

SISSA Lecture Notes

Methods of
Electronic Structure Theory

Raffaele Resta

<resta@sissa.it>

Very preliminary draft,
incomplete and partly obsolete
(run through L^AT_EX on January 17, 1996)

Contents

1	Theory of the electronic ground state	4
1.1	Introduction	4
1.2	Density-functional theory	6
1.3	Local-density approximation	8
1.4	Minimization strategies	10
1.5	Basis functions and pseudopotentials	12
1.6	Total energy expansions	15
1.7	References	16
2	Screening and electrostatics	21
2.1	Electronic screening	21
2.2	Diagonal screening	23
2.3	Microscopic model screening	25
2.3.1	Metals	25
2.3.2	Semiconductors	26
2.3.3	Insulators	26
2.3.4	Clausius-Mossotti vs. diagonal models	27
2.4	Macroscopics	31
2.4.1	Phenomenology	31
2.4.2	Piezoelectricity	32
2.4.3	Ionic screening	33
2.5	References	35
3	Formal linear-response theory	38
3.1	Density-functional formulation	38
3.2	Independent-particle polarizability	39
3.3	Physical response operators	40
3.4	Density-functional perturbation theory	42
3.5	Role of symmetry	43
3.6	References	44

4	Electron-gas results	46
4.1	Ground state	46
4.2	Dielectric function	47
4.3	References	50
5	Pseudopotential perturbation theory	52
5.1	Basic concepts	52
5.2	Formalism	53
5.3	Crystalline metals	56
5.4	The pairwise interaction	57
5.5	Liquid metals	60
5.6	Phonons vs. elasticity: a paradox	61
5.7	References	63
6	Ground state of periodic solids (incomplete)	65
6.1	Kohn-Sham equation	65
6.2	Total energy per cell	66
6.3	Forces	69
6.4	Macroscopic stress	71
6.5	The stress theorem	72
6.6	Elastic constants	75
6.7	References	75
7	Macroscopic electric polarization	77
8	Linear response in periodic solids	78
9	Lattice dynamics (preliminary draft)	79
9.1	Definitions	79
9.2	Central two-body forces	81
9.3	Metals	82
9.4	Ionic crystals	84
9.5	Effective charges	86
9.6	Ion-electron-ion interaction	86
9.7	Ab-initio force constants	87
9.8	Outline of the results	88
9.9	References	89
10	Computational alchemy: superlattices and alloys	107

Chapter 1

Theory of the electronic ground state

1.1 Introduction

The static quantities which describe a state of matter at zero temperature are its quantum-mechanical ground-state properties. For a system of nuclei and electrons the Born-Oppenheimer (or adiabatic) approximation allows to decouple the electronic and ionic degrees of freedom. In most circumstances the nuclei behave as classical particles, while the electrons constitute a quantum many-body system. The assumption that the electrons are in their ground state is a very good one: for densities of interest the electronic system is in fact close to complete degeneracy at room temperature,¹ and a detailed account of Fermi-Dirac electronic distribution is seldom necessary.[†]

The changes in the electronic ground state in presence of static “external” influences determine the static compressibility, elastic and piezoelectric constants, dielectric susceptibilities, and other linear and non-linear response functions. By “external” I mean external to the electronic system; but the perturbation can be nevertheless “internal” to the piece of matter. This is *e.g.* the case of strain and of any lattice deformation.

In addition to the above genuine ground state properties, the dynamics of the ionic motions can be treated in several circumstances by requiring that the electrons follow the ions adiabatically, being in their ground state for every instantaneous configuration of the ions. Therefore the theory of the electronic ground state suffices to describe a much larger class of properties of matter. These include phonon dispersion curves, anharmonicity, low-frequency dielectric behavior, and most of finite-temperature thermodynamic properties.

In order to ensure adiabaticity of a given dynamical phenomenon one mainly requires the frequencies of the ionic motions to be much smaller than the frequencies of the lowest electronic transitions. This is easy to achieve in an insu-

[†]In some cases—related to structural phase transitions in metals—the entropic contribution of electronic excitations is nonnegligible.^{2,3}

lator in a neighborhood of its ground state; but the condition certainly breaks down in metals, where the electronic spectrum has excitations of vanishingly small energy. Despite this fact, it can be shown^{4,5} that the adiabatic approximation is a very good one for ionic motion even in the lightest of metals, Li. This accounts for some decades of successes⁶ in using the *static* electronic response in the description of lattice-dynamical or finite-temperature⁷ properties of simple (*sp*-bonded) metals in their solid and liquid state. The adiabatic approximation is generally adopted even for solids where the zero-point motion is nonnegligible.[†]

The field of ground-state properties of real materials has a long history,¹⁰ which I will not try to sketch here. Generally speaking, theories where the quantum many-body nature of the electronic system is accounted for, and providing sensible quantitative predictions, have first been developed for simple metals, and only more recently for semiconductors, insulators, transition metals, and more complex materials. An historical breakthrough occurred in the early sixties, when pseudopotential perturbation theory was developed for simple metals;⁶ an abridged account of it is given here in Chapter 5. This theory can be considered as the archetypical “total-energy” method for solids, and several of its features are relevant in order to discuss the modern developments of electronic structure theory.

Nowadays it is possible to perform calculations for real materials truly from *first principles* (*i.e.* without any experimental input), and which allow meaningful comparison with the experimental measurements in many interesting situations. The relationships between theory and experiment in the last fifteen years have changed to such an extent that it is now possible to perform “computer experiments”, *i.e.* to simulate situations not occurring, or nonmeasurable, in real life. This gives access to physical quantities and mechanisms which could only be postulated in earlier ages. On several occasions, the result of a computer experiment prompted for the existence of a new “theorem”. Modern computational physics leads to insight, more than merely answers.

The error bar of a given simulation can be partitioned into two contributions: systematic error and computational noise. The former is due to some well defined *physical* approximations used to make the many-body problem tractable, while the latter comes from truncation, finite basis-set expansions and other *numerical* approximations.

The enormous know how which has been accumulated in recent years allows often accurate estimates of the systematic error. This kind of error must be separately checked property by property, but is remarkably uniform over a wide range of different materials. As for the computational noise, it can be kept arbitrarily small, within affordable computer bills, for systems of moderate complexity. In full analogy with laboratory physics, even in computational physics

[†]Ref. 8; when studying phase transitions at finite temperature, the vibronic interactions between electronic states cannot be always neglected: this is a well known feature in *e.g.* narrow-band-gap ferroelectrics (see Ch. 14 in Ref. 9 for a discussion).

it is possible to perform *differential* experiments, where systematic errors are almost eliminated.

The understanding of the cooperative behavior of $\sim 10^{24}$ strongly interacting electrons and ions is possible thanks to a panacea which has been available since the early days of condensed matter physics: symmetry. But, besides it, two more fundamental tools have made possible the recent spectacular advances. The first is the enormous increase of computational power, made available to a large scientific community. The second tool is the density-functional theory (DFT) for exchange and correlation in many-electron systems:^{11,12} this is the starting point for almost every work which concerns, at first-principle level, ground-state properties of matter. The only few attempts outside the scope of DFT have been within the variational Hartree-Fock method.¹³

DFT is an exact theory: as such, it would be only formal and useless to attack real systems. Its power stems from the fact that a very natural approximation to it, the local-density approximation (LDA), gives a workable scheme and proves to be very accurate for a great deal of different materials and properties.¹⁴⁻¹⁶ It happens that, in most of the recent work within this area, LDA is the *only* uncontrolled approximation, *i.e.* the unique source of systematic errors.

1.2 Density-functional theory

The system of N interacting electrons is described by the Hamiltonian

$$H = T + U + V_{ext}, \quad (1.1)$$

where T is the electronic kinetic energy, U is the electron-electron interaction, and V_{ext} is the one-body "external" potential

$$V_{ext} = \sum_{i=1}^N V_{ext}(\mathbf{r}_i), \quad (1.2)$$

generated by the ions and by any possible additional source.

The theory focuses on the electronic density

$$n(\mathbf{r}) = N \int d\mathbf{r}_2 d\mathbf{r}_3 \dots d\mathbf{r}_N |\Psi(\mathbf{r}, \mathbf{r}_2, \mathbf{r}_3 \dots \mathbf{r}_N)|^2 \quad (1.3)$$

as the basic independent variable for the description of the system ground state. This is allowed by an important theorem, due to Hohenberg and Kohn,¹⁷ stating that $n(\mathbf{r})$ uniquely determines the external potential, hence the wavefunction Ψ and all other electronic properties of the system. The uniqueness proof is a very simple *reductio ad absurdum*, under the only hypothesis of a nondegenerate ground state. We then isolate in Eq. (1.1) the terms T and U ; since Ψ is a functional of $n(\mathbf{r})$, these terms define

$$F[n(\mathbf{r})] = \langle \Psi | T + U | \Psi \rangle, \quad (1.4)$$

which is *the* density functional after which the theory is named. For more mathematical rigor in defining the functional, see the contribution of Levy and Perdew to Ref. 12, p. 11.

Using the above results, the total energy of the piece of matter takes the form:

$$E_{tot}[n(\mathbf{r})] = E_{ext} + \int d\mathbf{r} V_{ext}(\mathbf{r})n(\mathbf{r}) + F[n(\mathbf{r})], \quad (1.5)$$

where E_{ext} is the classical energy of the ions: for an isolated system, this term amounts to the ion-ion Coulomb repulsion.

It is customary to extract from $F[n]$ the classical Coulomb energy of the electronic system, known as the Hartree term:

$$E_H[n] = \frac{e^2}{2} \int d\mathbf{r} d\mathbf{r}' \frac{n(\mathbf{r})n(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|} = \frac{1}{2} \int d\mathbf{r} V_H(\mathbf{r})n(\mathbf{r}). \quad (1.6)$$

To proceed further, we extract another term defined in a more complicate way. We consider a fictitious *noninteracting* system, whose density is equal to $n(\mathbf{r})$: this condition is realized by a *different* external potential, named after Kohn and Sham¹⁷ V_{KS} . The orbitals of the fictitious system obey the single-particle Schrödinger equation (in atomic units):

$$\left(-\frac{1}{2}\nabla^2 + V_{KS}\right)\varphi_i = \epsilon_i\varphi_i, \quad (1.7)$$

and are related to the density in the obvious way

$$n(\mathbf{r}) = \sum_i f_i |\varphi_i(\mathbf{r})|^2, \quad (1.8)$$

where the occupancy factor f_i selects (in the simple cases) the $N/2$ lowest-energy doubly occupied KS orbitals. This single-particle mapping uniquely defines a kinetic energy functional:

$$T_s[n] = -\frac{1}{2} \sum_i f_i \langle \varphi_i(\mathbf{r}) | \nabla^2 | \varphi_i(\mathbf{r}) \rangle, \quad (1.9)$$

and we recast the density functional in the form

$$F[n] = T_s[n] + E_H[n] + E_{xc}[n], \quad (1.10)$$

which defines $E_{xc}[n]$, known as the exchange-correlation functional. Despite its name, it also involves the kinetic energy of the interacting system, since $T_s[n]$ must *not* be confused with the real kinetic energy of the system [T term in Eq. (1.4)].

If we now assign a given $V_{ext}(\mathbf{r})$, then the total energy, Eq. (1.5), as a functional of $n(\mathbf{r})$ has a variational minimum at its physical ground state value, subject to the constraint $\int d\mathbf{r} n(\mathbf{r}) = N$: the Euler equation is

$$\frac{\delta T_s[n]}{\delta n(\mathbf{r})} + V_{ext}(\mathbf{r}) + V_H(\mathbf{r}) + V_{xc}(\mathbf{r}) - \mu = 0, \quad (1.11)$$

where the exchange-correlation potential is the functional derivative

$$V_{xc}(\mathbf{r}) = \frac{\delta E_{xc}[n]}{\delta n(\mathbf{r})}. \quad (1.12)$$

When using the KS orbitals—under orthonormality constraints—as variational parameters for the density, Eq. (1.8), then the Euler equations assume exactly the form of Eq. (1.7), where the eigenvalues ϵ_i are identified with the Lagrange multipliers and the KS potential is

$$V_{KS} = V_{ext} + V_H + V_{xc}. \quad (1.13)$$

A respelling of this main conclusion is the following: solution of the KS equation

$$\left(-\frac{1}{2}\nabla^2 + V_{ext} + V_H + V_{xc}\right)\varphi_i = \epsilon_i\varphi_i \quad (1.14)$$

provides the *exact* density of the system.

In this way we obtain a formulation where the “simple” part of the problem assumes a familiar single-particle form, while the “difficult” many-body part is swept under the rug of V_{xc} , the exchange-correlation term in the KS potential. Some theorems are known about the behavior of the exact KS potential (Almblad and von Barth, in Ref. 12, p. 209) as well as about the ensemble of densities where the functional is defined (Levy and Perdew, in Ref. 12, p. 11). From a practical viewpoint, the exchange-correlation potential is a clever place where to condense all our ignorance about the many-body problem.

A final warning is appropriate at this point. It is tempting to identify the KS eigenvalues ϵ_i with the electronic excitation spectrum of the system: this identification is wrong in general, and in particular the KS gap in a semiconductor does *not* coincide with the optical gap.¹⁸

1.3 Local-density approximation

A very natural approximated form of DFT is obtained upon assumption that $E_{xc}[n]$ a local function of \mathbf{r} :

$$E_{xc}[n] = \int d\mathbf{r} \epsilon_{xc}(n(\mathbf{r})) n(\mathbf{r}), \quad (1.15)$$

where $\epsilon_{xc}(n)$ is the exchange and correlation energy per particle in a uniform electron gas (alias “jellium”) of density n , *i.e.* in a homogeneous and isotropic system of interacting electrons in a uniform neutralizing background.^{19,20} The exchange-correlation potential is written, within LDA, as

$$V_{xc}(\mathbf{r}) = \left. \frac{d}{dn} (\epsilon_{xc}(n) n) \right|_{n=n(\mathbf{r})} \equiv \mu_{xc}(n(\mathbf{r})), \quad (1.16)$$

where $\mu_{xc}(n)$ is the exchange and correlation contribution to the chemical potential of a uniform system. It fulfills the inequalities:

$$\mu_{xc}(n) < 0 \quad , \quad \frac{d\mu_{xc}(n)}{dn} < 0 . \quad (1.17)$$

The LDA provides, after Eq. (1.14) a selfconsistent single-particle scheme for calculating ground-state properties; this scheme is implemented and solved—usually in an iterative way—for real materials. The calculations are completely *ab-initio*, in the sense that no empirical information is fed in at any stage. The form of $\epsilon_{xc}(n)$ to be used in calculations can be borrowed from various sources; for densities of interest in condensed matter, the most accurate electron-gas data on the market are those of the quantum Monte Carlo simulation of Ceperley and Alder,²¹ illustrated below in Chapter 4.

Once a selfconsistent solution of the KS equation is achieved, the total energy of the system follows from the results of the previous Section as:

$$E_{tot}[n] = E_{ext} + \int d\mathbf{r} V_{ext}(\mathbf{r})n(\mathbf{r}) + T_s[n] + E_H[n] + E_{xc}[n] . \quad (1.18)$$

Here only the term T_s requires explicit knowledge of the KS orbitals, the other terms being functions of them via the electron density only.

In some known pathological cases,²² the electronic state of minimum total energy is *not* obtained upon occupancy of the lowest-energy KS single-particle orbitals. This fact does not invalidate anyhow the DFT formulation, since the Euler equation leading to Eq. (1.14) only provides a set of stationary points, and the absolute minimum has to be searched amongst them.

The LDA is exact (by construction) for a uniform system, and is expected to be good for systems of slowly varying density: indeed, it works much better than expected in a wide range of different materials,^{14–16} provided only that the electronic system is not too strongly correlated. Some reasons for this have been *a posteriori* found,¹¹ while some attempts within the scope of DFT, but beyond LDA, do not seem to have opened new avenues up to date.

The outstanding performance of LDA is demonstrated very perspicuously in a recent benchmark calculation by Ballone *et al.*,²³ where finite (N up to 20) jellium spheres are studied. For a given sphere volume \mathcal{V} , the (positive) density of background charge is constant, equal to $-en_0 = -eN/\mathcal{V}$, up to the sphere radius, and drops to zero at the boundary. In such a system, the amazing finding of Ballone *et al.* is that the electron density obtained from a fully correlated quantum Monte Carlo calculation is practically indistinguishable from the one obtained from LDA: remarkably, this happens everywhere, *i.e.* in the bulk, surface, and tail regions. This is reported in Fig 1.1 for a 20-electron system, whose “bulk” density n_0 is close to the valence-electron density in Al ($r_s=2$ in jellium units, see Ch. 4).

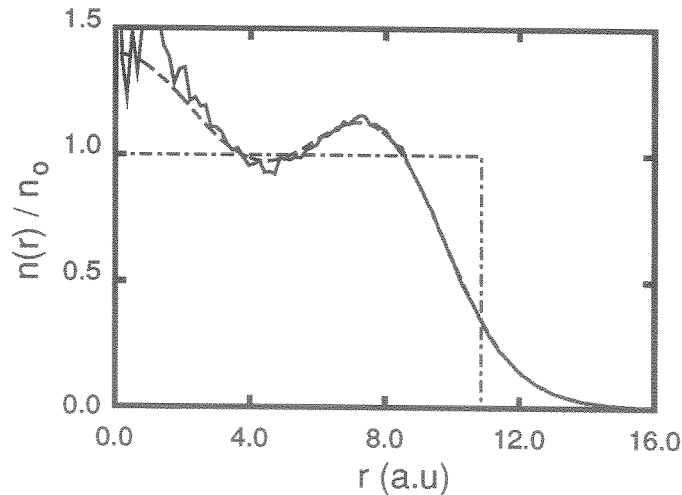


Figure 1.1. Electron density $n(\mathbf{r})$ as a function of the distance from the center of the jellium sphere, normalized to n_0 , see text. Solid: quantum Monte Carlo; dashed: LDA; dash-dotted: background. Unpublished figure, from the calculations of Ref. 23; courtesy of P. Ballone.

1.4 Minimization strategies

The total energy of the system E_{tot} , Eq. (1.5), is variational in the trial KS orbitals. Therefore one can directly search for the minimum of Eq. (1.18) using the set of the occupied orbitals $\{\varphi_i(\mathbf{r})\}$ as variational parameters, subject to the orthonormality constraint, and without making use of the associated Euler equation, *i.e.* of the KS equation. This direct path has become in modern times a fashionable alternative to the more traditional minimization scheme (dating back to Hartree), where a selfconsistent solution of the KS equation is found iteratively. Very efficient schemes can be obtained by adapting the well known method of conjugate-gradient minimization to the direct minimization of the electronic energy functional within LDA.^{24,25}

So far we have discussed the electronic many-body problem, at a fixed configuration of ionic positions $\{\mathbf{R}_I\}$, and in the spirit of the Born-Oppenheimer approximation. However, if the aim is finding the configuration of structural equilibrium at $T=0$, one can directly search for the global minimum of Eq. (1.18), where both $\{\varphi_i(\mathbf{r})\}$ and $\{\mathbf{R}_I\}$ are chosen on the same ground as variational parameters. The drawback of this approach is that—even for systems of moderate size—the functional has many different minima in the parameter space, and the traditional minimization algorithms will likely fall into a local minimum instead than into the global one. A smart solution to complex optimization problems

comes from the simulated-annealing concept, introduced in 1983 by Kirkpatrick *et al.*²⁶. Basically, an objective function $E(\lambda)$ is minimized by generating a succession of configurations λ 's (*i.e.* points in the parameter space) with a Boltzmann-type probability $\exp(-\beta E(\lambda))$, via a standard Monte Carlo sampling procedure.²⁷ At high values of β (*i.e.* of the inverse fictitious temperature) the configuration of lowest $E(\lambda)$ is reached, while wider regions of the parameter space are visited at lower β values. Slow simulated annealing helps the system to sample several local minima before "choosing" to fall into the lowest one. A further appealing feature of the method is the fact that the writing of a computer code performing simulated annealing is a very easy task.²⁸

The simulated-annealing concept has been brought in 1985 to electronic structure theory by Car and Parrinello,²⁹ providing the most significant advance of the 80's in electronic structure theory; tutorial accounts of the method can be found in Refs. 30-33. The simulated annealing in the Car-Parrinello method is *not* performed via Monte Carlo sampling, and *dynamical* simulated annealing—via molecular dynamics³⁴ algorithms—is performed instead. To this purpose, Car and Parrinello define a fictitious dynamical system by considering the parameters $\varphi_i(\mathbf{r})$ and \mathbf{R}_i in the E_{tot} functional to be dependent on time. A generalized classical Lagrangian is introduced, where the ions have their physical mass, while the "electronic" masses are arbitrary parameters, which control the rate of variation of the electronic degrees of freedom; the orthonormality is translated into simple holonomic constraints. A simulation run is started at high temperature; after thermal equilibration, where a large portion of the configuration space is sampled, the system is cooled by slowing down the velocities (electronic and ionic). The nature of the thermal motion evolves in time, switching from random walk to a steep approach to the closest minimum. Eventually, and for sufficiently slow cooling rates, the system is caught in the absolute minimum at low T.

During a global optimization search the system visits in general nonphysical regions in the $\{\varphi_i(\mathbf{r})\}$, $\{\mathbf{R}_i\}$ parameter space, *i.e.* configurations which are *not* on the ground-state Born-Oppenheimer surface. This is irrelevant as far as one is interested in the end product only, *i.e.* the absolute T=0 minimum. Nonetheless, the Car-Parrinello method is capable of doing much more than this: in fact, their fictitious dynamics provides the possibility of simulating real dynamics of the nuclei on the Born-Oppenheimer surface, thus opening new avenues to quantum-mechanical studies of finite-temperature properties. The tunable knob allowing this performance is the electronic mass: playing with this parameter, it is possible to hinder thermal equilibration between the ionic and electronic degrees of freedom. While the ionic temperature is chosen as the physical temperature of interest, the (fictitious) electronic temperature is kept close to zero, implying that the trajectories lie on the Born-Oppenheimer surface. It has been found that such a metastable configuration has indeed a lifetime comparable with the duration of the simulation run (typically a few picoseconds), at least for nonmetallic systems.

A large number of results, on many different physical properties, has been obtained by several groups via the Car-Parrinello scheme, upon exploiting its unique features particularly in dealing with low-symmetry systems.³⁰⁻³³ A (nonexhaustive) list of the materials and systems studied includes: semiconductors and simple metals (in crystalline, liquid and amorphous states), clusters, defects, grain boundaries, and surfaces.

1.5 Basis functions and pseudopotentials

The single-particle orbitals of a periodic solid have the Bloch form: they are expanded—at a given \mathbf{q} (quasimomentum) vector—over a finite basis set in order to solve variationally the KS Schrödinger equation.

A first obvious choice would be a plane-waves (PW) basis set: besides conceptual simplicity, PW's have invaluable *numerical* advantages (unbiased convergence behavior, simple form of matrix elements, simple integration of Poisson equation, use of fast-Fourier transform techniques, &c.). Unfortunately PW's cannot be straightforwardly used because of the fast oscillations of the orbitals in a neighborhood of the nuclei, which would demand an enormous basis size to be described with acceptable resolution.

The second obvious choice is in a sense the opposite extreme, emphasizing the atomic-like nature of the crystal states: expansion over a localized basis, or tight-binding method. This is feasible, but computationally very heavy, owing to huge numbers of multi-center integrals and slow convergence of Coulomb sums. The problems are the same as in variational Hartree-Fock, and a good account of them can be found in Ref. 13.

Several kinds of different expansions have been proposed over the years,³⁵⁻³⁷ but very few methods survive in the context of first-principle calculations. Two extensively used methods are the LMTO (linear combination of muffin-tin orbitals)^{38,39} and the LAPW (linear augmented plane-wave method):^{40,42} these methods are also called “all-electron”, because they provide an explicit expansion of *all* the KS orbitals of the system.

A different viewpoint is provided by pseudopotentials.¹⁰ This concept supplies smooth wavefunctions, which are comfortably expanded in either PW's or localized orbitals. In many situations, the core electrons do not contribute to the properties of interest and can be “frozen” in their free-atom configuration, while the chemistry and the physics are dominated by the behavior of valence electrons. The idea is to map the all-electron, frozen-core problem, onto an equivalent problem involving valence electrons only, and where the orbitals are smooth.⁴⁴ The formal transformation of the Hamiltonian is orthogonalization to core states, giving pseudowavefunctions and pseudopotentials. It turns out that the orthogonalization “cancels” to a large extent the ionic attraction in the core region:⁴⁵ this cancellation principle is often referred to as a “theorem”. The pseudowavefunction of the lowest valence state is nodeless, and all the in-

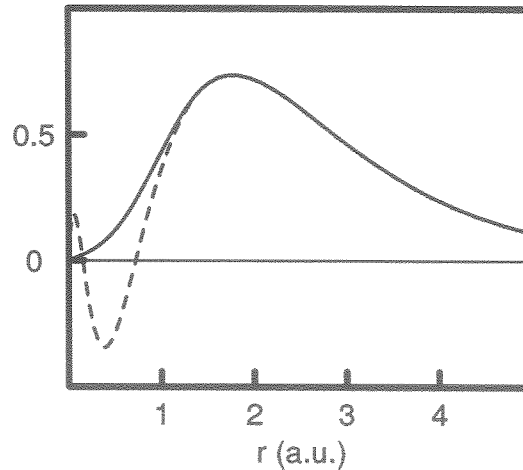


Figure 1.2. Comparison of the 3s radial wavefunctions of Si, pseudo (solid) vs. all-electron (dashed). Pseudization produces a nodeless function, which coincides with the all-electron one beyond a “core” radius; The two functions plotted have the same norm. After the tabulation of Ref. 43.

interesting ones are (hopefully) smooth as well. The ionic pseudopotential^{46,47} consists of a Coulomb attractive term, whose charge is the valence one, plus a short-range part, mostly due to nonclassical orthogonalization repulsion, but which also has contributions from the true interactions with core electrons.

The pseudopotential approach is useful only in situations where core overlap between neighboring ions is negligible. But even in these situations the frozen-core *ansatz* has some subtleties unexpected at first glance: its validity relies upon an important result of von Barth and Gelatt.^{48,47}

Several recipes were proposed for generating pseudopotentials,⁴⁶ and all along the sixties and the seventies the use of pseudopotentials has been an invaluable approximation, providing a great number of very remarkable results in solid and liquid state physics. But, after 1979, the pseudopotential concept made a transition: instead of being an *uncontrolled* approximation, the pseudopotentials of the modern generation are a mathematical transformation of the frozen-core Hamiltonian, being essentially exact in a wide energy range: these pseudopotentials enjoy a large transferability to different chemical environments. The key idea making this step possible is norm conservation, first proposed in quantum chemistry^{49,50} and brought to condensed matter physics by Hamann, Schlüter and Chiang.⁵¹ The concept of norm-conservation is illustrated in Fig. 1.2, where it is shown that the pseudocharge distribution exactly reproduces the full valence charge beyond the “core” radius. The pseudopotentials of the old generation,⁴⁶ instead, were typically tailored after eigenvalue

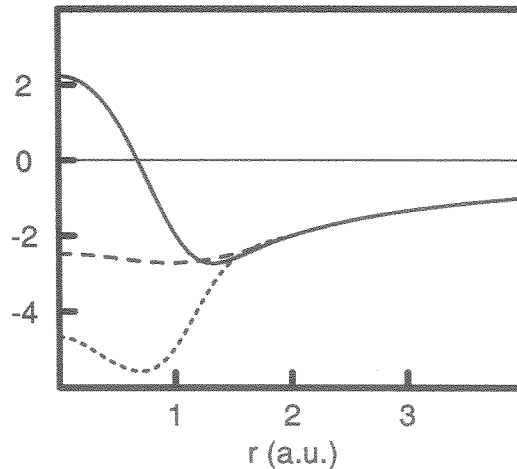


Figure 1.3. l -dependent radial pseudopotential (a.u.) of the Si ion core: s (solid), p (dashed), and d (dotted); all of them coincide with $-4e^2/r$ beyond the core radius. After the tabulation of Ref. 43.

properties only. The tradeoff to be paid for norm conservation is nonlocality, *i.e.* the radial ionic pseudopotential explicitly depends on the angular momentum l : this is illustrated in Fig. 1.3; such a weak form of nonlocality is often called “quasilocality”. Since nonlocality is mainly a technical problem, in this book most of the formal results will be presented in the local form, although state-of-the-art implementations use norm-conserving pseudopotentials with almost no exception. Useful tabulations of modern pseudopotentials are available for applications.^{43,52} A very good account of theory and applications of modern pseudopotentials—updated to 1989—can be found in Ref. 47.

The atomic species where norm-conserving pseudopotentials show their best performances are those where the bonding is of sp kind (typically group I metals, group IV and III-V semiconductors), with the exception of first-row elements. In other atoms a good transferability can be achieved only by pseudopotentials which are very “hard”, meaning with this that they need a large PW expansion in order to achieve convergence in a given calculation: typical “unfriendly” atoms—from a pseudopotential viewpoint—are transition metals and first-row elements. We illustrate the case of Cu, whose $3d$ all-electron wavefunction is shown in Fig. 1.4, dashed line. The wavefunction is nodeless, and very localized. Since the core radius must be smaller than half a typical bondlength, any norm-conserving pseudowavefunction needs large PW expansions. These unfriendly atoms are mastered within the ultrasoft pseudopotential scheme proposed by Vanderbilt,⁵⁴ tailored to generate pseudowavefunctions which are smooth, although *not* norm-conserving; both features are easily detected from the plot of Fig. 1.4, where the pseudowavefunction is shown as a solid line. The missing charge is accurately accounted for by an *ad-hoc* procedure in the selfconsistent

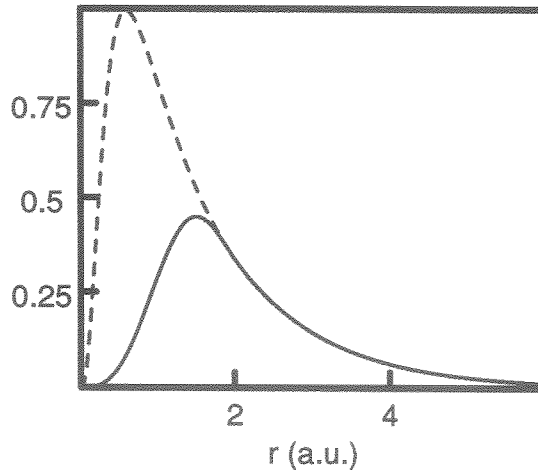


Figure 1.4. Cu 3d all-electron radial wavefunction (dashed); pseudowavefunctions generated from the “ultrasoft” Vanderbilt pseudopotential (solid). After the calculation of Ref. 53; courtesy of A. Pasquarello.

scheme: as a result, the ultrasoft pseudopotentials achieve both high transferability and fast PW convergence. For details about the implementation and for a review of some applications, see Ref. 33.

1.6 Total energy expansions

Suppose you can modify the ground state of the system through some tunable knobs which act on the electrons via the external one-body potential, Eq. (1.2), and you wish to know how turning the knobs affects the total energy E_{tot} . For the sake of simplicity we summarize all the tunable parameters of V_{ext} in a single one. At this point we choose to remain completely general about the physical nature of the perturbation; we only assume that $V_{ext}(\lambda)$ is analytic around the unperturbed $\lambda=0$ situation. In a typical implementation, the perturbation could be a lattice distortion, and the parameter λ a phonon coordinate.

The total energy is

$$E_{tot}(\lambda) = E_{ext}(\lambda) + \langle \Psi_\lambda | T + U + V_{ext}(\lambda) | \Psi_\lambda \rangle, \quad (1.19)$$

where the first term is purely classic and the second one is the electronic energy $E_{el}(\lambda)$. We take its derivative

$$E'_{el}(\lambda) = \langle \Psi_\lambda | V'_{ext}(\lambda) | \Psi_\lambda \rangle + E_{el}(\lambda) (\langle \Psi'_\lambda | \Psi_\lambda \rangle + \langle \Psi_\lambda | \Psi'_\lambda \rangle), \quad (1.20)$$

and using normalization we recognize the familiar expression, going under the name of Hellman-Feynman theorem^{55,56}

$$E'_{el}(\lambda) = \langle \Psi_\lambda | V'_{ext}(\lambda) | \Psi_\lambda \rangle. \quad (1.21)$$

We now expand the energy in powers of λ : $E_{tot}(\lambda) = E_{tot}^{(0)} + E_{tot}^{(1)} + E_{tot}^{(2)} + \dots$, and analogously for the perturbation $V_{ext}(\lambda)$. The zero-order term is the unperturbed energy; the first order term is got from Eq. (1.21) at $\lambda = 0$:

$$E_{tot}^{(1)} = E_{ext}^{(1)} + \langle \Psi_0 | V_{ext}^{(1)} | \Psi_0 \rangle = E_{ext}^{(1)} + \int d\mathbf{r} V_{ext}^{(1)}(\mathbf{r}) n^{(0)}(\mathbf{r}). \quad (1.22)$$

The message of Eq. (1.22) is that knowledge of the unperturbed electronic density is enough to evaluate first-order energy corrections induced by a given perturbation.

Taking one more derivative from Eq. (1.21), and evaluating it at $\lambda = 0$, we get the second order term

$$\begin{aligned} E_{el}^{(2)} &= \frac{\lambda^2}{2} (\langle \Psi_0 | V''_{ext}(0) | \Psi_0 \rangle + \langle \Psi'_0 | V'_{ext}(0) | \Psi_0 \rangle + \langle \Psi_0 | V'_{ext}(0) | \Psi'_0 \rangle) = \\ &= \langle \Psi_0 | V_{ext}^{(2)} | \Psi_0 \rangle + \frac{\lambda}{2} (\langle \Psi'_0 | V_{ext}^{(1)} | \Psi_0 \rangle + \langle \Psi_0 | V_{ext}^{(1)} | \Psi'_0 \rangle). \end{aligned} \quad (1.23)$$

Finally, the term in brackets involves only the first-order potential and the first derivative of the density with respect to λ , as it is easily recognized after a glance to Eq. (1.3). Therefore the second-order term in the total energy expansion is:

$$E_{tot}^{(2)} = E_{ext}^{(2)} + \int d\mathbf{r} V_{ext}^{(2)}(\mathbf{r}) n^{(0)}(\mathbf{r}) + \frac{1}{2} \int d\mathbf{r} V_{ext}^{(1)}(\mathbf{r}) n^{(1)}(\mathbf{r}). \quad (1.24)$$

The content of Eq. (1.24) is that knowledge of the first-order density polarization induced by the perturbation is enough to evaluate second-order corrections to the total energy. This is a very general and old theorem, rediscovered several times in different contexts; within harmonic lattice dynamics the formulation is apparently due to De Cicco and Johnson,⁵⁷ whose proof is furthermore limited to the Hartree-Fock approximation. It has been shown here that Eq. (1.24) is an exact result in a many-body formulation.

Iteration of the previous path would easily show that the lowest n -th order density response of the electronic system determines the $(n+1)$ -th order energy correction. A more powerful result is the so-called “ $2n+1$ ” theorem, stating that the lowest n -th order corrections to the KS *wavefunctions* determine the energy correction up to the $(2n+1)$ -th order. Although even this result is not new,^{58,59} an elegant proof has appeared recently.⁶⁰

1.7 References

- [1.1] C. Kittel, *Introduction to Solid State Physics*, 6th. edition (Wiley, New York, 1986).

- [1.2] R.E. Watson and M. Weinert, *Phys. Rev. B* **30**, 1641 (1984).
- [1.3] F. Willaime, and C. Massobrio, *Phys. Rev. Lett.* **63**, 2244 (1989).
- [1.4] G.V. Chester, *Adv. Phys.* **10**, 357 (1961).
- [1.5] E.G. Brovman and Yu.M. Kagan, in: *Dynamical Properties of Solids*, edited by G.K. Horton and A.A. Maradudin, vol. 1 (North-Holland, Amsterdam, 1974), p. 191.
- [1.6] W.A. Harrison, *Pseudopotentials in the Theory of Metals* (Benjamin, New York, 1966).
- [1.7] J. Hafner, *From Hamiltonians to Phase Diagrams* (Springer, Berlin, 1987).
- [1.8] N.W. Ashcroft, *Phys. Rev. B* **39**, 10552 (1989).
- [1.9] M.E. Lines and A.M. Glass, *Principles and Applications of Ferroelectrics and Related Materials* (Clarendon Press, Oxford, 1977).
- [1.10] M.L. Cohen, V. Heine, and J.C. Phillips, *The Quantum Mechanics of Materials*, *Scientific American* **246**(6), 66 (1982).
- [1.11] *Theory of the Inhomogeneous Electron Gas*, edited by S. Lundqvist and N.H. March (Plenum, New York, 1983).
- [1.12] *Density Functional Methods in Physics*, edited by R.M. Dreizler and J. da Providencia (Plenum, New York, 1985).
- [1.13] C. Pisani, R. Dovesi, and C. Roetti, *Hartree-Fock Ab-Initio Treatment of Crystalline Systems* (Springer, Berlin, 1988).
- [1.14] See for instance: R.M. Martin, in *Electronic Structure, Dynamics and Quantum Structural Properties of Condensed Matter*, edited by J.T. Devreese and P. Van Camp (Plenum, New York, 1985), p. 175.
- [1.15] M.L. Cohen, in: *Highlights in Condensed Matter Theory*, edited by F. Bassani, F. Fumi and M.P. Tosi (North Holland, Amsterdam, 1985).
- [1.16] R.O. Jones and O. Gunnarson, *Rev. Mod. Phys.* **61**, 689 (1989).
- [1.17] P. Hohenberg and W. Kohn, *Phys. Rev.* **136**, B864 (1964); W. Kohn and L.J. Sham, *Phys. Rev.* **140**, A1133 (1965).
- [1.18] J.P. Perdew and M. Levy, *Phys. Rev. Lett.* **51**, 1884 (1983); L.J. Sham and M. Schlüter, *ibid.* **51**, 1888 (1983).
- [1.19] A. Fetter and J.D. Walecka, *Quantum Theory of Many-Particle Systems* (McGraw-Hill, New York, 1971).

- [1.20] K.S. Singwi and M.P. Tosi, in *Solid State Physics*, edited by H. Ehrenreich, F. Seitz and D. Turnbull, vol 36 (Academic, New York, 1981), p.177.
- [1.21] D.M. Ceperley and B.J. Alder, *Phys. Rev. Lett.* **45**, 566 (1980); J. Perdew and A. Zunger, *Phys. Rev. B* **23**, 5048 (1981).
- [1.22] Ki-Soo Sohn, S. Lee, D.M. Bylander, and L. Kleinman, *Phys. Rev. B* **39**, 9983 (1989).
- [1.23] P. Ballone, C.J. Umrigar, and P. Delaly, *Phys. Rev. B* **45**, 6293 (1992).
- [1.24] I. Stich, R. Car, M. Parrinello, and S. Baroni, *Phys. Rev. B* **39**, 4997 (1989).
- [1.25] M.C. Payne, M.P. Teter, D.C. Allan, T.A. Arias, and J.D. Joannopoulos, *Rev. Mod. Phys.* **64**, 51 (1992).
- [1.26] S. Kirkpatrick, G.D. Gelatt jr., and M.P. Vecchi, *Science* **220**, 671 (1983).
- [1.27] M.P. Allen and D.J. Tildesley, *Computer Simulations of Liquids* (Clarendon, Oxford, 1987).
- [1.28] W.H. Press, B.P. Flannery, S.A. Teukolsky, and W.T. Vetterling, *Numerical Recipes* (Cambridge University Press, Cambridge, 1986).
- [1.29] R. Car and M. Parrinello, *Phys. Rev. Lett.* **55**, 2471 (1985).
- [1.30] R. Car and M. Parrinello, in: *Simple Molecular Systems at Very High Density*, edited by P. Loubeyre, A. Polian, and N. Boccara, NATO ASI Series B, vol. 186 (Plenum Publishing, New York, 1988), p. 455.
- [1.31] G. Pastore, E. Smargiassi, and F. Buda, *Phys. Rev. B* **44**, 6334 (1991).
- [1.32] G. Galli and M. Parrinello, in *Computer Simulation in Material Science*, edited by M. Meyer and V. Pontikis (Kluwer, Dordrecht, 1991).
- [1.33] G. Galli and A. Pasquarello, in *New Perspectives on Computer Simulations in Chemical Physics*, edited by M. Allan and D. Tildesley (Kluwer, Dordrecht, 1992).
- [1.34] *Molecular Dynamics Simulations of Statistical mechanical Systems*, edited by G. Ciccotti and W.G. Hoover (North-Holland, Amsterdam, 1985).
- [1.35] F. Bassani and G. Pastori Parravicini, *Electronic States and Optical Transitions in Solids* (Pergamon, Oxford, 1975).
- [1.36] G. Grosso, in: *Crystalline Semiconducting Materials and Devices*, edited by P. N. Butcher, N. H. March and M. P. Tosi (Plenum Publishing, New York, 1986), p. 55.

- [1.37] P. Giannozzi, G. Grosso, and G. Pastori Parravicini, *Rivista del Nuovo Cimento* **13** (3), 1 (1990).
- [1.38] H.L. Skriver, *Muffin-Tin Orbitals and Electronic Structure* (Springer, Berlin, 1984).
- [1.39] O.K. Andersen, O. Jepsen and M. Sob, in: *Electronic Band Structure and Its Applications*, edited by M. Yussouf (Springer, Berlin, 1987), p. 1.
- [1.40] O.K. Andersen, *Phys. Rev. B* **12**, 3060 (1975).
- [1.41] M. Weinert, M. Wimmer, and A.J. Freeman, *Phys. Rev. B* **26**, 4571 (1982).
- [1.42] H.J.F. Jansen and A.J. Freeman, *Phys. Rev. B.* **30**, 561 (1984).
- [1.43] G.B. Bachelet, D.R. Hamann, and M. Schlüter, *Phys. Rev. B* **26**, 4199 (1982).
- [1.44] M.H. Cohen and V. Heine, *Phys. Rev.* **122**, 1821 (1961).
- [1.45] J.C. Phillips and L. Kleinman, *Phys. Rev.* **116**, 287 (1959).
- [1.46] V. Heine, in *Solid State Physics*, edited by H. Ehrenreich, F. Seitz, and D. Turnbull, vol **24** (Academic, New York, 1970), p.1; M.L. Cohen and V. Heine, *ibid.* p.250.
- [1.47] W.E. Pickett, *Computer Phys. Reports* **9**, 115 (1989).
- [1.48] U. von Barth and C.D. Gelatt, *Phys. Rev. B* **21**, 2222 (1980).
- [1.49] L.R. Kahn and W.A. Goddard III, *J. Chem. Phys.* **56**, 2685 (1972).
- [1.50] C.F. Melius and W.A. Goddard III, *Phys. Rev. A* **10**, 1528 (1974).
- [1.51] D.R. Hamann, M. Schlüter and C. Chiang, *Phys. Rev. Lett.* **43**, 1494 (1979).
- [1.52] R. Stumpf, X. Gonze, and M. Scheffler, *Fritz-Haber-Institut Research Report*, Berlin 1990, unpublished.
- [1.53] K. Laasonen, A. Pasquarello, R. Car, C. Lee, and D. Vanderbilt, *Phys. Rev. B*, in press.
- [1.54] D. Vanderbilt, *Phys. Rev. B* **41**, 7892 (1990).
- [1.55] P. Ehrenfest, *Z. Phys.* **45**, 455 (1927); R.P. Feynman, *Phys. Rev.* **56**, 340 (1939).
- [1.56] B.M. Deb, *Rev. Mod. Phys.* **43**, 22 (1973).

- [1.57] P.D. De Cicco and F.A. Johnson, Proc. Roy. Soc. London Ser. A **310**, 111 (1969).
- [1.58] R. Moccia, Chem. Phys. Lett. **5**, 260 (1970).
- [1.59] R.D. Amos, in: *Ab-Initio Methods in Quantum Chemistry - I*, edited by K.P. Lawley (Wiley, New York, 1987), p. 99.
- [1.60] X. Gonze and J.P. Vigneron, Phys. Rev. B **39**, 13120 (1989).

Chapter 2

Screening and electrostatics

2.1 Electronic screening

Since we intend to study the effect of perturbations upon the electronic ground state of the system, we split the external potential as

$$V_{ext}(\mathbf{r}) = V_{ext}^{(0)}(\mathbf{r}) + \delta V_{ext}(\mathbf{r}), \quad (2.1)$$

where $V_{ext}^{(0)}(\mathbf{r})$ is the electron-ion potential in the isolated system at equilibrium geometry, and $\delta V_{ext}(\mathbf{r})$ is the perturbation due to some other source.

In order to make contact with traditional electrostatics, we may write

$$\delta V_{ext}(\mathbf{r}) = e\phi_0(\mathbf{r}), \quad (2.2)$$

where the potential $\phi_0(\mathbf{r})$ describes a static field:

$$\mathbf{E}_0(\mathbf{r}) = -\nabla\phi_0(\mathbf{r}). \quad (2.3)$$

This *microscopic* field is the unscreened (or “bare”) field: it is in fact, by definition, the field that would be generated by the source in the absence of electronic response. As usual in electrostatics,¹ one could perform the average of $\mathbf{E}_0(\mathbf{r})$ over a macroscopic region of the system to get the *macroscopic* field. The macroscopic average of the bare field must *not* be identified, in general, with the electric displacement vector \mathbf{D} : the former is, according to Eq. (2.3), longitudinal while the curl of \mathbf{D} may be nonzero.

For a given bare perturbation, the electronic systems polarizes, *i.e.*

$$\Delta n(\mathbf{r}) = n(\mathbf{r}) - n^{(0)}(\mathbf{r}) \neq 0, \quad (2.4)$$

where $n^{(0)}(\mathbf{r})$ and $n(\mathbf{r})$ are the densities of the unperturbed and perturbed crystals, respectively. Using Poisson equation, the *screened* electrostatic potential inside the piece of matter is

$$\phi(\mathbf{r}) = \phi_0(\mathbf{r}) + e \int d\mathbf{r}' \frac{\Delta n(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|}. \quad (2.5)$$

The microscopic local field is

$$\mathbf{E}(\mathbf{r}) = -\nabla\phi(\mathbf{r}); \quad (2.6)$$

this is by definition the field acting on a *classical* and *infinitesimal* test charge in the system. The macroscopic average of $\mathbf{E}(\mathbf{r})$ yields the familiar macroscopic electric field.¹

The basic problem in screening is therefore to get the polarization electronic density $\Delta n(\mathbf{r})$: this is a fundamental many-body problem, where quantum mechanics plays a major role. Unless differently stated, I will always refer to purely electronic, or clamped-nuclei, polarization. This is the basic quantity; contributions due to the relaxation of the nuclear coordinates can be added to it via lattice dynamics (see below, Section 2.4.3). Most of the electronic polarization, furthermore, comes in simple materials from the relaxation of valence states: core polarization can therefore be neglected and a pseudopotential approach is adequate in most cases.

The polarization density $\Delta n(\mathbf{r})$ is a functional of the perturbation $\delta V_{ext}(\mathbf{r})$, and we are interested here in the term $n^{(1)}(\mathbf{r})$ which is *linear* in it. This is written

$$n^{(1)}(\mathbf{r}) = \int d\mathbf{r}' \chi(\mathbf{r}, \mathbf{r}') \delta V_{ext}(\mathbf{r}'), \quad (2.7)$$

which defines the linear operator χ , known as the density response of the system. Eq. (2.7) is quite general, and applies to atoms, molecules or condensed matter. The density response is the functional derivative of the electronic density with respect to the external potential, evaluated at $\delta V_{ext}(\mathbf{r}) = 0$, and is therefore a ground-state property of the unperturbed system. As such, it is totally symmetric under the operations which leave the unperturbed system invariant.

When linear response is valid, *i.e.* $\Delta n(\mathbf{r}) \simeq n^{(1)}(\mathbf{r})$, Eq. (2.5) reads

$$\phi(\mathbf{r}) = \int d\mathbf{r}' \left[\delta(\mathbf{r} - \mathbf{r}') + e^2 \int d\mathbf{r}'' \frac{\chi(\mathbf{r}'', \mathbf{r}')}{|\mathbf{r} - \mathbf{r}''|} \right] \phi_0(\mathbf{r}'). \quad (2.8)$$

Introducing the following definition of the inverse dielectric response operator:

$$\varepsilon^{-1}(\mathbf{r}, \mathbf{r}') = \delta(\mathbf{r} - \mathbf{r}') + e^2 \int d\mathbf{r}'' \frac{\chi(\mathbf{r}'', \mathbf{r}')}{|\mathbf{r} - \mathbf{r}''|}, \quad (2.9)$$

the basic equation of linear dielectric screening reads

$$\phi(\mathbf{r}) = \int d\mathbf{r}' \varepsilon^{-1}(\mathbf{r}, \mathbf{r}') \phi_0(\mathbf{r}'). \quad (2.10)$$

Once the operator ε^{-1} known, any perturbation can be screened within the linear approximation, and everything reduces to a matter of electrostatics; but the evaluation of ε^{-1} is a complex many-body problem. The operator ε^{-1} , as defined here, is also called the test-charge test-charge inverse dielectric response.

In fact, it is easy to realize that the mutual interaction energy between two classical and infinitesimal test charges Q_1 and Q_2 , located at \mathbf{r}_1 and \mathbf{r}_2 , is given in terms of it as

$$E_{int}(\mathbf{r}_1, \mathbf{r}_2) = Q_1 Q_2 \int d\mathbf{r}' \epsilon^{-1}(\mathbf{r}_1, \mathbf{r}') / |\mathbf{r}' - \mathbf{r}_2|. \quad (2.11)$$

We also notice that E_{int} must be symmetric under interchange of the coordinates \mathbf{r}_1 and \mathbf{r}_2 , while Eq. (2.11) does not appear symmetric *in form*: the fact is that $\epsilon^{-1}(\mathbf{r}, \mathbf{r}')$ is in general *not* a symmetric operator. Instead, it is easy to prove that symmetry of E_{int} implies that the density response $\chi(\mathbf{r}, \mathbf{r}')$ for an arbitrary system *is* a symmetric operator.

2.2 Diagonal screening

The simplest system where one can study screening is an homogeneous and isotropic medium. The basic screening operators χ and ϵ^{-1} are then translationally invariant, *i.e.* functions of $|\mathbf{r} - \mathbf{r}'|$ only, and our previous expressions take the form of convolution products. Switching to reciprocal space we get *e.g.* from Eq. (2.7):

$$n^{(1)}(\mathbf{k}) = \chi(k) \delta V_{ext}(\mathbf{k}); \quad (2.12)$$

i.e. the operator χ is *diagonal* in reciprocal space. In this representation a product of operators is a simple algebraic product, and analogously for the inverse: Eq. (2.9) becomes

$$\epsilon^{-1}(k) = \frac{1}{\epsilon(k)} = 1 + \frac{4\pi e^2}{k^2} \chi(k). \quad (2.13)$$

The function $\epsilon(k)$ is known as the (static) dielectric function of the homogeneous system: the transformation of Eq. (2.10) is

$$\phi(\mathbf{k}) = \phi_0(\mathbf{k}) / \epsilon(k). \quad (2.14)$$

The content of Eq. (2.14) is that perturbations of different wavelengths are differently screened; a reasonable guess is that the longer is the wavelength the most effective is the screening. There is a typical screening length, such that at wavelengths much smaller than it the medium is unable to respond; the actual value of this screening length depends on the physical mechanism involved. I anticipate that for purely electronic screening in a crystalline material at zero temperature such length is of the order of the interatomic distance in either metals, semiconductors or insulators. Therefore, as a general feature, $\epsilon(k)$ is a monotonically decreasing function of k , going to 1 for large values of k . Typical and very simple model dielectric functions for a metal (Al) and for a semiconductor (Ge) are shown in Fig 2.1(a).

We illustrate diagonal screening, and the differences between metals and nonmetals, upon a specific perturbation: a point charge at the origin. The bare potential $\phi_0(r) = Q/r$ has a Coulomb long-range tail, and its Fourier transform, as a consequence, has a k^{-2} singularity around $k=0$:

$$\phi_0(k) = 4\pi Q/k^2. \quad (2.15)$$

The short-range behavior of the potential $\phi_0(r)$ is related instead to the power-law decrease of $\phi_0(k)$ at large k .

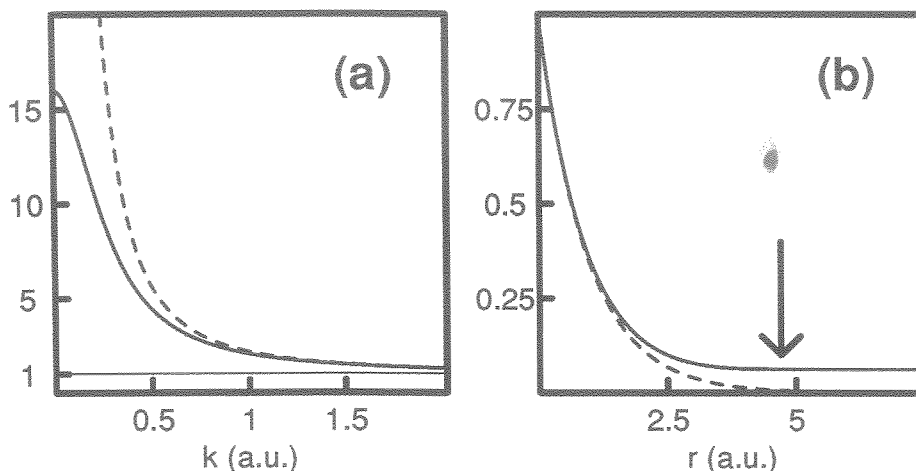


Figure 2.1. (a): Diagonal dielectric functions $\epsilon(k)$ for Al (dashed) and Ge (solid) (b): Ratio $\phi(r)/\phi_0(r)$ between the screened and unscreened potentials of a point charge, from the dielectric functions shown in (a).

Let me consider now the screened potential

$$\phi(r) = \frac{1}{(2\pi)^3} \int d\mathbf{k} e^{i\mathbf{k}\mathbf{r}} \phi(k) = \frac{1}{(2\pi)^3} \int d\mathbf{k} e^{i\mathbf{k}\mathbf{r}} \frac{4\pi Q}{\epsilon(k)k^2} : \quad (2.16)$$

The ratio $\phi(r)/\phi_0(r)$ has the typical behavior shown in Fig. 2.1(b), whose main features are easy to understand as follows. In order to simplify matters, I assume in this section that $\epsilon(k)$ is an analytic function for $0 < k < \infty$.

The small r behavior of a function is dominated by the large k behavior of its Fourier transform, and reciprocally, under the hypothesis of analyticity. According to the above considerations $\phi(k) \simeq \phi_0(k)$ at large k and hence $\phi(r) \simeq \phi_0(r)$ at small r . The potential is essentially unscreened at distances from the point charge much smaller than the screening length, both in metals and nonmetals, in agreement with common sense and with Fig. 2.1(b).

Metals and nonmetals manifest a *qualitative* difference at large r , or equivalently at small k , as it is evident in Fig. 2.1. In a *conductor* the screened potential is short-range, *i.e.* no Coulomb tail survives at macroscopic distances and the bare point charge is completely screened by the medium: this stems in fact from elementary Faraday screening in conductors. This implies that the screened potential $\phi(k)$ is regular in reciprocal space, and therefore that the k^{-2} divergence in $\phi_0(k)$ is cancelled in Eq. (2.14) by a divergence of the same kind in the dielectric function.

In a *nonmetal* a point charge is incompletely screened: its potential, over macroscopic distances, is simply the Coulomb one screened by the macroscopic dielectric constant of the medium. This is seen in Fig 2.1(b), where the asymptotic value of the solid line is in fact the inverse macroscopic dielectric constant ($=1/16$ for Ge). It follows that the screened potential must have a k^{-2} divergence: hence $\epsilon(k)$ is regular at small k , and the value of $\epsilon(0)$, Fig. 2.1(a) is in fact the dielectric constant of the medium. Since purely electronic screening is our main concern, in a polar crystal the relevant constant is ϵ_∞ , also called the “static high frequency” dielectric constant: *i.e.* measured at a frequency much higher than the vibrational motions and much smaller than the optic transitions.²

2.3 Microscopic model screening

2.3.1 Metals

The simplest model for an alkali metal is the electron gas, or “jellium”: the ion cores are smeared into a classical, nonpolarizable, uniform neutralizing background, across which the valence electrons travel freely, while interacting amongst themselves via Coulomb repulsion. The electron gas (or jellium) is a well-defined many-body system, discussed in more detail below, Chapter 4; at densities of interest in condensed matter physics, its ground state is homogeneous and metallic.

The dielectric function of the electron gas has been much studied over the years.^{3,4} In this section I only discuss the simplest level of approximation; more sophisticated treatments are briefly outlined in Chapter 4. I have already observed above that electronic screening is dominated by quantum many-body effects: roughly speaking by indetermination and Pauli principles; the simplest scheme which accounts for both is the Thomas-Fermi (TF) approximation. The TF dielectric function has been first proposed in 1936 by N.F. Mott; elegant derivations are reported in almost every textbook in solid state physics. Here I only quote the result:

$$\epsilon(k) = 1 + k_{TF}^2/k^2, \quad (2.17)$$

where k_{TF} , the TF inverse screening length, is a simple function of the density. This is indeed the dielectric function plotted in Fig. 2.1(a), dashed line, at the

valence electronic density of Al. Using this dielectric function, the potential of a point charge is exponentially screened:

$$\phi(r) = Qe^{-k_{TF}r}/r; \quad (2.18)$$

this is shown in Fig. 2.1(b). At metallic densities the screening length is of the order of interparticle spacing, and this makes shielding very effective.

2.3.2 Semiconductors

A pure semiconductor at zero temperature is of course an insulator, but a special kind of one, having a high value of the electronic dielectric constant ($\epsilon_\infty \sim 10$), a small gap, and a valence electronic distribution which is not too inhomogeneous. It is then, in a sense, “close” to a metal: a simple model for electronic screening is therefore the homogeneous and isotropic one (or “semiconducting electron gas”), which allows using a diagonal dielectric function.

Given that the electron gas is indeed a metal, some prescription to introduce semiconductorlike behavior has to be introduced *ad hoc* into the model. It is remarkable that several rather different prescriptions give nonetheless very similar dielectric functions. During the sixties, the most popular semiconductor dielectric function has been provided by the Penn model,⁵ which however can be solved only numerically. Nowadays the most elegant approach in this class is the TF model of Ref. 6. It generalizes the historical TF screening theory for metals, and furthermore it is simply soluble in closed form: this dielectric function is plotted in Fig. 2.1(a), solid line, for Ge. Incidentally, Ge and Al happen to have almost the same valence density, hence the same k_{TF} .

The screened potential induced by a point-charge impurity within this diagonal model is the one reported in Fig. 2.1(b), solid line; the arrow indicates one bond length. It happens that in all group IV semiconductors the microscopic potential recovers its macroscopic value almost exactly at this nearest-neighbor distance; remarkably, such feature is common essentially to all diagonal models on the market.^{5,7}

When comparing diagonal models to real materials, the first obvious difference is that a solid is *not* homogeneous and isotropic on a microscopic scale: the effects due to lattice periodicity, which are outside the scope of such models, are known under the name of “local-field effects”.^{8,9} Some comments about the importance of local-field effects are given below, Section 2.3.4.

2.3.3 Insulators

Electronic screening in a strong insulator (like a solid rare gas, a ionic or a molecular crystal) is strongly inhomogeneous and *cannot* reasonably be described—on a microscopic scale—by any homogeneous model. Instead, insulators are at the opposite extreme: they can be modeled as an assembly of localized, independently polarizable units (atoms, molecules, ions), which selfconsistently respond

to the local fields generated by the bare perturbation *and* by the induced electric multipoles at the other sites. In this picture the role of local-field effects is emphasized, and no diagonal dielectric function is a sound approximation.

The response of a model system of this kind is very simple in the case where the external perturbation is a constant field, generated *e.g.* by a plane capacitor polarizing the dielectric sample. The problem has been solved for isotropic materials in the nineteenth century:¹⁰ the electronic dielectric constant is found from the classic Clausius-Mossotti selfconsistency relationship

$$\frac{\epsilon_{\infty} - 1}{\epsilon_{\infty} + 2} = \frac{4\pi\alpha}{3\Omega}, \quad (2.19)$$

where a primitive lattice is assumed, Ω is the cell volume, and α is the dipolar polarizability of each unit.

The generalization of Eq. (2.19) to crystals of symmetry lower than cubic is not widely known, but it is feasible indeed: the calculation of local fields requires then explicit evaluation of suitable Ewald sums. All of the tools needed are the ones which are routinely used to cope with electric fields in classical lattice dynamics.¹¹ Finally, I mention that when the actual perturbation is different from a constant field, a selfconsistent solution may be worked out case by case from a generalized Clausius-Mossotti model: this approach goes under the name of Mott-Littleton model.^{12,13}

2.3.4 Clausius-Mossotti vs. diagonal models

I have stated, upon plausibility arguments, that electronic screening in semiconductors is rather well described by a diagonal model, while the Clausius-Mossotti model is much more appropriate in insulators. The two are really alternative extreme pictures: what about screening in real materials?

Why not describing an insulator through a diagonal model, or a covalent material through Clausius-Mossotti? The latter possibility is particularly tempting: one uses the experimental value of ϵ_{∞} in Eq. (2.19), and attributes the resulting polarizability α either to the atoms or to the bonds. I show here, using the specific examples of Ge and GaAs, that both pictures are completely *wrong*. Insight into the physical mechanisms in real materials has been obtained from computer experiments:^{8,14} the discussion given here follows Ref. 15.

Suppose we embed in our dielectric a charged thin capacitor plate, bearing a bare (or unscreened) charge per unit surface σ_0 . The dielectric polarizes, and straightforward macroscopic electrostatics¹ predicts in a neighborhood of the plate an induced electronic charge whose value per unit surface is:

$$\sigma_{pol} = \sigma_0(1/\epsilon_{\infty} - 1). \quad (2.20)$$

Our concern here is the *microscopic* distribution in space of the linearly induced electronic charge $\rho_{pol}(\mathbf{r})=en^{(1)}(\mathbf{r})$: I shall discuss this point first within diagonal

screening, then within the Clausius-Mossotti picture, and finally illustrating the results of a quantum-mechanical detailed calculation. The plate is taken normal to the (001) direction (orientation is irrelevant only for diagonal screening); because of technical reasons, the plate has a *finite*—although small—thickness.

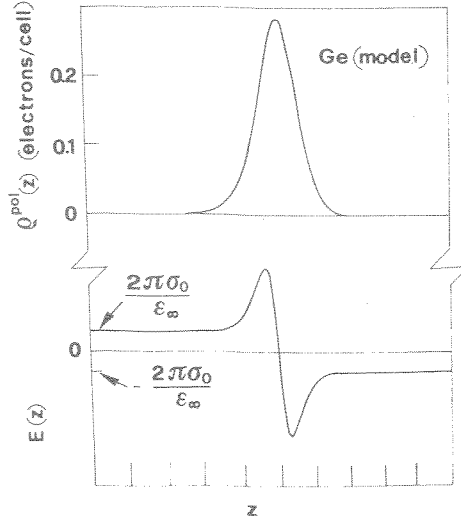


Figure 2.2. Upper panel: Diagonal screening of a negatively charged thin plate in Ge. The bars on the z -axis indicate positions of atomic planes. Bottom panel: microscopic electric field.

Use of a diagonal dielectric function is equivalent to the hypothesis of a homogeneous and isotropic medium: since the perturbation is translationally invariant along xy , the response charge ρ_{pol} will also explicitly depend only upon z . The polarization charge given by the TF dielectric function⁶ of Fig. 2.1(a) for this perturbation is shown in Fig. 2.2: the integrated value of this charge

$$\sigma_{pol} = \int_{-\infty}^{+\infty} dz \rho_{pol}(z) \quad (2.21)$$

is in agreement with the macroscopic value, Eq. (2.20). The basic message of Fig. 2.2 is that the microscopic polarization charge vanishes beyond a distance from the plate which is very close to one bond length. The electric field is also shown in the same figure: this is determined only up to an arbitrary constant, which is fixed by the boundary conditions and *not* by the charge distribution; the symmetric choice is performed here. We further notice that, for an infinitely thin plate, the electric field would be monotonically increasing and discontinuous:

the region of negative slope in the bottom panel of Fig. 2.2 monitors therefore the actual thickness of our plate.

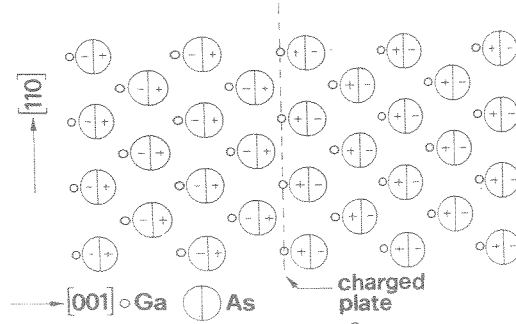


Figure 2.3. Microscopic polarization around a negatively charged thin plate in GaAs, within a Clausius-Mossotti picture (schematic). The projection of the ions in the (110) plane is given; only anions (big circles) are shown as polarized; cations are shown as small circles.

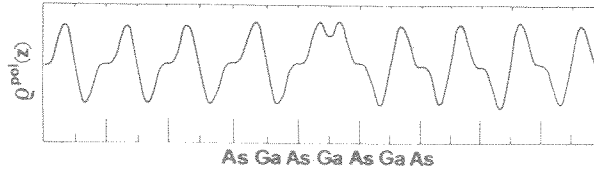


Figure 2.4. Planar average along xy of the Clausius-Mossotti polarization charge. Schematic plot in arbitrary units (longer bars: anionic planes; shorter bars: cationic).

Let us now switch to the alternative extreme picture: our charged plate is embedded in an assembly of individually polarizable units, which selfconsistently respond to the perturbation. We take the specific example of GaAs, schematically shown in Fig. 2.3. The induced charge does *not* vanish far from the plate, and is responsible for oscillations of the microscopic local fields, which reflect the periodicity of the screening medium. Such charge is strongly inhomogeneous, although lattice-periodic in the planes normal to z : taking its planar average $\bar{\rho}_{pol}(z)$, we get a function schematically shown in Fig. 2.4. The integrated value of $\bar{\rho}_{pol}(z)$ is the macroscopic surface charge σ_{pol} , as in Eq. (2.21), although integration of this oscillating function requires some care.¹⁵

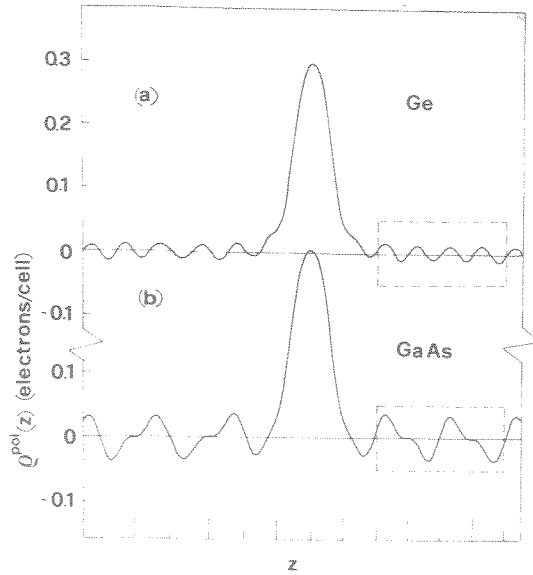


Figure 2.5. Planar average of the induced electronic charge around a thin charged plate embedded in real dielectrics.

Comparison of Fig. 2.2 with Fig. 2.4 shows at a glance the main features of electron-gas-like screening versus local-dipole screening: the next point is to ascertain to which paradigm are real materials closer. At present no experimental probe is powerful enough to settle this issue: but a very accurate understanding has been achieved over the years through computational physics, which often is the unique source of information about microscopic electrostatics. For the screening problem studied here, the quantum-mechanical calculations are performed in Refs. 14,15 : results for Ge and GaAs are shown in Fig. 2.5. Comparing this figure with the previous ones it is immediately realized that real materials are intermediate between the two extreme models and show the basic mechanisms of both: diagonal screening and local-field screening. The relative importance of the two contributions has a definite trend with ionicity: local-field effects are more important in GaAs than in Ge, in agreement with the fact that Clausius-Mossotti is the paradigmatic model in the extreme ionic limit. However, both in Ge and GaAs, local-field effects appear as a fairly small correction to diagonal screening, which is by far the most prominent mechanism for the perturbation discussed here.

The Clausius-Mossotti model tends to be proposed in most textbooks as the only possible paradigm for electronic screening, thus leading to incorrect viewpoints. Looking back at Eq. (2.19), it is tempting to use it for any material and in the *reverse* way: given the experimental ϵ_{∞} , one infers the corresponding value for the “local” polarizability α . In covalent solids the electron

density is—roughly speaking—more polarizable in the bond regions than at the atomic sites, therefore in such materials α has been often identified with a “bond” polarizability.^{16,17} The calculations and pictures reported here demonstrate very perspicuously that the concept of bond polarizability lacks of any physical basis and is incorrect: a covalent material is definitely *not* an assembly of independently polarizable dipoles.

2.4 Macroscopics

2.4.1 Phenomenology

The discussion so far has been about some very basic features of the *microscopic* polarization induced in the many-electron system by an arbitrary bare perturbation. In the rest of this Chapter we focus instead on some features of the *macroscopic* polarization in semiconductors and insulators, at a phenomenological level, regardless of the microscopic mechanisms responsible for such polarization. The microscopic quantum theory of macroscopic polarization is the subject of Chapter 7.

The macroscopic polarization \mathbf{P} induced (to linear order) by a constant field in a dielectric can be expressed as a function of the macroscopic *bare* field \mathbf{E}_0 , or alternatively of the *screened* one \mathbf{E} . The latter choice yields[†]

$$\mathbf{P}(\mathbf{E}) = \frac{\mathbf{D} - \mathbf{E}}{4\pi} = \frac{\epsilon_\infty - 1}{4\pi} \mathbf{E}; \quad (2.22)$$

in noncubic materials, ϵ_∞ is understood to be a tensor. The occurrence of ϵ_∞ indicates that Eq. (2.22) is a clamped-ions polarization, where only the electronic system responds to the perturbation. In other words Eq. (2.22) applies when the only perturbing agent is the macroscopic field, and the bare electron-ion potential coincides with the unperturbed one:

$$V_{ext}(\mathbf{r}) = V_{ext}^{(0)}(\mathbf{r}) - e\mathbf{E}_0\mathbf{r}. \quad (2.23)$$

We are interested into the more general case, where the macroscopic polarization is possibly due to a source *other* than an “external” field \mathbf{E}_0 , such as *e.g.* a lattice distortion. Even in such cases, the polarization may (or may not) be *accompanied* by a macroscopic field, depending on the boundary conditions chosen for the macroscopic sample.

To proceed further, we assume that the perturbation is measured by a parameter λ : to fix the ideas, λ could be identified either with macroscopic strain (Section 2.4.2), or with a phonon coordinate (Section 2.4.3). In the study of this problem, it is convenient to use λ and \mathbf{E} as the independent variables. The

[†]We do not discuss in this Chapter crystals in which *spontaneous* polarization is symmetry allowed.

screened field \mathbf{E} —as opposite to the bare one \mathbf{E}_0 —is in fact the one which is easiest to control in bulk dielectrics, both experimentally and theoretically: in the former case, the control of \mathbf{E} amounts to placing the dielectric inside a capacitor at a given voltage, and in the latter case to solving Poisson equation with assigned boundary conditions.

The thermodynamic potential¹⁸ per unit (unperturbed) volume $\tilde{F}(\lambda, \mathbf{E})$ provides the generalized force f and the electric displacement \mathbf{D} as conjugate variables:

$$f(\lambda, \mathbf{E}) = -\frac{\partial}{\partial \lambda} \tilde{F}(\lambda, \mathbf{E}) \quad (2.24)$$

$$\mathbf{D}(\lambda, \mathbf{E}) = -4\pi \frac{\partial}{\partial \mathbf{E}} \tilde{F}(\lambda, \mathbf{E}); \quad (2.25)$$

the latter equation generalizes Eq. (2.22) as

$$\mathbf{P}(\lambda, \mathbf{E}) = -\frac{\partial}{\partial \mathbf{E}} \tilde{F}(\lambda, \mathbf{E}) - \frac{1}{4\pi} \mathbf{E}. \quad (2.26)$$

Assuming the solid to be in equilibrium at $\lambda=0$ and at null field, a second-order expansion of the thermodynamic potential in both λ and \mathbf{E} yields:

$$f(\lambda, \mathbf{E}) \simeq -\frac{\partial^2 \tilde{F}(0, 0)}{\partial \lambda^2} \lambda - \frac{\partial^2 \tilde{F}(0, 0)}{\partial \lambda \partial \mathbf{E}} \mathbf{E} \quad (2.27)$$

$$\mathbf{P}(\lambda, \mathbf{E}) \simeq -\frac{\partial^2 \tilde{F}(0, 0)}{\partial \lambda \partial \mathbf{E}} \lambda + \frac{\epsilon_\infty - 1}{4\pi} \mathbf{E}. \quad (2.28)$$

These are the basic phenomenological equations which describe the lowest-order response of the macroscopic dielectric to a perturbation λ , within a given boundary condition for the field. Besides ϵ_∞ , two other material constants enter the phenomenological equations. The second derivative with respect to λ has the meaning of an harmonic force constant per unit volume, at vanishing field. This is not a new quantity; using the notations of Section 1.6 we identify

$$\frac{E_{tot}^{(2)}}{\mathcal{V}} = \frac{1}{2} \frac{\partial^2 \tilde{F}(0, 0)}{\partial \lambda^2} \lambda^2, \quad (2.29)$$

where \mathcal{V} is the volume of the macroscopic sample, and the thermodynamic limit $\mathcal{V} \rightarrow \infty$ is understood. The mixed derivative entering both Eqs. (2.27) and (2.28) is a novel material constant, which couples the mechanical and electrical degrees of freedom. Its occurrence is basic for a proper description of the perturbed dielectric: the two most important cases are discussed in the following of this Chapter.

2.4.2 Piezoelectricity

We apply the above macroscopic formulation to the case where λ is identified with a macroscopic strain of suitable symmetry. The generalized force f coin-

cides then with macroscopic stress, and the harmonic constant with a suitable macroscopic elastic constant, measured at vanishing electric field.

The constant which couples the mechanical and electrical variables—whenever such coupling is allowed by the symmetry of the crystal and of the strain¹⁹—measures the piezoelectric effect. The piezoelectric constant

$$\gamma = -\frac{\partial^2 \tilde{F}(0, 0)}{\partial \lambda \partial \mathbf{E}} \quad (2.30)$$

can be defined, after Eq. (2.27), as the stress linearly induced by a unit field at zero strain; an alternative definition is provided by Eq. (2.28), where the piezoelectric constant γ appears as the polarization linearly induced by a unit strain at zero field.

The latter definition is related to the experimental situation sketched in Fig. 2.6(a). The crystal is uniaxially strained—along a piezoelectric direction—while in a shorted capacitor: the measured quantity is the current flowing through the shorting wire. Of course, this is not the unique possible realization of the piezoelectric effect. An obvious alternative setup would be to keep the sample free-standing in vacuum while straining it, as sketched in Fig. 2.6(b). In this case some surface charge piles up at the crystal boundary, and a macroscopic field is generated inside the sample: the *longitudinal* boundary condition $\mathbf{E} = -4\pi\mathbf{P}$ is appropriate for Eq. (2.28), which then yields $\mathbf{P} = \gamma\lambda/\epsilon_\infty$, at variance with the previous case.

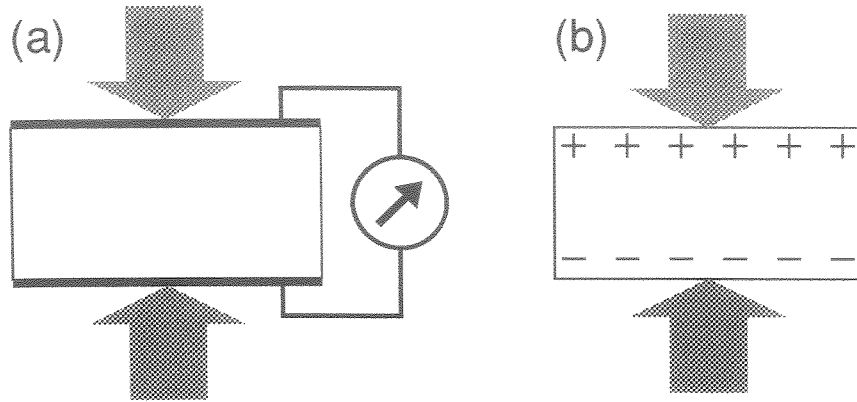


Figure 2.6. The piezoelectric effect: (a) for a sample within a shorted capacitor, in a null field; (b) for an isolated sample, in a depolarizing field $\mathbf{E} = -4\pi\mathbf{P}$.

2.4.3 Ionic screening

In order to simplify notations, we specialize to cubic binary crystals. A zone-center optic phonon is a macroscopic (*i.e.* lattice-periodical) mode, where the

two sublattices oscillate against each other. The phenomenological theory of these modes, based on Eqs. (2.27) and (2.28), is due to Huang:^{20,21} this theory is model-independent and therefore *exact* for harmonic modes, insofar as the effects of electromagnetic retardation can be neglected. Furthermore, the latter effects can be easily included leading thus to the polariton theory, which will not be discussed here.

We abandon in the following the scalar parameter λ , and we replace it with the relative coordinate \mathbf{u} as the mechanical variable. When the field is kept vanishing, the macroscopic mode coincides with the zone-center transverse-optic phonon, whose frequency ω_{TO} is one of the basic parameters of the theory. The harmonic force constant per unit volume is $m\omega_{TO}^2/\Omega$, where m is the reduced ionic mass, and Ω is the cell volume.

The mechanical-electrical coupling constant in Eqs. (2.27) and (2.28) is conveniently written

$$\frac{\partial \tilde{F}(0, 0)}{\partial \mathbf{u} \partial \mathbf{E}} = \frac{eZ^*}{\Omega}, \quad (2.31)$$

where Z^* is called the Born (or transverse) effective charge tensor: in cubic binary crystals it reduces to a scalar, while it vanishes in nonpolar (*i.e.* diamond-structure) materials. We thus recast Eqs. (2.27) and (2.28) in the form

$$\Omega f = -m\omega_{TO}^2 \mathbf{u} - eZ^* \mathbf{E} \quad (2.32)$$

$$\mathbf{P} = -\frac{eZ^*}{\Omega} \mathbf{u} + \frac{\epsilon_\infty - 1}{4\pi} \mathbf{E}, \quad (2.33)$$

where Z^* measures either the force linearly induced on the ions by a unit electric field at zero displacement, or alternatively the polarization linearly induced by a unit displacement at zero field.

In a given static field \mathbf{E} , the equilibrium value of the sublattice displacement \mathbf{u} is provided by Eq. (2.32)

$$\mathbf{u} = -\frac{eZ^*}{m\omega_{TO}^2} \mathbf{E} : \quad (2.34)$$

it is proportional to Z^* , and inversely proportional to the hardness of the restoring force. Upon replacing the equilibrium value of \mathbf{u} into Eq. (2.33), we get the total (ionic and electronic) polarization as

$$\mathbf{P}(\mathbf{E}) = \left[\frac{(eZ^*)^2}{m\Omega\omega_{TO}^2} + \frac{\epsilon_\infty - 1}{4\pi} \right] \mathbf{E}, \quad (2.35)$$

to be compared with the purely electronic polarization of Eq. (2.22). Thence the static dielectric constant is

$$\epsilon_0 = \epsilon_\infty + \frac{4\pi(eZ^*)^2}{m\Omega\omega_{TO}^2}. \quad (2.36)$$

So far about static screening; this same viewpoint can be extended to time-dependent macroscopic screening, provided the frequencies are in the quasistatic (or adiabatic) regime. In other words, the extension is correct in the range of infrared frequencies, but cannot be extrapolated to the optic region of the electromagnetic spectrum. In a driving applied field at frequency ω we have essentially a problem of forced oscillations; the equilibrium relationship Eq. (2.34) is replaced by the equation of motion:

$$-m\omega^2 \mathbf{u} = -m\omega_{TO}^2 \mathbf{u} - eZ^* \mathbf{E}. \quad (2.37)$$

Solving for \mathbf{u} one gets the polarization and eventually the ω -dependent dielectric constant

$$\epsilon(\omega) = \epsilon_\infty + \frac{4\pi(eZ^*)^2}{m\Omega(\omega_{TO}^2 - \omega^2)}; \quad (2.38)$$

for a nice experimental plot of $\epsilon(\omega)$ in the infrared regime see Fig 13b, p. 274 in Ref. 2. The ω -dependent dielectric constant has an obvious resonance at $\omega = \omega_{TO}$, which is the normal mode of the system: Eq. (2.37) confirms that the macroscopic field vanishes in this case.

It can be shown² that the zero of $\epsilon(\omega)$ is the frequency of the longitudinal macroscopic normal mode (alias zone-center longitudinal optic phonon): we get

$$\omega_{LO}^2 = \omega_{TO}^2 + \frac{4\pi(eZ^*)^2}{m\Omega\epsilon_\infty}. \quad (2.39)$$

This phonon is accompanied by a depolarizing field $\mathbf{E} = -4\pi\mathbf{P}$, which oscillates in phase with \mathbf{u} , and whose magnitude is easily found from the previous expressions: it is therefore a coupled mechanical-electrical mode, peculiar to ionic crystals. At frequencies $\omega_{TO} < \omega < \omega_{LO}$ the dielectric constant is negative and the propagation of electromagnetic waves in the polar crystal is forbidden.

The previous exact results can be recast in the Lyddane-Sachs-Teller form:

$$\epsilon_0/\epsilon_\infty = \omega_{LO}^2/\omega_{TO}^2, \quad (2.40)$$

which is valid for any cubic binary crystal; elegant extensions to complex crystals and to amorphous polar solids are available.^{22,23} All of the results of the macroscopic phenomenological theory reported here are recovered in the first-principle theory of lattice dynamics,²⁴ which is the subject of Chapter 9.

2.5 References

- [2.1] J.D. Jackson, *Classical Electrodynamics* (Wiley, New York, 1975).
- [2.2] C. Kittel, *Introduction to Solid State Physics*, 6th. edition (Wiley, New York, 1986).

- [2.3] A. Fetter and J.D. Walecka, *Quantum Theory of Many-Particle Systems* (McGraw-Hill, New York, 1971).
- [2.4] K.S. Singwi and M.P. Tosi, in *Solid State Physics*, edited by H. Ehrenreich, F. Seitz and D. Turnbull, vol **36** (Academic, New York, 1981), p.177.
- [2.5] D.R. Penn, Phys. Rev. **128**, 2093 (1962); G. Srinivasan, *ibid.* **178**, 1244 (1969); R.D. Grimes and E.R. Cowley, Can. J. Phys. **53**, 2549 (1975).
- [2.6] R. Resta, Phys. Rev. B **16**, 2717 (1977).
- [2.7] R. Resta, Phys. Rev. B **38**, 818 (1988).
- [2.8] A. Baldereschi and R. Resta, in: *Ab-Initio Calculation of Phonon Spectra*, edited by J.T. Devreese et al. (Plenum Publishing, New York, 1983), p. 1.
- [2.9] R. Resta, in: *Festkörperprobleme*, vol. **XXV**, edited by P. Grosse (Vieweg, Braunschweig, 1985), p. 183.
- [2.10] O.F. Mossotti, Memorie di Matematica e di Fisica della Società Italiana delle Scienze Residente in Modena, **24**, 49 (1850); R. Clausius, *Die Mechanische Behandlung der Electrica* (Vieweg, Berlin, 1879).
- [2.11] A.A. Maradudin, E.W. Montroll, G.H. Weiss, and I.P. Ipatova, *Theory of Lattice Dynamics in the Harmonic Approximation*, Solid State Physics, Suppl. **3** (Academic, New York, 1971).
- [2.12] N.F. Mott and M.J. Littleton, Trans. Faraday Soc. **34**, 45 (1938).
- [2.13] I.V. Abarenkov and I.M. Antonova, Phys. Status Solidi B **64**, 747 (1974); *ibid.* **65**, 325 (1974).
- [2.14] K. Kunc and R. Resta, Phys. Rev. Lett. **51**, 686 (1983).
- [2.15] R. Resta and K. Kunc, Phys. Rev. B **34**, 7146 (1986).
- [2.16] M. Wolkenstein, Compt. Rend. Acad. Sci. URSS **32**, 185 (1941).
- [2.17] M. Cardona, in: *Light Scattering in Solids II*, edited by M. Cardona and G. Güntherodt (Springer, Berlin, 1982), p. 19.
- [2.18] L.D. Landau and E.M. Lifshitz, *Electrodynamics of Continuous Media* (Pergamon Press, Oxford, 1984).
- [2.19] M. Lax, *Symmetry Principles in Solid State and Molecular Physics* (Wiley, New York, 1974).
- [2.20] K. Huang, Proc. Roy. Soc. **A203**, 178 (1950).

- [2.21] M. Born and K. Huang, *Dynamical Theory of Crystal Lattices* (Oxford University Press, Oxford, 1954).
- [2.22] A.S. Barker, Jr., Phys. Rev. B **12**, 4071 (1975).
- [2.23] T.W. Noh and J. Sievers, Phys. Rev. Lett. **63**, 1800 (1989).
- [2.24] R. Pick, M.H. Cohen, and R.M. Martin, Phys. Rev. B **1**, 910 (1970).

Chapter 3

Formal linear-response theory

3.1 Density-functional formulation

When the perturbation $\delta V_{ext}(\mathbf{r})$ is applied to the system, the electrons polarize and modify the KS potential, Eq. (1.13), through the density dependence of V_H and V_{xc} as well:

$$\delta V_{KS} = \delta V_{ext} + \delta V_H + \delta V_{xc}. \quad (3.1)$$

While δV_{ext} is the given external perturbation, the remaining terms are not known beforehand and have to be evaluated in a selfconsistent way; once this problem solved, one has access to the electron density $n^{(1)}(\mathbf{r})$ linearly-induced by the perturbed KS potential. The formal treatment of this problem—to first-order in δV_{ext} —is the subject of this Chapter. Throughout it, I will remain within a formally *exact* many-body formulation, in the DFT framework, and which applies to a general system (atom, molecule, cluster, condensed matter...). The LDA is needed only at the level of practical implementations.

There are basically two distinct ways of dealing with linear response: both of them will be reviewed here. The former way is inspired by the classic selfconsistent approach of Ehrenreich and Cohen,¹ also known as RPA (random-phase approximation), and dating back to 1959. This was originally formulated at the Hartree level of approximation (*i.e.* neglecting δV_{xc}); it has been generalized later to Hartree-Fock² or to LDA,^{3,4} essentially, the selfconsistent problem is solved in closed form via the inversion of suitable operators. I will give here a formally exact DFT formulation, based essentially upon the same algebra as in the approximated treatments.

The latter way for coping with selfconsistent linear response is the so-called “direct” method: it can be spelled out as an iterative scheme for finding the perturbed electronic ground state, selfconsistently to first-order in the bare perturbation. In atomic physics this method is very old;^{5,3} the Hartree-Fock modification of the same concept is also known in molecular physics, and goes under the name of “coupled Hartree-Fock”.⁶ The idea was brought to solid state

physics in 1987 by Baroni, Giannozzi, and Testa,⁷ and has provided since then the most effective and accurate way of dealing with linear-response properties and lattice dynamics in semiconductors.^{8,9} Some applications are discussed below, in Chapters 8 and 9. A somewhat different implementation of the same main idea is due to Gonze *et al.*¹⁰ In the following, the direct approach of Ref. 7 will be referred to as the density-functional perturbation theory (DFPT).

3.2 Independent-particle polarizability

Within the RPA-like approach,¹ one starts introducing an auxiliary (and non-physical) linear response operator χ_0 , defined as the functional derivative of the electron density with respect to the *total* KS potential, i.e.

$$n^{(1)}(\mathbf{r}) = \int d\mathbf{r}' \chi_0(\mathbf{r}, \mathbf{r}') \delta V_{KS}(\mathbf{r}'). \quad (3.2)$$

This auxiliary operator will be called the independent-particle polarizability, and will be eventually related to the physical response operators.

The first-order variation of the i -th KS orbital is obtained from straightforward perturbation theory

$$\varphi_i^{(1)} = \sum_{j \neq i} \varphi_j^{(0)} \frac{\langle \varphi_j^{(0)} | \delta V_{KS} | \varphi_i^{(0)} \rangle}{\epsilon_i - \epsilon_j}, \quad (3.3)$$

and the first-order electronic density is, from Eq. (1.8)

$$n^{(1)}(\mathbf{r}) = \sum'_{i,j} f_i \left[\varphi_i^{(0)*}(\mathbf{r}) \varphi_j^{(0)}(\mathbf{r}) \frac{\langle \varphi_j^{(0)} | \delta V_{KS} | \varphi_i^{(0)} \rangle}{\epsilon_i - \epsilon_j} + c.c. \right], \quad (3.4)$$

where *c.c.* stands for the complex conjugate, and the primed sum is over all i, j with $i \neq j$. A trivial transformation leads to the more symmetric form:

$$n^{(1)}(\mathbf{r}) = \sum'_{i,j} \frac{f_i - f_j}{\epsilon_i - \epsilon_j} \langle \varphi_j^{(0)} | \delta V_{KS} | \varphi_i^{(0)} \rangle \varphi_i^{(0)*}(\mathbf{r}) \varphi_j^{(0)}(\mathbf{r}). \quad (3.5)$$

The occupancy factor, for a closed shell system at zero temperature, assumes the values of either 2 (occupied KS orbitals) or 0 (empty KS orbitals): Eq. (3.5) is then recast as

$$n^{(1)}(\mathbf{r}) = -4 \sum_{v,c} \frac{\langle \varphi_c^{(0)} | \delta V_{KS} | \varphi_v^{(0)} \rangle}{\epsilon_c - \epsilon_v} \varphi_v^{(0)*}(\mathbf{r}) \varphi_c^{(0)}(\mathbf{r}), \quad (3.6)$$

where it is understood that the index v (like valence) runs only over the occupied KS orbitals and c (like conduction) over the unoccupied ones.

We have purposely not explicitated until here the matrix elements of δV_{KS} , because $n^{(1)}(\mathbf{r})$ in the form of Eq. (3.6) will be useful below, Section 3.4. If we now exploit the *locality* of the KS potential, then Eq. (3.6) is identical to Eq. (3.2), once made the identification

$$\chi_0(\mathbf{r}, \mathbf{r}') = -4 \sum_{v,c} \frac{\varphi_c^{(0)*}(\mathbf{r}') \varphi_v^{(0)}(\mathbf{r}') \varphi_v^{(0)*}(\mathbf{r}) \varphi_c^{(0)}(\mathbf{r})}{\epsilon_c - \epsilon_v}. \quad (3.7)$$

It can be explicitly checked that χ_0 is a real symmetric operator, which is easily built from knowledge of the *unperturbed* KS Hamiltonian only; it has been proved that χ_0 is negative definite.¹¹

3.3 Physical response operators

The next step is to get the physical response operators χ and ϵ^{-1} in terms of the independent-particle polarizability χ_0 .

The selfconsistent variation of the Hartree potential is, upon linearization,

$$\delta V_H(\mathbf{r}) \simeq V_H^{(1)}(\mathbf{r}) = e^2 \int d\mathbf{r}' n^{(1)}(\mathbf{r}') / |\mathbf{r} - \mathbf{r}'|, \quad (3.8)$$

or in shorthand operator notation $\delta V_H \simeq V_H^{(1)} = v_c n^{(1)}$, where

$$v_c(\mathbf{r}, \mathbf{r}') = e^2 / |\mathbf{r} - \mathbf{r}'|. \quad (3.9)$$

Linearization of the exchange-correlation term gives analogously $\delta V_{xc} \simeq V_{xc}^{(1)} = f_{xc} n^{(1)}$, where f_{xc} is the functional derivative

$$f_{xc}(\mathbf{r}, \mathbf{r}') = \frac{\delta V_{xc}(\mathbf{r})}{\delta n(\mathbf{r}')}, \quad (3.10)$$

evaluated at the unperturbed ground state. From Eq. (3.2) we get

$$n^{(1)} = \chi_0 V_{KS}^{(1)} = \chi_0 (\delta V_{ext} + v_c n^{(1)} + f_{xc} n^{(1)}). \quad (3.11)$$

The density response χ is straightforwardly obtained solving Eq. (3.11) with respect to $n^{(1)}$:

$$n^{(1)} = (1 - \chi_0 v_c - \chi_0 f_{xc})^{-1} \chi_0 \delta V_{ext}; \quad (3.12)$$

$$\chi = (1 - \chi_0 v_c - \chi_0 f_{xc})^{-1} \chi_0 = (\chi_0^{-1} - v_c - f_{xc})^{-1}, \quad (3.13)$$

which incidentally displays that χ is a real symmetric operator; general stability considerations require χ to be negative definite.¹¹

Identification of δV_{ext} with $e\phi_0$ shows that the basic screening operator ϵ^{-1} of Eq. (2.10) is:

$$\epsilon^{-1} = (1 - f_{xc} \chi_0) (1 - v_c \chi_0 - f_{xc} \chi_0)^{-1}. \quad (3.14)$$

It is convenient to express ϵ^{-1} through an equivalent expression for the operator ϵ , its inverse:

$$\epsilon = 1 - v_c \chi_0 (1 - f_{xc} \chi_0)^{-1}. \quad (3.15)$$

The historical RPA expressions¹ are recovered in the Hartree approximation, *i.e.* taking $f_{xc}=0$; for the density response this gives

$$\chi_{RPA} = (\chi_0^{-1} - v_c)^{-1}, \quad (3.16)$$

showing that even χ_{RPA} is a negative definite operator. The *exact* density response can be formally expressed as a “correction” to the RPA one, from Eq. (3.13), as:

$$\chi = (\chi_{RPA}^{-1} - f_{xc})^{-1} = \chi_{RPA} (1 - f_{xc} \chi_{RPA})^{-1}, \quad (3.17)$$

but it must be kept in mind that in this case the RPA has to be evaluated with the KS one-particle states.

The selfconsistent screening problem has thus been eventually solved: the key point, where selfconsistency has been achieved, is the *inversion* of the relevant operators. All of the previous results, like Eqs. (3.13) and (3.15), are formally exact; but the difficult many-body part is of course hidden in the operator f_{xc} .

The static dielectric response is a ground-state property: as such, it is certainly within the scope of DFT. Nonetheless it could seem disturbing that the exact expressions involve, through the operator χ_0 , Eq. (3.7), one-particle eigenvalues and eigenfunctions (both occupied and empty) of the unperturbed system: the denominators particularly take the form of “transition” energies. There is no paradox: on one side the one-particle KS states are *not* to be related to electronic excited states of the real physical system;¹² on the other side these same states are quantities generated by the ground-state KS Hamiltonian, and they enter the formulation because of simple *mathematical* reasons (*i.e.* use of perturbation theory in the study of linear response).

If the previous theory is implemented within LDA, the operator f_{xc} is local, and is a simple function of the local density:

$$f_{xc}^{(LDA)}(\mathbf{r}, \mathbf{r}') = \delta(\mathbf{r} - \mathbf{r}') \left. \frac{d\mu_{xc}(n)}{dn} \right|_{n=n^{(0)}(\mathbf{r})}. \quad (3.18)$$

The response operators take a form which can be worked out for real systems; the results thus obtained yield therefore the linear term in the response, and are *exact* at this level, *i.e.* no further approximation besides LDA has been introduced at any stage. Owing to Eq. (1.17) $f_{xc}^{(LDA)}$ is negative: this implies (excluding pathological circumstances) that that $\chi_{LDA} - \chi_{RPA}$ is negative definite, after Eq. (3.17). In other words the f_{xc} correction to the RPA result *enhances* screening.¹³

3.4 Density-functional perturbation theory

The linear operator χ (or equivalently ϵ^{-1}) is all what is needed to describe, to linear order, the response to an *arbitrary* bare perturbation $\delta V_{ext}(\mathbf{r})$. It is therefore a very useful quantity, containing a lot of microscopic information about the system. But exactly because of this reason, a *complete* calculation of χ within the above formalism can be a difficult task: it often amounts to compute much more than what is actually needed to solve a specific physical problem.

Some typical linear-response problems of condensed matter physics have been alternatively attacked in a brute-force way. The perturbed system, with a given “frozen-in” bare perturbation, is considered on the same ground as the unperturbed one, and two independent selfconsistent calculations are performed: then the physical quantities of interest are obtained by difference. Harmonic phonons are the linear property most studied in this way, either via so-called “frozen phonon” straightforward calculations,^{14,15} or via more elaborated approaches;¹⁶ other dielectric properties have been studied as well.^{17,18}

Baroni, Giannozzi and Testa⁷ have proposed and implemented DFPT for periodic solids, which combines some advantages of both (frozen-perturbation and linear-response) approaches, while eliminating most of their drawbacks. The idea is to start with the unperturbed reference system in its ground state; then a *specific* bare perturbation is introduced, and the new selfconsistent ground state is evaluated keeping only the terms which are linear in the perturbation. The logics is the same as in some well known methods in atomic^{5,3} and molecular⁶ physics.

It is convenient to start again with Eq. (3.6), and then write also the independent equation relating, to linear order, $n^{(1)}$ and δV_{KS} :

$$V_{KS}^{(1)} = \delta V_{ext} + v_c n^{(1)} + f_{xc} n^{(1)}. \quad (3.19)$$

The two linear equations together give a selfconsistent scheme, which can be solved iteratively for $n^{(1)}$ and $V_{KS}^{(1)}$, starting at iteration zero with the unscreened situation: $V_{KS}^{(1)} = \delta V_{ext}$ and $n^{(1)} = 0$. Iteration of the linearized equations replaces thus the operator inversion which was an essential step in the RPA-like scheme described in Section 3.2. To state the concept in different words: only the relevant portion of the inverse operator is built, and this is performed iteratively step by step.

A further essential improvement in DFPT is avoiding complete diagonalization of the unperturbed Hamiltonian, using the identity:

$$\sum_c \frac{|\varphi_c^{(0)}\rangle\langle\varphi_c^{(0)}|}{\epsilon_c - \epsilon} = -P_c G^{(0)}(\epsilon) P_c, \quad (3.20)$$

where $P_c = 1 - P_v$ is the projector over the unoccupied KS manifold and $G^{(0)}(\epsilon) = (\epsilon - H_{KS}^{(0)})^{-1}$ is the unperturbed single-particle Green’s function. This

allows to recast Eq. (3.6) in a form where the (slowly convergent) sum over the empty c states formally disappears:

$$n^{(1)}(\mathbf{r}) = 4 \sum_v \varphi_v^{(0)*}(\mathbf{r}) \langle \mathbf{r} | P_c G^{(0)}(\epsilon_v) P_c V_{KS}^{(1)} | \varphi_v^{(0)} \rangle. \quad (3.21)$$

In order to proceed further we rewrite Eq. (3.21) in the form

$$n^{(1)}(\mathbf{r}) = -4 \sum_v \varphi_v^{(0)*}(\mathbf{r}) \langle \mathbf{r} | P_c | \tilde{\varphi}_v^{(1)} \rangle, \quad (3.22)$$

where $\tilde{\varphi}_v^{(1)}$ fulfills the equation

$$(H_{KS}^{(0)} - \epsilon_v) | \tilde{\varphi}_v^{(1)} \rangle = P_c V_{KS}^{(1)} | \varphi_v^{(0)} \rangle. \quad (3.23)$$

The solution $\tilde{\varphi}_v^{(1)}$ of Eq. (3.23), with the condition of being orthogonal to the v manifold, is unique and is efficiently computed solving the linear system; the DFPT iteration is performed in fact using the equations (3.23), (3.22) and (3.19).

This method has invaluable advantages when dealing with extended condensed matter systems using periodic boundary conditions; this point will be discussed in Chapter 8. Another essential advantage is manifest if a modern pseudopotential scheme is adopted in calculations. In fact, being the bare ionic pseudopotentials nonlocal, perturbations involving ionic displacements (*e.g.* phonons, strain) are nonlocal as well, and the linear-response operators of Section 3.2 are useless. DFPT can instead be (and actually is) implemented in a modern pseudopotential context without any major problem: for this reason DFPT is the state of the art for lattice dynamics in semiconductors.⁹

3.5 Role of symmetry

The symmetry of the system, in its unperturbed ground state, plays an overwhelming role and allows fundamental decoupling of the linear response. Suppose that the system is totally symmetric under the operations of a group of transformations: then if the bare perturbation belongs to a given irreducible representation, the linear response to it belongs to the same representation. I will illustrate this concept over three simple and very important examples.

The first one is the homogeneous system, already discussed in Section 2.2: the ground state is invariant under the continuous group of translations. The irreducible representations of this group are one-dimensional and are labeled by the quantum number \mathbf{k} (momentum); a perturbation of given \mathbf{k} only induces—to linear order—a response at the same \mathbf{k} . The decoupling, or “diagonality” of the response, follows therefore from the general symmetry argument.

The second example is a spherical (*i.e.* closed-shell) atom. It is invariant under the continuous group of rotations, whose irreducible representations are

labeled by the quantum number l (angular momentum). A dipolar ($l=1$) perturbation, like *e.g.* a constant field, only induces—to linear order—a dipole moment; an irreducible quadrupolar ($l=2$) perturbation does the same, and so forth. Linearity allows therefore decoupling of the problem, multipole by multipole.

The third example is a periodic solid: it is invariant under the group of discrete translations, whose irreducible representations are labeled by the quantum number \mathbf{q} (quasimomentum). Components of different \mathbf{q} do not couple to linear order: this allows to introduce the concept of response matrices in reciprocal space, like *e.g.* the dielectric matrix. A thorough analysis is given in Chapter 8.

3.6 References

- [3.1] H. Ehrenreich and M.H. Cohen, *Phys. Rev.* **115**, 786 (1959).
- [3.2] A.D. McLachian and M.A. Ball, *Rev. Mod. Phys.* **36**, 844 (1964).
- [3.3] G.D. Mahan, *Phys. Rev. A* **22**, 1780 (1980).
- [3.4] R.M. Martin and K. Kunc, in: *Ab-Initio Calculation of Phonon Spectra*, edited by J.T. Devreese et al. (Plenum Publishing, New York, 1983), p. 49.
- [3.5] R.M. Sternheimer, *Phys. Rev.* **96**, 951 (1954); *ibid.* **107**, 1565 (1957); *ibid.* **115**, 1198 (1959); *ibid.* **183**, 112 (1969); *Phys. Rev. A* **1**, 321 (1970).
- [3.6] R.D. Amos, in: *Ab-Initio Methods in Quantum Chemistry - I*, edited by K.P. Lawley (Wiley, New York, 1987), p. 99.
- [3.7] S. Baroni, P. Giannozzi, and A. Testa, *Phys. Rev. Lett.* **58**, 1861 (1987).
- [3.8] S. de Gironcoli, S. Baroni, and R. Resta, *Phys. Rev. Lett.* **62**, 2853 (1989).
- [3.9] P. Giannozzi, S. de Gironcoli, P. Pavone, and S. Baroni, *Phys. Rev. B* **43**, 7231 (1991).
- [3.10] X. Gonze, D.C. Allan, and M.P. Teter, *Phys. Rev. Lett.* **68**, 3603 (1992).
- [3.11] R. Car, E. Tosatti, S. Baroni and S. Leelaprute, *Phys. Rev. B* **24**, 985 (1981).
- [3.12] J.P. Perdew and M. Levy, *Phys. Rev. Lett.* **51**, 1884 (1983); L.J. Sham and M. Schlüter, *ibid.* **51**, 1888 (1983).

- [3.13] S. Baroni and R. Resta, *Phys. Rev. B* **33**, 7017 (1986).
- [3.14] See for instance: R.M. Martin, in *Electronic Structure, Dynamics and Quantum Structural Properties of Condensed Matter*, edited by J.T. Devreese and P. Van Camp (Plenum, New York, 1985), p. 175.
- [3.15] M.L. Cohen, in: *Highlights in Condensed Matter Theory*, edited by F. Bassani, F. Fumi and M.P. Tosi (North Holland, Amsterdam, 1985).
- [3.16] K. Kunc, in *Electronic Structure, Dynamics and Quantum Structural Properties of Condensed Matter*, edited by J.T. Devreese and P. Van Camp (Plenum, New York, 1985), p. 227.
- [3.17] K. Kunc and R. Resta, *Phys. Rev. Lett.* **51**, 686 (1983).
- [3.18] R. Resta and K. Kunc, *Phys. Rev. B* **34**, 7146 (1986).

Chapter 4

Electron-gas results

4.1 Ground state

The electron gas (alias jellium) is the paradigmatic many-electron system in condensed matter physics; there have been six decades of scientific literature about it.^{1,2} A single parameter characterizes the electron-gas properties: the density n_0 , or equivalently the Fermi momentum $k_F = (3\pi^2 n_0)^{1/3}$, or the dimensionless parameter r_s , which defines the interelectronic distance through $1/n_0 = 4\pi r_s^3 a_0^3/3$. At densities of interest in condensed matter physics (*i.e.* r_s between 2 and 7) the electron-gas ground state is homogeneous and paramagnetic (alias closed-shell), and the KS potential is constant. The single-particle KS orbitals are therefore PW's of momentum \mathbf{k} , whose eigenvalues are

$$\epsilon(\mathbf{k}) = \frac{1}{2}k^2 + V_{KS}; \quad (4.1)$$

here the two terms are each referred to the zero of the potential, which is completely arbitrary in extended Coulomb systems; it is convenient to choose this reference to be the mean electrostatic potential. Using such a convention, $V_{KS} = \mu_{xc}(n_0)$, while the total chemical potential is:

$$\mu = \epsilon(k_F) = \frac{1}{2}k_F^2 + \mu_{xc}. \quad (4.2)$$

Using the results of Chapter 1, the total energy of the system is written as:

$$E_{tot} = E_{ext} + \int d\mathbf{r} V_{ext}(\mathbf{r})n(\mathbf{r}) + E_H + T_s + E_{xc}. \quad (4.3)$$

The first three terms are classical Coulomb energies, accounting for background-background, background-electron and electron-electron interaction energies, respectively: each of these terms individually diverges, but it is easy to show that their sum cancels exactly for an homogenous electron distribution.³ The two

remaining terms scale like the size of the system; the total energy *per electron* can then be written as:

$$\frac{1}{N} E_{tot} = \frac{1}{N} (T_s + E_{xc}) = \frac{3}{10} k_F^2 + \epsilon_{xc}. \quad (4.4)$$

This is shown in Fig. 4.1 as a function of r_s (solid line), where the quantum Monte Carlo electron-gas results of Ceperley and Alder have been used, after Ref. 4. The two individual contributions are also separately shown: kinetic energy of the noninteracting system (dashed), and exchange correlation energy (dotted). The system is stable at $r_s=4.18$, not very far from the average valence electron density of metallic Na ($r_s=3.93$).

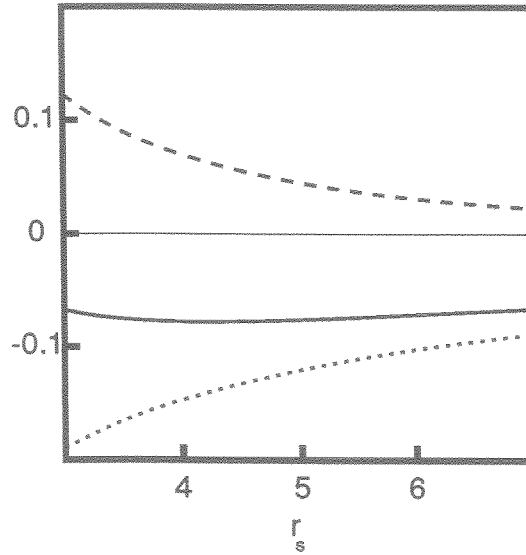


Figure 4.1. Total energy (a.u.) per electron of the paramagnetic electron gas (solid); kinetic energy of the fictitious noninteracting system T_s (dashed); exchange-correlation term ϵ_{xc} (dotted).

4.2 Dielectric function

Some main features of the electron-gas response have been anticipated previously, in Chapter 2; here we reexamine this system in the light of the general theory of Chapter 3. For the densities of interest, the system is homogeneous and metallic; since translational symmetry uniquely determines the eigenstates,

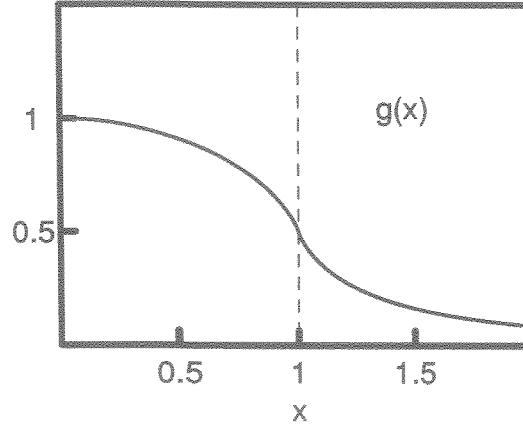


Figure 4.2. The Lindhard function $g(x)$. Notice the singular derivative at $x=1$.

the independent-particle polarizability coincides for this system with the free-particle polarizability. The general expression for χ_0 , Eq. (3.7) reads then:

$$\chi_0(\mathbf{r}, \mathbf{r}') = -\frac{4}{(2\pi)^6} \int d\mathbf{k}_v d\mathbf{k}_c \frac{e^{i(\mathbf{k}_c - \mathbf{k}_v)(\mathbf{r} - \mathbf{r}')}}{\epsilon(\mathbf{k}_c) - \epsilon(\mathbf{k}_v)}, \quad (4.5)$$

where \mathbf{k}_v is inside the Fermi sphere ($k_v < k_F$) and \mathbf{k}_c outside. Eq. (4.5) shows the dependence upon $\mathbf{r} - \mathbf{r}'$, typical of any ground-state operator in the homogeneous system; switching to reciprocal space, where these operators are diagonal, the integrations of Eq. (4.5) can be performed exactly:

$$\chi_0(k) = -\frac{k_F}{\pi^2} g\left(\frac{k}{2k_F}\right); \quad (4.6)$$

$$g(x) = \frac{1}{2} + \frac{1-x^2}{4x} \ln \left| \frac{1+x}{1-x} \right|. \quad (4.7)$$

The function $g(x)$ is shown in Fig. 4.2; this result is usually attributed to Lindhard,⁵ although it is at least implicit in a 1937 paper of Bardeen.⁶ At the RPA level (*i.e.* neglecting the occurrence of f_{xc}) the dielectric function is

$$\epsilon_{RPA}(k) = 1 + \frac{k_{TF}^2}{k^2} g\left(\frac{k}{2k_F}\right), \quad (4.8)$$

where the g factor appears as a correction to the TF result, Eq. (2.17); the TF inverse screening length is given by:

$$k_{TF}^2 = \frac{4k_F}{\pi a_0} \simeq \frac{2.44}{r_s a_0^2}. \quad (4.9)$$

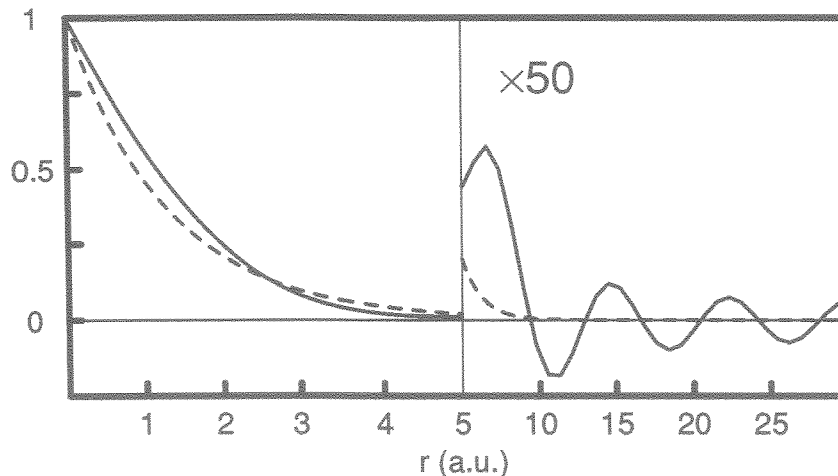


Figure 4.3. Ratio $\phi(r)/\phi_0(r)$ between the screened and unscreened potentials of a point charge in an electron gas at $r_s=4$: solid, RPA; dashed, TF. The Friedel oscillations are magnified (by a factor 50) in the right panel; their wavelength is $\pi/k_F=6.55$ a.u.

Since $g(k) \simeq 1$ at small k , one could naively expect that at large r the screened potential $\phi(r)$ of a point charge behaves like the TF one, Eq. (2.18). This is false, because the small- $k \leftrightarrow$ large- r relationship holds under the hypothesis of analyticity, while the RPA dielectric function has a logarithmic singularity at $k = 2k_F$. As a fingerprint of this singularity, the asymptotic behavior is $\phi(r) \sim \cos(2k_F r)/r^3$; an elegant proof can be found *e.g.* in Ref. 3, p.178. The cosine behavior goes under the name of “Friedel oscillations”. The TF and RPA screened potentials of a point charge are compared in Fig. 4.3.

The logarithmic singularity is a consequence of the sharpness of the Fermi surface used in calculating χ_0 : one could suspect an artifact of RPA, guessing this singularity to be cancelled either by f_{xc} or by finite temperature effects. As for the latter, they can be ruled out: at metallic densities and at room temperature the electronic Fermi-Dirac distribution *is* really very sharp.⁷ As a matter of fact the singularity at $k = 2k_F$ is no artifact at all, and manifests itself very strikingly in the phonon spectra of real metals: this was experimentally detected only *after* the theory of the effect, due to W. Kohn.⁸ This is discussed below, Section 9.3.

The exchange-correlation correction f_{xc} in the homogeneous electron gas is diagonal as well, and the exact dielectric function is formally written, after Eq. (3.15), as

$$\epsilon(k) = 1 - \frac{4\pi e^2}{k^2} \frac{\chi_0(k)}{1 - f_{xc}(k)\chi_0(k)}. \quad (4.10)$$

As I have already stated, the electron-gas literature is abundant and several different calculations of $f_{xc}(k)$ are available.^{1,2} Usually these are written as

$f_{xc}(k) = -4\pi e^2 G(k)/k^2$, where the dimensionless function $G(k)$ in the electron-gas jargon is the “local-field correction”. This is somewhat confusing, because in the context of screening in nonmetals (and therefore in this book, see Section 2.3) the term “local field” is used for a rather different concept. Within LDA f_{xc} is a negative constant. Several different calculations for $G(k)$ are available on the market,^{1,2} and the corresponding f_{xc} for some of them are shown in Fig. 4.4 at $r_s=4$: the different calculations are only in rough qualitative agreement, particularly in the interesting region around k_F ($k_F=0.48$ a.u.). Within LDA f_{xc} is a negative constant, shown in Fig. 4.4 as a thin line. The $G(k)$ of Ichimaru and Utsumi⁹ (solid line) coincides at small k with the LDA value, since it has been fitted to the same electron-gas data⁴ as our LDA ones; but the figure shows that the different theories disagree even in the sign of the second derivative of f_{xc} at $k=0$. As for the large- k behavior, it has been demonstrated¹⁰ that f_{xc} tends to a negative constant (and *not* to zero) at large k : here again, the existing calculations often violate this requirement. Unfortunately, a quantum Monte Carlo calculation of the linear response of the electron gas is not available yet, at least in three dimensions (while it has been recently performed in the two-dimensional case¹¹).

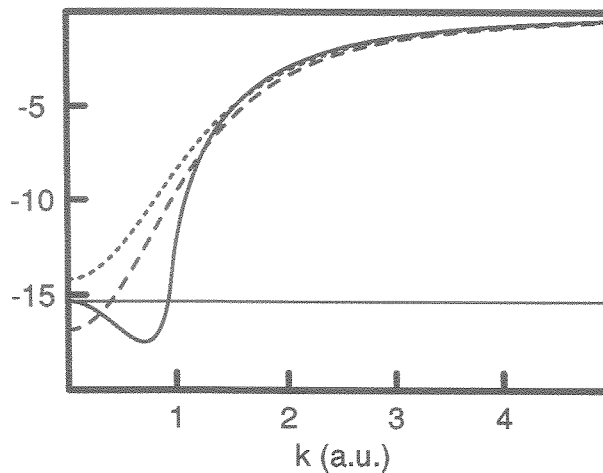


Figure 4.4. The exchange-correlation correction $f_{xc}(k)$ in a.u. for the homogeneous electron gas at $r_s=4$. The thin line indicates the LDA constant value; solid line after Ref. 9; dotted line after Ref. 12; dashed line after Ref. 13.

4.3 References

- [4.1] K.S. Singwi and M.P. Tosi, in *Solid State Physics*, edited by H. Ehrenreich, F. Seitz and D. Turnbull, vol **36** (Academic, New York, 1981), p.177.

- [4.2] S. Ichimaru, *Rev. Mod. Phys.* **54**, 1017 (1982).
- [4.3] A. Fetter and J.D. Walecka, *Quantum Theory of Many-Particle Systems* (McGraw-Hill, New York, 1971).
- [4.4] D.M. Ceperley and B.J. Alder, *Phys. Rev. Lett.* **45**, 566 (1980); J. Perdew and A. Zunger, *Phys. Rev. B* **23**, 5048 (1981).
- [4.5] J. Lindhard, *Kgl. Danske Mat. Fys. Medd.* **28**, 8, 1954.
- [4.6] J. Bardeen, *Phys. Rev.* **52**, 688 (1937).
- [4.7] C. Kittel, *Introduction to Solid State Physics*, 6th. edition (Wiley, New York, 1986).
- [4.8] W. Kohn, *Phys. Rev. Lett.* **2**, 393 (1959).
- [4.9] S. Ichimaru and K. Utsumi, *Phys. Rev. B* **24**, 7385 (1981).
- [4.10] A. Holas, in *Strongly Coupled Plasma Physics*, edited by F.J. Rogers and H. Dewitt, NATO ASI Series B, vol. 154 (Plenum Publishing, New York, 1989), p. 463.
- [4.11] S. Moroni, D. Ceperley, and G. Senatore, *Phys. Rev. Lett.* **69**, 1837 (1992).
- [4.12] K.S. Singwi, A. Sjölander, M.P. Tosi, and R.H. Land, *Phys. Rev. B* **1**, 1044 (1970).
- [4.13] P. Vashista and K.S. Singwi, *Phys. Rev. B* **6**, 875 (1972).

Chapter 5

Pseudopotential perturbation theory

5.1 Basic concepts

The pseudopotential perturbation theory^{1,2} provides a universal description of simple metals—*i.e.* those which are bonded by *sp* electrons—ranging from static crystalline properties (cohesion, structural stability, stacking fault energies), through harmonic properties (elasticity, lattice dynamics), and up to extremely anharmonic—albeit adiabatic—dynamical properties (finite-temperature properties, melting, phase diagrams).^{3–5} Although being based upon (very little) empirical parametrization, the theory can be considered as the archetypical “total energy” method in condensed matter physics.

The theory gives a realistic description of the electronic ground state of a simple metal (both solid or liquid), through a perturbation expansion. The unperturbed reference system is chosen as the uniform electron gas, at a density equal to the average valence density of the real system under study, which is assumed to be *macroscopically* homogeneous. The unperturbed external potential $V_{ext}^{(0)}(\mathbf{r})$ is therefore the potential of the neutralizing uniform background; the perturbation potential $\delta V_{ext}(\mathbf{r})$ switches the background off, while switching the electron-ion interaction on. A low-order perturbation expansion will certainly hold if the perturbing potential is weak, but this is generally *not* the case for bare electron-ion interactions. Nonetheless, the expansion holds as well provided only that the scattering of valence electrons by ions is weak, thus allowing the use of weak pseudopotentials which provide the same scattering power.²

We consider here only second order perturbation theory, with local pseudopotentials. This level of treatment is still (1993) the state of the art for finite-temperature simulations of simple metals. As for crystalline solids at $T=0$, a theory of this kind provided a number of outstanding results all along the sixties,^{1,2,6} but is nowadays of historical and tutorial interest only. It is reviewed here to the purpose of introducing some important concepts used in subsequent Chapters.

For a given ionic configuration, the total energy of the system is identically written as:

$$E_{tot} = E_{ext} + E_{el} = E_{tot}^{(0)} + \Delta E_{ext} + \Delta E_{el}, \quad (5.1)$$

where the total energy of the electron gas—discussed in Section 4.1—is used for $E_{tot}^{(0)}$ at the same density n_0 as the valence electron density of the system under study. Our next task is then to evaluate the two (structure-dependent) terms ΔE_{ext} and ΔE_{el} . An essential simplifying feature is the fact that the perturbation is neutral in average, such as to avoid Coulomb divergences in both terms. I anticipate the basic feature that ΔE_{el} , when expanded to second order, is equivalent to an indirect ion-electron-ion interaction, mediated by the polarizability of the electron gas.

5.2 Formalism

I give here an outline of the basic treatment; several variants to it have been proposed in the literature.² For the sake of simplicity, I only consider the case of an elemental condensed system, whose ion cores have valence Z : the average electronic density is therefore $n_0 = Z/\Omega$, where the atomic volume Ω coincides in the crystalline case with the cell volume for primitive lattices; a finite system is made of N_c cores (at the sites \mathbf{R}_l) and $N = ZN_c$ electrons.

Starting with the classical term:

$$\Delta E_{ext} = \frac{1}{2} \left(\sum_{l \neq m} \frac{Z^2 e^2}{|\mathbf{R}_l - \mathbf{R}_m|} - \int d\mathbf{r} d\mathbf{r}' \frac{n_0^2}{|\mathbf{r} - \mathbf{r}'|} \right), \quad (5.2)$$

we notice that in the thermodynamic limit this expression becomes equal to the electrostatic energy of point charges of magnitude Ze immersed in a uniform compensating background. This classical energy takes the name of Ewald (or Madelung) energy, and scales—like all of the terms in Eq. (5.1)—linearly with the number of sites:

$$\gamma_{Ewald} = \frac{\Delta E_{ext}}{N_c} = \frac{Z^2 e^2}{2} \left(\sum_{l \neq 0} \frac{1}{R_l} - \frac{1}{\Omega} \int d\mathbf{r} \frac{1}{r} \right). \quad (5.3)$$

For any given periodic structure, this conditionally convergent expression is efficiently calculated partly in real space and partly in reciprocal space, using Ewald-Fuchs techniques. For primitive lattices it is sometimes expressed in terms of the atomic-sphere radius $R_\Omega = (3\Omega/4\pi)^{1/3} = Z^{1/3} r_s$, as:

$$\gamma_{Ewald} = -\frac{Z^2 e^2}{R_\Omega} \alpha_{Ew}, \quad (5.4)$$

where α_{Ew} is an adimensional structural constant, which equals *e.g.* 1.79175 (for the fcc lattice), or 1.79186 (for the bcc lattice). Disordered systems are

simulated upon periodically repeating large unit cells. Short-range interactions of the Born-Mayer kind are sometimes added to Eq. (5.2) in order to account for repulsion amongst bare ion cores.⁷

We now switch to the quantum-mechanical term ΔE_{el} , which is evaluated starting from a power expansion in the strength of the bare perturbation

$$\delta V_{ext}(\mathbf{r}) = \sum_I v(|\mathbf{r} - \mathbf{R}_I|) + e^2 \int d\mathbf{r}' \frac{n^{(0)}}{|\mathbf{r} - \mathbf{r}'|}, \quad (5.5)$$

where the second term accounts for subtraction of the uniform background, and v is the local electron-ion pseudopotential:

$$v(r) = -Ze^2/r + v^{(sr)}(r), \quad (5.6)$$

whose short-range term $v^{(sr)}$ mimicks the nonclassic repulsion due to core orthogonalization.

Using Eqs. (1.22) and (1.24) one gets:

$$\Delta E_{el} \simeq E_{el}^{(1)} + E_{el}^{(2)} \simeq \int d\mathbf{r} \delta V_{ext}(\mathbf{r}) n^{(0)}(\mathbf{r}) + \frac{1}{2} \int d\mathbf{r} \delta V_{ext}(\mathbf{r}) n^{(1)}(\mathbf{r}), \quad (5.7)$$

which is correct up to second order. Only the last term in Eq. (5.7) involves the actual electronic structure of the perturbed system: because of this, it is usually called the “band-structure” term E_{bs} . As for the first term in Eq. (5.7), one has

$$\begin{aligned} \int d\mathbf{r} \delta V_{ext}(\mathbf{r}) n^{(0)}(\mathbf{r}) &= \frac{N_c Z}{\Omega} \int d\mathbf{r} v^{(sr)}(r) + \\ &+ \frac{Ze^2}{\Omega} \int d\mathbf{r} \left(-Z \sum_I \frac{1}{|\mathbf{r} - \mathbf{R}_I|} + \int d\mathbf{r}' \frac{n^{(0)}}{|\mathbf{r} - \mathbf{r}'|} \right). \end{aligned} \quad (5.8)$$

In the thermodynamic limit, the (divergent) electrostatic terms identically cancel owing to translational invariance; the first term provides the only surviving contribution. Defining the average of the noncoulombic electron-ion interaction as

$$\alpha_1 = \frac{1}{\Omega} \int d\mathbf{r} v^{(sr)}(r), \quad (5.9)$$

the results found so far allow recasting Eq. (5.1) as:

$$E_{tot} \simeq N_c (ZU^{(0)} + Z\alpha_1) + N_c \gamma_{Ewald} + E_{bs}, \quad (5.10)$$

where we have indicated with $U^{(0)} = E_{tot}^{(0)}/N$ the energy per electron of the uniform electron gas. The first term in Eq. (5.10) depends on the atomic volume but not on structure, while the last two are structure-dependent. The α_1 contribution to Eq. (5.10) has a physically transparent meaning: its volume

derivative represents a positive pressure on the valence-electron gas, which is due to the nonclassical effect of orthogonalization repulsion from the ion cores.

Last but not least, we have now to evaluate the term E_{bs} , whose expression in reciprocal space is:

$$E_{bs} = \frac{1}{2(2\pi)^3} \int d\mathbf{k} \delta V_{ext}^*(\mathbf{k}) n^{(1)}(\mathbf{k}); \quad (5.11)$$

owing to the fact that $n^{(1)}$ is a neutral distribution, and δV_{ext}^* a neutral potential, the integral is regular around $\mathbf{k}=0$. At this point all of the “band-structure” information which is needed can be fed in via linear-response theory:

$$E_{bs} = \frac{1}{2(2\pi)^3} \int d\mathbf{k} \delta V_{ext}^*(\mathbf{k}) \chi(k) \delta V_{ext}(\mathbf{k}). \quad (5.12)$$

Introducing then the structure factor per ion

$$S(\mathbf{k}) = \frac{1}{N_c} \sum_i e^{-i\mathbf{k}\mathbf{R}_i}, \quad (5.13)$$

the band-structure term is transformed into

$$E_{bs} = \frac{N_c^2}{2(2\pi)^3} \int' d\mathbf{k} |S(\mathbf{k})|^2 v^2(k) \chi(k), \quad (5.14)$$

where the primed integral accounts for the subtraction of the background potential in order to get δV_{ext}^* from the full pseudopotential $v(k)$. This is a rather subtle point: in the thermodynamic limit $N_c |S(\mathbf{k})|^2$ has a singular $\mathbf{k}=0$ term, equal to $(2\pi)^3 \delta(\mathbf{k})/\Omega$, both for ordered and disordered ionic configurations (see below), and with the prime we indicate that this δ -like term has to be omitted from the integration. It is customary to use the dielectric function $\epsilon(k)$ instead of the density response $\chi(k)$; by defining the dimensionless function:

$$\mathcal{G}(k) = \left(\frac{4\pi Z e^2}{k^2}\right)^{-2} v^2(k) \left[1 - \frac{1}{\epsilon(k)}\right], \quad (5.15)$$

known as the “energy-wave-number characteristic”, we write the total energy of the system as:

$$E_{tot} = N_c(ZU^{(0)} + Z\alpha_1 + \gamma_{Ewald}) - \frac{N_c^2}{2(2\pi)^3} \int' d\mathbf{k} |S(\mathbf{k})|^2 \mathcal{G}(k) \frac{4\pi Z^2 e^2}{k^2}. \quad (5.16)$$

This formal expression is quite general and holds for both ordered or disordered configurations of the nuclei.

5.3 Crystalline metals

The above formalism straightforwardly provides the total energy per atom in a crystal at zero temperature, where the configuration of the ions is lattice-periodical. Starting from Eqs. (5.13) and (5.16), and assuming for simplicity a primitive lattice, the total energy per atom is given by:

$$\frac{E_{tot}}{N_c} = ZU^{(0)} + Z\alpha_1 + \gamma_{Ewald} - \frac{1}{2\Omega} \sum_{\mathbf{G} \neq 0} \mathcal{G}(\mathbf{G}) \frac{4\pi Z^2 e^2}{\mathbf{G}^2}, \quad (5.17)$$

where we have used the identity

$$\sum_{lm} e^{i\mathbf{k}(\mathbf{R}_l - \mathbf{R}_m)} = N_c \sum_l e^{i\mathbf{k}\mathbf{R}_l} = N_c \frac{(2\pi)^3}{\Omega} \sum_{\mathbf{G}} \delta(\mathbf{k} - \mathbf{G}), \quad (5.18)$$

and $\mathbf{G}=0$ has to be omitted from the sum, as explained above.

Such a simple expression provides excellent results for the binding energy of simple metals. It can be used in conjunction with several different recipes for generating model pseudopotentials.² A crude, although effective, pseudopotential in this class is the Ascroft empty-core model:⁸

$$v(r) = 0, \quad r < r_c; \quad v(r) = -Ze^2/r, \quad r > r_c, \quad (5.19)$$

depending on the single parameter r_c , the core radius; this is fixed once for all by fitting a single empirical datum (typically the equilibrium lattice constant), and then several physical properties are investigated “almost” from first principles. The energy-wave-number characteristic assumes in this case the simple form

$$\mathcal{G}(k) = \cos^2(kr_c) \left[1 - \frac{1}{\epsilon(k)} \right]; \quad (5.20)$$

since $\epsilon(k) \rightarrow 1$ at large k , the reciprocal sum in Eq. (5.17) converges. We furthermore notice that the cosine oscillations are the fingerprint of the pseudopotential discontinuity at $r=r_c$ in Eq. (5.19); faster reciprocal-space convergence is easily achieved with pseudopotentials which are smooth in r -space.

	r_c (a.u.)	Theory	Expt.
Na	1.67	0.45	0.46
K	2.14	0.38	0.39
Rb	2.61	0.34	0.37
Al	1.12	1.41	1.38

Table 5.1 Modulus of the binding energy (Ry) for some simple metals, as obtained from pseudopotential perturbation theory, and empty core model pseudopotentials. The core radii are reported as well. After Ref. 9.

The binding energies found from Eq. (5.17), and using the model potential of Eq. (5.19), are reported in Table 5.1: these are compared to the experimental binding energy obtained as the sum of the cohesive energy and of the appropriate ionization energy of the free atom.

The macroscopic elastic constants can be obtained either from numerical differentiation, using Eq. (5.17) with unit cells of various size and shape, or equivalently from some more elegant procedure.¹⁰ There are three independent elastic constants in cubic materials: these are indicated as c_{11} , c_{12} , and c_{44} . A calculation for the alkali metals is reported in Table 5.2: this is performed within pseudopotential perturbation theory and the Ashcroft empty-core pseudopotentials, Eq. (5.19), but with core radii slightly different from the calculation reported in Table 5.1.

	r_c (a.u.)	$\frac{1}{2}(c_{11} - c_{12})$	c_{44}	$B = \frac{1}{3}(c_{11} + 2c_{12})$
Na	1.76	0.083 (0.069 - 0.085)	0.637 (0.58 - 0.65)	0.897 (0.54 - 0.88)
K	2.13	0.034 (0.038 - 0.044)	0.268 (0.26 - 0.29)	0.388 (0.37 - 0.40)
Rb	2.26	0.024 (0.026)	0.190 (0.16)	0.251 (0.26)

Table 5.2 Calculated elastic constants of the alkali metals, in 10^{11} dyn/cm², after Ref. 10. The corresponding experimental low-temperature data are shown in parentheses.

5.4 The pairwise interaction

In this Section we get more insight into the physical meaning of Eq. (5.16) for a general nonperiodical configuration of the ions; to this aim, it is convenient to start expressing also the classical Ewald energy in terms of the structure factor. We straightforwardly obtain:

$$\frac{\Delta E_{ext}}{N_c} = \gamma_{Ewald} = \frac{1}{2(2\pi)^3} \int' d\mathbf{k} [N_c |S(\mathbf{k})|^2 - 1] \frac{4\pi Z^2 e^2}{k^2}, \quad (5.21)$$

where the second term in square brackets accounts for removal of self-interaction terms [$l=m$ terms in Eq. (5.2)], and the primed integral has the same meaning as above.

It proves convenient to explicitly separate $l=m$ terms also in Eq. (5.16). The band-structure energy per atom is then written as:

$$\frac{E_{bs}}{N_c} = U_{self}^{(el)} - \frac{N_c}{2(2\pi)^3} \int' d\mathbf{k} [|S(\mathbf{k})|^2 - \frac{1}{N_c}] \mathcal{G}(k) \frac{4\pi Z^2 e^2}{k^2}, \quad (5.22)$$

where the self-interaction electronic term is given by:

$$U_{self}^{(el)} = -\frac{1}{2(2\pi)^3} \int d\mathbf{k} \mathcal{G}(k) \frac{4\pi Z^2 e^2}{k^2}. \quad (5.23)$$

This term is structure-independent: owing to the cutoff provided by $\mathcal{G}(k)$, the integral is convergent at large k , and the prime has been dropped because the integrand has no δ -like terms at $\mathbf{k}=0$. We thus arrive at

$$\frac{E_{tot}}{N_c} = ZU^{(0)} + Z\alpha_1 + U_{self}^{(el)} + \frac{1}{2(2\pi)^3} \int' d\mathbf{k} [N_c |S(\mathbf{k})|^2 - 1] \phi(k), \quad (5.24)$$

where the last term collects all of the structure-dependent effects, and

$$\phi(k) = \frac{4\pi Z^2 e^2}{k^2} [1 - \mathcal{G}(k)] \quad (5.25)$$

plays the role of an effective ion-ion screened interaction. In order to see this, we first transform the primed integral in Eq. (5.24) into an unprimed one, explicitly subtracting off the δ -like term in $N_c |S(\mathbf{k})|^2$:

$$\begin{aligned} & \frac{1}{2(2\pi)^3} \int' d\mathbf{k} [N_c |S(\mathbf{k})|^2 - 1] \phi(k) = \\ & = -\frac{1}{2\Omega} \phi(k=0) + \frac{1}{2(2\pi)^3} \int d\mathbf{k} [N_c |S(\mathbf{k})|^2 - 1] \phi(k). \end{aligned} \quad (5.26)$$

At this point, Eq. (5.24) can be straightforwardly transformed back to real space, yielding for the energy of the whole system:

$$E_{tot} = N_c [ZU^{(0)} + Z\alpha_1 + U_{self}^{(el)} - \frac{1}{2\Omega} \phi(k=0)] + \frac{1}{2} \sum_{l \neq m} \phi(|\mathbf{R}_l - \mathbf{R}_m|). \quad (5.27)$$

This is a very important result, usually attributed to M.H. Cohen:¹¹ the total energy of a metal, when expanded to second order in the electron-ion interaction, is the sum of two contributions. The first contribution depends on the average density only, while the second has the explicit form of a central pairwise interaction. Most of the literature is somewhat reticent in detailing the various terms which appear in the structure-independent contribution: the four terms in square bracket in Eq. (5.27) are consistent with those given in Ref. 4 (Section 6.5).

We pause at this point to illustrate the basic features of the effective pairwise interaction $\phi(r)$, whose Fourier transform is given in Eq. (5.25). This interaction is the sum of two terms: the first term can be interpreted as a direct ion-ion interaction, and the second as an indirect ion-electron-ion interaction, mediated by the polarizable valence-electron gas. Incidentally, we notice that $U_{self}^{(el)}$ has the simple meaning of the value of the indirect interaction at $r=0$, after Eq. (5.23).

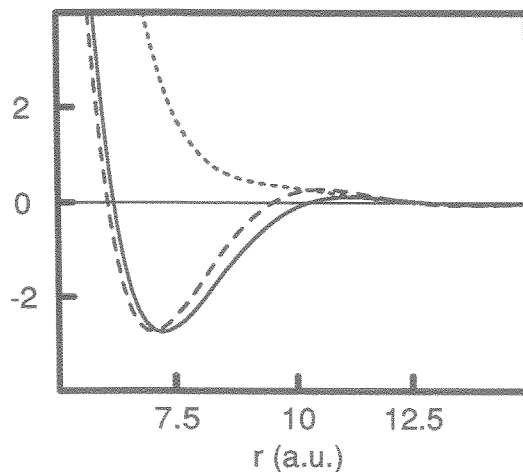


Figure 5.1. The pairwise interaction, in 10^{-3} a.u., for solid Na ($r_s=3.93$). The empty core pseudopotential of Eq. (5.19) has been used, with $r_c=1.69$ a.u., and dielectric functions as follows: Ref. 12 (solid); RPA (dotted); LDA (dashed).

The meaning of Eqs. (5.15) and (5.25) is more transparent if we neglect for a while the short-range term in the pseudopotential $v(k)$, Eq. (5.6): we obtain then

$$\phi(k) \simeq \frac{4\pi Z^2 e^2}{\epsilon(k)k^2}, \quad (5.28)$$

i.e. the effective pairwise interaction becomes simply a screened Coulomb potential, whose Fourier antitransform is shown in Fig. 4.3 for two simple choices of the dielectric function. When the short-range term in the pseudopotential is included, \mathcal{G} has the limiting value of 1 at small k anyhow: hence the pairwise potential of Eq. (5.25) is short-range, in the sense that the indirect term exactly cancels the Coulomb singularity of the direct one, and $\phi(k=0)$ is well defined. Nonetheless, the Fourier antitransform of Eq. (5.25) has long-range oscillations of the Friedel type, which are due to the sharpness of the Fermi surface. The effect of such oscillations upon the structural stability of simple metals has been widely discussed.² Typical pairwise potentials for crystalline Na are shown in Fig. 5.1; the one drawn in solid line, after the dielectric function of Singwi *et al.*,¹² provides remarkably accurate phonon spectra (see below). The picture also shows that plain RPA screening provides a very poor pairwise interaction, while a simple LDA dielectric function—where f_{xc} in Eq. (4.10) is taken as constant—compares rather well with the more sophisticated approaches as far as the pairwise interaction is concerned.

A final comment about the density- and structure-dependent terms in the expressions for the total energy, Eqs. (5.24) and (5.27). $U^{(0)}$, α_1 , and $\phi(k=0)/2\Omega$

have an obvious dependence upon the average electron density; besides them, the $\mathcal{G}(k)$ function depend on density as well, through the density dependence of the electron-gas dielectric function. Therefore $U_{self}^{(el)}$ depends on density as well, and even the pairwise interaction potential ϕ , accounting for all of the structural effects, has an *implicit* density dependence.

5.5 Liquid metals

The energy of any desordered configuration is given by Eq. (5.24), with the appropriate structure factor; one is quite often interested into some kind of configurational average $\langle E_{tot} \rangle$ of Eq. (5.16) or Eq. (5.24). For a liquid, and making contact with the traditional notations,¹³ the average structure factor is to be identified with:

$$\bar{S}(k) = \frac{1}{N_c} \left\langle \sum_{im} e^{i\mathbf{k}(\mathbf{R}_i - \mathbf{R}_m)} \right\rangle = N_c \langle |S(k)|^2 \rangle. \quad (5.29)$$

The behavior of this function in the thermodynamic limit is most easily investigated starting from the pair-correlation function, in terms of which the structure factor reads

$$\bar{S}(k) = 1 + \frac{1}{\Omega} \int d\mathbf{r} e^{i\mathbf{k}\mathbf{r}} g(r). \quad (5.30)$$

Since $g(r) \rightarrow 1$ at large r , $\bar{S}(k)$ has a δ -like singularity at $k=0$, as anticipated above: we recast the previous equation as

$$\bar{S}(k) = 1 + \frac{(2\pi)}{\Omega} \delta(\mathbf{k}) + \frac{1}{\Omega} \int d\mathbf{r} e^{i\mathbf{k}\mathbf{r}} [g(r) - 1], \quad (5.31)$$

and now the integral yields a regular continuous function of k .

The configurational average energy per atom is then, after Eqs. (5.24) and (5.26):

$$\frac{\langle E_{tot} \rangle}{N_c} = ZU^{(0)} + Z\alpha_1 + U_{self}^{(el)} - \frac{1}{2\Omega} \phi(k=0) + \frac{1}{2(2\pi)^3} \int d\mathbf{k} [\bar{S}(k) - 1] \phi(k); \quad (5.32)$$

a straightforward transformation provides the equivalent expression:

$$\frac{\langle E_{tot} \rangle}{N_c} = ZU^{(0)} + Z\alpha_1 + U_{self}^{(el)} - \frac{1}{2\Omega} \phi(k=0) + \frac{1}{2\Omega} \int d\mathbf{r} g(r) \phi(r); \quad (5.33)$$

this latter expression coincides of course with the configurational average of Eq. (5.27), as directly evaluated in real space.

At finite temperatures, the effective pairwise interaction ϕ provides, after Eq. (5.27), the total adiabatic potential for ionic motion at fixed volume.³⁻⁵ When T is larger than the harmonic Debye frequency (~ 100 - 200 K), the ionic

motion is classical and the time evolution can be simulated through standard molecular dynamics. It is then possible to get anharmonic properties, thermal expansion, melting of the solid, structural and other physical properties of the fluid phase either from approximate analytical theories or from computer experiments. The results are very good for simple metals.^{14,15}

5.6 Phonons vs. elasticity: a paradox

When considering deviations from the equilibrium positions, a second order expansion of Eq. (5.16) or Eq. (5.27) in terms of the ionic positions (at constant atomic volume) provides the spectrum of the harmonic phonons. The detailed formalism leading to the dynamical matrix will be illustrated below, Chapter 9. Here we anticipate some results, showing in Fig. 5.2 the phonon spectrum of Na, where the same ingredients as in Ref. 16 have been used: the corresponding pairwise interaction has been shown in Fig. 5.1 (solid line).

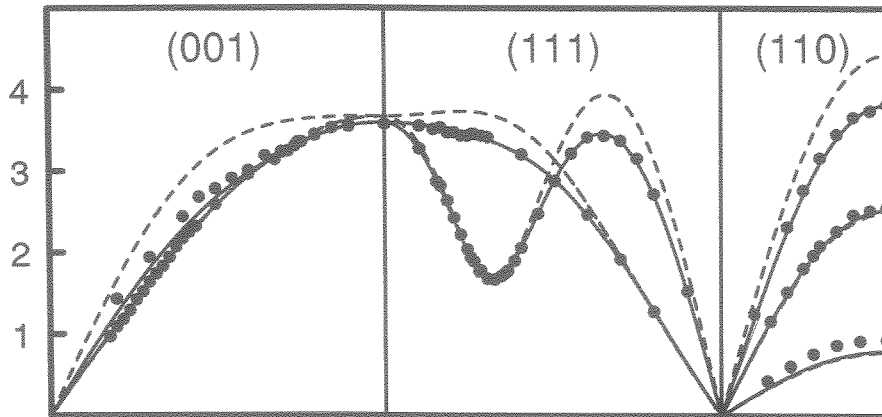


Figure 5.2. Calculated phonon-dispersion curves for bcc Na, in THz. The empty core pseudopotential of Eq. (5.19) has been used, with $r_c=1.69$ a.u., and dielectric functions as follows: Ref. 12 (solid); RPA (dashed). The neutron scattering data are also shown (dots), after Ref. 17.

The agreement amongst theory and experiment appears as excellent; nonetheless the zone-center slopes of the phonon dispersion curves given by pseudopotential perturbation theory are *incorrect*. Such slopes determine the macroscopic elastic constants, whose values determined in this way are incorrect as well.

To better state the issue, the elastic constants may be determined in two ways, either via homogeneous deformations of the whole solid or via elastic waves of long wavelength, and the two measurements must provide the same

values for a real material. We now discuss this same issue within the framework of pseudopotential perturbation theory. As for the homogeneous deformations, some results obtained in this way are reported in the previous Section (Table 5.2), and found in good agreement with the experiment; in particular, the calculated values of c_{12} and c_{44} do not coincide: for instance in the case of Na Table 5.2 implies $c_{12}=0.842$ and $c_{44}=0.637$ (in units of 10^{11} dyn/cm²). On the contrary, it is easy to show that long-wavelength phonons obtained from Eq. (5.27) always imply $c_{12}=c_{44}$. This is because the phonon frequencies are obtained from a difference of two ionic configurations having the same atomic volume: hence only the last term of Eq. (5.27) enters lattice dynamics, and the remaining (volume-dependent) term is irrelevant. Therefore the dynamical matrix of a simple metal—within pseudopotential perturbation theory to second order—is indistinguishable from the dynamical matrix of a classical system with genuine pairwise central forces: for such a system, a venerable theorem, due to Cauchy,¹⁸ implies in cubic symmetry $c_{12}=c_{44}$.

We have therefore found an important paradox, and identified its source: within pseudopotential perturbation theory, lattice vibrations sample only the last (two-body) term in Eq. (5.27), while a homogeneous deformation of the whole solid samples the first (volume-dependent) term as well. It follows that second-order pseudopotential perturbation theory, whose total energy is given in Eqs. (5.16) and (5.27), has an internal inconsistency. Given the fact that the bare ionic pseudopotential in reciprocal space is particularly strong at low- k , it is no surprise that the paradox manifests itself in the long-wavelength limit, although microscopic (*i.e.* short-wavelength) properties like the phonon spectra, Fig. 5.2, are predicted pretty well. It must also be observed that a long acoustic wave *really* samples large regions of the solid where the average atomic volume is varied, *i.e.* where the effective n_0 varies as well.

The point where an inconsistent logical step has been performed is easy to detect. The theory describes, *for a given configuration of the ions*, the ground state as a perturbation over the electron gas, chosen as a reference system, and where only the second order terms in the bare pseudopotentials have been kept. In order to perform lattice dynamics, we now need to perturb the ionic configuration and to evaluate second-order terms in the ionic displacements. If we use Eqs. (5.16) and (5.27) for lattice dynamics, we actually perform a second-order expansion over a second-order expansion: it is then clear that some terms are missing, while harmonic properties of the solid can be consistently described only if a fourth-order pseudopotential perturbation theory is adopted. This is in principle possible, but it generates three- and four-body terms in the effective ion-electron-ion interaction;¹⁹ in this case the hypothesis of Cauchy's theorem no longer apply, and we recover the equality for the elastic constants evaluated from either homogeneous deformations or long acoustic waves.

5.7 References

- [5.1] W.A. Harrison, *Pseudopotentials in the Theory of Metals* (Benjamin, New York, 1966).
- [5.2] V. Heine, in *Solid State Physics*, edited by H. Ehrenreich, F. Seitz, and D. Turnbull, vol **24** (Academic, New York, 1970), p.1; M.L. Cohen and V. Heine, *ibid.* p.250.
- [5.3] N.W. Ashcroft and D. Stroud, in *Solid State Physics*, edited by H. Ehrenreich, F. Seitz and D. Turnbull, vol **33** (Academic, New York, 1978), p. 1.
- [5.4] N.H. March and M.P. Tosi, *Coulomb Liquids* (Academic, New York, 1984).
- [5.5] J. Hafner, *From Hamiltonians to Phase Diagrams* (Springer, Berlin, 1987).
- [5.6] D.C. Wallace, *Thermodynamics of Crystals* (Wiley, New York, 1972).
- [5.7] D.C. Wallace, *Phys. Rev.* **176**, 832 (1968).
- [5.8] N.W. Ashcroft, *Phys. Lett.* **23**, 48 (1966).
- [5.9] N.W. Ashcroft and D.C. Langreth, *Phys. Rev.* **155**, 683 (1967).
- [5.10] W.M. Shyu and G.D. Gaspari, *Phys. Rev.* **170**, 687 (1968).
- [5.11] M.H. Cohen, *J. Phys. Radium* **23**, 643 (1962).
- [5.12] K.S. Singwi, A. Sjölander, M.P. Tosi, and R.H. Land, *Phys. Rev. B* **1**, 1044 (1970).
- [5.13] J.P. Hansen and I.R. McDonald, *Theory of Simple Liquids* (Academic, New York, 1986).
- [5.14] M.P. Tosi, in *Electron Correlation in Solids, Molecules, and Atoms*, edited by J.T. Devreese and F. Brosens (Plenum, New York, 1983), p. 235.
- [5.15] See for instance: D.C. Wallace, in *Electronic Structure, Dynamics and Quantum Structural Properties of Condensed Matter*, edited by J.T. Devreese and P. Van Camp (Plenum, New York, 1985), p. 521.
- [5.16] D.L. Price, K.S. Singwi, and M.P. Tosi, *Phys. Rev. B* **2**, 2983 (1970)
- [5.17] A.D.B. Woods, B.N. Brockhouse, R.H. March, A.T. Stewart, and R. Bowers, *Phys. Rev.* **128**, 1112 (1962).
- [5.18] M. Born and K. Huang, *Dynamical Theory of Crystal Lattices* (Oxford University Press, Oxford, 1954).

- [5.19] E.G. Brovman and Yu.M. Kagan, in: *Dynamical Properties of Solids*, edited by G.K. Horton and A.A. Maradudin, vol. 1 (North-Holland, Amsterdam, 1974), p. 191.

Chapter 6

Ground state of periodic solids (incomplete)

6.1 Kohn-Sham equation

We apply here the general results of Chapter 1 to an infinite crystalline solid in the thermodynamic limit, where the electron density is lattice-periodical. In such circumstances, the KS potential is periodic as well, and the KS orbitals have the Bloch form. Such a seemingly obvious statement amounts indeed to an additional assumption, *i.e.* the boundary condition of vanishing macroscopic electric field is implicitly chosen for integrating Poisson equation. It is worth noticing that—for a dielectric solid—other choices are in principle allowed: the macroscopic field does *not* depend on the periodic charge and represents in fact an arbitrary boundary condition.

The methods of electronic structure theory discussed in this chapter—and more generally throughout these lecture notes—refer to implementations in a plane-waves basis set. In order to greatly simplify the algebra, local pseudopotentials, in the form of Eq. (5.6), are used throughout. All of the results and theorems are trivially generalizable to nonlocal pseudopotentials as well, although this requires lengthy expressions. In fact, all the calculations which have been performed for real materials using the methods illustrated in the present lecture notes have been indeed performed using *ab-initio* norm-conserving pseudopotentials.

We only consider here the case of insulators, where the KS bands are either completely filled or empty. The KS equation is then

$$\left(-\frac{1}{2}\nabla^2 + \sum_{\mathbf{G}} V_{KS,\mathbf{G}} e^{i\mathbf{G}\mathbf{r}}\right)\varphi_{v,\mathbf{q}}(\mathbf{r}) = \epsilon_v(\mathbf{q})\varphi_{v,\mathbf{q}}(\mathbf{r}), \quad (6.1)$$

where $\varphi_{v,\mathbf{q}}(\mathbf{r})$ are the KS orbitals of the v -th occupied band, corresponding to the energies $\epsilon_v(\mathbf{q})$, and $V_{KS,\mathbf{G}}$ are the Fourier coefficients of the KS potential,

given by:

$$V_{KS,\mathbf{G}} = \frac{1}{\Omega} \sum_s v_s(\mathbf{G}) e^{-i\mathbf{G}\mathbf{R}_s} + \frac{4\pi e^2}{G^2} n_{\mathbf{G}} + V_{xc,\mathbf{G}}. \quad (6.2)$$

Here the subscript s indicates different atoms in the unit cell at sites \mathbf{R}_s , Ω is—at variance with the notations used in the previous chapter—the *cell* volume, $v_s(\mathbf{G})$ are the Fourier *transforms* of the local pseudopotentials evaluated at reciprocal-lattice vectors, $n_{\mathbf{G}}$ are the Fourier coefficients of the periodic electron density, and $V_{xc,\mathbf{G}}$ are the coefficients of the exchange-correlation potential. The latter is—within LDA—a universal function of the local density: $V_{xc}(\mathbf{r}) = \mu_{xc}(n(\mathbf{r}))$, where the form of $\mu_{xc}(n)$ is usually borrowed from electron-gas results.¹ The electronic density is simply obtained from Eq. (1.8), summing on the occupied bands and on the \mathbf{q} -vectors in the Brillouin zone:

$$n(\mathbf{r}) = \frac{2}{\Omega} \sum_{v,\mathbf{q}} |\varphi_{v,\mathbf{q}}(\mathbf{r})|^2 = \frac{2}{(2\pi)^3} \sum_v \int_{BZ} d\mathbf{q} |\varphi_{v,\mathbf{q}}(\mathbf{r})|^2, \quad (6.3)$$

where the Bloch states have a plane-wave-like normalization. This equation, coupled with Eq. (6.1), provides a selfconsistent scheme which can be straightforwardly implemented and solved, usually via iteration.²

It is now important to discuss the $\mathbf{G}=0$ coefficient of V_{KS} , which would fix the absolute scale for one-electron band energies. I stress that this coefficient is completely *arbitrary*, and no such absolute scale exists for infinite solids, due to the presence of Coulomb interactions. Such arbitrariness stems indeed from a constant which can be arbitrarily added when integrating Poisson equation in an extended system. A careful analysis of the thermodynamic limit shows very elegantly³ that no physical argument can be invoked to fix this zero in an infinite periodic solid as a bulk quantity. The computer codes which actually perform selfconsistent calculations for solids—either of the LDA or of the Hartree-Fock⁴ type—set a conventional reference for the one-electron energy zero, which is in general different from code to code, and has no fundamental physical meaning. This arbitrary constant in the one-electron energy levels has no effect on the genuine bulk physical properties, like *e.g.* the total energy per cell, discussed below.

6.2 Total energy per cell

We will find a closed-form expression for the total energy per cell within LDA; the only delicate part in this derivation is the proper cancellation of Coulomb divergences, which will be performed here in an original way. The main results derived in this section have been first published in Ref. 5, although they have been probably known earlier to different authors.

We start again from the general expression for the total energy of any system given in Eq. (1.18), which is reported here for convenience:

$$E_{tot}[n] = E_{ext} + \int d\mathbf{r} V_{ext}(\mathbf{r})n(\mathbf{r}) + T_s[n] + E_H[n] + E_{xc}[n]. \quad (6.4)$$

If the crystal is made of N_c cells, then

$$T_s = \frac{2}{\Omega} \sum_{v, \mathbf{q}} \langle \mathbf{q}, v | \frac{1}{2} p^2 | \mathbf{q}, v \rangle = \frac{2N_c}{(2\pi)^3} \sum_v \int_{BZ} d\mathbf{q} \langle \mathbf{q}, v | \frac{1}{2} p^2 | \mathbf{q}, v \rangle. \quad (6.5)$$

The classical term E_{ext} does not depend on electron density, being for an isolated system simply the Coulomb ion-ion repulsion, whose expression is:

$$E_{ext} = \frac{N_c e^2}{2} \sum'_{l, s, s'} \frac{Z_s Z_{s'}}{|\mathbf{R}_l + \mathbf{R}_s - \mathbf{R}_{s'}|}, \quad (6.6)$$

where l is a cell index. The Hartree term E_H has the simple form of a classical Coulomb repulsion energy as well:

$$E_H[n] = \frac{e^2}{2} \int d\mathbf{r} d\mathbf{r}' \frac{n(\mathbf{r})n(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|}. \quad (6.7)$$

The term involving the bare pseudopotentials can be written as:

$$\begin{aligned} \int d\mathbf{r} V_{ext}(\mathbf{r})n(\mathbf{r}) &= \sum_{ls} \int d\mathbf{r} n(\mathbf{r})v_s(\mathbf{r} - \mathbf{R}_l - \mathbf{R}_s) = \\ &= -e^2 \sum_{ls} \int d\mathbf{r} n(\mathbf{r}) \frac{Z_s}{|\mathbf{r} - \mathbf{R}_l - \mathbf{R}_s|} + \sum_{ls} \int d\mathbf{r} n(\mathbf{r})v_s^{(sr)}(\mathbf{r} - \mathbf{R}_l - \mathbf{R}_s), \end{aligned} \quad (6.8)$$

where the short-range term has been explicitly separated as in Eq. (5.6); exploiting then the lattice periodicity, this last term becomes:

$$\sum_{ls} \int d\mathbf{r} n(\mathbf{r})v_s^{(sr)}(\mathbf{r} - \mathbf{R}_l - \mathbf{R}_s) = N_c \sum_{s, \mathbf{G}} n_{\mathbf{G}}^* v_s^{(sr)}(\mathbf{G}) e^{-i\mathbf{G}\mathbf{R}_s}. \quad (6.9)$$

The $\mathbf{G}=0$ coefficient of the density is $n_0 = Z/\Omega$, where Z is the total ionic charge per cell; a suitable generalization of the definition given in the previous chapter allows to define α_1 as the cell-average of the noncoulombic interaction: the $\mathbf{G}=0$ contribution to Eq. (6.9) can therefore be written in the form $N_c Z \alpha_1$. I now consider the whole (ionic and electronic) charge density altogether:

$$\rho(\mathbf{r}) = e \left[n(\mathbf{r}) - \sum_{ls} Z_s \delta(\mathbf{r} - \mathbf{R}_l - \mathbf{R}_s) \right], \quad (6.10)$$

and whose classical Coulomb interaction energy is:

$$E_{Coul} = \frac{1}{2} \int d\mathbf{r} d\mathbf{r}' \frac{\rho(\mathbf{r})\rho(\mathbf{r}')}{|\mathbf{r}-\mathbf{r}'|}. \quad (6.11)$$

The three terms in the total energy involving Coulomb interactions can be cast then as:

$$E_{ext} + E_H + \int d\mathbf{r} V_{ext}(\mathbf{r})n(\mathbf{r}) = E_{Coul} + N_c \sum_{s, \mathbf{G}} n_{\mathbf{G}}^* v_s^{(sr)}(\mathbf{G}) e^{-i\mathbf{G}\mathbf{R}_s}. \quad (6.12)$$

The Coulomb energy of an extended and neutral system is the integrated energy density, which is most elegantly expressed using the electric field:⁶

$$E_{Coul} = \frac{N_c}{8\pi} \int_{cell} d\mathbf{r} |\mathbf{E}(\mathbf{r})|^2 = \frac{N_c \Omega}{8\pi} \sum_{\mathbf{G}} |\mathbf{E}_{\mathbf{G}}|^2, \quad (6.13)$$

where $\mathbf{E}_{\mathbf{G}}$ are the Fourier coefficients of the microscopic electric field. Given the boundary condition of vanishing macroscopic field assumed since the very beginning, the $\mathbf{G}=0$ term must be excluded from the Coulomb sum.

The exchange-correlation contribution E_{xc} to the total energy is very simple within LDA:

$$E_{xc} = N_c \Omega \sum_{\mathbf{G}} n_{\mathbf{G}}^* \epsilon_{xc, \mathbf{G}}, \quad (6.14)$$

where $\epsilon_{xc, \mathbf{G}}$ are the Fourier coefficients of $\epsilon_{xc}(n(\mathbf{r}))$. Collecting all the previous results together, the total energy per cell is:

$$\frac{E_{tot}}{N_c} = \frac{T_s}{N_c} + Z\alpha_1 + \Omega \sum_{\mathbf{G}} n_{\mathbf{G}}^* \epsilon_{xc, \mathbf{G}} + \sum_{\mathbf{G} \neq 0} \left[\frac{\Omega}{8\pi} |\mathbf{E}_{\mathbf{G}}|^2 + n_{\mathbf{G}}^* \sum_s v_s^{(sr)}(\mathbf{G}) e^{-i\mathbf{G}\mathbf{R}_s} \right], \quad (6.15)$$

where the divergent $\mathbf{G}=0$ Coulomb term has been properly removed. At this point we may replace again the microscopic field with the electron density and the bare pseudopotentials, yielding:

$$\begin{aligned} \frac{E_{tot}}{N_c} = & \frac{T_s}{N_c} + Z\alpha_1 + \Omega \sum_{\mathbf{G}} n_{\mathbf{G}}^* \epsilon_{xc, \mathbf{G}} + \sum_{\mathbf{G} \neq 0} \left[\frac{\Omega 4\pi e^2}{2G^2} |n_{\mathbf{G}}|^2 + \right. \\ & \left. + n_{\mathbf{G}}^* \sum_s v_s(\mathbf{G}) e^{-i\mathbf{G}\mathbf{R}_s} + \frac{4\pi e^2}{2\Omega G^2} \sum_{ss'} Z_s Z_{s'} e^{i\mathbf{G}(\mathbf{R}_{s'} - \mathbf{R}_s)} \right]. \end{aligned} \quad (6.16)$$

We easily identify the last contribution in square brackets:

$$\gamma_{Ewald} = \frac{1}{2\Omega} \sum_{\mathbf{G} \neq 0} \frac{4\pi e^2}{G^2} \sum_{ss'} Z_s Z_{s'} e^{i\mathbf{G}(\mathbf{R}_{s'} - \mathbf{R}_s)} \quad (6.17)$$

with the classical energy (per cell) of point charges immersed in a uniform compensating background, already discussed—for primitive lattices—in the previous chapter. The final result is therefore:

$$\begin{aligned} \frac{E_{tot}}{N_c} = & \frac{2}{(2\pi)^3} \sum_v \int_{BZ} d\mathbf{q} \langle \mathbf{q}, v | \frac{1}{2} p^2 | \mathbf{q}, v \rangle + Z\alpha_1 + \gamma_{Ewald} + \\ & + \Omega \sum_{\mathbf{G}} n_{\mathbf{G}}^* \epsilon_{xc, \mathbf{G}} + \frac{\Omega}{2} \sum_{\mathbf{G} \neq 0} \frac{4\pi e^2}{G^2} |n_{\mathbf{G}}|^2 + \sum_{\mathbf{G} \neq 0} n_{\mathbf{G}}^* \sum_s v_s(\mathbf{G}) e^{-i\mathbf{G}\mathbf{R}_s}. \end{aligned} \quad (6.18)$$

An alternative expression is derived upon exploiting Eq. (6.1) to replace the diagonal elements of $p^2/2$, and using explicitly the sum of the KS eigenvalues; since they are *arbitrary* on an absolute scale, we set their zero at the average value of the Coulomb potential within the solid. The total energy per cell is then:

$$\begin{aligned} \frac{E_{tot}}{N_c} = & \frac{2}{(2\pi)^3} \sum_v \int_{BZ} d\mathbf{q} \epsilon_v(\mathbf{q}) + \gamma_{Ewald} + \\ & + \Omega \sum_{\mathbf{G}} n_{\mathbf{G}}^* (\epsilon_{xc, \mathbf{G}} - V_{xc, \mathbf{G}}) - \frac{\Omega}{2} \sum_{\mathbf{G} \neq 0} \frac{4\pi e^2}{G^2} |n_{\mathbf{G}}|^2, \end{aligned} \quad (6.19)$$

where the two last terms have the obvious meaning of double-counting corrections for exchange-correlation and Coulomb interactions. In case a different zero is chosen for the band structure, an extra term must be added, and the total energy per cell is in fact *independent* of the arbitrary choice of the zero of V_{KS} in Eq. (6.1).

The two expressions given above for the total energy are equivalent when applied to the *selfconsistent* density and KS orbitals; if the solution is found via the usual iterative method, at a given cycle before achieving selfconsistency the two expressions no longer provide the same value. The first expression, Eq. (6.18), has the virtue of being variational, providing therefore at any cycle an upper bound to the converged value; even more important, its error is quadratic in the “distance” from selfconsistency, while different expressions produce a linear error.^{2,7} The direct minimization methods^{8,9} discussed in Chap. 1 make of course use of Eq. (6.18).

6.3 Forces

The total energy plays a crucial role in structural stability of materials. In addition, from total energy calculations performed at nonequilibrium geometry, the forces on atoms can be determined via numerical differentiation. From this information, the lattice dynamics of crystals can be investigated, as well as the structural relaxation paths in low symmetry situations (such as *e.g.* defects, surfaces, interfaces).

The Hellman-Feynman theorem^{10,11} circumvents the need for numerical differentiation, and allows the direct evaluation of forces, in the form of analytic derivatives, from a single-shot selfconsistent calculation. In its original (all-electron) formulation the theorem has an immediate electrostatic meaning: the total force acting over a given atomic nucleus in a molecule or in a solid is just equal to the electrostatic force originating from the electron distribution and from all the other nuclei. In a pseudopotential formalism, this purely electrostatic picture is lost, although the theorem maintains its validity and usefulness.

Using the results of Chap., 1, the force acting on the atom at the site $\mathbf{R}_{ls} = \mathbf{R}_l + \mathbf{R}_s$ is straightforwardly obtained upon displacing this atom and evaluating the first-order change of the total energy, while all the other atoms are kept fixed. This is a straightforward application of Eq. (1.21) or equivalently Eq. (1.22), where the parameter λ is identified with the displacement:

$$\mathbf{F}_{ls,\alpha} = Z_s e^2 \sum'_{l's'} Z_{s'} \frac{\mathbf{R}_{ls} - \mathbf{R}_{l's'}}{|\mathbf{R}_{ls} - \mathbf{R}_{l's'}|^3} + \int d\mathbf{r} n(\mathbf{r}) v'_{s,\alpha}(\mathbf{r} - \mathbf{R}_{ls}); \quad (6.20)$$

the primed sum is carried over all the sites other than ls , α is a cartesian component, and the obvious notation $v'_{s,\alpha}(\mathbf{r}) = r_\alpha v'_{s,\alpha}(|\mathbf{r}|)/r$ has been used for the pseudopotential gradient.

In this Section we are of course interested into nonequilibrium configurations, such as to provide nonvanishing forces. However, for the time being only *lattice-periodical* configurations are considered. We mention incidentally that the case of nonperiodical—albeit commensurate—configurations of a given solid can be studied much in the same way, but at the price of using larger elementary cells: this is called the “supercell” approach. Since we limit ourselves to the case where $n(\mathbf{r})$ is periodic, the force in Eq. (6.20) does not depend on the cell index l ; using explicitly the Fourier series of the periodic charge, we get:

$$\begin{aligned} \mathbf{F}_{s,\alpha} &= \mathbf{F}_{s,\alpha}^{(Ew)} + \sum_{\mathbf{G}} n_{\mathbf{G}}^* \int d\mathbf{r} e^{-i\mathbf{G}\mathbf{r}} v'_{s,\alpha}(\mathbf{r} - \mathbf{R}_s) = \\ & \mathbf{F}_{s,\alpha}^{(Ew)} + i \sum_{\mathbf{G}} G_\alpha n_{\mathbf{G}}^* e^{-i\mathbf{G}\mathbf{R}_s} v_s(\mathbf{G}). \end{aligned} \quad (6.21)$$

The classic term $\mathbf{F}_{s,\alpha}^{(Ew)}$ accounts for nuclear repulsion and can be evaluated using Ewald techniques. The generalization to modern quasilocal pseudopotentials requires a modest effort.⁵

So far, everything looks quite simple; given the ground-state density corresponding to an arbitrary (nonequilibrium) configuration, the forces on the atoms can be easily evaluated. But the important point to notice is that the Hellmann-Feynman theorem, Eqs. (1.21) and (1.22) holds if the *exact* many-electron wavefunction is used to build the density; but in general it does not hold for *approximate* wavefunctions. It is rather straightforward to prove that the theorem still holds exactly if all quantities are consistently evaluated at the

LDA level; nonetheless the problem remains a severe one when a finite-basis expansion is adopted, as usual, to find a variational approximation. This main problem has hampered a massive use of the theorem in quantum chemistry, and basis sets have been seldom optimized specifically to the aim.^{12,13} When plane waves are used, such as for the treatment given here, the problem simply does not show up, and the Hellman-Feynman theorem can be safely applied to any finite-basis expansion, provided only that the calculation is very well converged towards selfconsistency. The key feature of plane waves allowing such performance is the following: there is no “basis error” when evaluating the gradient of a given function projected over the basis manifold. This is due to the fact that a finite set of plane waves spans the same manifold as the one spanned by their \mathbf{r} -gradients. Other basis are much less comfortable: experience from quantum chemistry has been carried over successfully to localized basis in solid state physics, but more sophisticated basis sets have long escaped a proper calculation of forces. In particular, a direct calculation of forces within the LAPW method has been achieved only in 1991,¹⁴ and requires a rather heavy formalism.

Applications: equilibrium in low-symmetry situations, Karel, fononi (anche anarmonici), CP...

6.4 Macroscopic stress

For a large, although finite, system the macroscopic stress is a simple function of the forces acting on individual atoms; for instance, the system is in equilibrium and stress-free if and only if all the forces vanish, including those acting on the surface atoms. When studying an extended solid in the thermodynamic limit, it is useful to consider a bulk region of it as a *nonisolated* system. In this case, the macroscopic stress must be regarded as an extensive physical observable, *independent* from the forces: for instance all the forces may vanish in the bulk region of a solid which is nonetheless subject to macroscopic stress.

A toy model to illustrate this concept is a one-dimensional chain of balls and springs. A finite chain is in equilibrium when all the forces on the balls vanish; upon application of external forces on the two end balls only, the whole chain will be under tension. It is then evident that an infinite chain—in the periodic configuration where all the springs are elongated the same amount—is in a state of tensile stress; nonetheless, the forces on individual balls are zero.

For an infinite periodic solid, forces and stress can be considered as complementary physical variables. In fact the forces provide the first-order change of the total energy under variation of the ionic positions in the unit cell; the stress tensor provides the energy change due to macroscopic strain, *i.e.* to variations in size and shape of the unit cell. A free-standing solid is in equilibrium only if forces *and* stress are vanishing. A simple illustration can be given taking *e.g.* any cubic crystal: the forces upon the atoms vanish by symmetry, at any value of the lattice constant a . But of course there is only one equilibrium value for

α ; other values are possible only if an external pressure is applied to the sample.

After this introduction, the main issue can be stated. For a given periodic configuration of a solid in the Born-Oppenheimer approximation, the forces can be evaluated—as shown above in Sec. 6.3—from a single calculation of the ground-state electronic wavefunction. Is it possible to evaluate even the macroscopic stress tensor from a single ground-state calculation? This question was positively answered in 1983 by Nielsen and Martin,^{15,16} who also provide an explicit expression for the stress tensor within LDA. Implementations of this expression within first-principle theory of real materials has proved very useful in two ways: as an efficient method for calculating physical observables,^{17,18} and as a practical tool when searching for the theoretical structural equilibrium (where forces *and* stress must vanish) in low-symmetry solids: we quote Ref. 19 as a single example. The stress theorem has been recently generalized beyond LDA.²⁰

We start considering a macroscopic and finite solid of volume \mathcal{V} , for which the internal stress is balanced by external forces applied at the sample boundary. Such stress is defined as the derivative of the total energy with respect to macroscopic strain:

$$T_{\alpha\beta} = \frac{\partial E_{tot}}{\partial \varepsilon_{\alpha\beta}} = \frac{\partial E_{ext}}{\partial \varepsilon_{\alpha\beta}} + \int_{\mathcal{V}} d\mathbf{r} n(\mathbf{r}) \frac{\partial V_{ext}(\mathbf{r})}{\partial \varepsilon_{\alpha\beta}}, \quad (6.22)$$

where the Hellmann-Feynman theorem, Eq. (1.21), has been used in the last equality.

6.5 The stress theorem

The derivation of the stress theorem given here is somewhat different from the original Nielsen-Martin one,^{15,16} and follows instead Ref. 20.

In nonprimitive lattices, and for low-symmetry strain tensors, the displaced positions are *not* determined by symmetry only. Starting from the equilibrium positions, the ionic displacements in a strained crystal are, to linear order in the strain:

$$R_{l,\alpha} \rightarrow R'_{l,\alpha} = R_{l,\alpha} + R_{ls,\beta} \varepsilon_{\alpha\beta} + \Gamma_{s,\alpha\beta\gamma} \varepsilon_{\beta\gamma}, \quad (6.23)$$

where Γ is the so-called internal strain tensor (sum over repeated cartesian indices is implicitly understood). The internal strain, however, is lattice-periodical (l -independent) and equivalent to a zone-center optic mode: therefore its contribution to macroscopic stress is directly obtained from a suitable force calculation.¹⁵ In order to simplify the algebra, we deal in these notes with the case of a primitive lattice only, where the effect of macroscopic strain is a simple rigid (homothetical) scaling: the displaced ionic positions are

$$R_{l,\alpha} \rightarrow R'_{l,\alpha} = R_{l,\alpha} + R_{l,\beta} \varepsilon_{\alpha\beta}. \quad (6.24)$$

We thus formally get the strain derivative of the total energy from Eq. (6.22) as:

$$T_{\alpha\beta} = \frac{\partial E_{ext}}{\partial \varepsilon_{\alpha\beta}} - \int_{\mathcal{V}} d\mathbf{r} n(\mathbf{r}) \sum_l R_{l,\beta} v'_\alpha(\mathbf{r} - \mathbf{R}_l), \quad (6.25)$$

This expression is correct for any given finite system, but it is boundary-sensitive, and therefore unusable as such for a macroscopic system in the thermodynamic limit: we transform it into an equivalent expression, whose thermodynamic limit can be safely performed. To this aim we exploit the commutator

$$[H_{KS}, r_\beta \frac{\partial}{\partial r_\alpha}] = p_\alpha p_\beta - r_\beta \frac{\partial}{\partial r_\alpha} (V_{ext} + V_H + V_{xc}), \quad (6.26)$$

which has a vanishing mean value over the orbitals. Taking then the sum over the occupied states we get the identity:

$$- \sum_i f_i \langle \varphi_i | p_\alpha p_\beta | \varphi_i \rangle + \int d\mathbf{r} n(\mathbf{r}) r_\beta \frac{\partial}{\partial r_\alpha} (V_{ext} + V_H + V_{xc}) = 0; \quad (6.27)$$

the Hartree term is conveniently evaluated in the symmetric form:

$$\int_{\mathcal{V}} d\mathbf{r} n(\mathbf{r}) r_\beta \frac{\partial}{\partial r_\alpha} V_H = -\frac{e^2}{2} \int_{\mathcal{V}} d\mathbf{r} d\mathbf{r}' n(\mathbf{r}) n(\mathbf{r}') \frac{(r_\alpha - r'_\alpha)(r_\beta - r'_\beta)}{|\mathbf{r} - \mathbf{r}'|^3}, \quad (6.28)$$

while we define the XC stress as

$$T_{\alpha\beta}^{(xc)} = \int_{\mathcal{V}} d\mathbf{r} n(\mathbf{r}) r_\beta \frac{\partial}{\partial r_\alpha} V_{xc}. \quad (6.29)$$

At the LDA level, the XC stress is easily evaluated: in fact, using $V_{xc}(\mathbf{r}) = \mu_{xc}(n(\mathbf{r}))$ this term can be integrated by parts over a volume enclosing the finite system, to yield:²¹

$$T_{\alpha\beta}^{(xc)} = \delta_{\alpha\beta} \int_{\mathcal{V}} d\mathbf{r} n(\mathbf{r}) [\epsilon_{xc}(n(\mathbf{r})) - \mu_{xc}(n(\mathbf{r}))], \quad (6.30)$$

and is an effective isotropic pressure (owing to the $\delta_{\alpha\beta}$). We notice that beyond LDA, and for a low-symmetry solid, the XC stress is anisotropic.²⁰

Using the previous expressions, and substituting $R_{l,\beta} = r_\beta - (\mathbf{r} - \mathbf{R}_l)_\beta$, Eq. (6.25) is recast as:

$$T_{\alpha\beta} = \frac{\partial E_{ext}}{\partial \varepsilon_{\alpha\beta}} + T_{\alpha\beta}^{(xc)} + \int_{\mathcal{V}} d\mathbf{r} n(\mathbf{r}) \sum_l (\mathbf{r} - \mathbf{R}_l)_\beta v'_\alpha(\mathbf{r} - \mathbf{R}_l) - \sum_i f_i \langle \varphi_i | p_\alpha p_\beta | \varphi_i \rangle - \frac{e^2}{2} \int_{\mathcal{V}} d\mathbf{r} d\mathbf{r}' n(\mathbf{r}) n(\mathbf{r}') \frac{(r_\alpha - r'_\alpha)(r_\beta - r'_\beta)}{|\mathbf{r} - \mathbf{r}'|^3}, \quad (6.31)$$

where now all the terms are manifestly boundary-insensitive and extensive in form. The integrals are cast in terms of *relative* coordinates of the system

particles (electron and ions), and therefore have a well defined thermodynamic limit for any extended system (either crystalline or disordered). The average stress is defined as $\bar{\sigma}_{\alpha\beta} = T_{\alpha\beta}/\mathcal{V}$.

Another equivalent form of Eq. (6.31) can be useful. Let us separate within the pseudopotential the pure Coulomb term from the short-range repulsive term as in Eq. (5.6): then we may collect in Eq. (6.31) all the pure Coulomb terms (ion-ion, ion-electron, and Hartree) as the integrated Maxwell stress,⁶ i.e.

$$T_{\alpha\beta}^{(M)} = \frac{1}{4\pi} \int_{\mathcal{V}} d\mathbf{r} [E_{\alpha}(\mathbf{r})E_{\beta}(\mathbf{r}) - \frac{1}{2}\delta_{\alpha\beta}E^2(\mathbf{r})] = \quad (6.32)$$

$$= -\frac{1}{2} \int_{\mathcal{V}} d\mathbf{r}d\mathbf{r}' \rho(\mathbf{r})\rho(\mathbf{r}') \frac{(r_{\alpha} - r'_{\alpha})(r_{\beta} - r'_{\beta})}{|\mathbf{r} - \mathbf{r}'|^3}, \quad (6.33)$$

where $\mathbf{E}(\mathbf{r})$ is the *microscopic* electric field and $\rho(\mathbf{r})$ is the total (electronic and ionic) charge density, as in Eq. (6.10). An useful form of Eq. (6.31) is therefore

$$T_{\alpha\beta} = T_{\alpha\beta}^{(xc)} + T_{\alpha\beta}^{(M)} - \sum_i f_i \langle \varphi_i | p_{\alpha} p_{\beta} | \varphi_i \rangle + \int_{\mathcal{V}} d\mathbf{r} n(\mathbf{r}) \sum_l (\mathbf{r} - \mathbf{R}_l)_{\beta} v_{\alpha}^{(sr)'}(\mathbf{r} - \mathbf{R}_l), \quad (6.34)$$

where now all the terms have a transparent physical meaning. The third one is the kinetic stress, while the fourth is a “pseudopotential stress”, due to core orthogonalization repulsion.

For a crystalline solid the density $n(\mathbf{r})$ is a periodic function, and the average stress

$$\bar{\sigma}_{\alpha\beta} = \frac{1}{N_c\Omega} \frac{\partial E_{tot}}{\partial \varepsilon_{\alpha\beta}} = \frac{1}{\Omega} \frac{\partial \gamma_{Ewald}}{\partial \varepsilon_{\alpha\beta}} + \frac{1}{N_c\Omega} \frac{\partial E_{el}}{\partial \varepsilon_{\alpha\beta}} \quad (6.35)$$

can then be evaluated upon performing the integrals in Eq. (6.31) over a single periodicity cell. When a plane-wave expansion is adopted, these integrals are easily cast as reciprocal-lattice sums, whose expressions are reported in Refs. 15,16. As an example, I only give the expression for the Maxwell average stress in terms of the Fourier coefficients of the microscopic electric field:

$$\bar{\sigma}_{\alpha\beta}^{(M)} = \frac{1}{4\pi} \sum_{\mathbf{G}} [E_{\mathbf{G},\alpha} E_{\mathbf{G},\beta} - \frac{1}{2}\delta_{\alpha\beta} |\mathbf{E}_{\mathbf{G}}|^2], \quad (6.36)$$

and the $\mathbf{G}=0$ term vanishes whenever no macroscopic field is present. In case nonlocal ionic pseudopotentials are used, an extra term must be considered in the commutator of Eq. (6.26).

The above proof of the stress theorem is a typical example of a rather general strategy to be used in evaluating macroscopic properties of an extended system in the thermodynamic limit. One starts from a large but finite system and writes a formal expression for the quantity of interest (in the previous example the macroscopic stress tensor); this expression, being boundary-sensitive, cannot be used as such. A suitable transformation leads then to a boundary-insensitive

form, where only *relative* coordinates of the system particles appear. Only at this point the thermodynamic limit is taken and lattice periodicity is exploited. Similar strategies have been used in other theorems concerning macroscopic physical quantities. Early examples concern deformation potentials²²⁻²⁴ and elastic constants;²⁵ other important cases are piezoelectricity,^{26,18} and macroscopic polarization. The latter case is discussed in the next Chapter.

6.6 Elastic constants

The search for structural equilibrium in a given solid consists traditionally in performing several total energy calculations at different values of the lattice constants and—for low-symmetry solids—at different ionic positions in the ionic cell. Starting from the values thus obtained at selected points in configuration space, one interpolates with a continuous function which provides at a time the structural minimum of E_{tot} , as well as some of its second derivatives, *i.e.* the macroscopic elastic constants.

6.7 References

- [6.1] D.M. Ceperley and B.J. Alder, *Phys. Rev. Lett.* **45**, 566 (1980); J. Perdew and A. Zunger, *Phys. Rev. B* **23**, 5048 (1981).
- [6.2] W.E. Pickett, *Computer Phys. Reports* **9**, 115 (1989).
- [6.3] L. Kleinman, *Phys. Rev. B* **24**, 7412 (1981).
- [6.4] C. Pisani, R. Dovesi, and C. Roetti, *Hartree-Fock Ab-Initio Treatment of Crystalline Systems* (Springer, Berlin, 1988).
- [6.5] J. Ihm, A. Zunger, and M.L. Cohen, *J. Phys. C* **12**, 4409 (1979); **13**, 3095(E) (1980).
- [6.6] J.D. Jackson, *Classical Electrodynamics* (Wiley, New York, 1975).
- [6.7] J.R. Chelikowsky and S.G. Louie, *Phys. Rev. B* **29**, 3470 (1984).
- [6.8] R. Car and M. Parrinello, *Phys. Rev. Lett.* **55**, 2471 (1985).
- [6.9] I. Stich, R. Car, M. Parrinello, and S. Baroni, *Phys. Rev. B* **39**, 4997 (1989).
- [6.10] P. Ehrenfest, *Z. Phys.* **45**, 455 (1927); R.P. Feynman, *Phys. Rev.* **56**, 340 (1939).
- [6.11] B.M. Deb, *Rev. Mod. Phys.* **43**, 22 (1973).

- [6.12] R. Moccia, Chem. Phys. Lett. **5**, 260 (1970).
- [6.13] P. Pulay, in: *Applications of Electronic Structure Theory*, edited by H.F. Schaefer (Plenum, New York, 1977), p.153.
- [6.14] Rici Yu, D. Singh, and H. Krakauer, Phys. Rev. B **43**, 6411 (1991).
- [6.15] O.H. Nielsen and R.M. Martin, Phys. Rev. Lett. **50**, 697 (1983).
- [6.16] O.H. Nielsen and R.M. Martin, Phys. Rev. B **32**, 3780 (1985).
- [6.17] O.H. Nielsen and R.M. Martin, Phys. Rev. B **32**, 3792 (1985).
- [6.18] S. de Gironcoli, S. Baroni, and R. Resta, Phys. Rev. Lett. **62**, 2853 (1989).
- [6.19] R.J. Needs and R.M. Martin, Phys. Rev. B **30**, 5390 (1984).
- [6.20] A. Dal Corso and R. Resta, Phys. Rev. B **50**, 4327 (1994).
- [6.21] J.F. Janak, Phys. Rev. B **9**, 3985 (1974)
- [6.22] E.I. Blount, Phys. Rev. **114**, 418 (1959).
- [6.23] F.S. Khan and P.B. Allen, Phys. Rev. B **29**, 3341 (1984).
- [6.24] R. Resta, Phys. Rev. B **44**, 11035 (1991).
- [6.25] C. Herring, unpublished; reported by M. Lax in *Proceedings of 1963 International Conference on Lattice Dynamics, Copenhagen* (Pergamon, London, 1965), p. 583.
- [6.26] R.M. Martin, Phys. Rev. B **5**, 1607 (1972).

Chapter 9

Lattice dynamics (preliminary draft)

Commento: cominciare con una adiabatic dynamic piu' generale. Affrontare il problema partendo dalla velocity-velocity correlation function, etc.. solo dopo passare all'approx. armonica.

9.1 Definitions

Let us consider a solid at zero temperature in the adiabatic approximation: the ionic positions at equilibrium are $\mathbf{R}_{ls} = \mathbf{R}_l + \mathbf{R}_s$, where l is a cell index and s is (in nonprimitive lattices) a basis index in the unit cell of n atoms. A distorted configuration of the solid is described by the set of the ionic displacements $\{\mathbf{u}_{ls}\}$; greek subscript will be used for cartesian components.

The total energy is expanded in the displacements around equilibrium as

$$E_{tot}(\{\mathbf{u}_{ls}\}) = E_{tot}^{(0)} + E_{tot}^{(2)}(\{\mathbf{u}_{ls}\}) + \dots \quad (9.1)$$

and the harmonic force constant are second derivatives, defined through

$$E_{tot}^{(2)}(\{\mathbf{u}_{ls}\}) = \frac{1}{2} \sum_{l'ss',\alpha\beta} c_{ss',\alpha\beta}(l,l') u_{ls,\alpha} u_{l's',\beta}. \quad (9.2)$$

Because of lattice periodicity, the force constants depend only on $\mathbf{R}_{l'} - \mathbf{R}_l$ and hence $c_{ss',\alpha\beta}(0,l')$ contain all of the information. Besides this, there are further important constraints imposed by translational and rotational invariance:¹ we consider explicitly only the former. If the solid is translated as a whole, the energy is unchanged, and hence the energy expansion vanishes to all orders; this is easily shown to imply

$$c_{ss,\alpha\beta}(0,0) = -\sum_{l's'}' c_{ss',\alpha\beta}(0,l'), \quad (9.3)$$

with the usual meaning of the primed sum.

Owing to the adiabatic approximation,^{2,3} the ionic motion is classical in the potential energy provided by E_{tot} . The harmonic oscillations around equilibrium are then governed by the equation of motion:

$$M_s \ddot{u}_{ls,\alpha} = - \sum_{l',\beta} c_{ss',\alpha\beta}(l,l') u_{l's',\beta}, \quad (9.4)$$

for which normal-mode solutions (within periodic boundary conditions) are phonons of \mathbf{q} wavevector:

$$u_{ls,\alpha} = u_{s,\alpha}(\mathbf{q}) e^{i\mathbf{q}\cdot(\mathbf{R}_l + \mathbf{R}_s)} e^{i\omega(\mathbf{q})t}. \quad (9.5)$$

Owing to the transformation, the secular problem factorizes for different \mathbf{q} 's (chosen by convenience within the first Brillouin zone). Using then the reciprocal form of the force constants:

$$c_{ss',\alpha\beta}(\mathbf{q}) = e^{i\mathbf{q}\cdot(\mathbf{R}_{s'} - \mathbf{R}_s)} \sum_{l'} c_{ss',\alpha\beta}(0,l') e^{i\mathbf{q}\cdot\mathbf{R}_{l'}}, \quad (9.6)$$

and introducing the auxiliary quantities:

$$u_{s,\alpha}(\mathbf{q}) = \frac{1}{\sqrt{M_s}} e_{s,\alpha}(\mathbf{q}); \quad (9.7)$$

$$D_{ss',\alpha\beta}(\mathbf{q}) = (M_s M_{s'})^{-\frac{1}{2}} c_{ss',\alpha\beta}(\mathbf{q}), \quad (9.8)$$

where M_s are the ionic masses, the secular problem becomes:

$$\omega^2(\mathbf{q}) e_{s,\alpha}(\mathbf{q}) = \sum_{s',\beta} D_{ss',\alpha\beta}(\mathbf{q}) e_{s',\beta}(\mathbf{q}). \quad (9.9)$$

The dynamical matrix is hermitian, and has (in stable systems) nonnegative eigenvalues: at every \mathbf{q} there are $3n$ normal modes, whose frequencies are the square roots of the eigenvalues; the eigendisplacements are related to the eigenvectors by a trivial mass-dependent factor.

At this point I stress the importance—both in model and first-principle lattice dynamics—of a standard technical trick. It is convenient not to bother explicitly with the actual value of the on-site force constant $c_{ss,\alpha\beta}(0,0)$. To this effect one uses an auxiliary set of force constants, indicated with an overbar, which in general violate translational invariance and are therefore nonphysical. The physical force constants are recovered from

$$c_{ss',\alpha\beta}(\mathbf{q}) = \bar{c}_{ss',\alpha\beta}(\mathbf{q}) - \delta_{ss'} \sum_{s''} \bar{c}_{ss'',\alpha\beta}(0), \quad (9.10)$$

which ensures translational invariance in form. It is easy to verify that $\bar{c}_{ss,\alpha\beta}(0)$ cancels in this expression, and hence the value of the on-site force constant $c_{ss,\alpha\beta}(0,0)$ is irrelevant, whenever Eq. (9.10) is explicitly used.

9.2 Central two-body forces

We discuss in this section the contribution to lattice dynamics of central two-body interactions. In the first-principle context, such treatment only concerns the bare ion-ion Coulomb repulsion, being purely classical: the energy of this term is indicated as E_{ext} in Chapter 1. However, for pedagogical purposes, we will previously discuss simpler treatments, like pseudopotential perturbation theory (in simple metals), and the rigid-ion model (in alkali halides). In both these cases, the whole interactions involved are indeed central pairwise.

Suppose that the central two-body potential between ions of species s and s' is $\Phi_{ss'}(r)$; then its contribution to the lattice-dynamical force constants is

$$c_{ss',\alpha\beta}(l,l') = - \left. \frac{\partial^2 \Phi_{ss'}(|\mathbf{r}|)}{\partial r_\alpha \partial r_\beta} \right|_{\mathbf{r}=\mathbf{R}_{l'}+\mathbf{R}_{s'}-\mathbf{R}_l-\mathbf{R}_s} \quad (9.11)$$

Explicit evaluation of the second derivatives of a radial function yields

$$\frac{\partial^2 \Phi_{ss'}(|\mathbf{r}|)}{\partial r_\alpha \partial r_\beta} = \frac{1}{r} \Phi'_{ss'}(\mathbf{r}) (\delta_{\alpha,\beta} - \frac{r_\alpha r_\beta}{r^2}) + \Phi''_{ss'}(\mathbf{r}) \frac{r_\alpha r_\beta}{r^2}, \quad (9.12)$$

and we only need to calculate $\Phi'(\mathbf{r})/r$ and $\Phi''(\mathbf{r})$ for each shell of neighbors. These quantities are often called “tangential” and “radial” force constants.

The above expressions only apply to $l \neq l'$ and $s \neq s'$; using an *arbitrary* value for the on-site force constant $c_{ss,\alpha\beta}(0,0)$, we get the analogous of Eq. (9.6) as

$$\bar{c}_{ss',\alpha\beta}(\mathbf{q}) = e^{i\mathbf{q}\cdot(\mathbf{R}_{s'}-\mathbf{R}_s)} \sum_{l'} c_{ss',\alpha\beta}(0,l') e^{i\mathbf{q}\cdot\mathbf{R}_{l'}}, \quad (9.13)$$

and the physical force constants are obtained from Eq. (9.10). The expression in Eq. (9.13) is usable as such only for short-range interactions, where the sum over l' is rapidly convergent; the resulting force constants are then analytic functions of \mathbf{q} .

It is now convenient to transform the force constants into an equivalent expression involving the Fourier transform $\Phi_{ss'}(k)$ of the pairwise interaction. Standard manipulations, starting from Eq. (9.13), give

$$\bar{c}_{ss',\alpha\beta}(\mathbf{q}) = \frac{1}{\Omega} \sum_{\mathbf{G}} (\mathbf{q} + \mathbf{G})_\alpha (\mathbf{q} + \mathbf{G})_\beta \Phi_{ss'}(|\mathbf{q} + \mathbf{G}|) e^{i\mathbf{G}\cdot(\mathbf{R}_{s'}-\mathbf{R}_s)}. \quad (9.14)$$

Specializing to a pure Coulomb interaction between point charges, its reciprocal space expression is $\Phi_{ss'}(k) = 4\pi Q_s Q_{s'}/k^2$. The force constants are nonanalytic, as a fingerprint of long-range Coulomb interactions: we cast the force constants repulsion in the form:

$$\bar{c}_{ss',\alpha\beta}^{(Coul)}(\mathbf{q}) = \frac{4\pi Q_s Q_{s'}}{\Omega} \left[\frac{q_\alpha q_\beta}{q^2} + \sum_{\mathbf{G} \neq 0} \frac{(\mathbf{q} + \mathbf{G})_\alpha (\mathbf{q} + \mathbf{G})_\beta}{|\mathbf{q} + \mathbf{G}|^2} e^{i\mathbf{G}\cdot(\mathbf{R}_{s'}-\mathbf{R}_s)} \right], \quad (9.15)$$

where the nonanalytic term has been expressly separated. This expression is valid at $\mathbf{q} \neq 0$: a major problem is the fact that $\bar{c}_{ss',\alpha\beta}(\mathbf{q} = 0)$ does *not* cancel (for $s \neq s'$) in Eq. (9.10). What is the correct value to be used? There are several ways to reach a rigorous answer: the final result can however be summarized into a working recipe, first proposed by Pines in the fifties. Under the hypothesis that the extended system is *overall* neutral, one gets always correct results upon assumption that the Fourier transform of Coulomb interaction is $4\pi/k^2$ at $k \neq 0$ and vanishes at $k = 0$. Therefore the first term in square brackets must be taken as zero when \mathbf{q} is exactly zero. When $\mathbf{q} \rightarrow 0$, *i.e.* in the long-wavelength limit, this term gives in general a finite contribution. In fact $q_\alpha q_\beta / q^2$ acts as a projector on the \mathbf{q} -direction, thus discriminating between longitudinal and transverse modes. Basically, this is the algebra responsible for LO-TO splitting of zone-center frequencies in polar crystals.

Still, a minor problem remains: the reciprocal space sum over \mathbf{G} in Eq. (9.15) does not converge. The solution is straightforward: one evaluates the Coulomb term giving a small gaussian spread to the ionic charges, in order to ensure convergence. When the spread is much smaller than the typical interionic distances, the force constants assume their physical (spread-independent) value. Alternatively, more efficient Ewald techniques can be used.¹ There is one simple case where the sum can be evaluated in closed form: let us consider $\bar{c}_{12,\alpha\beta}^{(Coul)}(0)$ in a cubic binary crystal. Because of symmetry, this must be diagonal in $\alpha\beta$ and equal to one third of the trace:

$$\bar{c}_{12,\alpha\beta}^{(Coul)}(0) = \delta_{\alpha\beta} \frac{4\pi Q_1 Q_2}{3\Omega} \sum_{\mathbf{G} \neq 0} e^{i\mathbf{G} \cdot (\mathbf{R}_2 - \mathbf{R}_1)}. \quad (9.16)$$

Since we know that such a sum over all \mathbf{G} 's (including $\mathbf{G} = 0$) vanishes, we get the result

$$\bar{c}_{12,\alpha\beta}^{(Coul)}(0) = -\delta_{\alpha\beta} \frac{4\pi Q_1 Q_2}{3\Omega}. \quad (9.17)$$

9.3 Metals

Accurate phonon spectra in simple metals can be obtained from pseudopotential perturbation theory to second order. Within this framework, the total energy is the sum of two terms: the former depends only upon the average density, while the second is in fact a pairwise central interaction. Phonon modes do not affect the average density, and therefore the full dynamical matrix is straightforwardly obtained from the formalism of the previous section, where the two body interaction is given by Eqs. (5.15) and (5.25). For a primitive lattice we omit the s subscripts; using Eq. (9.10) we get the dynamical matrix in the form:

$$D_{\alpha\beta}(\mathbf{q}) = \frac{1}{M\Omega} \sum_{\mathbf{G}} [(\mathbf{q} + \mathbf{G})_\alpha (\mathbf{q} + \mathbf{G})_\beta \Phi(|\mathbf{q} + \mathbf{G}|) - G_\alpha G_\beta \Phi(G)]. \quad (9.18)$$

This form for the dynamical matrix of a simple metal was first found by Toya,⁴ with a somewhat different derivation than proposed here. A typical phonon spectrum obtained from the diagonalization of Eq. (9.18) was already displayed previously, Fig. 5.2.

I illustrate now an important qualitative feature. As explained in the previous section, the Coulomb interaction can be assumed to vanish at $k=0$. Therefore we write Eq. (9.18) as:

$$D_{\alpha\beta}(\mathbf{q}) = \frac{q_\alpha q_\beta}{M\Omega} \Phi(q) + \frac{1}{M\Omega} \sum'_{\mathbf{G} \neq 0} [(\mathbf{q} + \mathbf{G})_\alpha (\mathbf{q} + \mathbf{G})_\beta \Phi(|\mathbf{q} + \mathbf{G}|) - G_\alpha G_\beta \Phi(G)]. \quad (9.19)$$

The sum in the second term obviously vanishes at small \mathbf{q} ; as for the first ($\mathbf{G}=0$) term in Eq. (9.19), it vanishes as well, but such behavior is less obvious and *crucially* depends on the metallic nature of the screening electronic system. To leading order Eqs. (5.15) and (5.25) imply:

$$\Phi(q) \simeq \frac{4\pi Z^2 e^2}{\varepsilon(q) q^2}, \quad (9.20)$$

where Z is the bare charge of the ion. Suppose we *neglect* screening, assuming therefore $\varepsilon(q)=1$. Then we get

$$D_{\alpha\beta}(\mathbf{q}) \simeq \frac{4\pi Z^2 e^2}{M\Omega} \frac{q_\alpha q_\beta}{q^2}; \quad (9.21)$$

the zone-center normal modes of the system are obtained from diagonalization of such dynamical matrix. One immediately gets two transverse modes of vanishing frequency and a longitudinal mode of *finite* frequency, which therefore is *not* an acoustic phonon. Its frequency is given by:

$$\omega_L^2(0) = \frac{4\pi Z^2 e^2}{M\Omega}, \quad (9.22)$$

and coincides in fact with the plasma frequency ω_p of a system of classical charges in a neutralizing rigid (it i.e. nonpolarizable) background. Such system is a paradigmatic one in the theory of classical liquids, being known as the one-component plasma.⁵ However, such model system has nothing to do with real metals, where in fact long-wavelength longitudinal modes *do* have vanishing frequency.

Within the present formulation, it can be clearly seen that the key feature which accounts for the existence of acoustic phonons is the q^{-2} divergence in the dielectric function, *i.e.* its metallic character. In the following discussion I will neglect the second term in Eq. (9.19), which amounts to neglecting the discrete nature of the lattice and to dealing with the long-wavelength modes in a macroscopic way. We thus obtain for the longitudinal mode:

$$\omega_L^2(q) \simeq \frac{\omega_p^2}{\varepsilon(q)}, \quad (9.23)$$

which is correctly quadratic in q . The coefficient is nothing else than the (squared) speed of sound: we evaluate it using the TF dielectric function, Eq. (2.17), which is exact to leading order in q . The speed is then given by $\sqrt{\omega_p}/k_{TF}$, thus recovering an important result found in 1950 upon macroscopic arguments by Bohm and Staver and derived in several textbooks.⁶ The important message of all this can be summarized saying that acoustic propagation in metals is allowed by electronic screening; even a rough account of it yields a speed of sound in semiquantitative agreement with the experiment.

The electron-gas dielectric function $\varepsilon(k)$ has a logarithmic singularity for $k=2k_F$: therefore the dynamical matrix of Eq. (9.18) is singular at the \mathbf{q} points (in the first Brillouin zone) which fulfill $|\mathbf{q} + \mathbf{G}| = 2k_F$ (for some \mathbf{G}). The surfaces defined by this relationship can be found from a simple geometrical construction: when \mathbf{q} crosses one of these surfaces, one expects a “kink” in the experimental phonon dispersion curves, as first proposed by W. Kohn.⁷ Shortly after the theory, the effect was observed in the spectra of lead,⁸ and goes under the name of “Kohn anomaly” since then.

9.4 Ionic crystals

The lattice dynamics of ionic crystals has important qualitative features which are better illustrated—for pedagogical purposes—starting from the simple rigid-ion model, which was first applied to the lattice dynamics of alkali halides in 1940.⁹ The basic assumption is that lattice dynamics is governed by two different kinds of forces: (i) a pure Coulomb force acting between ions, schematized as rigid point charges; and (ii) a short range pairwise interaction, which mimics the Pauli repulsion. Such a model was historically introduced by Born and Mayer to explain the lattice stability of ionic crystals (mettere formuletta & citare Ashcroft, Kittel??).

The dynamical matrix is found summing two contributions, which are separately evaluated; we specialize to a cubic binary system in the following. The Coulomb term is given by Eq. (9.15), using for the ionic charges the values $Q_s = (-1)^s Q$, where $s=1$ (2) labels the anion (cation); in the original version of the model $Q=|e|$ (*i.e.* full ionicity) is assumed for alkali halides. The short range term is evaluated directly in the form of Eq. (9.13), where only a small number of terms is summed: within the simplest version the sum includes indeed only nearest neighbors (six terms for the rocksalt structure). Therefore the short-range $\bar{c}_{ss',\alpha\beta}^{(s,r.)}(\mathbf{q})$ vanishes when $s=s'$, while $\bar{c}_{12,\alpha\beta}^{(s,r.)}(\mathbf{q})$ depends on two parameters only, given in Eq. (9.12). The two parameters of the rigid-ion model are fitted to a pair of empirical data (typically the equilibrium lattice constant and the bulk modulus) Explicit diagonalization of the dynamical matrix provides the phonon spectrum. A typical result is shown in Fig.....The overall agreement.....except optic modes

It proves useful to study in detail the diagonalization at the zone center.

Keeping only terms of order zero in \mathbf{q} , and using Eq. (9.10), we get:

$$c_{11,\alpha\beta}^{(s.r.)}(\mathbf{q}) \simeq c_{22,\alpha\beta}^{(s.r.)}(\mathbf{q}) \simeq -\bar{c}_{12,\alpha\beta}^{(s.r.)}(0) ; c_{12,\alpha\beta}^{(s.r.)}(\mathbf{q}) \simeq \bar{c}_{12,\alpha\beta}^{(s.r.)}(0); \quad (9.24)$$

furthermore bulk cartesian tensors are diagonal in a cubic material, hence $\bar{c}_{12,\alpha\beta}^{(s.r.)}(0) = -R_0\delta_{\alpha\beta}$, where the constant R_0 completely accounts for the short range forces at the zone center. Using the same path for the Coulomb term, Eq. (9.15), we find a nonanalytic term (homogeneous of degree zero in \mathbf{q}), plus an analytic term which behaves in all respects like an effectively short-range additional interaction. From Eqs. (9.15) and (9.17) we cast the complete dynamical matrix at the zone center as

$$D_{ss',\alpha\beta}(\mathbf{q}) \simeq \left[\frac{4\pi Q^2}{\Omega} \frac{q_\alpha q_\beta}{q^2} + \left(R_0 - \frac{4\pi Q^2}{3\Omega} \right) \delta_{\alpha\beta} \right] \mathcal{M}_{ss'}, \quad (9.25)$$

where we define the 2×2 matrix \mathcal{M} as :

$$\mathcal{M} = \begin{pmatrix} \frac{1}{M_1} & -\frac{1}{\sqrt{M_1 M_2}} \\ -\frac{1}{\sqrt{M_1 M_2}} & \frac{1}{M_2} \end{pmatrix}. \quad (9.26)$$

One of the eigenvalues of \mathcal{M} is zero, and the other is $1/m = 1/M_1 + 1/M_2$, *i.e.* the inverse reduced mass; the corresponding eigenvectors are easily recognized to be the acoustic and optic modes, respectively. Going back to the zone-center dynamical matrix, the acoustic mode is threefold degenerate with frequency zero, while the optic modes require further diagonalization over the Cartesian coordinates. There are two transverse modes—where the polarization is perpendicular to \mathbf{q} —with

$$\omega_{TO}^2 = \frac{1}{m} \left(R_0 - \frac{4\pi Q^2}{3\Omega} \right), \quad (9.27)$$

and one longitudinal mode—where the polarization is parallel to \mathbf{q} —whose frequency is

$$\omega_{LO}^2 = \omega_{TO}^2 + \frac{4\pi Q^2}{m\Omega}. \quad (9.28)$$

The longitudinal-transverse (LT) splitting of the zone-center optic modes takes the simple form of a plasma frequency, where the mass density is in fact a density of *reduced* mass. The LT splitting is substantially overestimated with respect to the experiment, and this drawback is simply due to the fact that no mechanism within the rigid-ion model allows for electronic screening of the Coulomb interaction. Models correcting this problem are well known in the literature: they agree better with the experimental spectra, at the price of using more empirical parameters. In the following we focus on the first-principle approach to lattice dynamics.

9.5 Effective charges

We have introduced the Born effective charges as basic phenomenological material constants in cubic binary materials; the concept is easily extended to a general crystal. The effective-charge tensor $Z_{s,\alpha\beta}^*$ measures the α -component of the dipole linearly induced by a unit displacement of an ion of kind s in the β -direction, while all the other ions are kept fixed, *and* the field is kept vanishing. An equivalent definition is via the force acting on ions of kind s in a constant macroscopic field. Of course, if the s site is of cubic symmetry (at equilibrium), then $Z_{s,\alpha\beta}^* = Z_s^* \delta_{\alpha\beta}$.

Mettere in relazione la Eq. (9.28) con la Eq. (2.39). Usare la formulazione come in Axe, con la definizione first-principles. Usare anche (forse) l'ultimo lavoro di Gonze. N.B.: c'è un chapter9.old da saccheggiare.

9.6 Ion-electron-ion interaction

In order to solve the problem of lattice dynamics we look for the force constants arising from a second order expansion of the total energy. Using the notations of Chap. 1, this expansion is written as $E_{tot}^{(2)} = E_{ext}^{(2)} + E_{el}^{(2)}$. The contribution of the first term has been evaluated indeed in the previous section: the result goes under the name of “direct” ion-ion force constants. We remain with the quantum-mechanical term:

$$E_{el}^{(2)} = \int d\mathbf{r} V_{ext}^{(2)}(\mathbf{r}) n^{(0)}(\mathbf{r}) + \frac{1}{2} \int d\mathbf{r} V_{ext}^{(1)}(\mathbf{r}) n^{(1)}(\mathbf{r}), \quad (9.29)$$

responsible for “indirect”, or ion-electron-ion force constants.

At this point we need to express V_{ext} in more detail. For the sake of simplicity all of the formulation will be given in a local pseudopotential framework, where the unperturbed “external” (*i.e.* electron-ion) potential $V_{ext}^{(0)}$ has the form:

$$V_{ext}^{(0)}(\mathbf{r}) = \sum_{ls} v_s(|\mathbf{r} - \mathbf{R}_{ls}|). \quad (9.30)$$

Alternatively, the same formulation can be consistently interpreted in the all-electron sense, and the v 's are then full Coulomb point-charge potentials.

When the lattice distortion $\{\mathbf{u}_{ls}\}$ is present in the crystal, the bare perturbation is:

$$\delta V_{ext}(\mathbf{r}) = \sum_{ls} [v_s(|\mathbf{r} - \mathbf{R}_{ls} - \mathbf{u}_{ls}|) - v_s(|\mathbf{r} - \mathbf{R}_{ls}|)]. \quad (9.31)$$

The first order term in the displacements is;

$$V_{ext}^{(1)}(\mathbf{r}) = - \sum_{ls,\alpha} u_{ls,\alpha} v'_{s,\alpha}(|\mathbf{r} - \mathbf{R}_{ls}|), \quad (9.32)$$

with obvious notations for the gradient of the pseudopotential. The linear density response to such perturbation is

$$n^{(1)}(\mathbf{r}) = - \sum_{ls,\alpha} u_{ls,\alpha} \int d\mathbf{r}' \chi(\mathbf{r}, \mathbf{r}') v'_{s,\alpha}(|\mathbf{r}' - \mathbf{R}_{ls}|), \quad (9.33)$$

such that we get for the second-order electronic energy:

$$E_{el}^{(2)} = \int d\mathbf{r} V_{ext}^{(2)}(\mathbf{r}) n^{(0)}(\mathbf{r}) + \frac{1}{2} \sum_{l'l'ss',\alpha\beta} u_{ls,\alpha} u_{l's',\beta} \int d\mathbf{r} d\mathbf{r}' v'_{s,\alpha}(|\mathbf{r} - \mathbf{R}_{ls}|) \chi(\mathbf{r}, \mathbf{r}') v'_{s',\beta}(|\mathbf{r}' - \mathbf{R}_{l's'}|). \quad (9.34)$$

Instead of explicating the first term as well, we start making the important observation that $E_{el}^{(2)}$ is translationally invariant, while the two terms in Eq. (9.34) individually are not such. This is easily checked by considering a uniform translation, *i.e.* \mathbf{u}_{ls} independent of ls . We further observe that the second derivatives

$$\frac{\partial^2 V_{ext}(\mathbf{r})}{\partial u_{ls,\alpha} \partial u_{l's',\beta}} = \delta_{ll'} \delta_{ss'} v''_{s,\alpha\beta}(|\mathbf{r} - \mathbf{R}_{ls}|) \quad (9.35)$$

give a contribution only to the on-site force constants $c_{ss,\alpha\beta}(l, l)$. Therefore there is no need for the explicit evaluation of the $V_{ext}^{(2)}$ term, provided the translationally invariant form of the force constants, Eq. (9.10), is used. As for the off-site force constants, comparison of Eq. (9.34) with Eq. (9.2) gives:

$$c_{ss',\alpha\beta}^{(el)}(l, l') = \int d\mathbf{r} d\mathbf{r}' v'_{s,\alpha}(|\mathbf{r} - \mathbf{R}_{ls}|) \chi(\mathbf{r}, \mathbf{r}') v'_{s',\beta}(|\mathbf{r}' - \mathbf{R}_{l's'}|). \quad (9.36)$$

After straightforward manipulation, the ion-electron-ion indirect term in the reciprocal-space force constants is written as:

$$\begin{aligned} \bar{c}(el)_{ss',\alpha\beta}(\mathbf{q}) &= \frac{1}{\Omega} \sum_{\mathbf{G}, \mathbf{G}'} (\mathbf{q} + \mathbf{G})_{\alpha} (\mathbf{q} + \mathbf{G}')_{\beta} e^{i(\mathbf{G}' \cdot \mathbf{R}_{s'} - \mathbf{G} \cdot \mathbf{R}_s)} \times \\ &\times v_s(|\mathbf{q} + \mathbf{G}|) \chi(\mathbf{q} + \mathbf{G}, \mathbf{q} + \mathbf{G}') v_{s'}(|\mathbf{q} + \mathbf{G}'|). \end{aligned} \quad (9.37)$$

9.7 Ab-initio force constants

We now collect together the two terms in the force constants: the direct ion-ion interaction, Eq. (9.15), and the ion-electron-ion interaction, Eq. (9.37). The final expression to be used in *ab-initio* lattice dynamics is then:

$$\bar{c}_{ss',\alpha\beta}(\mathbf{q}) = \frac{1}{\Omega} \sum_{\mathbf{G}, \mathbf{G}'} (\mathbf{q} + \mathbf{G})_{\alpha} (\mathbf{q} + \mathbf{G}')_{\beta} e^{i(\mathbf{G}' \cdot \mathbf{R}_{s'} - \mathbf{G} \cdot \mathbf{R}_s)} \times \quad (9.38)$$

$$\times \left[\frac{4\pi e^2 Z_s Z_{s'}}{|\mathbf{q} + \mathbf{G}||\mathbf{q} + \mathbf{G}'|} \delta_{\mathbf{G}\mathbf{G}'} + v_s(|\mathbf{q} + \mathbf{G}|) \chi(\mathbf{q} + \mathbf{G}, \mathbf{q} + \mathbf{G}') v_{s'}(|\mathbf{q} + \mathbf{G}'|) \right].$$

An alternative (and equivalent) expression is easily written in term of the inverse dielectric matrix ϵ^{-1} . This is a celebrated result due to Pick, Cohen and Martin:¹⁰ it is an exact formulation of harmonic lattice dynamics in the adiabatic approximation, in terms of the linear response of the electronic system.

The role of the offdiagonal χ elements and their analytic properties in the neighborhood of $\mathbf{q} = 0$ is essential. Here I only mention that a diagonal approximation to the χ matrix implies in insulators (or semiconductors) a strongly unphysical result: the acoustic modes have nonvanishing frequency at the zone center. The χ matrix elements fulfill in fact a relationship, known as the “acoustic sum rule”, which is violated by any diagonal approximation in insulators. The theory also provides an exact microscopic definition of the phenomenological Born effective charges in polar insulating crystals.

In the case of metals, the diagonal screening approximation does not show such a severe shortcoming. Incidentally, I point out that a diagonal (and isotropic) screening approximation within Eq. (9.38) gives a set of force constants identical in form to Eq. (9.14), *i.e.* equivalent to a system with central pairwise interactions.

The PCM theory, as formulated above, cannot be used in a modern pseudopotential context.¹¹ In fact the bare perturbations generated by ionic displacements are in this case nonlocal, and the standard dielectric operators (χ , ϵ^{-1}) are therefore unable to cope with lattice dynamics. If we limit ourselves to LDA as usual, then the BGT approach¹² solves the problem. The harmonic force constants can be directly evaluated, at an arbitrary \mathbf{q} -vector, even within a nonlocal pseudopotential framework. The BGT expression reduces in principle to Eq. (9.38) in the simple case of local ionic potentials, but it is worth noticing that the full χ operator is *not* explicitly evaluated within BGT, thus substantially reducing the computational effort.

Espressione fenomenologica-like. Furbata della traccia. LST generalizzato.

9.8 Outline of the results

The basic expression for the force constants, Eq. (9.38) applies to any periodic solid (metal, semiconductor, insulator). Harmonic lattice dynamics in metals within LRT has a long history, much elder than the milestone PCM paper. Reviews of important results can be found *e.g.* in Refs. 13,3. As for insulators, I am unaware of any LRT result at the first-principle level: to the best of my knowledge, only standard “frozen-phonon” calculations are available.

Harmonic lattice dynamics in semiconductors, within the PCM formulation of Eq. (9.38) and using local pseudopotentials, has been studied by some authors: an account of the most interesting results can be found in Ref. 14. State-of-the-art calculations, in a nonlocal pseudopotential framework, have been possible

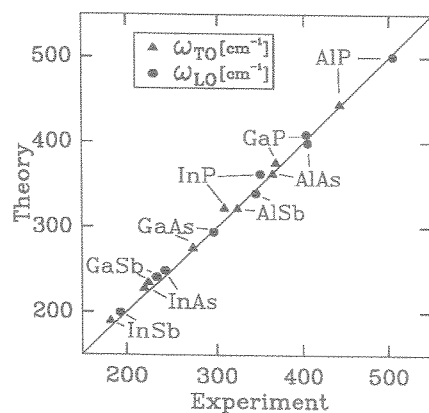


Figure 1.1. Comparison between the calculated and experimental zone-center phonon frequencies in *III – V* semiconductors.

only after BGT. Calculations for a few materials at the zone center were first reported in the original BGT paper.¹² A systematic calculation for all the *III – V* semiconductors is reported in Ref. 15: the results are shown here in Fig. 1.1. More complicated materials, like superlattices and alloys, have been studied within the same approach.^{16,17} Very recently, the method has been implemented for calculating the force constants at arbitrary \mathbf{q} -vectors in the Brillouin zone, with excellent results.¹⁸

Aggiornare le referenze, e inserire disegno con gli spettri.

9.9 References

- [9.1] A.A. Maradudin, E.W. Montroll, G.H. Weiss, and I.P. Ipatova, *Theory of Lattice Dynamics in the Harmonic Approximation*, Solid State Physics, Suppl. **3** (Academic, New York, 1971).
- [9.2] G.V. Chester, *Adv. Phys.* **10**, 357 (1961).
- [9.3] E.G. Brovman and Yu.M. Kagan, in: *Dynamical Properties of Solids*, edited by G.K. Horton and A.A. Maradudin, vol. 1 (North-Holland, Amsterdam, 1974), p. 191.
- [9.4] T. Toya, *J. Res. Inst. Catalysis, Hokkaido Univ.* **6**, 161 and 183 (1958).
- [9.5] J.P. Hansen and I.R. McDonald, *Theory of Simple Liquids* (Academic, New York, 1986).

- [9.6] N.W. Ashcroft and N.D. Mermin, *Solid State Physics* (Saunders, Philadelphia, 1976).
- [9.7] W. Kohn, Phys. Rev. Lett. **2**, 393 (1959).
- [9.8] Brockhouse et al, Phys. Rev. Lett. **7**, 93 (1961).
- [9.9] E.W. Kellermann, Phil. Trans. Roy. Soc., **A238**, 513 (1940).
- [9.10] R. Pick, M.H. Cohen, and R.M. Martin, Phys. Rev. B **1**, 910 (1970).
- [9.11] W.E. Pickett, Computer Phys. Reports **9**, 115 (1989).
- [9.12] S. Baroni, P. Giannozzi, and A. Testa, Phys. Rev. Lett. **58**, 1861 (1987).
- [9.13] W.A. Harrison, *Pseudopotentials in the Theory of Metals* (Benjamin, New York, 1966).
- [9.14] A. Baldereschi and R. Resta, in: *Ab-Initio Calculation of Phonon Spectra*, edited by J.T. Devreese et al. (Plenum Publishing, New York, 1983), p. 1.
- [9.15] S. de Gironcoli, S. Baroni, and R. Resta, Phys. Rev. Lett. **62**, 2853 (1989).
- [9.16] S. Baroni, P. Giannozzi and E. Molinari, in: *Proceedings of the 19th International Conference on the Physics of Semiconductors*, edited by W. Zawadzky, (Institute of Physics, Polish Academy of Sciences, Wroclaw 1988), p. 795.
- [9.17] S. Baroni, P. Giannozzi and E. Molinari, Phys. Rev. B, in press (15 february 1990).
- [9.18] P. Giannozzi, S. de Gironcoli, P. Pavone, and S. Baroni, Phys. Rev. B **43**, 7231 (1991).