum_{n\alpha i}\frac{M_\alpha}{2}\dot{s}_{n\alpha i}^2+\frac{1}{2}\sum_{\substack{n\alpha i\\ n'\alpha' i'}}\Phi_{n\alpha i}^{n'\alpha' i'}s_{n\alpha i}s_{n'\alpha' i'}\,. \tag{3.120}$$

The $s_{n\alpha i}(t)$ are linear combinations of the particular solutions (3.108)

$$s_{n\alpha i}(t)=\frac{1}{\sqrt{NM_\alpha}}\sum_{jq}Q_j(\boldsymbol{q},t)e_{\alpha i}^{(j)}(\boldsymbol{q})\exp(\mathrm{i}\boldsymbol{q}\cdot\boldsymbol{R}_n) \tag{3.121}$$

where the time-dependent exponential factor in (3.108) has been included in the $Q_i(\boldsymbol{q},t)$ and a factor $1/\sqrt{N}$ has been separated out.

By putting (3.121) into (3.120), the Hamiltonian can be expressed in terms of the *normal coordinates* Q_j. For the conversion we use

$$\sum_n\exp[\mathrm{i}(\boldsymbol{q}-\boldsymbol{q}')\cdot\boldsymbol{R}_n]=N\delta_{\boldsymbol{q}'-\boldsymbol{q},\boldsymbol{K}_m} \tag{3.122}$$

where $\delta_{\boldsymbol{q},\boldsymbol{K}_m}$ is equal to unity, when $\boldsymbol{q}$ equals zero or a reciprocal lattice vector, and otherwise disappears.

Since the $s_{n\alpha i}(t)$ must be real, we have

$$e_{\alpha i}^{*(j)}(\boldsymbol{q})Q_j^*(\boldsymbol{q},t)=e_{\alpha i}^{(j)}(-\boldsymbol{q})Q_j(-\boldsymbol{q},t)\,. \tag{3.123}$$

We satisfy this by the requirements

$$e_{\alpha i}^{*(j)}(\boldsymbol{q})=e_{\alpha i}^{(j)}(-\boldsymbol{q})\quad\text{and}\quad Q_j^*(\boldsymbol{q},t)=Q_j(-\boldsymbol{q},t)\,. \tag{3.124}$$

Here we have complex $e_{\alpha i}$. Following (3.106), the $\boldsymbol{e}_\alpha$ can be chosen mutually orthogonal. We must then have further for the $e_{\alpha i}$

$$\sum_{\alpha i}e_{\alpha i}^{*(j)}(\boldsymbol{q})e_{\alpha i}^{(j')}(\boldsymbol{q})=\delta_{jj'}\,. \tag{3.125}$$

After much calculation, (3.122–125) yield

$$H=\frac{1}{2}\sum_{j\boldsymbol{q}}[\dot{Q}_j^*(\boldsymbol{q},t)\dot{Q}_j(\boldsymbol{q},t)+\omega_j^2Q_j^*(\boldsymbol{q},t)Q_j(\boldsymbol{q},t)]\,. \tag{3.126}$$

By introducing normal coordinates, the Hamiltonian resolves into a sum of $3rN$ individual terms. The coupled individual oscillations of the ions are formally replaced by decoupled collective oscillations. The normal coordinates used here are complex. One can also choose real normal coordinates instead.

One finds the conjugate momentum P to the Q^* from the Lagrange function $L=T-V$ according to

$$P_j(\boldsymbol{q},t)=\frac{\partial L}{\partial\dot{Q}_j^*(\boldsymbol{q},t)}=\dot{Q}_j(\boldsymbol{q},t)\,. \tag{3.127}$$

We then find that

$$H=\frac{1}{2}\sum_{j\boldsymbol{q}}[P_j^*(\boldsymbol{q},t)P_j(\boldsymbol{q},t)+\omega_j^2(\boldsymbol{q})Q_j^*(\boldsymbol{q},t)Q_j(\boldsymbol{q},t)]\,. \tag{3.128}$$

The Hamilton equations yield ($\dot{P}=-\partial H/\partial Q^*$)

$$\dot{P}_j(\boldsymbol{q},t)=\ddot{Q}_j(\boldsymbol{q},t)=-\omega_j^2(\boldsymbol{q})Q_j(\boldsymbol{q},t)\,. \tag{3.129}$$

The equation of motion in normal coordinates is then

$$\ddot{Q}_j(\boldsymbol{q},t)+\omega_j^2(\boldsymbol{q})Q_j(\boldsymbol{q},t)=0 \tag{3.130}$$

which is formally identical with the equation of motion of a harmonic oscillator of frequency $\omega_j(\boldsymbol{q})$.

The transition to a quantum mechanical description is now easy to execute. We have only to interpret the P and Q as operators which have been subjected to the commutation relations

$$[Q_j(\boldsymbol{q}),P_{j'}(\boldsymbol{q}')]=\mathrm{i}\hbar\delta_{\boldsymbol{q}\boldsymbol{q}'}\delta_{jj'}\,. \tag{3.131}$$

The Hamiltonian operator (3.128) together with the commutation relations (3.131) corresponds exactly to (A.1) and (A.2) of the Appendix. We can therefore interpret the quantized collective oscillations as elementary excitations. They are called *phonons*. The introduction of creation and annihilation operators leads to a Hamiltonian of the form

$$H = \sum_{jq} \hbar\omega_j(q)\left[a_j^+(q)a_j(q) + \frac{1}{2}\right]. \tag{3.132}$$

Each of the states defined by a pair (q, j) is occupied by $n_j(q)$ phonons of energy $\hbar\omega_j(q)$, where the $n_j(q)$ are the eigenvalues of the operator $a_j^+(q)a_j(q)$ appearing in (3.132). The contribution which such a state (a normal mode) makes to the total energy is $n_j(q)\hbar\omega_j(q)$ and the total energy (including the zero point energy) is

$$E = \sum_{jq} \hbar\omega_j(q)\left[n_j(q) + \frac{1}{2}\right]. \tag{3.133}$$

In the harmonic approximation used here, the phonon gas described by (3.132) consists of noninteracting phonons. It is therefore appropriate to make a comparison with the noninteracting electron gas of Section 2.1. The basic difference between the two cases lies in the fact that electrons are *fermions* and phonons are *bosons*. Each state in the lattice vibration spectrum can therefore be occupied by any number of (indistinguishable) phonons. In addition, the number of phonons depends on the energy content of the lattice vibrations, i.e., on the temperature. At $T = 0$, no phonons are excited and the lattice has its zero point energy only.

Consequently the statistics are different. We are not interested, as in Section 2.1.4, in the distribution of N indistinguishable fermions among given energy states at a given temperature, but in the number of excited bosons in the oscillator states as a function of temperature. The probability P_n that a particular oscillator has the excitation energy $E_n = \hbar\omega(n + \frac{1}{2})$ is proportional to $\exp(-E_n/k_BT)$. Since $\sum_n P_n = 1$, this probability is given explicitly by

$$P_n = \frac{\exp(-E_n/k_BT)}{\sum_n \exp(-E_n/k_BT)} = \frac{\exp(-n\hbar\omega/k_BT)}{\sum_n \exp(-n\hbar\omega/k_BT)}. \tag{3.134}$$

Since $\sum_n x^n = (1 - x)^{-1}$ (all summations run from $n = 0$ to $n = \infty$), it follows that

$$P_n = \exp(-n\hbar\omega/k_BT)\{1 - \exp[-\hbar\omega/k_BT)]\}. \tag{3.135}$$

We then find for the mean energy of the oscillator

$$\bar{E} = \sum_n E_nP_n = E_0 + \sum_n n\hbar\omega P_n. \tag{3.136}$$

Since $\sum_n nx^n = x/(1 - x)^2$, we finally have

$$\bar{E} = \frac{\hbar\omega}{\exp\left(\frac{\hbar\omega}{k_BT}\right) - 1} + \frac{\hbar\omega}{2}. \tag{3.137}$$

The mean occupation probability of an oscillator, i.e., the mean number of phonons in state j, q is then given by

$$\bar{n}_j(q) = \frac{1}{\exp[\hbar\omega_j(q)/k_BT] - 1} \qquad (\textit{Bose distribution}). \tag{3.138}$$

Following Section 3.3.2, the spectrum of the lattice vibrations resolves into $3r$ branches (index j), each of which can be represented as a function in q-space. Since there is a state associated with each (quasi-discrete) point q of each branch, we have to distinguish between phonons of the different branches. Depending of the behaviour of a branch at $q = 0$ and the polarization of the normal mode, we distinguish between *acoustic* and *optical*, *longitudinal* and *transverse phonons*. Since they show different properties when interacting with other quasi-particles and collective excitations, we shall note the different types, as far as necessary: TA-, TO-, LA-, LO-phonons.

We close this section with two fundamental remarks. The transition to normal coordinates, and the decoupling of the Hamiltonian into independent normal modes thereby achieved, was made possible by the fact that the Hamiltonian (3.120) is a positive definite quadratic form. Such a form can be diagonalized. We could therefore have already made the transition to quantum mechanics and introduced phonons by following on from (3.103), and before the explicit consideration of lattice periodicity. The appearance of elementary excitations at this point is not therefore linked to the properties of the lattice. The division of all ω_j into branches which can be presented in one Brillouin zone of q-space, on the other hand, is a consequence of the periodicity of the lattice.

If one had not terminated the expansion (3.96) at the second term, it would not have been possible to diagonalize. The consideration of higher, anharmonic terms therefore brings *phonon-phonon interaction* into the picture (Chap. 7).

3.3.4 The Energy Content of the Lattice Vibrations, Specific Heat

The total energy of the lattice vibrations at a temperature T is given from (3.133) and (3.138) by

$$E = \sum_{jq}\left[\frac{1}{\exp(\hbar\omega_j/k_BT) - 1} + \frac{1}{2}\right]\hbar\omega_j(q). \tag{3.139}$$

The summation over all q and over all branches j can be readily evaluated in two limiting cases.

1) *High temperature.* When $k_B T$ is large compared with $\hbar\omega_j$, the exponential function in the denominator can be expanded and it follows that

$$E = \sum_{jq} k_B T\left[1 + \frac{1}{12}\left(\frac{\hbar\omega_j(q)}{k_B T}\right)^2 + \cdots\right] \approx 3rNk_B T. \tag{3.140}$$

To a first approximation each of the $3rN$ oscillators makes a contribution $k_B T$ to the total energy (Dulong-Petit law). This is the classical result. Quantum effects do not appear.

2) *Low temperature.* In this case we cannot put $k_B T$ small in comparison with $\hbar\omega_j$ since all values of frequency from zero up to $\hbar\omega_j \gg k_B T$ appear. On the other hand, (3.139) allows us to conclude that frequencies for which $\hbar\omega_j \gg k_B T$ make no contribution. We can therefore limit the summation to the three acoustic branches. Here too only the lowest frequencies will be important. For these the dispersion relation $\omega_j = \omega_j(q)$ can be approximated by the linear expression $\omega_j(q) = s_j(\theta, \varphi)q$. We then have

$$E = \sum_{jq} \frac{\hbar s_j q}{\exp(\hbar s_j q/k_B T) - 1} + \text{zero point energy } E_0\,. \tag{3.141}$$

For sufficiently large V_g we can replace the summation over discrete q-points by an integration in q-space: $\sum_q = [V_g/(2\pi)^3]\int d\tau_q$. Eq. (3.141) then becomes

$$\begin{aligned} E - E_0 &= \frac{V_g}{(2\pi)^3}\sum_j \int \frac{\hbar s_j q}{\exp(\hbar s_j q/k_B T) - 1}\, d\tau_q \\ &= \frac{V_g}{(2\pi)^3}\frac{(k_B T)^4}{\hbar^3}\sum_j \int \frac{d\Omega}{s_j^3}\int_0^\infty \frac{x^3 dx}{e^x - 1}\,. \end{aligned} \tag{3.142}$$

The upper integration limit can be made infinity since large values of x make no contribution to the integrand. If we now average s_j^{-3} over all directions and branches we find

$$E = \frac{\pi^2}{10}\frac{V_g}{\hbar^3}\frac{(k_B T)^4}{\bar{s}^3} + \text{zero point energy} \tag{3.143}$$

in which the value $\pi^4/15$ has been inserted for the integral over all x.

These approximations are insufficient for the intermediate temperature range. We note first of all that a function of $\omega_j(q)$ is involved in the sum over q in (3.139). By replacing the summation with an integration one can immediately rearrange it as an integration over ω_j. To do this we introduce, as in Section

2.2.10, a *density of states* $g(\omega)$. Thus

$$g(\omega)d\omega = \frac{V_g}{(2\pi)^3}\int_{\omega=\text{const}} \frac{df_q}{|\nabla_q \omega_j(q)|}\, d\omega. \tag{3.144}$$

In contrast to Section 2.2.10, we have not normalized $g(\omega)$ here to V_g. With (3.144), we have for any function $F(\omega)$

$$\sum_q F(\omega) = \frac{V_g}{(2\pi)^3}\int F(\omega)d\tau_q = \int_0^\infty F(\omega)g(\omega)d\omega\,. \tag{3.145}$$

For the approximation $\omega_j(q) = \bar{s}_j q$ (with s_j already averaged over all angles) we have for example

$$g_j(\omega_j)d\omega_j = \frac{V_g}{2\pi^2}\frac{\omega_j^2 d\omega_j}{\bar{s}_j^3} \tag{3.146}$$

and

$$\begin{aligned} E - E_0 &= \sum_j \int_0^\infty \frac{\hbar\omega_j}{\exp(\hbar\omega_j/k_B T) - 1}\, g_j(\omega_j)d\omega_j \\ &= \sum_j \frac{V_g}{2\pi^2}\frac{(k_B T)^4}{\hbar^3}\frac{1}{\bar{s}_j^3}\int_0^\infty \frac{x^3}{e^x - 1}dx\,. \end{aligned} \tag{3.147}$$

But this is exactly the result we have just obtained above for the low temperature limit. In this formulation we can at least correct it in one respect. The integration of the density of states (3.144) over all ω has to give exactly N q-values for a branch [(3.145) with $F = 1$].

In the linear approximation we must therefore terminate the spectrum $\omega_j(q)$ at a frequency ω_D (*Debye frequency*) in order that this condition be met. This leads to the equation

$$\frac{V_g}{2\pi^2}\frac{1}{\bar{s}_j^3}\frac{\omega_{jD}^3}{3} = N, \tag{3.148}$$

or with $q_{jD} = \omega_{jD}/\bar{s}_j$ independent of j

$$q_{jD} = \left(\frac{6\pi^2 N}{V_g}\right)^{1/3} = (6\pi^2 n)^{1/3}\,. \tag{3.149}$$

Now N/V_g is the reciprocal volume of a Wigner-Seitz cell and is therefore equal to the volume of a Brillouin zone divided by $(2\pi)^3$. If we insert this into (3.149) it follows immediately that q_D is the radius of a sphere with volume

uivalent to that of the Brillouin zone. The Debye approximation used here erefore consists of three approximations to the spectrum $\omega_j(q)$: 1) neglect of ıtical branches, 2) linear approximation of acoustic branches, and 3) replace- ent of the Brillouin zone by a sphere of equivalent volume and assumption of rectional-independence of the linear approximation in this sphere.

In our approximation the Debye correction means replacing the upper nit ∞ in the integral (3.142) by $\hbar\omega_D/k_BT$, and hence the multiplication of .143) with a temperature-dependent factor

$$\frac{15}{\pi^4}\int_0^{\hbar\omega_D/k_BT}\frac{x^3}{e^x-1}dx = f\left(\frac{\hbar\omega_D}{k_BT}\right) = f\left(\frac{\theta_D}{T}\right). \tag{3.150}$$

ere we have also introduced the *Debye temperature* θ_D through $k_B\theta_D = \hbar\omega_D$.

We want to touch here on a model introduced by Einstein, in preparation : its application later. In this it is assumed that only one oscillation frequency curs: $\omega_j(q) = \omega_E$. Then for the density of states

$$g(\omega_j)d\omega_j = N\delta(\omega_j - \omega_E)d\omega_j, \tag{3.151}$$

d from (3.147) follows that

$$E - E_0 = \frac{N\hbar\omega_E}{\exp(\hbar\omega_E/k_BT) - 1}. \tag{3.152}$$

Coarse as this approximation appears to be, it is nevertheless important for :ending the Debye approximation. From Fig. 3.7 we see that, for the linear ain with a basis, the Debye approximation gives a very good description for : acoustic branch. For the optical branch, however, it is clear that an assump- n of constant frequency for all optical phonons will be more accurate than a ear approximation of the Debye type.

In making a comparison between theory and experiment, the change of total ergy with temperature, i.e., the *specific heat*, will prove to be more important ın the total energy itself.

We do not have to distinguish between c_p and c_v in the harmonic approxima- n since this approximation does not include thermal expansion of the lattice. : obtain the specific heat from the Debye approximation by differentiating 147), (3.150) with respect to the temperature.

$$c_D(T) = 3Nk_Bf_D\left(\frac{\theta_D}{T}\right) \quad \text{with} \quad f_D(x) = \frac{3}{x^3}\int_0^x \frac{y^4e^y dy}{(e^y-1)^2}. \tag{3.153}$$

r large θ_D/T (low temperatures), $f_D(x)$ is approximately given by $4\pi^4/5x^3$. this approximation the specific heat is then proportional to T^3 (Debye's -law).

If we further introduce the Einstein temperature θ_E corresponding to the Debye temperature and a factor 3 for the three optical branches, we then find in the Einstein approximation that

$$c_E(T) = 3Nk_Bf_E\left(\frac{\theta_E}{T}\right) \quad \text{with} \quad f_E(x) = \frac{x^2e^x}{(e^x-1)^2}. \tag{3.154}$$

The temperature dependence of both approximations is given in Fig. 3.8.

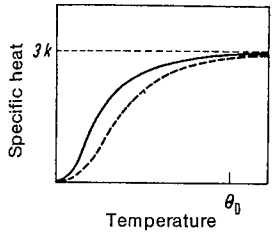


Fig. 3.8. Specific heat according to the Debye (upper curve) and Einstein (lower curve) approximations.

3.3.5 Calculation of Phonon Dispersion Relations

As already pointed out in Section 3.3.2, the function $\omega_j(q)$ can be represented in a Brillouin zone of q-space in a way corresponding to that in which a band structure $E_n(k)$ is represented in a Brillouin zone of k-space. In particular the same symmetry requirements apply.

Before we consider the calculation of such dispersion curves, let us draw attention to a few important points by taking the simple case of a two-dimensional square lattice.

The function $\omega_j(q)$ follows from (3.106) when the $3r \times 3r$ matrix $D_{\alpha i}^{\alpha' i'}(q)$ is known. The latter in turn is found from the force constants according to the equation

$$D_{\alpha i}^{\alpha' i'}(q) = \sum_{n'}\frac{1}{\sqrt{M_\alpha M_{\alpha'}}}\Phi_{n\alpha i}^{n'\alpha' i'}\exp[iq\cdot(R_n - R_{n'})]. \tag{3.155}$$

In our model we assume elastic (spring) forces between neighbouring atoms, as in the one-dimensional case in Section 3.3.2. In accordance with the assumption that the forces are strongest between nearest neighbours and diminish with increasing distance to the next-nearest neighbour, third-nearest neighbour, etc., we limit ourselves to consideration of the nearest and next-nearest (Fig. 3.9). It would not be realistic to restrict our consideration to nearest neighbours only, if only because a square lattice (just like a cubic lattice in space) would not be stable against shear forces if elastic forces only exist between directly adjacent neighbours.

We consider now a particular atom $n = 0$. If the n'th atom is moved from its equilibrium position by a vector distance $s_{n'}$, the ith component of force acting

on the atom is given by

$$F_{n'\to 0} = a_{n'} e_{n'} (e_{n'} \cdot s_{n'}) . \tag{3.156}$$

Here $a_{n'}$ is the force constant of the particular spring and $e_{n'}$ is the unit vector which gives the direction of $R_{n'}$. If we compare this expression with (3.97), in which we can omit the index α for the consideration of Bravais lattices, we find for the force constants

$$\Phi_{0i}^{n'i'} = -a_{n'} e_{n'i} e_{n'i'} \quad (n' \neq 0) . \tag{3.157}$$

The force constant $\Phi_{0i}^{0i'}$, which describes the force on the atom 0 resulting from the displacement of this atom by a vector s_0, follows directly as the sum over all the forces which would arise if all other atoms were displaced by a vector $-s_0$

$$\Phi_{0i}^{0i'} = -\sum_{n'(\neq 0)} \Phi_{0i}^{n'i'} = \sum_{n'(\neq 0)} a_{n'} e_{n'i} e_{n'i'} . \tag{3.158}$$

From (3.157) and (3.158) we can readily give all the force constants for our particular example. Let the force constants to the four nearest neighbours be f_1 and to the four next-nearest neighbours be f_2, and let the lattice constant be a. We then have

$$\begin{aligned}
&\Phi_{01}^{11} = \Phi_{01}^{21} = \Phi_{02}^{32} = \Phi_{02}^{42} = -f_1 , \\
&\Phi_{0i}^{5i'} = \Phi_{0i}^{6i'} = -\frac{f_2}{2} \qquad (i,i' = 1, 2) , \\
&\Phi_{0i}^{7i} = \Phi_{0i}^{8i} = -\frac{f_2}{2} \qquad (i = 1, 2) , \\
&\Phi_{0i}^{7j} = \Phi_{0i}^{8j} = +\frac{f_2}{2} \qquad (i \neq j = 1, 2) \\
&\Phi_{01}^{01} = \Phi_{02}^{02} = 2(f_1 + f_2) \\
&\text{all other } \Phi_{0i}^{n'i'} = 0 .
\end{aligned} \tag{3.159}$$

The numbering of the individual neighbours is shown in Fig. 3.9.

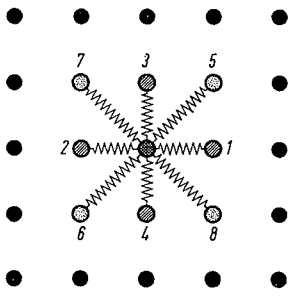


Fig. 3.9. Elastic forces to the nearest and next-nearest neighbours in a two-dimensional square lattice.

The $D_i^{i'}(q)$ can be calculated through (3.159) and inserted into (3.106). Eq. (3.106) is a system of equations which can only be solved if the determinant

$$|D_i^{i'}(q) - \omega^2 \delta_{ii'}| = 0 \tag{3.160}$$

vanishes. This leads in our example to

$$\begin{vmatrix} [f_1(1 - \cos q_x a) + f_2(1 - \cos q_x a \cos q_y a)] - \frac{M}{2}\omega^2 & \sin q_x a \sin q_y a \\ \sin q_x a \sin q_y a & [f_1(1 - \cos q_y a) + f_2(1 - \cos q_x a \cos q_y a)] - \frac{M}{2}\omega^2 \end{vmatrix} = 0 . \tag{3.161}$$

q_x and q_y are the components of q in the (two-dimensional) Brillouin zone. For the two-dimensional square lattice this is itself a square of side length π/a.

The determinantal condition (3.161) leads to two solutions of $\omega_j(q)$ corresponding to the two branches possible in the two-dimensional Bravais lattice. We shall give them for a few points and lines of symmetry. The symmetry symbols are those of Fig. 2.22a (Brillouin zone of the cubic lattice in k-space) which are contained within the square $k_z = 0$. The midpoint is therefore Γ; the Δ-axis leads to the midpoint (X) of a side of the square, and the Σ-axis to a corner of the square (M). The Z-axes are the sides of the square X-M-X-. . . . In the two-dimensional case, the point groups associated with the individual points and lines are of course others.

From (3.161) we have

$$\begin{aligned}
&\Delta\text{-axis } (\Gamma \to X): \quad \omega_1 = \left[\frac{2}{M}(f_1 + f_2)(1 - \cos q_x a)\right]^{1/2} \\
&\qquad\qquad\qquad\quad \omega_2 = \left[\frac{2}{M} f_2 (1 - \cos q_x a)\right]^{1/2} \\
&\Sigma\text{-axis } (\Gamma \to M): \quad \omega_1 = \left[\frac{2}{M}[f_1(1 - \cos q_x a) + f_2(1 - \cos 2q_x a)]\right]^{1/2} \\
&\qquad\qquad\qquad\quad \omega_2 = \left[\frac{2}{M} f_1 (1 - \cos q_x a)\right]^{1/2} \\
&Z\text{-axis } (X \to M): \quad \omega_1 = \left[\frac{2}{M}[2f_1 + f_2(1 - \cos q_y a)]\right]^{1/2} \\
&\qquad\qquad\qquad\quad \omega_2 = \left[\frac{2}{M}[f_1 + f_2 - (f_1 - f_2)\cos q_y a]\right]^{1/2} .
\end{aligned} \tag{3.162}$$

The dispersion curves are shown in Fig. 3.10 for an arbitrary choice of the ratio f_2/f_1 ($f_2/f_1 = 2$). Fig. 3.10 further shows the symbols of the irreducible

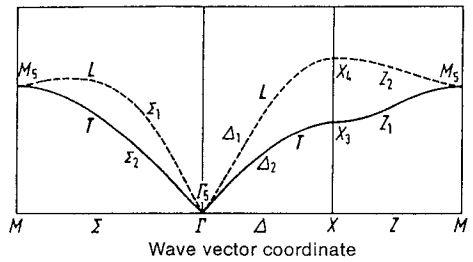


Fig. 3.10. Dispersion curves for the square net, considering spring forces between nearest and next-nearest neighbours ($f_1/f_2 = 2$). Ordinate: frequency in arbitrary units, abscissa: wave vector along symmetry lines of the Brillouin zone.

epresentations which are associated with the ω_i at points and lines of high symnetry. We shall not involve ourselves in the group theory side here. It is analoous to the example treated in Section 2.2.14. On the other hand, one further emark is important.

If we insert the solutions $\omega_j(\boldsymbol{q})$ into (3.106), we can determine the $c_i^{(j)}$. Then or example we find on the Δ-axis for ω_1: $c_x^{(1)} = 1$, $c_y^{(1)} = 0$. The ω_1-branch thus escribes longitudinal lattice vibrations. Correspondingly we find transverse ibrations for ω_2 ($c_x^{(2)} = 0$, $c_y^{(2)} = 1$). On the Σ-axis, too, the upper branch is ssociated with longitudinal vibrations and the lower branch with transverse.)n the other hand, the vibrations with a $\boldsymbol{q}$-vector on the Z-axes cannot be lassified in this way. The division into longitudinal and transverse vibrations is ıerefore possible only for particular $\boldsymbol{q}$-values.

Calculations have been carried out for many solids in an analogous way to ıe model treated above. The method is however limited in its applicability by ıe following facts:

) A model in which the chemical bond between atoms is simulated by the ɔrces of a spring can only be successful if the bonding forces are central. This is ıe case in ionic crystals where electrostatic forces between the lattice ions are ɛsponsible for the chemical bond. For a lattice based on covalent bonds the onding forces are directional and therefore dependent on the bond angle. uch forces can only be poorly represented by simple springs. In spite of this, uite good results are obtained here.

) The replacement of lattice ions by fixed spheres connected by springs leaves ıe polarizability of the ions and their compressibility out of the picture.

) Bonding forces do not act only between nearest neighbours. In principle ıey must also be considered to involve even ions which are very far apart. lowever, this introduces more and more parameters into the results and these ave to be determined by matching with experimental observations. It is then ointless to refine the model with more free parameters than can be determined y experimental measurements.

Within the limits of the spring representation, the simplest model is that of *gid ions*. Here, however, the number of parameters needed is very large if ne cannot restrict oneself to nearest neighbours only. For germanium, for ɛample, calculations show that contributions up to the sixth-nearest neighbour ave to be included in order to obtain satisfactory results.

The theory is substantially improved when one also takes account of the polarizability of lattice ions. This can be modelled by describing the valence electrons of the ions by a massless, negatively charged shell which is bound to the positive ions by isotropic elastic forces (*shell model*). By considering in Ge nearest neighbours only, satisfactory agreement with experiment was obtained using the shell model with only five free parameters. However here too the number of parameters rises rapidly when one attempts to explain details of the vibrational spectrum.

A further important improvement to the model is brought about when one takes account of the compressibility of lattice ions by assuming a compressible shell (*"breathing" shell model*). In many instances this refinement would lead to an almost quantitative agreement with experiment.

In fitting values to the parameters, we can first of all make use of the elastic constants and the dielectric constant (Secs. 3.3.7 and 3.3.8). These will not usually prove to be sufficient to determine all the free parameters. One therefore looks to experimentally determined vibration spectra and matches parameter values to them. Dispersion curves can be measured by inelastic neutron scattering, the method being restricted to solids with not too large a scattering cross section.

Figure 3.11 shows the branches of the function $\omega_j(\boldsymbol{q})$ for diamond, along the most important axes within and on the surface of the Brillouin zone. The most significant features in Fig. 3.11 can be derived from symmetry considerations alone, as we learned in Section 2.2.14.

We do not want here to go into all the finer points which can arise in calculating dispersion curves. One such point would be the consideration of angle-dependent forces in covalent crystals.

In metals, screening of the ion-ion interaction by the electron gas of the

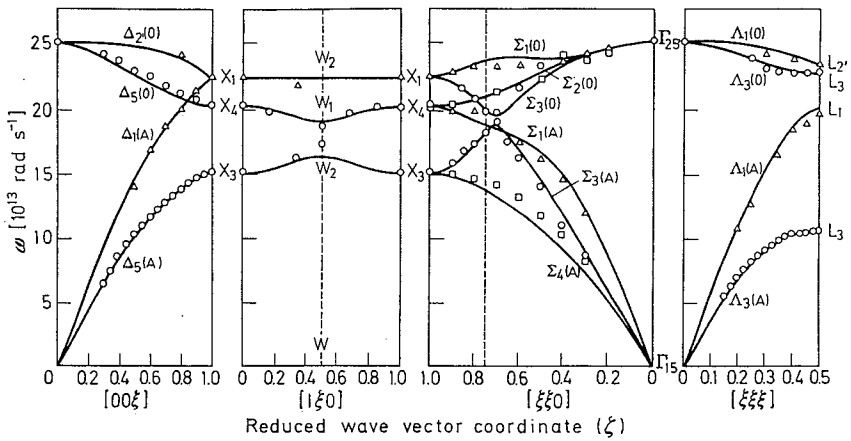


Fig. 3.11. Dispersion curves for diamond. Ordinate: frequency ν, abscissa: wave vector along high symmetry axes [from J. L. Warren et al.: Phys. Rev. 158, 805 (1967)]

valence electrons is important. This therefore represents one place where electron-ion interaction has to be incorporated into the theory.

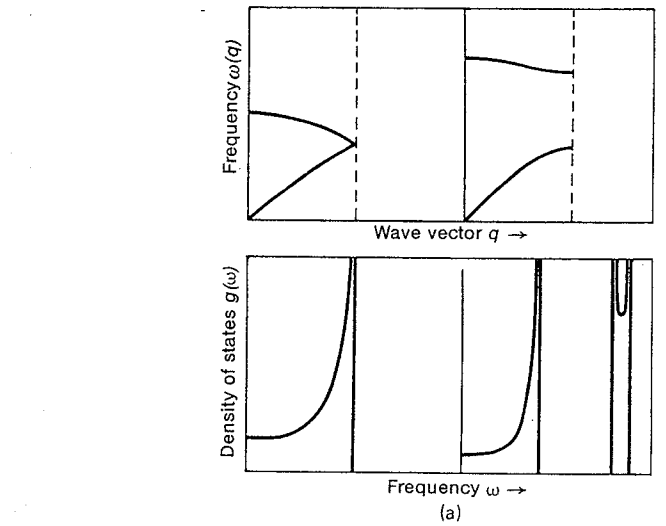
In principle the force constants can also be derived from the Schrödinger equation for the entire problem. For metals, the concept of pseudopotentials helps here to develop an approximation procedure for the calculation of dispersion curves which does not require the introduction of nonphysical spring forces. We can only point here to the literature covering this, in particular to *Harrison* [58].

3.3.6 The Density of States

Knowledge of the dispersion curves for the entire Brillouin zone, i.e., the function $\omega_j(\boldsymbol{q})$, allows us to calculate the density of states in accordance with (3.144).

For the one-dimensional case of the linear chain with two atoms in its basis, Fig. 3.12 shows the density of states for two different mass ratios. For equal masses (left), the density of states corresponds to the case of the simple linear chain. There is only one acoustic branch. The upper branch appearing in the figure is only the continuation of the lower acoustic branch, for, if one takes the linear chain with one atom per elementary cell to be a chain of double the lattice constant with two identical atoms per elementary cell, the associated Brillouin zone is only half as large. The acoustic branch in the usual representation has then to be reduced to the first Brillouin zone in the new representation, and forms an upper branch connected to the lower branch at the surface of the (one-dimensional) Brillouin zone. If one then assumes infinitesimally different masses for the two atoms in the basis, the branches separate and the upper branch becomes a "true optical branch". From this short discussion we can already see that the words "optical" and "acoustic", just like "transverse" and "longitudinal" are only meaningful in limiting cases.

In the two- and the three-dimensional cases, the density of states of different branches can overlap. The spectrum becomes more complicated in comparison with the one-dimensional case, where the density of states consists of the separate contributions of the two branches. As with the band structure $E_n(\boldsymbol{k})$, one needs the function $\omega_j(\boldsymbol{q})$ for the entire Brillouin zone for a calculation of the density of states. We shall not go into the methods of calculation here. Such spectra are often approximated by combining the density of states of the Debye and Einstein approximations. Fig. 3.12a shows that an acoustic branch can be roughly reproduced by a Debye spectrum [approximation $\omega_{ac}(q) \propto q$]. Optical branches are often very flat. If one totally neglects the frequency dependence, the optical branch can be approximated by a δ-function with an Einstein frequency. If one combines both, one can represent the density of states spectrum to the point where it can be used to calculate the temperature dependence of the specific heat. Fig. 3.12b presents such an approximation for tungsten and for lithium. In combining Debye and Einstein terms, the Debye temperature must



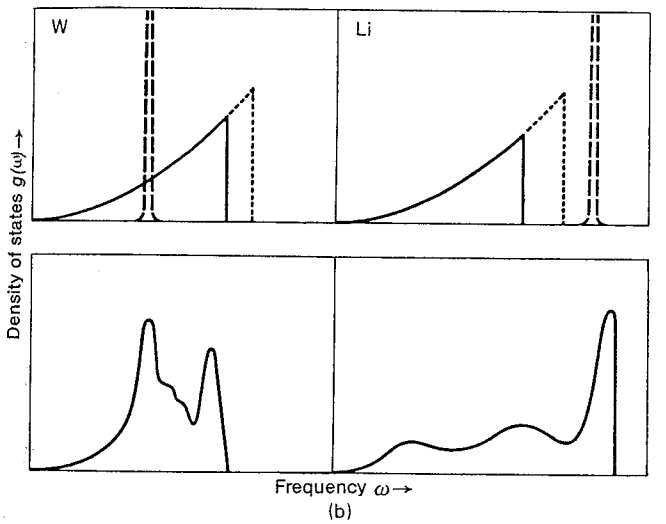


Fig. 3.12. (a) Dispersion curves $\omega(q)$ and density of states $g(\omega)$ for the linear chain with two atoms in the basis, for different mass relationships ($M_1 = M_2, M_1 \neq M_2$). (b) Approximation to the density of states of tungsten and lithium using Debye and Einstein terms (from *Leibfried* [106.VII/1]).

be lowered so that the summation over all states in both branches gives the total number of normal modes.

3.3.7 The Long Wavelength Limit: Acoustic Branch

The limiting case of long wavelengths (small q) is particularly interesting on several counts. We first consider the acoustic branch. We shall discuss the optical branch in the following section.

In the acoustic branch all atoms in the basis move in the same direction. For long wavelength the vibration amplitudes change only slowly from elementary cell to elementary cell. Atomic structure has then no part to play and a transition to a continuum is possible. We execute this as follows.

First we can think of the basis atoms as united at the centre of mass (total mass M). We have therefore only to consider a Bravais lattice and the equations of motion become

$$M\ddot{s}_{ni} = -\sum_{n'i'} \Phi_{ni}^{n'i'} s_{n'i'} . \tag{3.163}$$

We now define a (slowly changing) *displacement field* $s(r, t)$ which is identical with the discrete $s_n(t)$ at the lattice points

$$s(r = R_n, t) = s_n(t) . \tag{3.164}$$

We insert this displacement field into (3.163). We expand the field s at location $R_{n'}$ about location $R_n = 0$, and assume that contributions to the sum in (3.163) only come from regions where s does not alter much. We can then terminate the expansion at the first nonvanishing term

$$s_{n'i'} = s_{i'}(R_{n'}) = s_{i'}(0) + \sum_j \frac{\partial s_{i'}}{\partial r_j} R_{n'j} + \frac{1}{2} \sum_{kl} \frac{\partial^2 s_{i'}}{\partial r_k \partial r_l} R_{n'k} R_{n'l} . \tag{3.165}$$

Putting this expansion into (3.163), the first term disappears because of (3.100) and the second because of the condition $\Phi_{0i}^{n'i'} = \Phi_{0i}^{-n'i'}$. We are left with

$$M\ddot{s}_{ni} = -\frac{1}{2} \sum_{\substack{n'i' \\ kl}} \Phi_{ni}^{n'i'} R_{n'k} R_{n'l} \frac{\partial^2 s_{i'}}{\partial r_k \partial r_l} . \tag{3.166}$$

If we multiply this equation by the density $\rho = M/V_{WSC}$ (V_{WSC} = volume of a Wigner-Seitz cell) and introduce the abbreviation

$$C_{ii'kl} = \sum_{n'} \Phi_{ni}^{n'i'} R_{n'k} R_{n'l} \tag{3.167}$$

we finally find

$$\rho\ddot{s}_i = \sum_{i'kl} C_{ii'kl} \frac{\partial^2 s_{i'}}{\partial r_k \partial r_l} . \tag{3.168}$$

The $C_{ii'kl}$ possess a range of symmetries. In particular

$$C_{iklm} = C_{lmik} . \tag{3.169}$$

We do not want to derive this relation for the general case. If the forces between all lattice components are central, it follows from a general symmetry relation which we shall only state here. For central forces the potential can be written as

$$V = \sum_{ss'} v(|R_{s'} - R_s|) \tag{3.170}$$

and the force constants become

$$\Phi_{ni}^{n'i'} = g(|R_{n'} - R_n|) R_{n'i} R_{n'i'} . \tag{3.171}$$

If we put this in (3.167), we see that in this special case all the indices of the $C_{ii'kl}$ are interchangeable.

Hence we can also write (3.168) as

$$\rho\ddot{s}_i = \sum_k \sum_{mn} C_{ikmn} \frac{\partial}{\partial r_k} \frac{1}{2} \left(\frac{\partial s_m}{\partial r_n} + \frac{\partial s_n}{\partial r_m} \right) . \tag{3.172}$$

This however is precisely the equation of motion of an elastic continuum.

$(1/2)[(\partial s_m/\partial r_n) + (\partial s_n/\partial r_m)]$ is the strain tensor ε_{mn}, related by Hooke's law to the stress tensor

$$\sigma_{ik} = \sum_{mn} C_{ikmn} \varepsilon_{mn} . \tag{3.173}$$

We can then identify the *elastic tensor* C_{ikmn} of (3.173) with the corresponding tensor of (3.172) and obtain

$$\rho\ddot{s}_i = \sum_k \frac{\partial}{\partial r_k} \sigma_{ik} . \tag{3.174}$$

The connection with the mechanics of continua is thereby found.

We can use these results to match unknown parameters in the calculation of force constants (e.g., for the shell model), using the components of the elastic tensor which can be measured.

Stress and strain tensors, and hence also the elastic tensor, are symmetric. One usually combines the groups of indices 11, 22, 33, 23, 13, 12 into new indices 1, 2, 3, 4, 5, 6 and writes (3.173) then in the form

$$\sigma_\alpha = \sum_\beta C_{\alpha\beta}\varepsilon_\beta \,. \tag{3.175}$$

The (symmetric) 6 × 6-matrix of the $C_{\alpha\beta}$ has 21 independent elements (elastic constants). For reasons of symmetry many of these constants are zero or equal to each other.

A cubic crystal for example is described by three elastic constants $C_{11} = C_{22} = C_{33}$, $C_{12} = C_{23} = C_{31}$, $C_{44} = C_{55} = C_{66}$, while all other C_{ik} are zero.

An isotropic solid is described by two elastic constants. For these it is customary to use the three cubic elastic constants just mentioned, with the condition $C_{11} = C_{12} + 2C_{44}$.

More complicated crystals can have up to 21 different elastic constants. With central forces, i.e., where (3.171) applies, this number is reduced to 15 by further relations (*Cauchy relations*).

In the lattice dynamics of the previous section we looked particularly at solutions of the equations of motion which have wavelike character. We therefore seek here too solutions of (3.172) with the form

$$\boldsymbol{s} = \boldsymbol{e}\exp\left[\mathrm{i}(\boldsymbol{q}\cdot\boldsymbol{r} - \omega t)\right]. \tag{3.176}$$

For isotropic media this leads to the equation

$$\rho\omega^2\boldsymbol{e} = (C_{12} + C_{44})\boldsymbol{q}(\boldsymbol{q}\cdot\boldsymbol{e}) + C_{44}q^2\boldsymbol{e}\,. \tag{3.177}$$

It has one solution which corresponds to longitudinal waves ($\boldsymbol{q} \parallel \boldsymbol{e}$) and two which correspond to transverse waves ($\boldsymbol{q} \perp \boldsymbol{e}$). The dispersion relations for the two cases are

$$\rho\omega_\mathrm{L}^2 = (C_{12} + 2C_{44})q^2 = C_{11}q^2 \quad \text{and} \quad \rho\omega_\mathrm{T}^2 = C_{44}q^2\,. \tag{3.178}$$

In both cases ω is proportional to q. The waves propagate with the *longitudinal* or *transverse velocity of sound*, respectively,

$$c_\mathrm{L} = \sqrt{C_{11}/\rho}\,, \qquad c_\mathrm{T} = \sqrt{C_{44}/\rho}\,. \tag{3.179}$$

We note that the isotropic linear relation between ω and q in this approximation just corresponds to the Debye approximation. The constant $\bar{s}_j^{-3}$ used there becomes equal to c_L^{-3} or c_T^{-3} (as appropriate) for each of the three acoustic branches. Averaging over all three branches gives

$$\bar{s}^{-3} = c_\mathrm{L}^{-3} + 2c_\mathrm{T}^{-3}\,. \tag{3.180}$$

This is the expression to be inserted into the density of states (3.146).

We shall also discuss briefly the corresponding relations for cubic crystals. Here we have

$$\rho\ddot{s}_x = C_{11}\frac{\partial^2 s_x}{\partial x^2} + (C_{12} + C_{44})\left(\frac{\partial^2 s_y}{\partial x\partial y} + \frac{\partial^2 s_z}{\partial x\partial z}\right) + C_{44}\left(\frac{\partial^2 s_x}{\partial y^2} + \frac{\partial^2 s_x}{\partial z^2}\right). \tag{3.181}$$

The two further equations for the y- and z-components of $\boldsymbol{s}$ follow by cyclic interchange of x, y, z.

Here, too, longitudinal and transverse solutions only appear in particular directions, while in a general direction $\boldsymbol{e}$ is neither perpendicular nor parallel to $\boldsymbol{q}$.

3.3.8 The Long Wavelength Limit: Optical Branch

In the long wavelength limit of the optical branch, the basis atoms in the individual Wigner-Seitz cells vibrate relative to one another, while the motion in adjacent cells is practically identical. If we look in particular at a solid with two oppositely charged ions in its basis, the two ion sublattices vibrate rigidly opposite to each other.

We shall examine the instantaneous deviation $\boldsymbol{s}_\pm$ of an ion. Let its effective charge be $\pm e^*$. The displacement can then also be described formally by the introduction of a dipole with moment $\pm e^*\boldsymbol{s}_\pm$ (addition of a charge $\pm e^*$ at position $\boldsymbol{s}_\pm$ and a charge $\mp e^*$ at the original position to compensate the charge on the ion). Thus the polarization of a cell is given by $e^*(\boldsymbol{s}_+ - \boldsymbol{s}_-) \equiv e^*\boldsymbol{s}$. The charge displacement can however also lead to internal fields which induce further dipole moments in the lattice ions. This brings a further contribution to the polarization of the cell, of magnitude αE_eff ($\alpha = \alpha_+ + \alpha_-$). $\alpha_\pm$ is the polarizability of the ions and E_eff the effective field at the ion location. This local field is related to the macroscopic field for simple cubic lattices by the well-known relation

$$E_\mathrm{eff} = E + \frac{1}{3\varepsilon_0}P\,. \tag{3.182}$$

To derive this equation one takes a sphere of sufficiently large radius around the ion observed, and considers the region outside this sphere to be a homogeneous dielectric. Within the sphere the contributions of the other lattice ions to the effective field are summed. These contributions cancel for cubic lattices. Only the contribution given in (3.182) then remains.

Taking (3.182) to be valid, the total polarization for N Wigner-Seitz cells in the volume V_g is

$$\boldsymbol{P} = \frac{N}{V_\mathrm{g}}(e^*\boldsymbol{s} + \alpha\boldsymbol{E}_\mathrm{eff}) = \frac{N}{V_\mathrm{g}}\frac{e^*\boldsymbol{s} + \alpha\boldsymbol{E}}{1 - N\alpha/3V_\mathrm{g}\varepsilon_0}\,. \tag{3.183}$$

The equations of motion for the displacements are given by

$$\begin{aligned} M_+\ddot{s}_+ &= -k(s_+ - s_-) + e^*E_{\text{eff}} \\ M_-\ddot{s}_- &= +k(s_+ - s_-) - e^*E_{\text{eff}} \end{aligned} \tag{3.184}$$

where k is the proportionality constant of the restoring force. With the reduced mass $\overline{M} = M_+M_-/(M_+ + M_-)$ (3.184) leads to

$$\overline{M}\ddot{s} = -ks + e^*E_{\text{eff}}\,. \tag{3.185}$$

If we express the effective field here in terms of the macroscopic field, we obtain two equations linking s, E, and P. It is convenient to replace the vector by the vector $w = \sqrt{N\overline{M}/V_0}s$. Eqs. (3.183) and (3.184) then become

$$\begin{aligned} \ddot{w} &= b_{11}w + b_{12}E \\ P &= b_{21}w + b_{22}E \end{aligned} \tag{3.186}$$

with symmetric coefficient matrix ($b_{12} = b_{21}$).

We can relate the coefficients to parameters which can be measured experimentally. In the static case, $\ddot{w} = 0$ and hence

$$P = \left(b_{22} - \frac{b_{12}b_{21}}{b_{11}}\right)E = [\varepsilon(0) - 1]\varepsilon_0 E\,. \tag{3.187}$$

$\varepsilon(0)$ here is the *static dielectric constant*.

For externally applied electric fields of very high frequency the ions are no longer able to follow the rapidly changing forces. We then have $w = 0$ and, with $\varepsilon(\infty)$ the dielectric constant for this limiting case,

$$P = b_{22}E = [\varepsilon(\infty) - 1]\varepsilon_0 E\,. \tag{3.188}$$

We now look for solutions of (3.186) of the type $\exp[\mathrm{i}(\boldsymbol{q}\cdot\boldsymbol{r} - \omega t)]$. External fields are not involved. We divide w into an irrotational and a divergence-free component ($w = w_{\mathrm{L}} + w_{\mathrm{T}}$, $\nabla \times w_{\mathrm{L}} = 0$, $\nabla\cdot w_{\mathrm{T}} = 0$). The two components correspond exactly in our plane wave approximation to longitudinal and transverse waves, respectively. We further take account of the condition $\nabla\cdot\boldsymbol{D} = \nabla\cdot(\varepsilon_0 E + P) = 0$. From this we then have

$$\nabla\cdot(\varepsilon_0 E + b_{21}w + b_{22}E) = 0 \tag{3.189}$$

with solution

$$E = -\frac{b_{21}}{\varepsilon_0 + b_{22}}w_{\mathrm{L}}\,. \tag{3.190}$$

Putting this into the first (3.186) we find

$$\ddot{w}_{\mathrm{T}} + \ddot{w}_{\mathrm{L}} = b_{11}(w_{\mathrm{T}} + w_{\mathrm{L}}) - \frac{b_{12}b_{21}}{\varepsilon_0 + b_{22}}w_{\mathrm{L}}\,. \tag{3.191}$$

If we separate the equation into irrotational and divergence-free components we obtain equations

$$\ddot{w}_{\mathrm{T}} = b_{11}w_{\mathrm{T}}\,, \qquad \ddot{w}_{\mathrm{L}} = \left(b_{11} - \frac{b_{12}b_{21}}{\varepsilon_0 + b_{22}}\right)w_{\mathrm{L}} = b_{11}\frac{\varepsilon(0)}{\varepsilon(\infty)}w_{\mathrm{L}}\,. \tag{3.192}$$

If we call the frequencies of the transverse and longitudinal waves ω_{T} and ω_{L}, respectively (these are just the limiting frequencies of the corresponding optical branches as $\boldsymbol{q}$ goes to zero), we finally have

$$\begin{aligned} b_{11} &= -\omega_{\mathrm{T}}^2 = -\omega_{\mathrm{L}}^2\frac{\varepsilon(\infty)}{\varepsilon(0)}\,, \qquad b_{22} = [\varepsilon(\infty) - 1]\varepsilon_0\,, \\ b_{12} &= b_{21} = \omega_{\mathrm{T}}\{[\varepsilon(0) - \varepsilon(\infty)]\varepsilon_0\}^{1/2}\,. \end{aligned} \tag{3.193}$$

Therefore

$$\omega_{\mathrm{L}}^2 = \frac{\varepsilon(0)}{\varepsilon(\infty)}\omega_{\mathrm{T}}^2\,. \tag{3.194}$$

This last relationship is known as the *Lyddane-Sachs-Teller* relation. For ionic crystals with two atoms in the basis of a Wigner-Seitz cell, it relates the limiting frequencies of the two types of optical vibrations. In the form given here the relation is of course restricted to cubic crystals, in accordance with the remarks made following (3.182).

The classical treatment of optical vibrations given here is important for the interaction between phonons and photons discussed in Section 4.2.

3.4 Spin-Spin Interaction: Magnons

3.4.1 Introduction

We have mostly neglected the spin of the electrons and lattice ions in the elementary excitations examined so far. Apart from a short discussion in Section 2.2.15 of the effect of spin-orbit coupling on the band structure of a solid, spin has only entered our considerations through the Pauli principle. The Pauli principle is responsible for the exchange interaction (Sec. 1.3) which we took account of in the one-electron Schrödinger equation (1.32). On the other hand,

The equations of motion for the displacements are given by

$$M_+\ddot{s}_+ = -k(s_+ - s_-) + e^*E_{\text{eff}}$$
$$M_-\ddot{s}_- = +k(s_+ - s_-) - e^*E_{\text{eff}} \tag{3.184}$$

where k is the proportionality constant of the restoring force. With the reduced mass $\overline{M} = M_+M_-/(M_+ + M_-)$ (3.184) leads to

$$\overline{M}\ddot{s} = -ks + e^*E_{\text{eff}}\,. \tag{3.185}$$

If we express the effective field here in terms of the macroscopic field, we obtain two equations linking s, E, and P. It is convenient to replace the vector s by the vector $w = \sqrt{N\overline{M}/V_g}s$. Eqs. (3.183) and (3.184) then become

$$\ddot{w} = b_{11}w + b_{12}E$$
$$P = b_{21}w + b_{22}E \tag{3.186}$$

with symmetric coefficient matrix ($b_{12} = b_{21}$).

We can relate the coefficients to parameters which can be measured experimentally. In the static case, $\ddot{w} = 0$ and hence

$$P = \left(b_{22} - \frac{b_{12}b_{21}}{b_{11}}\right)E = [\varepsilon(0) - 1]\varepsilon_0 E\,. \tag{3.187}$$

$\varepsilon(0)$ here is the *static dielectric constant*.

For externally applied electric fields of very high frequency the ions are no longer able to follow the rapidly changing forces. We then have $w = 0$ and, with $\varepsilon(\infty)$ the dielectric constant for this limiting case,

$$P = b_{22}E = [\varepsilon(\infty) - 1]\varepsilon_0 E\,. \tag{3.188}$$

We now look for solutions of (3.186) of the type $\exp[i(\boldsymbol{q}\cdot\boldsymbol{r} - \omega t)]$. External fields are not involved. We divide w into an irrotational and a divergence-free component ($w = w_L + w_T$, $\nabla \times w_L = 0$, $\nabla\cdot w_T = 0$). The two components correspond exactly in our plane wave approximation to longitudinal and transverse waves, respectively. We further take account of the condition $\nabla\cdot\boldsymbol{D} = \nabla\cdot(\varepsilon_0\boldsymbol{E} + \boldsymbol{P}) = 0$. From this we then have

$$\nabla\cdot(\varepsilon_0\boldsymbol{E} + b_{21}w + b_{22}\boldsymbol{E}) = 0 \tag{3.189}$$

with solution

$$E = -\frac{b_{21}}{\varepsilon_0 + b_{22}}w_L\,. \tag{3.190}$$

Putting this into the first (3.186) we find

$$\ddot{w}_T + \ddot{w}_L = b_{11}(w_T + w_L) - \frac{b_{12}b_{21}}{\varepsilon_0 + b_{22}}w_L\,. \tag{3.191}$$

If we separate the equation into irrotational and divergence-free components we obtain equations

$$\ddot{w}_T = b_{11}w_T\,, \qquad \ddot{w}_L = \left(b_{11} - \frac{b_{12}b_{21}}{\varepsilon_0 + b_{22}}\right)w_L = b_{11}\frac{\varepsilon(0)}{\varepsilon(\infty)}w_L\,. \tag{3.192}$$

If we call the frequencies of the transverse and longitudinal waves ω_T and ω_L, respectively (these are just the limiting frequencies of the corresponding optical branches as $\boldsymbol{q}$ goes to zero), we finally have

$$b_{11} = -\omega_T^2 = -\omega_L^2\frac{\varepsilon(\infty)}{\varepsilon(0)}\,, \qquad b_{22} = [\varepsilon(\infty) - 1]\varepsilon_0\,,$$
$$b_{12} = b_{21} = \omega_T\{[\varepsilon(0) - \varepsilon(\infty)]\varepsilon_0\}^{1/2}\,. \tag{3.193}$$

Therefore

$$\omega_L^2 = \frac{\varepsilon(0)}{\varepsilon(\infty)}\omega_T^2\,. \tag{3.194}$$

This last relationship is known as the *Lyddane-Sachs-Teller* relation. For ionic crystals with two atoms in the basis of a Wigner-Seitz cell, it relates the limiting frequencies of the two types of optical vibrations. In the form given here the relation is of course restricted to cubic crystals, in accordance with the remarks made following (3.182).

The classical treatment of optical vibrations given here is important for the interaction between phonons and photons discussed in Section 4.2.

3.4 Spin-Spin Interaction: Magnons

3.4.1 Introduction

We have mostly neglected the spin of the electrons and lattice ions in the elementary excitations examined so far. Apart from a short discussion in Section 2.2.15 of the effect of spin-orbit coupling on the band structure of a solid, spin has only entered our considerations through the Pauli principle. The Pauli principle is responsible for the exchange interaction (Sec. 1.3) which we took account of in the one-electron Schrödinger equation (1.32). On the other hand,

'e have not yet paid any attention to the spin of the lattice ions. The exchange ıteraction can give rise to collective excitations of the ion spin system, known s spin waves. The associated quanta are called *magnons*.

Collective excitations are low-lying excited states above a ground state. The round state of the spin system is therefore important. If all spins are arranged in ıe same direction, the solid is a *ferromagnet*. In *ferrimagnets* and *antiferromag-ets*, spins are only aligned in different sublattices of the solid. In the next section e shall look at spin waves in ferromagnets and in this simple example become miliar with the basis of the magnon concept. We can then readily extend these sults to ferri- and antiferromagnetism. This we do in Section 3.4.3.

In examining elementary excitations in magnetic solids, we are only touching ı a fraction of the important magnetic phenomena. We therefore extend our nsiderations in Section 3.4.4 with a short description of the molecular field proximation, which is important in explaining the properties of a ferromagnet ar the Curie temperature.

It is not only the spin of the lattice ions which is significant where magnetism concerned. We saw earlier that the division into lattice ions and valence ctrons is an idealization which does not always hold. This is particularly so ıen *d*-electrons enter into the picture. In Section 3.4.5 we therefore discuss pects of ordered magnetism involving valence and conduction electrons.

Since we want to put the emphasis of the presentation on elementary excita-ns, we shall only cover a part of the wide field of magnetism in this chapter. r more comprehensive presentations we refer the reader to the general text-oks and monographs [78–83]. For more extensive treatments of spin waves in rticular, we recommend the article by *Keffer* [106, XVIII/2], the contribution *Elliott* in [117], and the relevant chapters of the book by *Harrison* [13] and *tel* [16]. There are numerous contributions on magnetism in transition metals the conference proceedings [102.37] (see also *Biondi* in [113a]). Finally we o point to the review articles by *Kittel* [101.22] and *Nagamiya* [101.20].

.2 Spin Waves in Ferromagnets: Magnons

our description of excited states we have so far always referred to a ground te in which the net spin of the valence electrons is zero. The reason for this s that each of the energy states in the one-electron approximation can be upied by two electrons with opposite spins. The ground state was regarded hat state in which all the energy levels were filled up to a limiting energy E_F, l all levels above that were empty. Such a state has neither net spin nor net mentum.

This argument need no longer apply when the mutual interaction of the trons is taken into account. This can be easily deduced from the Hartree-k approximation. From the result in Section 3.1.3, a Hartree-Fock electron a mean kinetic energy proportional to k_F^2 and a mean exchange energy portional to $-k_F$ (if the exchange integral itself is positive). The energy of

the spin-compensated ground state is then $N(ak_F^2 - bk_F)$. If we align all the spins parallel, the electrons occupy states in a sphere of double the volume in k-space. The energy of such a ferromagnetic state becomes $N(a2^{2/3}k_F^2 - b2^{1/3}k_F)$. This energy lies below the energy of the spin-compensated state if k_F is less than $0.44b/a$. The ferromagnetic state is therefore favoured in an electron gas of low density (small Fermi sphere radius).

Coulomb interaction has been ignored in this description. It prevents spin alignment when the density is low. In spite of this, the example shows that exchange interaction can be responsible for the spin correlation observed in *ferromagnets* (uniformly aligned spins), and in *antiferromagnets* and *ferrimagnets* (different alignment of the correlated spin system in different sublattices).

We take one further step in this consideration of the ferromagnetic Hartree-Fock electron gas. The energy of the ground state is given by (1.20), where all wave functions have identical spin. Since the spin functions are orthogonal, this means that they just drop out of (1.20) and the q_i can be replaced by r_i. We now examine an excited state in which the spin of *one* electron is reversed. The energy of this state is obtained from (1.20) by associating with $N-1$ electrons the spin function $\alpha(j)$, and with the ith electron the spin function $\beta(i)$. All the exchange integrals connecting this one electron with all other electrons then disappear. The difference in energy between the excited and ground states is

$$E_i - E_0 = \frac{e^2}{8\pi\varepsilon_0} \sum_{j(\neq i)} \int \frac{\varphi_j^*(r_1)\varphi_j(r_2)\varphi_i^*(r_2)\varphi_i(r_1)}{|r_1 - r_2|} d\tau_1 d\tau_2 \equiv \frac{1}{2} \sum_{j(\neq i)} J_{ij} \,. \tag{3.195}$$

The excited state $|i\rangle$ considered is degenerate with all states $|n\rangle$ in which another electron likewise has opposite spin. The solution of the Schrödinger equation for an excited state with this energy has therefore to be constructed as a linear combination of all $|n\rangle$: $\Phi = \sum_n a_n|n\rangle$.

We have here a behaviour analogous to that of lattice vibrations where kinetic energy fed to one lattice ion spreads itself by Coulomb interaction over all ions. The excitation which results can be described by wavelike states, and correspondingly the problem being considered here has wavelike solutions [$a_n \sim \exp(i\boldsymbol{k}\cdot\boldsymbol{r}_n)$]. The energy needed to reverse a spin is distributed throughout the entire spin system (*spin waves*, Fig. 3.13). Spin waves can be quantized like lattice waves. Thus we have the *magnons* as new collective excitations.

We do not however want to use the Hartree-Fock equations of the free electron gas to study this new type of elementary excitation, but rather a more general approach. The spins, whose correlation in ferromagnetism and related phenomena leads to a spontaneous magnetic moment, are mostly localized on the lattice ions. Furthermore several electrons contribute to the total ion spin. The ferromagnetic state is then the result of exchange interaction between the total spins of the different ions.

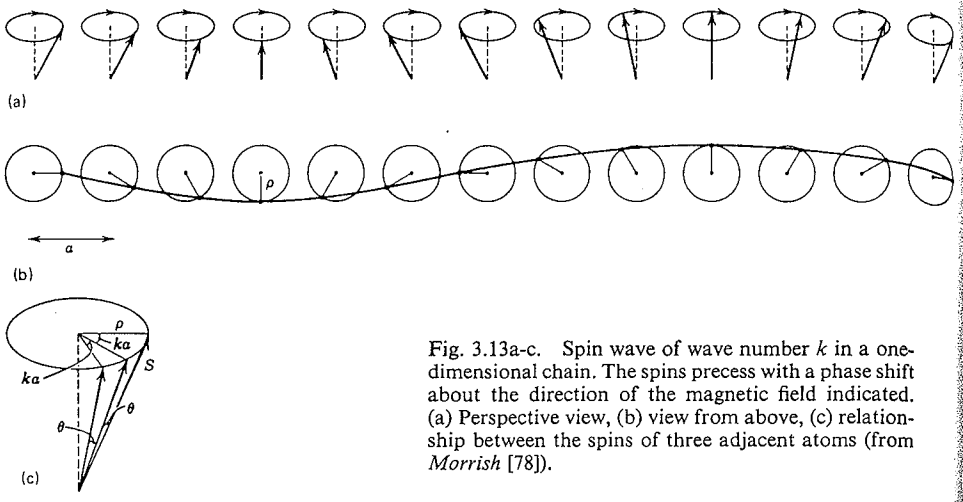


Fig. 3.13a-c. Spin wave of wave number k in a one-dimensional chain. The spins precess with a phase shift about the direction of the magnetic field indicated. (a) Perspective view, (b) view from above, (c) relationship between the spins of three adjacent atoms (from *Morrish* [78]).

The exchange Hamiltonian can be formally replaced by an operator introduced by Heisenberg

$$H = -\sum_{ij}{}' J_{ij} \boldsymbol{S}_i \cdot \boldsymbol{S}_j . \tag{3.196}$$

Here J_{ij} is an exchange integral. The $\boldsymbol{S}_i$ are the vector spin operators of the ith lattice ion. The sum is over all pairs of lattice ions.

For the case $s = 1/2$ the spin operators in (3.196) are given by the Pauli matrices

$$S_x = \frac{1}{2}\begin{vmatrix} 0 & 1 \\ 1 & 0 \end{vmatrix}, \qquad S_y = \frac{1}{2}\begin{vmatrix} 0 & -\mathrm{i} \\ \mathrm{i} & 0 \end{vmatrix}, \qquad S_z = \frac{1}{2}\begin{vmatrix} 1 & 0 \\ 0 & -1 \end{vmatrix} \tag{3.197}$$

with the commutation relations

$$[S_\lambda, S_\mu] = \mathrm{i} S_\nu \qquad \lambda,\mu,\nu = x, y, z \text{ and cycl.} \tag{3.198}$$

The Pauli matrices operate on spinors $\begin{vmatrix} \alpha \\ \beta \end{vmatrix}$. For the spin functions α, β follows

$$S_z\alpha = \frac{1}{2}\alpha, \qquad S_z\beta = -\frac{1}{2}\beta, \qquad S^2\alpha = \frac{3}{4}\alpha, \qquad S^2\beta = \frac{3}{4}\beta . \tag{3.199}$$

α thus describes the state with $s_z = +1/2$ and β the state with $s_z = -1/2$. The eigenvalue of the operator S^2 is $s(s + 1) = 3/4$.

For a total spin $s = n/2$, a spin operator S with $[S_\lambda, S_\mu] = iS_\nu$ can correspondingly be introduced. The matrices associated with its components are then $n + 1$-dimensional and there are $n + 1$ spin functions. The eigenvalues of S_z are $-s, -s + 1, \ldots, s - 1, s$ and of S^2: $s(s + 1)$.

In place of S_x and S_y, one often introduces the operators $S_+ = S_x + iS_y$ and $S_- = S_x - iS_y$. For the case $s = 1/2$ one then has

$$S_+\alpha = 0, \qquad S_+\beta = \alpha, \qquad S_-\alpha = \beta, \qquad S_-\beta = 0 . \tag{3.200}$$

S_+ thus converts a minus-spin into a plus-spin and vice versa. For $s > 1/2$, the S_+ raise the total spin by unity and the S_- lower it by unity.

With these spin operators we now calculate the expectation value of the operator (3.196) in the case $s = 1/2$ for a pair of indices i, j. We find

$$E_{\uparrow\uparrow} = -J_{ij}\langle\alpha_i\alpha_j|\boldsymbol{S}_i\cdot\boldsymbol{S}_j|\alpha_i\alpha_j\rangle, \qquad E_{\uparrow\downarrow} = -J_{ij}\langle\alpha_i\beta_j|\boldsymbol{S}_i\cdot\boldsymbol{S}_j|\alpha_i\beta_j\rangle , \tag{3.201}$$

or using (3.197) and taking account of the orthonormalization of the spin functions α and β,

$$E_{\uparrow\uparrow} = -\frac{1}{4}J_{ij}, \qquad E_{\uparrow\downarrow} = +\frac{1}{4}J_{ij} . \tag{3.202}$$

The difference in energy between these two possibilities is therefore $J_{ij}/2$ and the difference between a state in which all spins are aligned and one in which the ith spin is reversed is $\sum_{j(\neq i)} J_{ij}/2$. This agrees with (3.195).

The exchange interaction is therefore formally reproduced by the operator (3.196) just as though it were explicitly a spin-spin interaction.

Since the exchange interaction between nearest neighbours is dominant, one usually restricts consideration to these terms and therefore includes in the summation over j only those terms formed by an $\boldsymbol{R}_j = \boldsymbol{R}_i + \boldsymbol{R}_\delta$ where $\boldsymbol{R}_\delta$ is a vector to the nearest neighbour ($\delta = 1, 2, \ldots \nu$) of the ith ion. If further one immediately assumes $J_{i,i+\delta} = J$ for all δ (thereby limiting oneself to simple lattices) the effective interaction becomes

$$H = -J\sum_{i,\delta} \boldsymbol{S}_i\cdot\boldsymbol{S}_{i+\delta} . \tag{3.203}$$

We now assume that the ion spins in the ground state are so aligned that their z-components have the maximum values s. We describe the wave function for the ground state as a product of spin functions $|s\rangle_n$ which represent the spin of the nth ion in state s: $\Phi_0 = \prod_n |s\rangle_n$. We introduce spin-raising and spin-lowering operators S_+ and S_- so that the Hamiltonian (3.203) takes the form

$$H = -J\sum_{ij}{}' \left[S_{iz}S_{jz} + \frac{1}{2}(S_{i+}S_{j-} + S_{i-}S_{j+}) \right] \qquad (j = i + \delta) . \tag{3.204}$$

Application of this operator to the ground state yields

$$E_0 = -s^2 J \sum_{i,i+\delta} 1 = -Js^2 \nu N, \tag{3.205}$$

since the application of a spin-increase operator to a function with maximum spin must lead to zero. ν is again the number of nearest neighbours of the ion.

We consider now the state $\Phi_m = S_{m-} \prod_n |s\rangle_n$ in which the mth spin is diminished by one. We then have

$$H\Phi_m = -J \sum_{ij}{}' \left[S_{iz} S_{jz} S_{m-} + \frac{1}{2}(S_{i+} S_{j-} S_{m-} + S_{i-} S_{j+} S_{m-}) \right] \Phi_0 . \tag{3.206}$$

We can rearrange the product of spin operators on the right of (3.206) using the commutation relations which follow from (3.198): $[S_+, S_-] = 2S_z$, $[S_-, S_z] = S_-$, $[S_z, S_+] = S_+$. We then obtain

$$H\Phi_m = E_0 \Phi_m + 2Js \sum_{\delta} (\Phi_m - \Phi_{m+\delta}) . \tag{3.207}$$

Φ_m is not therefore an eigenstate of H. Such a state must rather be made up of all degenerate $\Phi_m = S_{m-}\Phi_0$: $\Phi = \sum_m a_m \Phi_m$. In view of the translation invariance of the lattice the a_m have the form $\exp(i\boldsymbol{k}\cdot\boldsymbol{R}_m)$. It then follows that

$$H\Phi = \sum_m \exp(i\boldsymbol{k}\cdot\boldsymbol{R}_m) H\Phi_m = [E_0 + 2Jvs(1 - \gamma_k)]\Phi , \tag{3.208}$$

where

$$\gamma_k = \frac{1}{\nu} \sum_{\delta} \exp(i\boldsymbol{k}\cdot\boldsymbol{R}_\delta) . \tag{3.209}$$

The energy of the excited state is therefore

$$E_k = E_0 + 2Jvs(1 - \gamma_k) , \tag{3.210}$$

where $\boldsymbol{k}$ (for cyclic boundary conditions) is limited to the N values inside a Brillouin zone in $\boldsymbol{k}$-space.

For small $\boldsymbol{k}$ this becomes

$$E_k = E_0 + Js \sum_{\delta} (\boldsymbol{k}\cdot\boldsymbol{R}_\delta)^2 . \tag{3.211}$$

Eqs. (3.210) and (3.211) are dispersion relations for spin waves.

To quantize the spin waves, we start out from the following line of thought: All spins are aligned in the ground state. Their z-components have the maximum value $s_z = s$. We can describe an excited state by specifying by how many units the s_z differ from the maximum value. If we call this number n and add the index of the ion in question, then each state is described by specifying the $n_1, n_2, \ldots n_N$ ($n_i = 0, 1, 2, \ldots, 2s$) and we can describe this state in an occupation number representation by a state vector $|n_1, n_2, \ldots, n_N\rangle$ for bosons (see Appendix). Correspondingly we can introduce creation and annihilation operators in accordance with (A.15). $a_j^+ a_j$ is then an operator whose eigenstates describe the departures of the spin of the jth ion from the maximum value.

The a^+ and a can readily be obtained from the S_+ and S_- introduced earlier. For the $S_\pm$ we find from the commutation relations

$$\begin{aligned} S_{j+}|n_j\rangle &= \sqrt{2s + 1 - n_j}\sqrt{n_j}|n_j - 1\rangle \\ S_{j-}|n_j\rangle &= \sqrt{2s - n_j}\sqrt{n_j + 1}|n_j + 1\rangle \\ S_{jz}|n_j\rangle &= (s - n_j)|n_j\rangle , \end{aligned} \tag{3.212}$$

where only the state of the jth ion has been given in the wave functions.

A comparison of (3.212) and (A.15) then shows the link between the a^+, a and the S_+, S_-, and S_z

$$S_+ = \sqrt{2s - a^+ a}\, a , \quad S_- = a^+ \sqrt{2s - a^+ a} , \quad S_z = s - a^+ a . \tag{3.213}$$

The expansion of the roots has to be inserted into the operators on the right side of the first two equations here.

We can use these relations to rewrite the Hamiltonian (3.204) in terms of creation and annihilation operators (*Holstein-Primakoff transformation*).

It is however convenient to go a step further. The a_j^+ and a_j change the spin of the jth ion. We saw already that through the exchange interaction such a change in spin is propagated through the whole spin system. We have therefore to include a transformation to creation and annihilation operators of the spin wave quanta. This corresponds to the transition from atomic to normal coordinates which we carried out for lattice vibrations before the quantization.

The corresponding transformation here is

$$a_j^* = \frac{1}{\sqrt{N}} \sum_k \exp(i\boldsymbol{k}\cdot\boldsymbol{R}_j) b_k^+ , \qquad a_j = \frac{1}{\sqrt{N}} \sum_k \exp(-i\boldsymbol{k}\cdot\boldsymbol{R}_j) b_k . \tag{3.214}$$

The new operators then satisfy the same commutation relations

$$[b_k, b_{k'}^+] = \delta_{kk'} , \qquad [b_k^+, b_{k'}^+] = [b_k, b_{k'}] = 0 \tag{3.215}$$

and we also have

$$\sum_j a_j^+ a_j = \sum_k b_k^+ b_k . \tag{3.216}$$

Transforming the Hamiltonian to the b_k^+, b_k presents the problem that in (3.213) $a_j^+ a_j$, i.e., sums over products of the b_k^+ and b_k, occur under the square

oot. If one restricts oneself to small departures from the ground state—and his is the only region where the concept of elementary excitations is reasonable –and hence to small n_j, one can terminate the expansion of the root early. $'_+$ is then a series with operators of the form b_k, $b_k^+ b_{k'} b_{k''}$, etc., S_- a series with perators b_k^+, $b_k^+ b_{k'}^+ b_{k''}$, etc. Additionally exponential functions appear containing summations over k, k', k'' in the exponent.

In the sum over the i, j in (3.204), relations between these k, k', k'' then ollow so that finally H can be written as a series, whose terms up to the fourth rder in the b_k have the following form for a lattice with a centre of inversion $\gamma_k = \gamma_{-k}$):

$$H = E_0 + \sum_k 2Jvs(1 - \gamma_k) b_k^+ b_k + \frac{vJ}{2N} \sum_{kk'\kappa} (\gamma_{k-\kappa} + \gamma_{k'} - 2\gamma_{k-\kappa-k'}) b_{k-\kappa}^+ b_{k'+\kappa}^+ b_{k'} b_k + \cdots , \tag{3.217}$$

here γ_k is again given by (3.209).

The first term is the energy of the ground state, the second is the energy ontained in the magnons. From (3.217) the energy of a magnon is the expresion already given in (3.210)

$$\hbar\omega_k = 2Jvs(1 - \gamma_k) . \tag{3.218}$$

$b_k^+ b_k$ is the magnon particle number operator.

The other terms of (3.217) describe the magnon-magnon interaction. The hird term in particular contains processes in which two magnons k, k' are anihilated and two magnons $k - \kappa$, $k' + \kappa$ created, with conservation of total omentum, or put another way, processes in which momentum κ is transferred rom one magnon to another.

This term also contains processes in which $\kappa = 0$ or $k' = k - \kappa$. Such terms ontribute to the magnon energy (3.218) and can be interpreted as renormalizaon of the magnon energy through exchange interaction.

As a simple application of the results, we want to estimate the energy contained in the magnons and hence their contribution to the specific heat. The nly difference compared with the case treated in Section 3.3.4 is the different orm of the dispersion relations for phonons and magnons. While phonon nergy increases linearly for small q, magnon energy increases as the square of . Restricting ourselves to the isotropic case $\hbar\omega_k \sim k^2$, the magnon energy ollows correspondingly as

$$E = \sum_k \frac{\hbar\omega_k}{\exp(\hbar\omega_k/k_B T) - 1} \propto \int d\tau_k \frac{k^2}{\exp(\alpha k^2/T) - 1} \propto \int_0^{k_{max}} \frac{k^4 dk}{\exp(\alpha k^2/T) - 1} . \tag{3.219}$$

Here we have again replaced the summation in k-space by an integral.

Since the estimate is only valid for low temperatures, when few magnons are excited, we can replace the upper limit k_{max} of the magnon dispersion spectrum in the integral with infinity. Rearrangement in terms of a dimensionless integration variable then leads to a factor $T^{5/2}$, and hence the specific heat is proportional to $T^{3/2}$, in agreement with experiment.

The temperature dependence of the magnetization can be calculated in a similar way. The departure of the magnetization from saturation $\Delta M = M(T) - M(0)$ is proportional to the average number of magnons $\sum_k \bar{n}_k$, i.e., from (3.219) it is proportional to an integral with k^2 rather then k^4 in the numerator of the integrand. This also leads here to a $T^{3/2}$ law. Corrections to both $T^{3/2}$-laws for specific heat and magnetization become necessary on several counts at higher temperatures. The range of validity is limited most of all by the magnon-magnon interaction, and by the replacement of (3.211) by an isotropic k^2-law. Since the concept of elementary excitations is only useful as long as mutual interactions between the excitations can be neglected, we shall not be interested here in better approximations. We shall turn our attention later to temperatures near the Curie temperature of a ferromagnet.

Our model so far restricts the results to solids which in the ground state have a spin system bound to the ions of a Bravais lattice. We therefore have to extend the theory in two directions.

1) Non-Bravais lattices. This includes the case of antiferro- and ferrimagnetism.
2) Ferromagnetic metals. In this case the spin of the nonlocalized valence electrons plays a major role.

3.4.3 Spin Waves in Lattices with a Basis, Ferri-, and Antiferromagnetism

For Bravais lattices, the dispersion relation (3.218) gives a k-dependence of the magnon energy which, similar to an acoustic branch of the phonon spectrum, starts with zero energy at $k = 0$ and increases up to the surface of the Brillouin zone. For *lattices with a basis* we must expect there to be other branches in the magnon spectrum, branches which correspond to the optical phonons. It will not be possible in such lattices to restrict the Heisenberg operator to exchange interactions between nearest neighbours. The different basis atoms form sublattices and, in addition to the interaction within a given sublattice, that between different sublattices is important. However, an extension of our model is also needed for other reasons. In most cases the ions in the individual sublattices will be different. The sublattices will then possess different total spin and often also different directions for the (individually parallel aligned) spin systems. The ground state will admittedly then show a magnetic moment. However, this will be the vector sum of the spins of the sublattices. For two sublattices with opposite spins it will therefore be the difference in the spins. Such a *ferrimagnet* reveals differences compared with a ferromagnet.

Before going into these questions we treat a simpler case which can already indicate what is important. Thus far we have assumed that the spins of nearest neighbours in a Bravais lattice of a ferromagnet are all aligned parallel in the ground state by exchange interaction. For this it is important that the exchange ntegral be positive. The case of negative exchange integral, on the other hand, s also possible. The preferred condition is then for nearest neighbours to have antiparallel spin. In the ground state—or so we assume at least initially—we hen find two sublattices of identical atoms but with opposite spin directions. This is the case of an *antiferromagnet* with mutually compensating spontaneous nagnetic moments of the two sublattices.

We can treat this model using the operator (3.203). Let the exchange integral between the lattice ions (assumed to be the same type) be negative. We denote ts absolute value by J and write

$$H = +J \sum_{i,\delta} \boldsymbol{S}_i \cdot \boldsymbol{S}_{j+\delta} . \tag{3.220}$$

We run into a difficulty in constructing the wave function for the ground tate. In the ferromagnetic case we could realize the ground state in only one way, namely by alignment of all spins in a preferred direction, which we chose as the -axis. A negligibly small magnetic field could serve to set up the ground state, nd this can be taken account of by an additive term in the Hamiltonian. In the ame way we can define a preferred direction here too and call it the z-axis. However, we still have to decide which ions in the uniform (apart from spin etting) lattice we want to combine into the sublattice with spin $+$, and which nto the sublattice with spin $-$. The two possibilities are degenerate, and in rder to accentuate one, and thereby stabilize the state, we have to introduce a mall finite magnetic field (*anisotropy field*), which is positive at the ions of the ne sublattice and negative at the ions of the other. Fields such as these, which re small compared with the other internal fields (see further below), are also bserved experimentally. One can take account of them in the Hamiltonian by n additive term of the sort

$$g\mu_B B_A \left(\sum_b S_{bz} - \sum_a S_{az} \right) . \tag{3.221}$$

Here we have designated the anisotropy field B_A and given the two sublattices he indices a and b. We shall not need the anisotropy field at first in our considerations, but we shall return to it later.

For each sublattice we must introduce creation and annihilation operators by means of a Holstein-Primakoff transformation. The transition to magnon operators is then possible, analogous to (3.214). Insertion into (3.220) and expansion of the roots finally yields instead of (3.217), with terms up to second order in the magnon operators,

$$H = -2NJvs^2 + 2Jvs\left[\sum_k (b_{ak}^+ b_{ak} + b_{bk}^+ b_{bk}) + \sum_k \gamma_k (b_{ak}^+ b_{bk}^+ + b_{ak} b_{bk}) \right] . \tag{3.222}$$

Here N is the number of ions in a sublattice. The first term is the energy of the undisturbed state. The second describes spin waves in the respective sublattices. The third corresponds to an interaction between the sublattices, in which a pair of magnons is respectively created or annihilated with zero effective spin change. This interaction can be removed by introducing magnon operators which describe combined spin waves in the two sublattices.

For this, one introduces the following operators:

$$\begin{aligned} c_{1k} &= u_k b_{ak} - v_k b_{bk}^+ , & c_{1k}^+ &= u_k b_{ak}^+ - v_k b_{bk} \\ c_{2k}^+ &= u_k b_{bk} - v_k b_{ak}^+ , & c_{2k} &= u_k b_{bk}^+ - v_k b_{ak} \end{aligned} \tag{3.223}$$

with real u_k, v_k, $u_k^2 - v_k^2 = 1$ and $[c_{1,2k}, c_{1,2k}^+] = 1$, $[c_{1k}, c_{2k}] = 0$.

The u_k and v_k are defined such that the coefficients of the mixed terms $c_{1k}c_{2k}$ and $c_{1k}^+ c_{2k}^+$ become zero. We are then left with

$$H = -2NvJs(s+1) + \sum_k \hbar\omega_k (c_{1k}^+ c_{1k} + c_{2k}^+ c_{2k} + 1) , \tag{3.224}$$

where

$$\hbar\omega_k = +2Jvs\sqrt{1 - \gamma_k^2} . \tag{3.225}$$

Along with the energy of the ground state, the Hamiltonian contains magnon particle number operators. For each k there are two different magnons, distinguished by the indices 1, 2. Clearly the ground state of the system (no magnons) is

$$E_0 = -2NvJs(s+1) + 2Jvs \sum_k \sqrt{1 - \gamma_k^2} . \tag{3.226}$$

If γ_k were zero, E_0 would be given by $E_0 = -2NJvs^2$, the energy of the strictly antiparallel-ordered lattice. Since $\gamma_k \neq 0$, the second term is less than $2NJvs$. The *ground state* is not therefore strictly aligned. Each sublattice has a small amount of disorder in its spin alignment.

Equation (3.225) is the dispersion relation for the antiferromagnetic magnons. The branches for the different magnons (indices 1 and 2) split up in an external magnetic field. For small k, $\sqrt{1 - \gamma_k^2} \propto k$ for simple lattices. In contrast to ferromagnetism the magnon energy increases here linearly with k.

Here we should point out a correction using the stabilizing anisotropy field. For the case in which an external magnetic field B and [in (3.225)] an anisotropy

field B_A are included, the dispersion relations (3.218) and (3.225) are

$$\hbar\omega_k = 2Js\nu(1 - \gamma_k) + 2\mu_B B \qquad \text{for ferromagnets} \qquad (3.227a)$$

$$\hbar\omega_k = 2Js\nu\sqrt{\left(1 + \frac{\mu_B B_A}{Js\nu}\right)^2 - \gamma_k^2} \pm 2\mu_B B \quad \text{for antiferromagnets.} \qquad (3.227b)$$

In both cases we can make B as small as we like. Furthermore, experimental results show B_A to be of the order of 1000 gauss, and therefore negligible compared with $Js\nu/\mu_B$ (order of 10^6 gauss). For $k = 0$, however, $\gamma_k = 1$ and (3.227b) becomes

$$\hbar\omega_0 = 2Js\nu\sqrt{\frac{\mu_B B_A}{Js\nu}\left(1 + \frac{\mu_B B_A}{Js\nu}\right)} \approx \sqrt{4Js\nu\mu_B B_A}\,. \qquad (3.228)$$

Since B_A appears here multiplied by $Js\nu$, $\hbar\omega_0$ can deviate markedly from zero. There is then an energy gap between ground state and the lowest excited state.

Specific heat and magnetization can now be calculated in a similar way to ferromagnets.

The theory of *ferrimagnetic magnons* can be correspondingly formulated. We restrict ourselves to the statement of the dispersion relations for the simplest

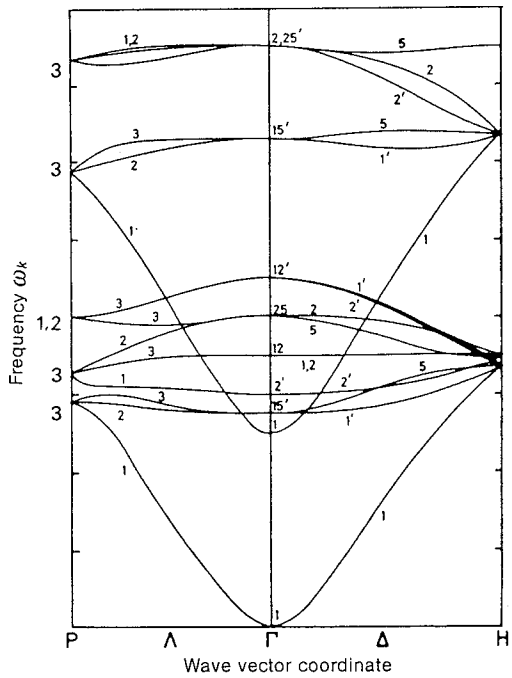


Fig. 3.14. Magnon dispersion curves for yttrium-iron-garnet ($Y_3Fe_5O_{12}$). The numbers indicate the symmetries of the individual branches [from W. Brinkman, R. J. Elliott: J. Appl. Phys. 37, 1458 (1966)].

case, that in the antiferromagnet considered above the spins of the sublattices have different values $s_a \neq s_b$. Again neglecting the anisotropy field we find

$$\hbar\omega_k = Jv[\sqrt{(s_a - s_b)^2 + 4s_a s_b(1 - \gamma_k^2)} \pm (s_a - s_b)]\,. \qquad (3.229)$$

Equation (3.225) follows from this when we put $s_a = s_b$. The dispersion spectrum has two branches which, for $k = 0$, have the values $\hbar\omega_0 = 0$ and $2Jv(s_a - s_b)$.

Ferrimagnets have in general a complicated lattice structure. In addition to the branches of the spectrum found here, we shall also find "optical" branches. Fig. 3.14 is an example of a complicated magnon dispersion spectrum. Consideration of the exchange between nearest neighbours in different sublattices and within a sublattice leads for YIG (yttrium iron garnet) to the fourteen branch spectrum shown. Group theory classifications have again to be used for such spectra. Space group symmetry is restricted here by the fact that similar ions with dissimilar spin alignment in the ground state are now regarded as different (*magnetic space groups*). To this are added symmetry operations in "spin-space" which leave the relative spin distribution of the lattice ions invariant. We cannot go into this aspect of group theory here.

3.4.4 Ferromagnetism Near the Curie Temperature

The concept of magnons as collective excitations without mutual interaction is of course only applicable to ferromagnetic problems when magnetization departs only a little from saturation magnetization. This is not, however, the only interesting region. The vicinity of the Curie temperature, above which spontaneous magnetiaztion vanishes, warrants particular attention. In this section, as an extension to spin wave theory, we want to show that the behaviour of a ferromagnet in this temperature range can likewise be explained by exchange interaction. The approximation used to do this is called the *molecular field approximation.*

The Hamiltonian for the exchange interaction (3.196), extended by an external magnetic field $\boldsymbol{B}$, is given by

$$H = -\sum_{ij}{}' J_{ij}\boldsymbol{S}_i \cdot \boldsymbol{S}_j - g\mu_B \boldsymbol{B} \cdot \sum_{i=1}^{N} \boldsymbol{S}_i\,. \qquad (3.230)$$

The problem in solving a Schrödinger equation with this Hamiltonian is the nonlinearity of the first term. In the spin wave case, the difficulty could be circumvented by the Holstein-Primakoff transformation, with subsequent expansion of the root operator and inclusion of only the first term. In the present case we would at least have to take further terms in the expansion. A simpler approximation for (3.230) is a linearization of the operator such that one of the two spin operators is replaced by its mean value

$$H = -\sum_{i=1}^{N}\left(g\mu_B \boldsymbol{B} + \sum_{\substack{j=1\\(\neq i)}}^{N} J_{ij}\langle \boldsymbol{S}_j\rangle\right)\cdot \boldsymbol{S}_i . \tag{3.231}$$

Along with the external magnetic field there thus appears an *internal field* $\boldsymbol{B}_M = (1/g\mu_B)\sum_j J_{ij}\langle \boldsymbol{S}_j\rangle$.

Such an internal field was introduced early by Weiss to explain ferromagnetism (*Weiss field*). In isotropic media $\boldsymbol{B}_M$ will not depend on the index of the exchange integral. The mean value $\langle \boldsymbol{S}_j\rangle$ will further have the same direction as the magnetization $\boldsymbol{M}$: $\boldsymbol{M} = g\mu_B\langle \boldsymbol{S}_j\rangle N$, so that assuming exchange interaction with nearest neighbours only, it follows that

$$\boldsymbol{B}_M = \lambda \boldsymbol{M} \qquad \text{with} \quad \lambda = \nu J/g^2\mu_B^2 N . \tag{3.232}$$

The *Weiss constant* λ is therefore directly proportional to the exchange integral J.

Equation (3.232) is sufficient to calculate the Curie temperature and thereby to link λ with values which can be measured experimentally. To do this we must first return to the theory of magnetization of a paramagnetic material. We assume that the ions of a solid possess a magnetic moment $\boldsymbol{\mu}$. The thermal motion of the ions about their equilibrium positions results in a statistical distribution of all moments. If we apply an external field $\boldsymbol{B}$, the magnetization $\boldsymbol{M}$ depends on the ratio of the magnetic energy $\boldsymbol{\mu}\cdot\boldsymbol{B} = g\mu_B M_j B$ ($M_j = j, j-1, \ldots, -j+1, -j$) and the thermal energy k_BT

$$M = N\frac{\sum_{-j}^{+j} g\mu_B M_j \exp\left(g\mu_B M_j B/k_BT\right)}{\sum_{-j}^{+j}\exp\left(g\mu_B M_j B/k_BT\right)} . \tag{3.233}$$

The series can be summed to give

$$M = Ng\mu_B j B_j(y) \tag{3.234}$$

with $y = g\mu_B jB/k_BT$ and with the Brillouin function $B_j(y)$ given by

$$B_j(y) = \frac{2j+1}{2j}\coth\left(\frac{2j+1}{2j}y\right) - \frac{1}{2j}\coth\frac{y}{2j} . \tag{3.235}$$

For small y (weak magnetic fields), $B_j(y)$ can be replaced by the first term of series: $B_j = (y/3)\cdot(j+1)/j$, and hence

$$M = \chi B \qquad \text{with} \quad \chi = \frac{Ng^2\mu_B^2 j(j+1)}{3k_BT} = \frac{C}{T} . \tag{3.236}$$

Eq. (3.236) is called the *Curie law* and C the *Curie constant*. The number $p_j = g\sqrt{j(j+1)}$ is designated the *effective number of magnetons*.

We can now apply (3.234) to the ferromagnetic case if we describe the interaction between the magnetic moments by an internal field. We have then only to add this internal field $\lambda \boldsymbol{M}$ to $\boldsymbol{B}$.

We obtain the spontaneous magnetic moment by putting the external field to zero

$$M = Ng\mu_B s B_s\left(\frac{g\mu_B s\lambda M}{k_BT}\right) . \tag{3.237}$$

For $T = 0$, since $\coth y = 1$ for $y = \infty$, we have from this $M = Ng\mu_B s$ as saturation magnetization. M decreases with increasing temperature and finally approaches zero. In this limit we can use the following expansion for the Brillouin function:

$$B_s(y) = \frac{s+1}{s}\frac{y}{3} - \frac{(2s+1)^4 - 1}{(2s)^4}\frac{y^3}{45} \quad \text{for small } y. \tag{3.238}$$

Inserting this into (3.237), we find for the saturation magnetization a law of the form

$$M \sim \sqrt{T_c - T}, \qquad T_c = N\frac{g^2\mu_B^2\lambda}{3k_B}s(s+1) = \frac{s(s+1)\nu J}{3k_B} . \tag{3.239}$$

Hence the magnetization disappears at the *Curie temperature* T_c. Above this temperature the solid is paramagnetic. Fig. 3.15 presents the temperature dependence of the spontaneous magnetization from (3.237), together with some experimental results for comparison.

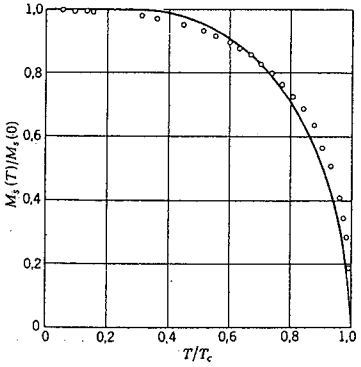


Fig. 3.15. Saturation magnetization according to (3.237) and experimental results on Ni for comparison (from *Kittel* [2]).

In the paramagnetic phase we can again describe the temperature dependence of magnetization by a Curie law of the form (3.236) if we include the internal

field along with the external. We then have

$$M = \frac{C}{T}(B + \lambda M) \quad \text{or} \quad M = \frac{C}{T - C\lambda}B. \tag{3.240}$$

From (3.236) and (3.239) the Curie constant to be inserted here is equal to the Curie temperature divided by the Weiss constant λ. Hence

$$M = \chi B = \frac{C}{T - T_c}B. \tag{3.241}$$

This is the *Curie-Weiss law*. Above T_c, magnetic susceptibility increases as $(T - T_c)^{-1}$.

We can consider antiferromagnets and ferrimagnets in a similar way. Instead of an internal Weiss field, we have then to take different internal fields for the individual sublattices. We shall not present these treatments since they add nothing new to the general theory. Just as spin-wave theory is less important than molecular field theory near the Curie point, the reverse is true at low temperatures where the latter is too coarse. The low temperature approximation derived from (3.237) yields a saturation magnetization temperature dependence of the form $M(T)/M(0) = 1 - (1/s)\exp[-3T_c/(s+1)T]$, which contradicts the experimentally proven $T^{3/2}$ law of spin-wave theory.

Thus we have to distinguish two areas of ordered magnetism, which have to be tackled by different methods. For small departures from the ground state, the method of elementary excitations is more important than all other approximations. At higher temperatures it is more advantageous to use semiclassical methods, which however can also be referred to the general exchange-interaction concept. This is not supposed to mean that the concept of elementary excitations is fundamentally inapplicable at high temperature. Many aspects of the behaviour of a ferromagnet near the Curie temperature can be clearly understood with magnons. One can even define *paramagnons* in the paramagnetic temperature region well above T_c.

A method which is in many respects better than the molecular field approximation chooses for the first term of the Hamiltonian (3.230), instead of a sum over products of spin operators $S_i \cdot S_j$, a sum over products of their z-components $S_{iz}S_{jz}$. This so-called *Ising model* plays an important role in the statistical theory of phase transitions, like the transition appearing here at T_c. This goes beyond the scope of this chapter.

3.4.5 Ordered Magnetism of Valence and Conduction Electrons, the Collective Electron Model

The model which we have considered so far is concerned with direct exchange interaction between localized spins of nearest neighbours. This assumes on the one hand that the electrons (of a lattice ion) contributing to the magnetic moment are sufficiently tightly bound for the ions to be taken as isolated, but on the other hand that the nearest neighbours are close enough for a noticeable exchange interaction to arise.

An exchange between magnetic ions often occurs over large distances in an insulator, the mechanism being that a paramagnetic ion in between communicates the interaction. If, for example, two metallic ions with unfilled d-shells are linked by an oxygen atom (e.g., MnO), each d-electron will interact with one of the two p-electrons of the outermost (spin-saturated) electron pair of the oxygen atom. Since the two spin directions of the two p-electrons are linked by the Pauli principle, this implies an effective interaction between the two d-electrons (*super exchange*).

A further possibility is *indirect exchange* in which localized spins of lattice ions interact with the conduction electrons of a metal. The information on the spin of a given ion is passed on by an electron to another ion. This indirect ion-ion interaction belongs to the so-called *Ruderman-Kittel interactions*. It plays the major role in the rare earths (Gd to Tm). These elements are distinguished by the fact that their ordered magnetism shows a multiplicity of different organizations. Excited states can also be described here by spin waves. For details we refer the reader to an article by *Cooper* [101.21].

The most important group of ferromagnetic metals is the transition metals (see here the conference volume [112.37]). Here the electrons whose spin gives rise to the ferromagnetism are *not* localized (*itinerant electrons*). The elements Fe, Co, Ni have per lattice atom 8, 9, 10 electrons in the uppermost bands. These are a superposition of 3d- and 4s-bands. We have already presented the density of states of Ni in Fig. 2.29. There are indications that one can assume approximately the same band structure, and hence the same density of states, for Cu and the transition metals. The elements are then distinguished only by the position of the Fermi energy relative to the band edges. Fig. 3.16 shows the density of states again schematically. For Cu, E_F lies above the filled d-bands.

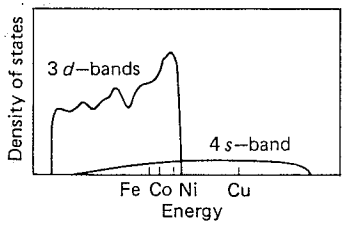


Fig. 3.16. Schematic representation of the density of states of the transition metals with the assumption that all these elements have approximately the same band structure (rigid band model). The d-bands are filled to higher energies in Fe, Co, Ni according to the number of valence electrons. In Cu, the Fermi energy lies above the d-bands, in the 4s-band.

For the transition metals, the *d*- and *s*-bands are only partly filled, Thus *d*-electrons, too, occupy states near the Fermi energy.

At the beginning of the chapter we saw that an electron gas in its ground state can become ferromagnetic if the reduction in energy through the exchange interaction of the aligned spins exceeds the increase in kinetic energy. We cannot use such a simple model in the present case since the valence electron gas almost fills the *d*-bands. The following simple model however can explain the main features of the ferromagnetism of transition metals (*Stoner's collective electron model*). We assume on the one hand that the Bloch states of the band model are maintained. However, we add the exchange energy to the energy of a Bloch state, describing it by an internal field. All states having one spin direction are then displaced in energy relative to the states having the other spin direction. If we further assume that the magnitude of this displacement is independent of k, and depends only on the *s*-, *p*-, *d*-character of the states, then this means a rigid displacement of the states in a band with one spin direction relative to the states with the opposite spin direction. If the Fermi energy lies within such a band, the displacement causes a preponderance of electrons with one spin direction, and hence a spontaneous magnetic moment in the ground state.

Applied to the case of nickel this means: As a result of the overlapping of the *s*-band and *d*-bands, of the ten valence electrons of a Ni atom on average 9.46 electrons are in the *d*-bands and 0.54 in the *s*-band. The exchange interaction practically involves only the *d*-electrons. The exchange-interaction displacement of the two *d*-subbands proves to be so large that the one subband is fully filled and the other contains 0.54 holes per atom. The saturation magnetization is then $I = 0.54N\mu_B$. If the magnetization were to be caused by p localized electrons, it would be given by $M = pN\mu_B$. Thus the combination of band model and exchange interaction leads to noninteger *effective magneton numbers* p. This model is further supported by the experimental result presented in Fig. 3.17. We assume -as already mentioned above—that the density of states in Fig. 3.16 applies for all transition metals (and their alloys!), and hence that only the position of the Fermi energy determines the effective magneton number. Thus p depends only on the number of electrons in the *d*-band, which in turn is determined by the constituents of an alloy. In particular the ferromagnetic properties must disappear as soon as the *d*-bands become completely filled. These predictions are remarkably well confirmed by the results presented in Fig. 3.17. On these questions see also an article by *Friedel* in [65].

Such a model is, of course, only a first step towards an understanding of ferromagnetism in metals. For extension of the theory we can only point to the literature. In particular we shall not discuss the spin waves which are also possible in this case. We shall just mention here an interesting feature of the ferromagnetism of conduction electrons.

We consider an electron gas and put $E = \hbar^2k^2/2m$ for the one-electron energy. On top of this energy we superimpose an exchange energy which displaces the states of one spin direction by a constant amount V relative to the

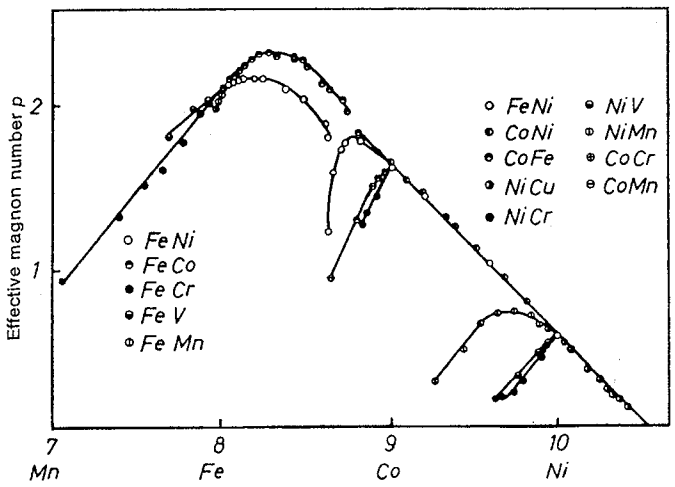


Fig. 3.17. Effective magneton number p for binary alloys of the transition metals as a function of the mean number of valence electrons (Slater-Pauling curve) [from J. Crangle, G. C. Hallam: Proc. Roy. Soc. A272, 119 (1963)].

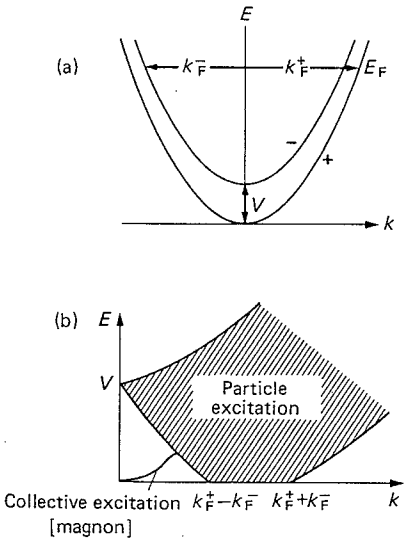


Fig. 3.18. (a) Shift of the free electron states having one spin direction relative to the states having the other spin direction by the exchange energy V. (b) Besides the excitation of an electron out of the Fermi sphere with conservation of spin, the spin can reverse in the transition. This is only possible if the relation between energy and momentum of the excitation is in the hatched area of the figure. Spin waves can disintegrate into these particle excitations if their energy lies above the threshold energy for the excitations. For a similar case see Fig. 3.5 (from *Elliott* [117]).

states of the other direction. This is illustrated in Fig. 3.18a. In the ground state let the states be filled up to energy E_F. It is then possible to set up excited states by excitation of a single electron, with conservation of electron spin or with simultaneous reversal of the electron spin. The first type of one-particle excitation has already been treated in Chapter 2. The other type leads to the spectrum presented in Fig. 3.18b. There the transition energy is shown as a function of the change of the electron k-vector in the transition. Transitions involving change of k are only possible at the energy V. Transitions involving no change of energy are only possible on the Fermi surface between the limits $k_F^+ \pm k_F^-$. In addition to these one-particle excitations, spin waves, which below a threshold energy do not fall in the region of pair excitations, are possible in the electron gas. Above this threshold energy they can disintegrate into these one-particle excitations. We have already met this simultaneous existence of one-particle excitations and collective excitations in the treatment of plasmons (Fig. 3.5) at the end of Section 3.1.4.

4. Electron-Phonon Interaction: Transport Phenomena

4.1 The Interaction Processes

4.1.1 Introduction

In the band model description, *electrons* in a solid are quasi-particles which occupy one-electron states. They are described by Bloch functions $|n, \boldsymbol{k}, \sigma\rangle$ where n is the band index, $\boldsymbol{k}$ is the wave vector of the electron, and σ is the spin. In this chapter we shall be almost entirely concerned with electrons of one band, the conduction band. Accordingly we need only give the band index explicitly in a few cases. Since for transitions within the conduction band electron spin direction is in most cases maintained, we shall on many occasions describe the electron by its wave vector only.

Phonons are collective excitations of the lattice. The vibrational state of the lattice is characterized by the number of phonons in the individual oscillator states defined by the wave vector $\boldsymbol{q}$ and the branch j of the dispersion spectrum $\omega_j(\boldsymbol{q})$.

The basic *electron-phonon interaction* process is the annihilation (absorption) or creation (emission) of a phonon $(\boldsymbol{q}, j)$ with simultaneous change of the electron state from $|\boldsymbol{k}, \sigma\rangle$ to $|\boldsymbol{k} \pm \boldsymbol{q}, \sigma\rangle$. These two processes are illustrated in the top row of Fig. 4.1.

The two graphs for phonon emission and phonon absorption can at the same time describe two further processes if one transforms them slightly (second row in Fig. 4.1). If one thinks of a time axis running from left to right in these graphs and regards electrons running backwards in time as holes running forward in time, then the graphs respectively describe the recombination of an electron-hole pair with emission of a phonon, and the creation of an electron-hole pair with absorption of a phonon.

These four basic processes can be described quantum mechanically by a first-order perturbation calculation. From this, conservation laws follow for the total energy and for the sum of the wave vectors of the elementary excitations involved in the process.

The contributions from the perturbation calculations of higher order can be described as multiple-stage processes, which are made up of a number of successive basic processes. In contrast to the initial and final states, the intermediate states are not "real" (stationary) states of the system. They can be regarded as traversed in such a short time that—in view of the energy-time uncertainty relation—the energy conservation law need not apply. It only applies between