Old and new aspects in lattice-dynamical theory

2001 J. Phys.: Condens. Matter 13 7593

(http://iopscience.iop.org/0953-8984/13/34/308)

INSTITUTE OF PHYSICS PUBLISHING JOURNAL OF PHYSICS: CONDENSED MATTER

J. Phys.: Condens. Matter **13** (2001) 7593–7610 PII: S0953-8984(01)25622-7

Old and new aspects in lattice-dynamical theory

Pasquale Pavone

Institut für Theoretische Physik, Universität Regensburg, D-93040 Regensburg, Germany

E-mail: pasquale.pavone@physik.uni-regensburg.de

Received 8 June 2001, in final form 8 June 2001
Published 9 August 2001
Online at stacks.iop.org/JPhysCM/13/7593

Abstract

An overview of the basics of theoretical lattice dynamics is presented. Starting from the classical formulation, a tutorial description of the evolution of the lattice-dynamical theories is reported. After a brief summary of the mostly used phenomenological models, the modern concepts of first-principles lattice dynamics are introduced, paying special attention to the methods based on the density-functional theory. Among these methods, the perturbative *ab initio* approach is highlighted. Furthermore, the application of this formalism to some cases of general interest is shown. In particular, we report on the controversy about the origin of a peak in the second-order Raman spectrum of diamond, on the phonon spectrum of the wurtzite nitrides and on the temperature-induced phase transitions in tin.

1. Introduction

The theory of lattice dynamics is one of the most confirmed and successful theories of solid state physics. Indeed, it can be considered a milestone in the comprehension of the properties of crystals. Ranging from infrared, Raman, neutron, and in recent years synchrotron spectra to non-linear properties such as thermal expansion and anharmonicity as well as to the electron–phonon interaction and superconductivity, there is little left in which lattice-dynamical phenomena do not play a relevant role. The formidable success of the theory of lattice vibrations implicitly constitutes the best proof of the validity of the generally accepted theoretical view of solid state physics.

The basic theory of lattice dynamics has been given by Born and Huang [1]. In their work, they mainly issued the general properties of the dynamical matrix, e.g., its symmetry and analyticity, without investigating its physical origin in terms of interactions between the constituents (electrons and nuclei) of a crystal. The first attempts to address a systematic study of the connection between electronic and dynamical properties were made in the seventies [2]. Nowadays, current methods in theoretical solid state physics, mainly facing the problem under the computational aspect, allow one to calculate the properties of simple materials by using techniques which do not rely on input from experimental informations. Such methods,

based on the quantum-mechanical theory, are commonly denoted as *first-principles* or *ab initio* methods. Furthermore, the development of algorithms and calculations based on the density-functional theory [3, 4] has supported in the last decade direct applications in the field of the response of electronic systems to external perturbations such as atomic displacements or static macroscopic electric fields [5, 6]. Thanks to these theoretical advances, it is possible now to consider the lattice-dynamical calculations as a complementary tool to the experiment, which is, in fact, completely independent from experimental inputs.

The aim of this work is to give a general classical introduction to the problem and at the same time to bring the reader into contact with the main aspects of the modern approaches to the lattice-dynamical theory. This paper can be formally divided into three parts. In sections 2 and 3 the classical basics of the theory of lattice dynamics are introduced together with a summary of the most frequently used lattice-dynamical models. The modern 'first-principles' approaches are presented to some extent in sections 4 and 5. Finally, some significant applications of the 'new' lattice-dynamical methods are reported in section 6.

2. Classical phonon theory

In this section, we introduce the problem of lattice dynamics of an infinite periodic lattice of classical objects, which will be called 'atoms' in what follows. Let us consider a crystal made of N_c cells with only one atom in the unit cell (the generalization to more than one atom is straightforward) and indicate by $\boldsymbol{R}$ the position of the atom of the generic cell in the undistorted crystal. Thus, the actual (instantaneous) position of this atom will be given as

$$\widetilde{\boldsymbol{R}}(t) = \boldsymbol{R} + \boldsymbol{u}(\boldsymbol{R}, t). \tag{1}$$

Equation (1) can be taken as the definition of the atomic displacements $\{\boldsymbol{u}(\boldsymbol{R}, t)\}$. The total potential energy $\mathcal{V}$ of the crystal will in general be a function of the atomic displacements, $\mathcal{V} = \mathcal{V}(\{\boldsymbol{u}\})$. The knowledge of this function allows one to write the equations of motion for the atoms in the crystal, which in general will be given by a set of coupled equations of the form

$$M\ddot{\boldsymbol{u}}(\boldsymbol{R}, t) = -\frac{\partial \mathcal{V}(\{\boldsymbol{u}\})}{\partial \boldsymbol{u}(\boldsymbol{R}, t)} \tag{2}$$

where M is the atomic mass.

2.1. Harmonic approximation

The potential energy of the crystal can be written as a Taylor expansion in terms of the atomic displacements around the minimum-energy positions,

$$\mathcal{V}(\{\boldsymbol{u}\}) = \mathcal{V}_0 + \frac{1}{2}\sum_{\boldsymbol{R},\boldsymbol{R}'} \boldsymbol{u}(\boldsymbol{R}, t) \cdot \boldsymbol{C}(\boldsymbol{R}, \boldsymbol{R}') \cdot \boldsymbol{u}(\boldsymbol{R}', t) + \mathcal{O}(\boldsymbol{u}^3). \tag{3}$$

The term linear in $\boldsymbol{u}$ is absent due to the equilibrium condition (vanishing force on the atoms). $\mathcal{V}_0$ is the equilibrium potential energy and the elements of $\boldsymbol{C}$ are the *interatomic force constants* defined by

$$\boldsymbol{C}(\boldsymbol{R}, \boldsymbol{R}') = \frac{\partial^2 \mathcal{V}}{\partial \boldsymbol{u}(\boldsymbol{R}) \partial \boldsymbol{u}(\boldsymbol{R}')} \tag{4}$$

where the derivatives denote gradients taken at the potential minimum positions $\{\boldsymbol{u}\} = \boldsymbol{0}$. The *harmonic approximation* consists in retaining in equation (3) only terms up to the quadratic order in the displacements. The atomic force constants defined in equation (4) are not

independent quantities, but they are connected to each other by relations due to the symmetry properties of the crystal. In particular, because of the translational invariance of the crystal, the force constants only depend on the difference $\boldsymbol{R} - \boldsymbol{R}'$ and satisfy the acoustic sum rule $\sum_{\boldsymbol{R}} \boldsymbol{C}(\boldsymbol{R}) = \boldsymbol{0}$, which expresses the fact that the potential energy remains unchanged for a uniform translation of the whole crystal. This property is related to the vanishing frequency of the acoustic modes at the Brillouin zone centre. Within the harmonic approximation, the equations of motion read

$$M\ddot{\boldsymbol{u}}(\boldsymbol{R}, t) = -\sum_{\boldsymbol{R}'} \boldsymbol{C}(\boldsymbol{R} - \boldsymbol{R}') \cdot \boldsymbol{u}(\boldsymbol{R}', t). \tag{5}$$

Translational invariance requires a solution that can be put in the Bloch-wave form

$$\boldsymbol{u}(\boldsymbol{R}, t) = \frac{1}{\sqrt{M}} \boldsymbol{w}\, \mathrm{e}^{\mathrm{i}\boldsymbol{q}\cdot\boldsymbol{R} - \mathrm{i}\omega t}. \tag{6}$$

The allowed values of $\boldsymbol{q}$ are chosen according to the Born–von Kármán periodic boundary conditions. By substitution of equation (6) in equation (5) we obtain

$$\boldsymbol{D}(\boldsymbol{q}) \cdot \boldsymbol{w} = \omega^2 \boldsymbol{w} \tag{7}$$

where we have introduced the discrete Fourier transform:

$$\boldsymbol{D}(\boldsymbol{q}) = \frac{1}{M} \sum_{\boldsymbol{R}} \mathrm{e}^{-\mathrm{i}\boldsymbol{q}\cdot\boldsymbol{R}} \boldsymbol{C}(\boldsymbol{R}). \tag{8}$$

The 3×3 Hermitian matrix $\boldsymbol{D}(\boldsymbol{q})$ defined by equation (8) is called the *dynamical matrix*. In the general case of N_a atoms per unit cell, the dynamical matrix has $3N_a \times 3N_a$ components and the eigenvalue problem in equation (7) has $3N_a$ solutions for ω^2 at each $\boldsymbol{q}$ point in the Brillouin zone; these will be denoted by $\omega_j^2(\boldsymbol{q})$, where $j = 1, 2, \ldots, 3N_a$, and can be interpreted as the branches of a multivalued function $\omega^2(\boldsymbol{q})$. The relations expressed by the equations $\omega = \omega_j(\boldsymbol{q})$ are known as *phonon dispersion relations*.

3. Lattice-dynamical models

The first approaches to the theoretical determination of the phonon dispersion relations were involving lattice-dynamical models in which the relevant parameters were fitted to the experimental frequencies. Indeed, these models were designed originally to interpolate between experimental phonon frequencies, which were mostly determined by inelastic neutron scattering. In the past, they have supported the understanding of neutron, Raman and other spectra.

In principle, these models try to construct the dynamical matrix starting from some approximation for the interatomic force constants or the potential energy of the crystal. The electrostatic (Coulomb) part of the effective ion–ion potential is calculated analytically. This part is indeed common to all model and quantum-mechanical procedures. The remaining part of the interaction potential is either parametrized (as in the 'force-constant model' and the 'Born–von Kármán model' [1]) or determined from *ad hoc* model potentials like the Born–Mayer potential for alkali halides or the Lennard-Jones potential for noble-gas crystals. While these 'rigid-ion models' do not account for electronic polarization at all, a fictitious electronic degree of freedom can be introduced to take into account dipolar effects. In the 'shell model', an atom is represented by a non-polarizable ion core and a *shell* of valence electrons. In this case an electric dipole is generated by the relative displacements of the shells with respect to the ion cores [7, 8]. Furthermore, to describe the peculiarity of covalent crystals, alternative models have been employed which use angular forces ('Keating models' [9]) or bond charges

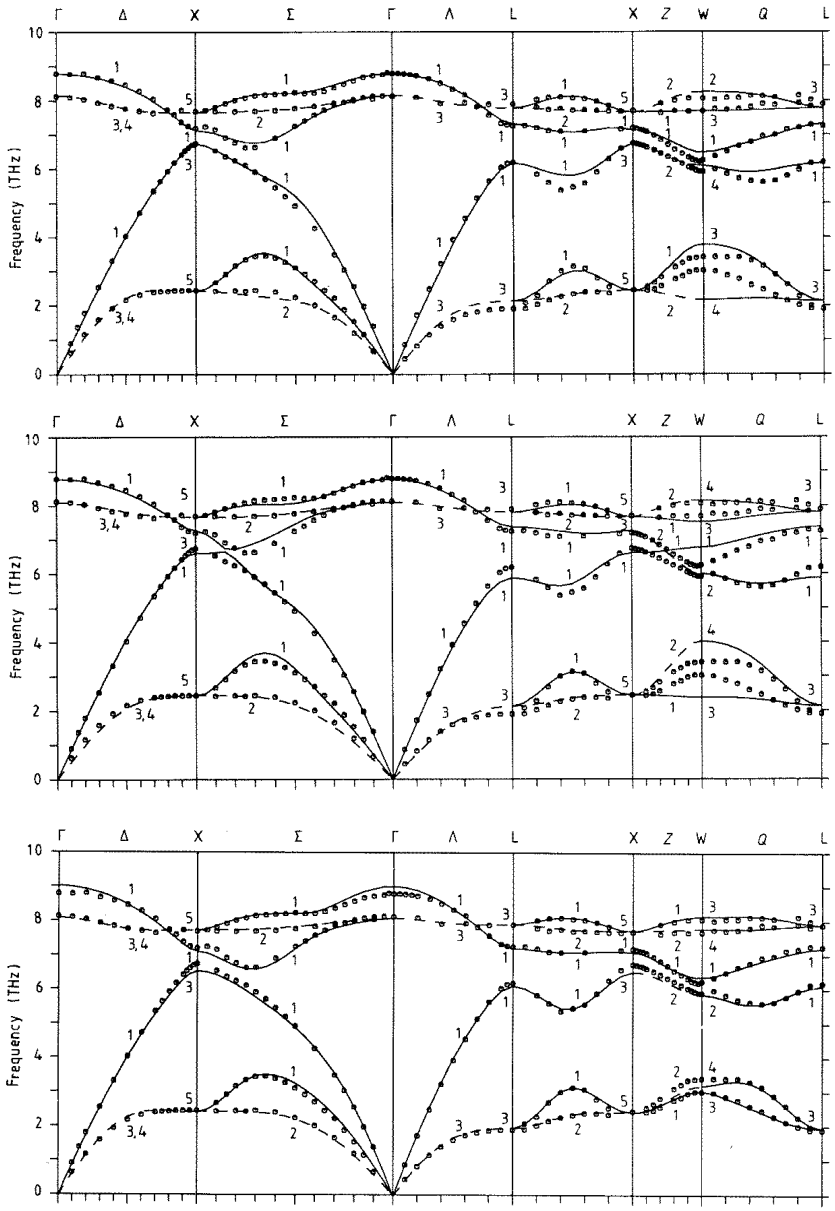

Figure 1. Phonon-dispersion curves of GaAs calculated within different schemes. From top to bottom: rigid-ion model, overlap-shell model and bond-charge model. Circles denote neutron-scattering data. From [11].

('bond-charge models' [10]) to simulate the effects of the highly anisotropic distribution of the electron density in these materials. An example of three different model results in comparison with experimental data is given in figure 1. This figure demonstrates clearly all advantages and disadvantages of lattice-dynamical models. All these models show almost perfect agreement with experiment along the lines with the highest symmetry, while they fail (with different trends) to reproduce the dispersions in lower-symmetry regions (e.g., along XWL). This is not

surprising, due to the fact that the experimental input consists usually of the frequencies at the high-symmetry points Γ, X and L.

In conclusion, the advantage of these mechanistic models consists in their simplicity as compared to quantum-mechanical computational efforts. In fact, they can be applied for generating phonon-dispersion curves of complex systems of considerable size. However, the models are an *a posteriori* construct: their predictive power is limited to those substances which are already partially known, and even in these cases the predictions for the unknown properties are not very reliable.

4. The role of electrons: quantum mechanics

In the previous sections, the explicit role of electrons has been either neglected or expressed by *ad hoc* contributions to the ion–ion interaction. However, the complete description of a system with electrons and nuclei cannot leave apart the quantum-mechanical nature of the electrons. Within quantum mechanics, the properties of a system of nuclei (described by the coordinate $R = \{\boldsymbol{R}_I\}$ and labelled by c for *core*) and electrons (with $r = \{\boldsymbol{r}_i\}$) are derived from the solution of the Schrödinger equation (SE):

$$\hat{H}_{\rm tot}\Phi(r, R) = \mathcal{E}\Phi(r, R) \tag{9}$$

with $\hat{H}_{\rm tot} = \hat{H} + \hat{T}_c + \hat{V}_{c-c}(R)$ and $\hat{H} \equiv \hat{H}_{\rm el} = \hat{T}_{\rm el} + \hat{V}_{\rm el-el}(r) + \hat{V}_{c-{\rm el}}(r, R)$, where the symbols have the standard meaning (T is the kinetic energy, V the interaction potential etc). The solution of equation (9) for a realistic crystal is out of the scope of any analytical or numerical treatment. However, the problem can be enormously simplified due to the very different dynamical-response times of electrons and nuclei.

4.1. Adiabatic approximation

The basic approximation which allows us to decouple the electron and phonon dynamics is the adiabatic approximation (also called the Born–Oppenheimer approximation). In the first step, the electron Hamiltonian is diagonalized for a given nuclear configuration (i.e., R is considered as a parameter)

$$\hat{H}\Psi(r; R) = E(R)\Psi(r; R). \tag{10}$$

Furthermore, one makes the following ansatz for the complete wavefunction

$$\Phi(r, R) = \chi(R)\Psi(r; R). \tag{11}$$

Inserting this ansatz in the SE (9), one obtains for $\chi(R)$ the following equation:

$$[\hat{T}_c + \hat{V}_c(R) + E(R)]\chi(R) \equiv [\hat{T}_c + \hat{V}(R)]\chi(R) = \mathcal{E}\chi(R) \tag{12}$$

provided the mixed terms $\frac{\hbar^2}{2M}\langle\Psi|\nabla_R^2|\Psi\rangle$ and $\frac{\hbar^2}{M}\langle\Psi|\nabla_R|\Psi\rangle\nabla_R$, where M is now the nuclear mass, can be neglected. This approximation appears reasonable due to the nuclear mass being much larger than the electronic one. As a consequence of this assumption, while dealing with the electronic motion, one can consider the nuclei in their equilibrium positions. Furthermore, as can be deduced from equation (12), the nuclear motion is determined by a potential field generated by the average motion of the electrons. The Schrödinger equation (12), moreover, delivers for the nuclear motion almost the same results as in the classical case. Thus, the results of section 2 remain formally valid (but with the exact effective potential!) also in the quantum-mechanical case.

5. First-principles lattice dynamics

Contrary to phenomenological models, *ab initio* calculations require an accurate and parameterless knowledge of the microscopic electronic response to frozen-in lattice vibrations. The basic idea common to all the first-principles methods is to determine the interatomic force constants via the *total* energy of a crystal with frozen nuclear coordinates. In terms of the quantities introduced in the previous section, this procedure has as central object the electronic energy $E(R)$ defined in equation (10). In the last decade, many theoretical advances have been made towards the application of these concepts to lattice dynamics. There are two commonly used approaches for this kind of calculation: the *direct (frozen-phonon) method* and the *perturbative approach*. Most of these modern methods are essentially based on the density-functional theory [3, 4], which will be discussed together with the two *ab initio* lattice-dynamical methods in the following.

5.1. Density-functional theory

The density-functional theory (DFT) is a modern method for dealing with many-particle systems. In several branches of physics and chemistry (e.g., solids, liquids, plasmas, molecules and nuclei) it has become a standard method of calculation and has led to new insights into physical and chemical concepts. In solid state physics alone there are, among many others, applications to electronic band structure, atom clusters, superconductivity, phonons, surfaces and defects [12]. From the general point of view, DFT can be regarded as a method for finding the ground-state properties of the Schrödinger equation (SE) of a many-body system. In comparison to the 'classic' methods of solving the SE, DFT presents itself as a more elegant and simpler approach to the problem (e.g., in comparison with the Hartree–Fock method).

The properties of a system of N interacting electrons described by a Hamiltonian $\hat{H}$ can be completely obtained through the solution of the Schrödinger equation

$$\hat{H}\Psi_i \equiv [\hat{T} + \hat{W} + \hat{V}]\Psi_i = E_i\Psi_i \tag{13}$$

where $\hat{T}$, $\hat{W} \equiv \hat{V}_{\mathrm{el-el}}$ and $\hat{V}$ are the kinetic-energy, the electron–electron-energy and the external-interaction-energy operators, respectively. The two operators $\hat{T}$ and $\hat{W}$ do not depend on the special system which is considered. For a given number of electrons N the system is specified by the external field $v(\boldsymbol{r})$ corresponding to the operator $\hat{V}$ (e.g., N electrons in an electric field, in a magnetic field, in the coulombic potential of a nucleus or in the periodic potential of an array of nuclei). For this reason the operators $\hat{T}$ and $\hat{W}$ are said to be *universal*, whereas the operator $\hat{V}$ is not.

Within quantum mechanics, a prominent role is played by the electronic wavefunctions of the ground ($i = 0$) and excited ($i > 0$) states $\Psi_i(\boldsymbol{r}_1, \ldots, \boldsymbol{r}_N)$. Each wavefunction Ψ_i is a quite messy object which is a function of *all* the $3N$ electronic coordinates. On the other hand, the corresponding electron density, $n_i(\boldsymbol{r}) = \langle\Psi_i|\hat{n}|\Psi_i\rangle$, is a simple function of *only* three variables. These considerations show that for the description of an electron system it would be much more convenient to use the electron density as fundamental quantity than the electronic wavefunction. This goal is achieved by the Hohenberg–Kohn (HK) [3] theorem, which states the existence of a one-to-one correspondence (up to a trivial additive constant in the potential) between the external potential $v(\boldsymbol{r})$ and the ground-state electron density $n_0(\boldsymbol{r}) \equiv n(\boldsymbol{r})$:

$$v(\boldsymbol{r}) + c \Longleftrightarrow n(\boldsymbol{r}). \tag{14}$$

As $v(\boldsymbol{r})$ defines the Hamiltonian of the system, every observable quantity must be determined by $v(\boldsymbol{r})$ only, i.e., in a more mathematical language, it must be a *functional* of $v(\boldsymbol{r})$. In fact, the HK theorem ratifies the possibility of determining every observable quantity as a functional of

the ground-state electron density $n(\boldsymbol{r})$. In particular this is true for the ground-state energy E_0 (in the following simply reported as E), which, due to a very important corollary of the HK theorem, has a global minimum at the exact ground-state electron density

$$E[n] = \min_{n'(\boldsymbol{r})} E[n']. \tag{15}$$

This equation *defines* the functional $E[n]$ but *does not give* any explicit construction rule for it. A formal expression of the ground-state energy functional can be obtained as follows:

$$E[n] = \langle\Psi_0|\hat{T} + \hat{W}|\Psi_0\rangle + V[n] \equiv F[n] + V[n] \tag{16}$$

where $F[n]$ is a *universal* functional of the ground-state electronic density, in the same sense as the sine function $\sin(x)$ is a *universal* function of its argument x. Moreover, in the formal expression of $E[n]$ we have introduced the external-energy functional

$$V[n] = \langle\Psi_0|\hat{V}|\Psi_0\rangle = \int \mathrm{d}\boldsymbol{r} v(\boldsymbol{r}) n(\boldsymbol{r}) \tag{17}$$

which is *not* a universal functional of n, depending—via $v(\boldsymbol{r})$—on the definition of the system under consideration.

5.1.1. Local-density approximation. Whereas the form of the functional $V[n]$ is known, the functional $F[n]$ can be formally written as the sum of the *unknown* kinetic-energy and electron–electron interaction-energy functionals

$$F[n] = \langle\Psi_0|\hat{T}|\Psi_0\rangle + \langle\Psi_0|\hat{W}|\Psi_0\rangle \equiv T[n] + W[n]. \tag{18}$$

The actual kinetic-energy functional $T[n]$ can be approximated by the kinetic-energy functional $T_s[n]$ of a *non-interacting* electron gas at density n. Furthermore, the electron–electron interaction energy functional $W[n]$ can be approximated by the Hartree functional $E_{\mathrm{H}}[n]$, which expresses the mean-field coulombic interaction related to the charge distribution $en(\boldsymbol{r})$. The remaining *unknown* term within these approximations *defines* the so-called exchange–correlation (XC) energy

$$E_{\mathrm{xc}}[n] = F[n] - T_s[n] - E_{\mathrm{H}}[n]. \tag{19}$$

In practice, this term contains what is missing from the sum of E_{H} and T_s to get F. The reason for introducing this representation lies in the fact that both T_s and E_{H} are *easy* functionals ($E_{\mathrm{H}}[n]$ is known and T_s has a known expression in terms of single-particle wavefunctions).

An actual calculation requires, nevertheless, an expression, obviously approximated, for $E_{\mathrm{xc}}[n]$. That mostly used in the literature is the local-density approximation:

$$E_{\mathrm{xc}}[n] \approx E_{\mathrm{xc}}^{\mathrm{LDA}}[n] = \int \epsilon_{\mathrm{xc}}^{h}(n(\boldsymbol{r}))n(\boldsymbol{r})\mathrm{d}\boldsymbol{r} \tag{20}$$

where $\epsilon_{\mathrm{xc}}^{h}(x)$ is the exchange–correlation energy per particle of the homogeneous electron gas with density x.

5.1.2. Kohn–Sham equations. Once the ground-state energy functional has been characterized by the choice of the approximation for the exchange–correlation energy, the remaining problem is to derive the ground-state electron density using the variational principle. The minimization of the ground-state energy functional is particularly easy (at least from the formal point of view) if it is performed in terms of *one-particle* orbitals. More details can be found in [13]. One obtains a set of self-consistent one-particle coupled equations, which are known as Kohn–Sham (KS) equations [4],

$$\left(-\frac{\hbar^2\nabla^2}{2m} + v_{\mathrm{KS}}[n]\right)\phi(\boldsymbol{r}) = \varepsilon\phi(\boldsymbol{r}) \tag{21}$$

where m is the electron mass and the effective KS potential is given as

$$v_{\mathrm{KS}}[n] = v(\boldsymbol{r}) + \frac{\delta E_{\mathrm{H}}[n]}{\delta n(\boldsymbol{r})} + \frac{\delta E_{\mathrm{xc}}[n]}{\delta n(\boldsymbol{r})} \tag{22}$$

where $n(\boldsymbol{r}) = \sum_{\mathrm{occ.orb.}} |\phi(\boldsymbol{r})|^2$. The self-consistent solution of the KS equations yields the ground-state electronic density $n(\boldsymbol{r})$.

5.1.3. Plane-wave expansion. The numerical implementation of the minimization of the ground-state energy requires the representation on a *finite* basis of the KS equations. The optimal choice of the basis set is problem dependent: for systems with different physical properties (in particular with different degree of localization of the single-particle wavefunction) different basis sets have to be chosen. For molecular systems the most natural choice is to expand the wavefunctions in atomic orbitals localized on the different atoms. On the other hand, the periodic character of the wavefunctions in a solid suggests in this case a plane-wave basis set. The latter is the most frequently used basis set in solid state calculations; nevertheless, in some special cases, mixed sets of plane waves and Gaussian functions are used to improve the description of partially localized states (such as those coming from d atomic states). When the plane-wave expansion of the wavefunctions is used, the dimension of the basis set is fixed by considering only plane waves with momentum lying inside a sphere of radius k_{max}. This constraint is usually expressed in terms of the so-called kinetic-energy cutoff defined by $E_{\mathrm{cut}} = \hbar^2 k_{\mathrm{max}}^2/2m$. Only plane waves which have a kinetic energy less than E_{cut} are included in the calculation.

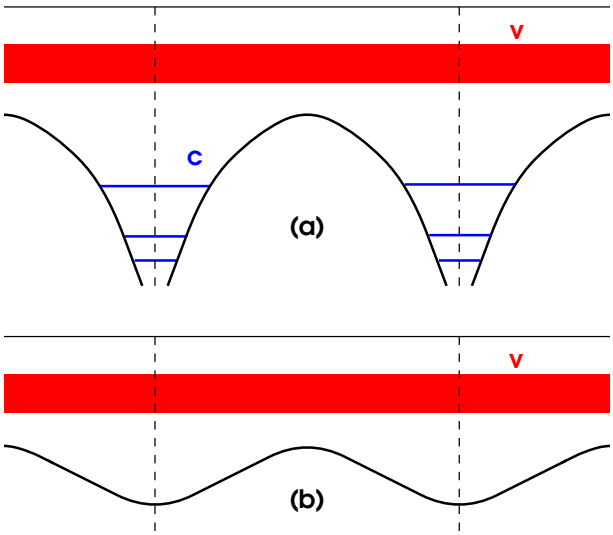

Figure 2. Schematic representation of the actual potential (a) and the pseudopotential (b) in a solid. The label v (c) indicates the valence (core) states.

5.1.4. Pseudopotentials. The actual periodic potential in a solid is schematically represented in figure 2. Two different types of single-particle eigenstate can be found: c, core states which are mainly atomic-like localized states, and v, valence states at higher energies which extend to the whole crystal. The description of the core states in terms of a plane-wave expansion

requires very large basis sets. However, the core states are supposed to have only a scarce influence on the chemical and bond-related properties of a solid. In fact, in this case the actual potential can be substituted with a *pseudo*-potential, which does not possess core states but reproduces exactly the energy of the valence (extended) states of the original potential (see figure 2(b)).

5.1.5. Brillouin-zone integrations: special points. The total ground-state energy of a periodic solid is often expressed in terms of reciprocal-space quantities which contain summations over the Brillouin zone (BZ) of the crystal. These terms have the general form

$$\mathcal{I} = \sum_{\boldsymbol{q}\in \mathrm{BZ}} f(\boldsymbol{q}) \tag{23}$$

where f is periodic in reciprocal space. The direct evaluation of $\mathcal{I}$ involves the calculation of $f(\boldsymbol{q})$ at many $\boldsymbol{q}$ points in the BZ. However, the calculation of f for a *single* $\boldsymbol{q}$ point can be numerically very expensive. This problem can be overcome by using the so-called special points. The integral $\mathcal{I}$ can be written

$$\mathcal{I} = \sum_{\boldsymbol{q}\in\mathcal{S}}^{L} w(\boldsymbol{q}) f(\boldsymbol{q}) + \mathcal{R}_L \tag{24}$$

where the set $\mathcal{S}$ is a set of special points—weighted by $w(\boldsymbol{q})$—if, for the given L, the remainder $\mathcal{R}_L$ is minimum. For practical purposes the most efficient sets of special points are those for which the remainder is close to zero, i.e., for $\mathcal{R}_L \approx 0$.

5.2. The 'frozen-phonon' method

Using the methods described in the previous section the electronic energy of the crystal can be computed as a function of a suitably chosen phonon coordinate (i.e., atomic displacements). In the frozen-phonon approach to lattice dynamics a distorted crystal is treated as a crystal in a new structure with a lower symmetry than the undistorted one [14]. The same method is used for dealing with both the undistorted and the perturbed crystals. Subsequently, the interatomic force constants can be obtained by numerical differentiation of the calculated energy. Phonon frequencies calculated with this method are shown to be very accurate [14, 15]. Although the frozen-phonon method allows in principle the calculation of the derivative of any order, the differentiation is done by taking differences between large numbers. Therefore, it requires in practice an increasing accuracy with increasing order of derivative. Furthermore, the symmetry reduction due to the perturbation increases drastically the computational effort. This fact restricts the computation to displacements corresponding to wavevectors $\boldsymbol{q}$ with high symmetry which lead to superstructures with not too many atoms in the corresponding unit cell.

5.3. Density-functional perturbation theory

The use of supercells can be avoided employing the perturbative approach to the density-functional theory. The basic DFT gives the ground-state energy of an electron system in an external potential. A generalization of DFT can be found to give also the response to many significant physical perturbations such as, e.g., the atomic displacement (phonons, lattice-dynamical properties) or the electric field (dielectric properties). This generalization is called density-functional perturbation theory [5, 6] and will be briefly described in the following.

Let us assume that the total external potential $V(r) = \sum_i v(\boldsymbol{r}_i)$ depends on some parameter (or set of parameters) λ. Let us also assume that this dependence can be expanded in a Taylor series,

$$V_\lambda = V_0 + V_1\lambda + V_2\lambda^2 + \mathcal{O}(\lambda^3) \tag{25}$$

where the explicit dependence on r has been omitted. The corresponding ground-state electron density can be written as

$$n_\lambda(\boldsymbol{r}) = n_0(\boldsymbol{r}) + \lambda n_1(\boldsymbol{r}) + \lambda^2 n_2(\boldsymbol{r}) + \mathcal{O}(\lambda^3). \tag{26}$$

Using the Hellmann–Feynman theorem [16], the derivative of the ground-state energy with respect to the parameter λ is given by

$$\frac{\partial E(\lambda)}{\partial \lambda} = \int n(\boldsymbol{r}, \lambda) \frac{\partial V(\lambda)}{\partial \lambda} \mathrm{d}^3 r. \tag{27}$$

The formal integration of the previous expression yields

$$E_\lambda = E_0 + E_1\lambda + E_2\lambda^2 + \cdots \tag{28}$$

where E_i is, in principle, a functional of the coefficients n_j (essentially derivatives calculated for the unperturbed system) of the expansion of the ground-state electron density up to the order i [17]. In particular, the second-order term E_2 of the expansion—corresponding to the electronic contribution to the force constant matrix if $\lambda = \{\boldsymbol{u}\}$ or to the dielectric constant if λ is an electric field—is found to be a functional of the electron density *only* up to the *first* order

$$E_2 = E_2[n_0, n_1]. \tag{29}$$

This result can be generalized by showing that the knowledge of the derivative of the ground-state density up to order ν gives the derivatives of the energy up to order $2\nu + 1$; the resulting theorem is known as the $2\nu + 1$ theorem [17].

In the specific case of a phonon perturbation, the explicit form of equation (29) allows us to write the electronic contribution to the interatomic force constants as

$$\boldsymbol{C}_{\mathrm{el}}(\boldsymbol{R}, \boldsymbol{R}') = \int \frac{\partial n(\boldsymbol{r})}{\partial \boldsymbol{u}(\boldsymbol{R})} \frac{\partial V}{\partial \boldsymbol{u}(\boldsymbol{R}')} \mathrm{d}^3 r + \int n_0(\boldsymbol{r}) \frac{\partial^2 V}{\partial \boldsymbol{u}(\boldsymbol{R}) \partial \boldsymbol{u}(\boldsymbol{R}')} \mathrm{d}^3 r \tag{30}$$

where the derivatives of the external (bare electron–ion) potential can be expressed analytically and the linear response of the electron density to a distortion of the nuclear geometry, $\partial n(\boldsymbol{r})/\partial \boldsymbol{u}(\boldsymbol{R})$, can be determined in terms of the linear response of the KS orbitals. These quantities can be obtained by applying the first-order perturbation theory to the KS equations (21). This procedure results in a system of coupled equations that should be solved self-consistently [6, 17].

6. A few interesting applications

First application of the perturbative method presented in section 5.3 has been made to derive the phonon-dispersion curves of some prototype tetrahedral semiconductors [6]. As an example of these calculations we show in figure 3 the results for GaAs. The agreement with the neutron scattering data [11] along *all* directions in the Brillouin zone is very good, especially in comparison with the results of the model calculation of figure 1. Various other applications, ranging from metals, semiconductor surfaces and oxides to superconductors, can be found in the literature [18, 19]. In the following, we will focus our attention on some peculiar cases which are of didactic and scientific interest.

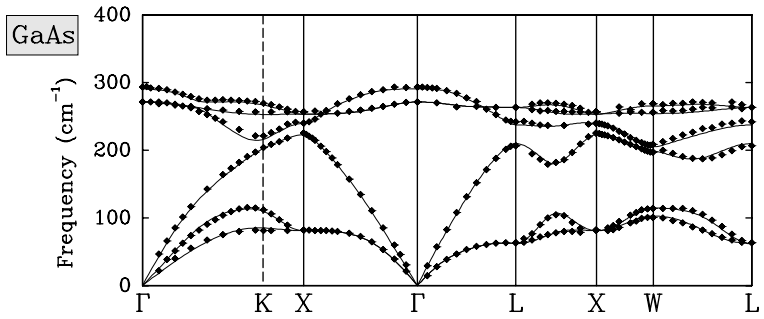

Figure 3. First-principles phonon dispersion curves of GaAs [6]. Experimental neutron-scattering data are denoted by diamonds [11].

6.1. The two-phonon Raman spectrum of diamond: a controversial interpretation

One of the most interesting features in the two-phonon Raman spectrum of diamond is the sharp peak near the high-frequency cutoff. Raman-scattering experiments show that this sharp peak occurs, at room temperature, approximately 3 cm^{-1} higher than twice the frequency of the first-order Raman line [20]. Several attempts have been made in the past for assigning the origin of this peak to different kinds of anomaly of diamond (for a complete review see [21] and references therein). First-order Raman scattering experiments, a two-phonon bound state and an anomaly in the bond polarizability are only a few examples of the many suggestive hypotheses that have been made and that have been shown to be erroneous.

The most accreditate explanation of this feature of the Raman spectrum can be related to a peculiarity in the phonon dispersion curves of diamond which would cause the phonon density of states to have a peak slightly above the optical frequency at the Γ point. In practice, the peak in the Raman spectrum would simply correspond to the peak of the density of states, without invoking *ad hoc* anomalies. Using the critical-point analysis, this would, in turn, mean that the dispersion of the uppermost branch in diamond has a strong overbending in any direction, resulting in a minimum frequency at the zone centre. In fact, first-principles calculations for diamond show that, at variance with the other tetrahedral semiconductors [6], the phonon dispersion of the uppermost (longitudinal) branch has a minimum at the Γ point [21, 22]. This can be directly seen in the three-dimensional plot in figure 4. As a consequence, the calculated phonon density of states in figure 5 shows a peak above $\omega_{\rm opt}(\Gamma)$.

In a recent work, Schwoerer-Böhning *et al* [23] presented results of inelastic x-ray scattering experiments which seem to contradict this interpretation. They found an overbending only along the ΓX direction, with more standard dispersion curves along the other high-symmetry lines. However, even more recent experimental data taken by inelastic neutron scattering [24] seem, on the other hand, to confirm the overbending as can be seen in figure 6. In conclusion, the discrepancy between the two experiments, and between the experiments and the calculations, needs more investigation.

6.2. Phonon dispersion of wurtzite-type nitrides

The wide-band-gap wurtzite semiconductors GaN and AlN are currently of great interest for optoelectronic applications at blue and near-ultraviolet wavelengths, as well as in high-temperature and high-frequency electronics. Since the behaviour of carriers in such devices is affected by their interaction with phonons, the lattice-dynamical properties of these nitrides

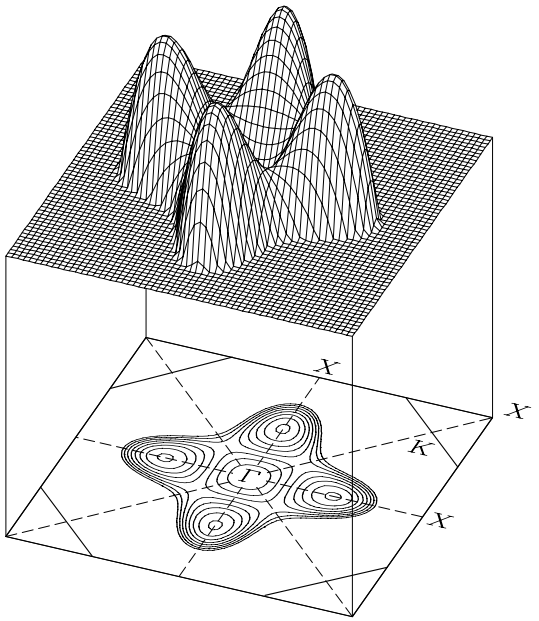

Figure 4. Phonon-dispersion sheet of diamond on the (100) plane. The zero level of the three-dimensional plot is the optical Γ point frequency; i.e., the displayed spectrum consists only of overbending frequencies. We also show the corresponding plane of the Brillouin zone with some special points and a contour plot of the phonon dispersion. The interval between contour lines is 3 cm^{-1}.

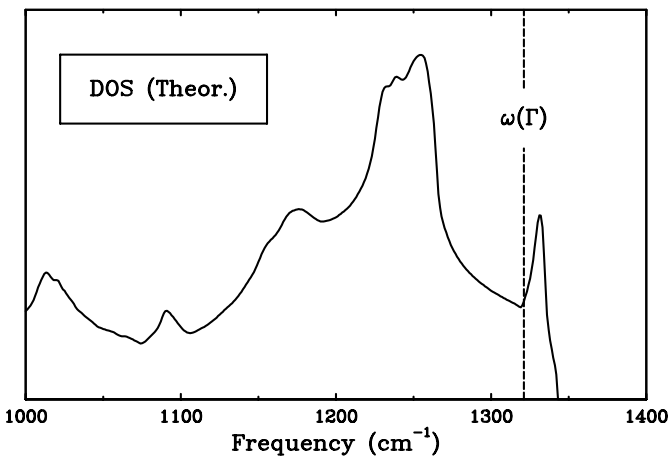

Figure 5. Calculated phonon density of states of diamond in the optical region. The vertical dashed line indicates the Γ point optical frequency.

are very important and have been studied intensively, mainly by Raman scattering and phenomenological models. However, a comprehensive study of the phonon dispersion curves of these materials has escaped for a long time both experimental and first-principles investigation. On the experimental side, the main reason for the lack of phonon-dispersion information on wurtzite GaN and AlN is that single crystals large enough for inelastic neutron

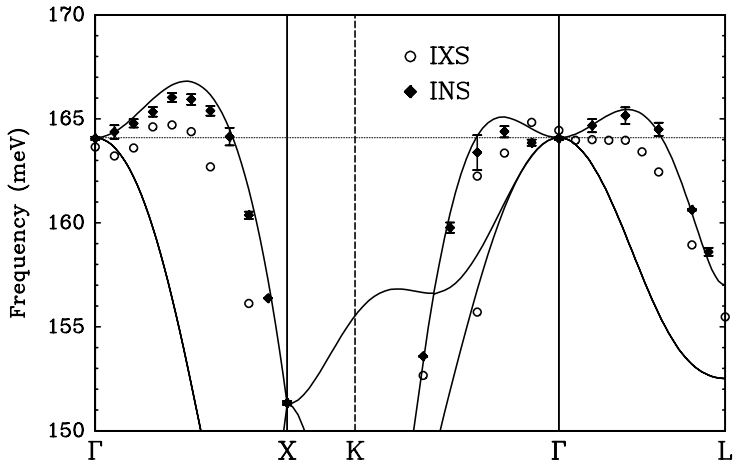

Figure 6. Phonon dispersion of diamond along the high-symmetry directions. The solid lines are the results of *ab initio* lattice-dynamical calculations [22], the circles are measured phonon energies from inelastic x-ray scattering [23] and the full diamonds are inelastic neutron scattering data [24]. In order to facilitate the comparison, the synchrotron data have been displaced 0.7 meV upwards.

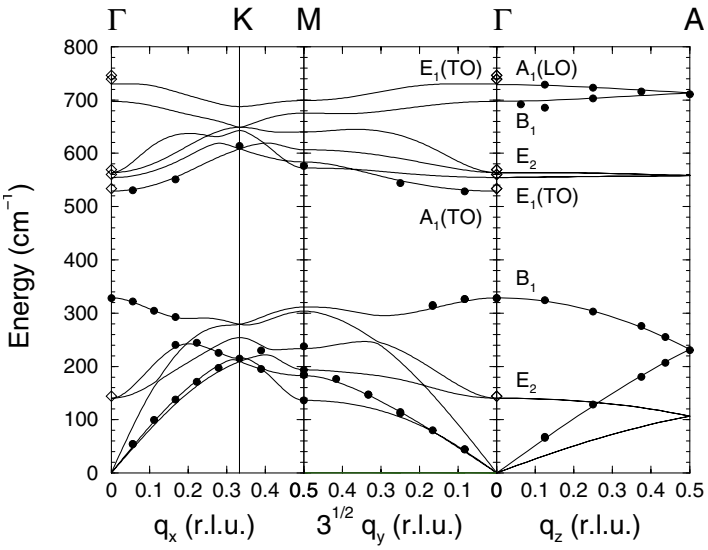

Figure 7. Phonon dispersion of wurtzite-type GaN along several high-symmetry directions [25]. The solid lines are the results of *ab initio* lattice-dynamical calculations, the filled circles are measured phonon energies from inelastic x-ray scattering and the open diamonds at Γ are the Raman data. The theoretical results have been scaled by a factor of 0.97 in order to obtain optimum agreement with the experiment.

scattering do not exist. From the theoretical point of view, the strong localization of the valence orbitals of the nitrogen atom makes the *ab initio* pseudopotential calculation particularly demanding. However, experimental limitation can now be overcome with inelastic *x-ray* scattering (IXS). With the availability of dedicated beamlines at third-generation synchrotrons, IXS has developed into a powerful alternative for studying dispersion effects of elementary

excitations in solids. On the other hand, the use of *soft* norm-conserving pseudopotentials [27] allows a drastic reduction of the computational efforts.

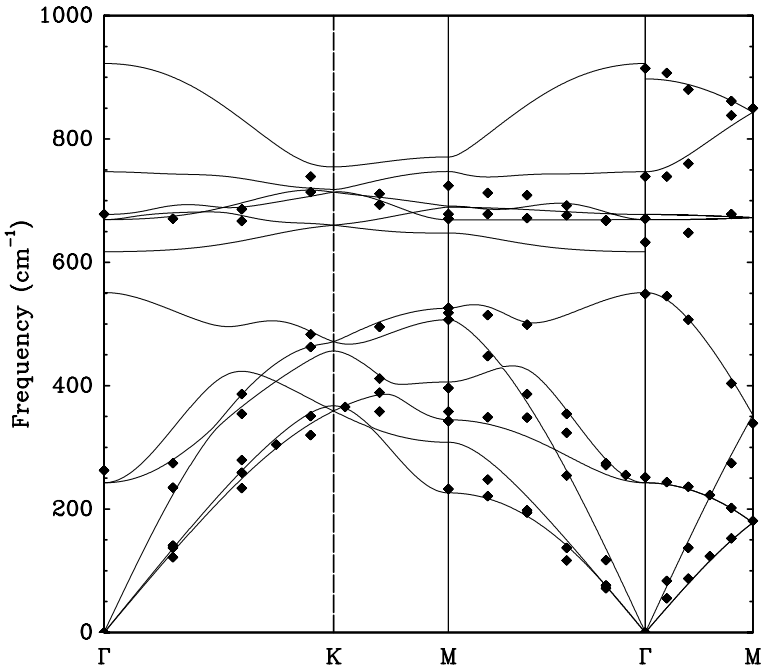

Figure 8. Phonon dispersion of wurtzite-type AlN [26]. Solid lines indicate first-principles calculations; the experimental data from inelastic x-ray scattering experiments are denoted by diamonds.

The results of a combined theoretical and experimental investigation of the lattice-dynamical properties of GAN and AlN [25, 26] are presented in figures 7 and 8, respectively. Nonlocal pseudopotentials (including nonlinear core corrections for Ga in GaN), 14 special points in the irreducible wedge of the Brillouin zone and a plane-wave basis set with a kinetic-energy cutoff of 50 Ryd, have been used in the calculation. Although general agreement is quite good, the calculations for GaN are systematically somewhat larger than the experimental data. For this compound, the difference between theory and experiment increases with increasing phonon frequency. In order to obtain optimum agreement with the experimental data, the theoretical results were scaled in figure 7 by a constant factor of 0.97. This correction is well within the overall reliability of the calculation. One main result of our study is the frequency of the two *silent* B_1 modes at Γ, since these data cannot be obtained either by first-order Raman or by infrared spectroscopy. These phonons are rather important input parameters for fits of phenomenological models.

6.3. Phase transitions in tin

Another interesting application of the *ab initio* phonon calculations described in section 5 is the investigation of the temperature-induced phase transition of tin [28]. Tin is commonly found in one of two allotropic forms: at ambient pressure, the stable phase at low temperature is α-Sn (grey tin), which is a zero-gap semiconductor having the diamond structure; when the temperature is raised above $T_c \approx 13\,^{\circ}$C, the crystal transforms into the β phase (white

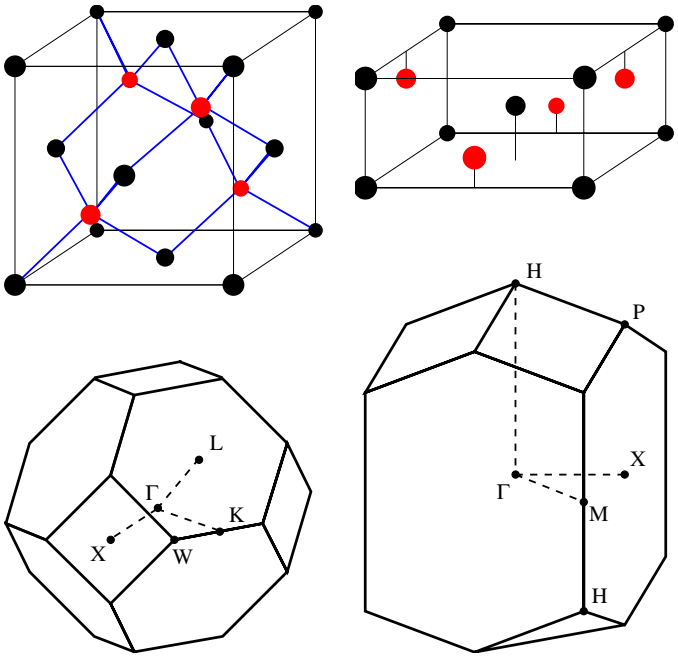

Figure 9. Crystal structure and first Brillouin zone for the diamond-structure α (left) and the bct-structure β (right) phases of tin.

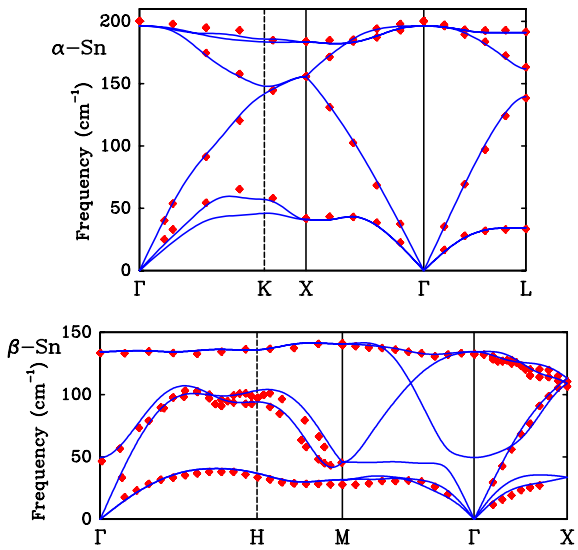

Figure 10. Calculated phonon dispersions (continuous lines) for the α (upper panel) and β (lower panel) phases of tin. Experimental data from neutron diffraction (from [30] and [31] for the α and β phases, respectively) are reported with diamonds.

tin), which is a body-centered tetragonal metal (see figure 9). The $\alpha \leftrightarrow \beta$ transition in tin is possibly the simplest and prototypical case of an entropy-driven structural transformation which is determined by the (harmonic) vibrational properties of the two phases of the material.

As the transition temperature of the $\alpha \leftrightarrow \beta$ transition in tin is $\approx$60% of the melting temperature, we do not expect, in this case, anharmonic effects to play any important role. The free energies of the two phases can be thus calculated in the harmonic approximation, which reads

$$F(\Omega, T) = E_0(\Omega) + k_B T \sum_{\nu} \ln \left[2 \sinh \left(\frac{\hbar \omega_\nu}{2 k_B T} \right) \right] \tag{31}$$

where $E_0(\Omega)$ is the static crystal energy—easily accessible to standard local-density-functional calculations, k_B is Boltzman's constant, Ω is the crystal volume and ω_ν is a normal-mode frequency which can be calculated using DFPT. Due to the metallic character of one of the two phases, special attention should be paid to the Brillouin-zone integrations. We employed the Gaussian-smearing special-point technique [29], using a Gaussian width of 0.01 Ryd and 0.02 Ryd for the α and β structures respectively, which at convergence require 60 and 163 special points in the irreducible wedge of the Brillouin zone.

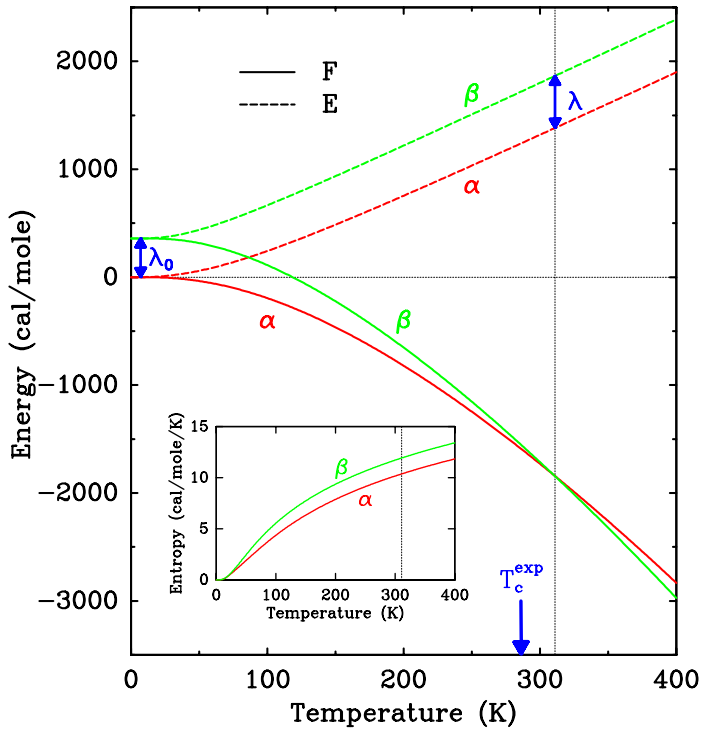

Figure 11. Zero-pressure free-energy (solid lines) and internal-energy (dashed lines) curves for the α and β phases of tin as functions of temperature. The thin vertical dotted line indicates the theoretical transition temperature, while the experimental value for T_c is shown by the arrow. $\lambda_0 = 359$ cal mol^{-1} is the $T = 0$ K free-energy difference—including the zero-point contribution—while $\lambda = 482$ cal mol^{-1} indicates the latent heat adsorbed in the $\alpha \leftrightarrow \beta$ transition. Finally, the inset displays the temperature dependence of the vibrational entropies of the two phases.

The calculated static energy of the α structure lies 516 cal mol^{-1} below that of the β phase. Taking into account zero-point contributions to the internal energy, at $T = 0$ K the α phase turns out in our calculations to be more stable than the β one by 359 cal mol^{-1}. In figure 10 we display our calculated phonon dispersion curves for the α and β phases of tin. Besides the excellent agreement with available neutron-diffraction experimental data, the main feature to be noticed is the different range spanned by the vibrational bands in the two phases, which

extends up to $\approx$200 cm^{-1} in grey tin, while it is limited to $\approx$140 cm^{-1} in white tin. The fact that the structure with a larger static energy has smaller vibrational frequencies indicates that a phase transition may occur when raising the temperature, as a consequence of the larger value of its entropy at high temperature. For the sake of simplicity, suppose that each structure (α or β) is characterized by a single vibrational frequency (Einstein model): the difference between the entropies of the two structures approaches for increasing temperature the value $\Delta S_\infty \approx 3k_B \ln(\omega_\beta/\omega_\alpha)$. Thus, at sufficiently high temperature the entropic contribution to the free-energy difference, $-T\Delta S_\infty$, takes over the difference between the internal energies, and the β phase becomes more stable (provided the transition temperature so obtained is well below the melting point of the two phases, so that the harmonic approximation is well justified).

The energetic behaviour is clearly illustrated in figure 11, where the internal energy and the free energies of the two structures are displayed. The internal energies of the two phases increase with temperature and their difference saturates at a value which is $\approx$35% larger than at $T = 0$ K. The free-energy curves bend down, and their difference decreases quasi-linearly with temperature, vanishing at $T_c = 38\,^\circ$C. The quality of the agreement between theory and experiment achieved for the $\alpha \leftrightarrow \beta$ transition in tin is such as to give confidence in the predictive power of free-energy calculations based on the harmonic approximation and vibrational frequencies calculated from first principles, and it indicates that this is the method of choice in all those cases where the relevant phenomena occur at temperatures well below the melting point.

Acknowledgments

I would like to thank here all the persons who have contributed to my work during the last years, but the list would probably need more space than reasonably allowed. Therefore I am obliged to do it in such an informal way. Among them, I owe special thanks to Dieter Strauch.

References

[1] Born M and Huang K 1954 *Dynamical Theory of Crystal Lattices* (Oxford: Oxford University Press)
[2] Pick R, Cohen M H and Martin R 1970 *Phys. Rev.* B **1** 910
[3] Hohenberg P and Kohn W 1964 *Phys. Rev.* **136** B864
[4] Kohn W and Sham L J 1965 *Phys. Rev.* **140** A1133
[5] Baroni S, Giannozzi P and Testa A 1987 *Phys. Rev. Lett.* **58** 1861
[6] Giannozzi P, de Gironcoli S, Pavone P and Baroni S 1991 *Phys. Rev.* B **43** 7231
[7] Dick B G and Overhauser A W 1958 *Phys. Rev.* **112** 90
Cochran W 1959 *Phys. Rev. Lett.* **2** 495
[8] Schröder U 1966 *Solid State Commun.* **4** 347
[9] Keating P N 1966 *Phys. Rev.* **145** 637
[10] Martin R M 1969 *Phys. Rev.* **186** 871
Weber W 1977 *Phys. Rev.* B **15** 4789
[11] Strauch D and Dorner B 1990 *J. Phys.: Condens. Matter* **2** 1457
[12] For a review of early applications of density-functional techniques to the structural properties of solids, see, e.g., Picket W E 1989 *Comput. Phys. Rep.* **9** 116
[13] Kohn W 1985 *Highlights of Condensed-Matter Theory* ed F Bassani, F Fumi and M P Tosi (Amsterdam: North-Holland)
[14] Kunc K and Martin R M 1982 *Phys. Rev. Lett.* **48** 406
[15] Yin M T and Cohen M L 1982 *Phys. Rev.* B **26** 3259
[16] Hellmann H 1937 *Einführung in die Quantenchemie* (Deuticke, Leipzig)
Feynman R P 1939 *Phys. Rev.* **56** 340
[17] Gonze X and Vigneron J P 1989 *Phys. Rev.* B **39** 13 120
[18] Gonze X 1997 *Phys. Rev.* B **55** 10 337

[19] An up-to-date and rather complete list of applications of DFPT to lattice dynamics can be found in Baroni S, de Gironcoli S, Dal Corso A and Giannozzi P 2001 *Rev. Mod. Phys.* at press
[20] Solin S A and Ramdas A K 1970 *Phys. Rev.* B **1** 1687
[21] Windl W, Pavone P, Karch K, Schütt O, Strauch D, Giannozzi P and Baroni S 1993 *Phys. Rev.* B **48** 3164
[22] Pavone P, Karch K, Schütt O, Windl W, Strauch D, Giannozzi P and Baroni S 1993 *Phys. Rev.* B **48** 3156
[23] Schwoerer-Böhning M, Macrander A T and Arms D A 1998 *Phys. Rev. Lett.* **80** 5572
[24] Kainzmaier H 2001 private communication
[25] Ruf T, Serrano J, Cardona M, Pavone P, Pabst M, Krish M, D'Astuto M, Suski T, Grzegory I and Leszczynski 2001 *Phys. Rev. Lett.* at press
[26] Schwoerer-Böhning M, Macrander A T, Pabst M and Pavone P 1999 *Phys. Status Solidi* **215** 177
[27] Troullier N and Martins J L 1991 *Phys. Rev.* B **43** 1993
[28] Pavone P, Baroni S and de Gironcoli S 1998 *Phys. Rev.* B **57** 10 421
[29] de Gironcoli S 1995 *Phys. Rev.* B **51** 6773
[30] Price D L, Rowe J M and Nicklow R M 1971 *Phys. Rev.* B **3** 1268
[31] Rowe J M 1967 *Phys. Rev.* **163** 547