Density Functional Theory and the Family of (L)APW-methods: a step-by-step introduction

S. Cottenier

August 6, 2004

ISBN 90-807215-1-4

Contents

Why this text ? **v**

Citing this text **vii**

Useful references for further reading **ix**

1 Density Functional Theory as a way to solve the quantum many body problem **1**
- 1.1 Level 1: The Born-Oppenheimer approximation . . . 1
- 1.2 Level 2: Density Functional Theory . . . 2
  - 1.2.1 The theorems of Hohenberg and Kohn . . . 2
  - 1.2.2 The Kohn-Sham equations . . . 4
  - 1.2.3 The exchange-correlation functional . . . 7
- 1.3 Level 3: Solving the equations . . . 8

2 The pseudopotential method (in brief) **11**

3 The APW method **15**

4 The LAPW method **21**
- 4.1 The regular LAPW method . . . 21
- 4.2 LAPW with Local Orbitals (LAPW+LO) . . . 22

5 The APW+lo method **25**
- 5.1 The 'pure' APW+lo basis set . . . 25
- 5.2 Mixed LAPW/APW+lo basis sets . . . 26
- 5.3 APW+lo with Local Orbitals (APW+lo+LO) . . . 26

6 Examples for WIEN2k **27**
- 6.1 Linearization energies in LAPW and APW+lo . . . 27
  - 6.1.1 Default initialization for LAPW . . . 27
  - 6.1.2 Choosing linearization energies for LAPW+LO . . . 28
  - 6.1.3 Choosing linearization energies for APW+lo . . . 34
  - 6.1.4 Automatic determination of linearization energies . . . 34
- 6.2 Finding the best $R_{mt}K_{max}$ and k-mesh . . . 36

A Fourier transforms, plane waves, the reciprocal lattice and Bloch's theorem 41
A.1 Fourier transforms 41
A.2 Plane waves 42
A.3 The reciprocal lattice 43
A.4 Bloch's theorem 45

B Quantum numbers and the Density Of States 47
B.1 Familiar examples 47
B.2 Crystalline solids 49

C The eigenvalue problem 53
C.1 Eigenvalues and eigenvectors 53
C.2 Basis transformations 55
C.3 A practical procedure to find eigenvalues and eigenvectors 56

D Solutions of the radial part of the Schrödinger equation 59

E The homogeneous electron gas 61

F Functionals 63
F.1 Definition and Examples 63
F.2 Functional derivatives 63

Why this text ?

As you can see in the reference section (page ix), there are more then a few excellent books, reviews and papers on all topics covered in the present text. Why would there be any need for yet another text? It is my feeling that much of the existing material has been written for an audience that was already rather familiar with theoretical and/or computational condensed matter physics. With the increasing availability of more and better general-purpose condensed matter codes – such as `WIEN2K` – increasingly more people with an explicit experimental background start to use these codes. Especially those new experimental users that have no theoretician nearby, might find the existing reading material either too short and fast or too voluminous and detailed to start with. This text tries to fill that gap: it is meant to give a rather detailed overview of DFT and the family of (L)APW-methods, but starts from the knowledge that every physicist, chemist and engineer has acquired at an undergraduate level. A large appendix section even brushes up the most important of these undergraduate items, with emphasis on those aspects that will play a role in the main text. From the basic material of the appendices, the reader is taken at the hand, and is step by step introduced into DFT and the (L)APW-methods. How the different elements explained in the text appear in the in- and output of the `WIEN2K`-code, is demonstrated in the example section.

Chapter 1 on Density Functional Theory can be read independently from the chapters 2 to 5 on (L)APW-based band structure methods. Readers that are particularly interested in one of both main subjects, can without much harm read only the relevant chapters.

This text is by no means a course to teach you how to use the `WIEN2K`-code from scratch. Such a course can be found in the User's Guide accompanying the code, and this text should be used parallel to it. The emphasis of the User's Guide is more on *how* to work with `WIEN2K`. This text explains the formalism behind the code, and should make clear *why* this or that quantity appears in the in- or output of `WIEN2K`, and what its influence on a calculation can be. This text is meant to help a beginner to acquire a deeper understanding about what she or he is doing with the code.

Much emphasis has been put on mathematical rigour. Therefore you will find many formulae that contain tons of indices. At first sight, this might scare some people. Nevertheless, this crowded notation is meant as a learning tool: try hard to understand the meaning and necessity of every piece of notation. This will prevent you from deviating into misconceptions. As a compensation for this loaded notation, steps between subsequent formulae are kept very small and should be really easy to understand.

This text is by no means complete or error-free. If you find an error or inconsistency, please e-mail `Stefaan.Cottenier@fys.kuleuven.ac.be`. If you – as a beginner – feel that important topics are not covered, you might suggest items to add. Or if you – as an expert – feel the same, you might even write them …

Stefaan Cottenier
Instituut voor Kern- en Stralingsfysica (IKS)
Katholieke Universiteit Leuven
Celestijnenlaan 200 D
B-3001 Leuven
Belgium

Stefaan.Cottenier@fys.kuleuven.ac.be
http://www.fys.kuleuven.ac.be/iks/nvsf/nvsf.html

Citing this text

If you found this text useful, it might occur that you want to cite it in your own work. In that case, please use exactly the following citation:

S. Cottenier, *Density Functional Theory and the family of (L)APW-methods: a step-by-step introduction* (Instituut voor Kern- en Stralingsfysica, K.U.Leuven, Belgium), 2002, ISBN 90-807215-1-4 (to be found at http://www.wien2k.at/reg_user/textbooks).

Useful references for further reading

For each of the main sections of this text, several good books, reviews or original papers exist in which you can read in-depth about the subject. Several of these sources are used in this text, a few times *verbatim*.

- On Density Functional Theory

  A good general review of DFT and its applications:

  Atomic-scale computational materials science
  J. Hafner, *Acta Materialia* 48 (2000) 71-92

  A book that covers DFT and many of the features that pop up in actual computational work (k-space, integration methods, computational aspects, examples of computational studies on structural properties, phonons, thermodynamical quantities,...) :

  Quantum-echanical ab initio calculation of the properties of crystalline materials
  C. Pisani (ed.), Springer-Verlag (1996), ISBN 3-540-61645-4

  Two books that go really in-depth:

  Density Functional Theory
  R.M. Dreizler and E.K.U. Gross, Springer-Verlag (1990), ISBN 3-540-51993-9

  Density Functional Theory of Atoms and Molecules
  R.G. Parr, W. Yang, Oxford University Press, New York (1989)

- On the pseudopotential method

  Pseudopotential methods in condensed matter applications
  Warren E. Pickett, *Computer Physics Reports* 9 (1989) 115-198

- On the APW method

  Three old but nice and detailed sources:

  Augmented Plane Wave method: a guide to performing electronic structure calculations
  T.L. Loucks, Benjamin, New York (1967)

  L.F. Mattheis, J.H. Wood, A.C. Switendick, *Methods in Computational Physics* 8 (1968) 64

  The calculation of electronic energy bands by the augmented plane wave method
  J.O. Dimmock in *Solid State Physics*, H. Ehrenreich, F. Seitz, D. Turnbull (editors), volume 26, Academic Press (1971), 103-273.

- On the LAPW method

  The basic LAPW-book that covers in detail many of the technical aspects:

  Planewaves, pseudopotentials and the LAPW-method,
  D. Singh, Kluwer Academic Publishing (1994), ISBN 0-7923-9421-7

  The original LAPW papers :

  Linear methods in band theory
  O. K. Andersen, *Physical Review B*, 12 (1975) 3060

  Use of energy derivative of the radial solution in an augmented plane wave method: application to copper
  D.D. Koeling, G.O. Arbman, *Journal of Physics F* 5 (1975) 2041

- The APW+lo method

  The original papers, that cover also (L)APW:

  An alternative way of linearizing the augmented plane-wave method
  E. Sjöstedt, L. Nordström, D.J. Singh, *Solid State Communications* 114 (2000) 15-20

  Efficient linearization of the augmented plane-wave method
  G.K.H. Madsen, P. Blaha, K. Schwarz, E. Sjöstedt, L. Nordström, *Physical Review B* 64 (2001) 195134

Electronic structure calculations of solids using the WIEN2k package for material science
K. Schwarz, P. Blaha, G.K.H. Madsen, *Computer Physics Communications* 147 (2002) 71

Chapter 1

Density Functional Theory as a way to solve the quantum many body problem

A solid is a collection of heavy, positively charged particles (nuclei) and lighter, negatively charged particles (electrons). If we have N nuclei, we are dealing with a problem of N+ZN electromagnetically interacting particles. This is a many-body problem, and because the particles are so light, quantum mechanics is needed: a quantum many body problem. The exact many-particle hamiltonian for this system is:

$$\begin{aligned}\hat{H} &= -\frac{\hbar^2}{2}\sum_i \frac{\nabla^2_{\vec{R}_i}}{M_i} - \frac{\hbar^2}{2}\sum_i \frac{\nabla^2_{\vec{r}_i}}{m_e} \\ &\quad - \frac{1}{4\pi\epsilon_0}\sum_{i,j}\frac{e^2\,Z_i}{|\,\vec{R}_i - \vec{r}_j|} + \frac{1}{8\pi\epsilon_0}\sum_{i\neq j}\frac{e^2}{|\,\vec{r}_i - \vec{r}_j|} + \frac{1}{8\pi\epsilon_0}\sum_{i\neq j}\frac{e^2\,Z_iZ_j}{|\,\vec{R}_i - \vec{R}_j|}\end{aligned} \tag{1.1}$$

The mass of the nucleus at $\vec{R}_i$ is M_i, the electrons have mass m_e and are at $\vec{r}_i$. The first term is the kinetic energy operator for the nuclei, the second for the electrons. The last three terms describe the Coulomb interaction between electrons and nuclei, between electrons and other electrons, and between nuclei and other nuclei. It is out of question to solve this problem exactly. In order to find acceptable approximate eigenstates, we will need to make approximations at 3 different levels.

1.1 Level 1: The Born-Oppenheimer approximation

The nuclei are much heavier and therefore much slower than the electrons. We can hence 'freeze' them at fixed positions and assume the electrons to be in instantaneous equilibrium with them. In other words: only the electrons are kept as players in our many body problem. The nuclei are deprived from this status, and reduced to a given source of positive charge, they become 'external' to the electron cloud. After having applied this approximation, we are left with a collection of NZ interacting negative particles, moving in the (now external or given) potential of the nuclei.

What are the consequences of the Born-Oppenheimer approximation on the hamiltonian 1.1? The nuclei do not move any more, their kinetic energy is zero and the first term disappears. The last term reduces to a constant. We are left with the kinetic energy of the electron gas, the potential energy due to electron-electron interactions and the potential energy of the electrons in the (now external) potential of the nuclei. We write this formally as:

$$\hat{H} \;=\; \hat{T} \;+\; \hat{V} \;+\; \hat{V}_{ext} \tag{1.2}$$

It is interesting to note here that the kinetic and electron-electron terms of 1.2 depend only on the fact that we are dealing with a many-electron system (and not with a many-proton system for instance, where the strong nuclear force would play a role). They are independent of the particular kind of many-electron system (Br_2 or H_2O molecule, Cu or Fe, bcc-Fe or fcc-Fe, ...). This part is universal. System-specific information (which nuclei, and on which positions) is given entirely by $\hat{V}_{ext}$.

1.2 Level 2: Density Functional Theory

The quantum many body problem obtained after the first level approximation (Born-Oppenheimer) is much simpler than the original one, but still far too difficult to solve. Several methods exist to reduce equation 1.2 to an approximate but tractable form. A historically very important one is the Hartree-Fock method (HF), described in many condensed matter textbooks. It performs very well for atoms and molecules, and is therefore used a lot in quantum chemistry. For solids it is less accurate, however. We will not treat HF, but explain a more modern and probably also more powerful method: Density Functional Theory (DFT)[1].

Although its history goes back to the early thirties of the 20^{th} century, DFT has been formally established in 1964 by two theorems due to Hohenberg and Kohn[2].

1.2.1 The theorems of Hohenberg and Kohn

The traditional formulation of the two theorems of Hohenberg and Kohn is as follows:

First theorem: *There is a* ***one-to-one*** *correspondence between the ground-state density $\rho(\vec{r})$ of a many-electron system (atom, molecule, solid) and the external potential V_{ext}. An immediate consequence is that the ground-state expectation value of any observable $\hat{O}$ is a unique functional of the exact ground-state electron density:*

$$<\Psi|\,\hat{O}\,|\,\Psi> \;=\; O[\rho] \tag{1.3}$$

Second theorem: *For $\hat{O}$ being the hamiltonian $\hat{H}$, the ground-state total energy functional $H[\rho] \equiv E_{V_{ext}}[\rho]$ is of the form*

$$E_{V_{ext}}[\rho] \;=\; \underbrace{<\Psi|\,\hat{T} \,+\, \hat{V}\,|\,\Psi>}_{F_{HK}[\rho]} \;+\; <\Psi|\,\hat{V}_{ext}\,|\,\Psi> \tag{1.4}$$

$$\;=\; F_{HK}[\rho] \;+\; \int \rho(\vec{r})\,V_{ext}(\vec{r})\,d\vec{r} \tag{1.5}$$

---

[1] As Hartree-Fock, DFT is a general method to solve the quantum many body problem. It can be applied not only to the electron gas as we use here, but also to the proton-neutron gas in order to construct nuclear models, or to the nucleus-electron gas without the Born-Oppenheimer approximation to describe solids with light elements

[2] P. Hohenberg and W. Kohn, *Physical Review* 136(3B) (1964) p. 864

where the Hohenberg-Kohn density functional $F_{HK}[\rho]$ is ***universal*** *for any many-electron system. $E_{V_{ext}}[\rho]$ reaches its* ***minimal*** *value (equal to the ground-state total energy) for the ground-state density corresponding to V_{ext}.*

We do not prove these theorems here, but ponder a few implications of the three keywords *invertibility* (one-to-one correspondence $\rho \leftrightarrow V_{ext}$), *universality* and *variational access* (minimal value).

First, the one-to-one correspondence between ground-state density and external potential is intriguing. It is obvious that a given many-electron system has a unique external potential, which by the hamiltonian 1.2 and the Schrödinger equation yields a unique ground-state many particle wave function. From this wave function, the corresponding electron density is easily found. An external potential hence leads in a well-defined way to a unique ground-state density corresponding to it. But intuitively it looks like the density contains less information then the wave function. If this would be true, it would not be possible to find a unique external potential if only a ground-state density is given. The first theorem of Hohenberg and Kohn tells exactly that this *is* possible! The density contains as much information as the wave function does (i.e. everything you could possibly know about an atom, molecule or solid). All observable quantities can be retrieved therefore in a unique way from the density only, i.e. they can be written as functionals of the density.

Second, the universality of $F_{HK}[\rho]$. Equation 1.5 is easily written down by using the density operator[3], and supposing the ground-state density is known, the contribution to the total energy from the external potential can be exactly calculated. An explicit expression for the Hohenberg-Kohn functional F_{HK} is not known. But anyway, because F_{HK} does not contain information on the nuclei and their position, it is a universal functional for any many-electron system. This means that in principle an expression for $F_{HK}[\rho]$ exists which can be used for every atom, molecule or solid which can be imagined.

Third, the second theorem makes it possible to use the variational principle of Rayleigh-Ritz in order to find the ground-state density. Out of the infinite number of possible densities, the one which minimizes $E_{V_{ext}}[\rho]$ is the ground-state density corresponding to the external potential $V_{ext}(\vec{r})$. Of course, this can be done only if (an approximation to) $F_{HK}]\rho]$ is known. But having found ρ, all knowledge about the system is within reach...

It is useful to stress the meaning of the energy functional $E_{V_{ext}}[\rho]$ once more. When it is

---

[3] The density operator $\hat{\rho}(\vec{r})$ for an N-particle system is defined as

$$\hat{\rho}(\vec{r}) \quad = \quad \sum_{i=1}^{N} \delta(\vec{r}_i - \vec{r}) \tag{1.6}$$

such that its evaluation for a many body wave function Ψ yields the density:

$$\begin{aligned}
\rho(\vec{r}) \quad &= \quad <\Psi(\vec{r}_1, \vec{r}_2, \ldots, \vec{r}_N)|\, \hat{\rho}(\vec{r})|\, \Psi(\vec{r}_1, \vec{r}_2, \ldots, \vec{r}_N)> && (1.7)\\
&= \quad <\Psi(\vec{r}_1, \vec{r}_2, \ldots, \vec{r}_N)|\, \sum_{i=1}^{N} \delta(\vec{r}_i - \vec{r})|\, \Psi(\vec{r}_1, \vec{r}_2, \ldots, \vec{r}_N)> && (1.8)\\
&= \quad \sum_{i=1}^{N} \int \Psi^*(\vec{r}_1, \vec{r}_2, \ldots, \vec{r}_i \equiv \vec{r}, \ldots \vec{r}_N)\Psi(\vec{r}_1, \vec{r}_2, \ldots, \vec{r}_i \equiv \vec{r}, \ldots \vec{r}_N)\, d\vec{r}_1\, d\vec{r}_2 \ldots \cancel{d\vec{r}_i} \ldots d\vec{r}_N && (1.9)
\end{aligned}$$

evaluated for the density ρ corresponding to the particular V_{ext} for this solid, it gives the ground state energy. When it is evaluated for any other density however, the resulting number has no physical meaning[4]!

1.2.2 The Kohn-Sham equations

The equations of Kohn and Sham, published in 1965, turn DFT into a practical tool[5]. They are a practical procedure to obtain the ground state density. Let us first rewrite the Hohenberg-Kohn functional. The *correlation energy* is defined as this part of the total energy which is present in the exact solution, but absent in the Hartree-Fock solution. The total energy functionals $E_e[\rho]$ and $E_{HF}[\rho]$ corresponding to the exact and Hartree-Fock Hamiltonians respectively, are[6]:

$$\begin{aligned} E_e &= T + V && (1.10) \\ E_{HF} &= T_0 + \underbrace{(V_H + V_x)}_{V} && (1.11) \end{aligned}$$

Here T and V are the exact kinetic and electron-electron potential energy functionals, T_0 is the functional for the kinetic energy of a non-interacting electron gas, V_H stands for the Hartree contribution and V_x for the exchange contribution. By subtracting 1.11 from 1.10, the functional for the correlation contribution appears to be:

$$V_c = T - T_0 \tag{1.12}$$

The exchange contribution to the total energy is defined as the part which is present in the Hartree-Fock solution, but absent in the Hartree solution. Obviously, with the Hartree functional given by

$$E_H = T_0 + V_H \tag{1.13}$$

it can be defined as

$$V_x = V - V_H \tag{1.14}$$

With this knowledge, we can rewrite the Hohenberg-Kohn functional in the following way:

$$\begin{aligned} F_{HK} &= T + V + T_0 - T_0 \\ &= T_0 + V + \underbrace{(T - T_0)}_{V_c} \\ &= T_0 + V + V_c + V_H - V_H \\ &= T_0 + V_H + V_c + \underbrace{(V - V_H)}_{V_x} \\ &= T_0 + V_H + \underbrace{(V_x + V_c)}_{V_{xc}} \end{aligned}$$

---

[4] This is too crude, but the details are subtle: any density $\rho_i(\vec{r})$, other than the ground-state density, which extremalizes $E_{V_{ext}}[\rho]$ is an excited-state density, with $E_i = E_{V_{ext}}[\rho = \rho_i]$ the corresponding energy. The inverse statement is not true: not all excited-state densities extremalize $E_{V_{ext}}[\rho]$. If $\rho_j(\vec{r})$ is such a density, $E_j = E_{V_{ext}}[\rho = \rho_j]$ is a lower bound for the energy of the excited state.

[5] W. Kohn and L. J. Sham, *Physical Review* 140(4A) (1965) p. 1133

[6] For easier notation, we drop for a while the square brackets.

Here V_{xc} is the exchange-correlation energy functional. We don't know it formally, as it contains the difficult exchange and correlation contributions only. If we assume for a while that we do know V_{xc}, we can write explicitly the energy functional:

$$E_{V_{ext}}[\rho] \;=\; T_0[\rho] \,+\, V_H[\rho] \,+\, V_{xc}[\rho] \,+\, V_{ext}[\rho] \tag{1.15}$$

One could use now the second Hohenberg-Kohn theorem to find the ground state density, but then we would have won nothing by our transformation. Instead, you can interpret the above expression also as the energy functional of a *non-interacting* classical electron gas, subject to two external potentials: one due to the nuclei, and one due to exchange and correlation effects. The corresponding Hamiltonian – called the Kohn-Sham Hamiltonian – is

$$\hat{H}_{KS} \;=\; \hat{T}_0 \,+\, \hat{V}_H \,+\, \hat{V}_{xc} \,+\, \hat{V}_{ext} \tag{1.16}$$

$$\;=\; -\frac{\hbar^2}{2m_e}\,\vec{\nabla}_i^2 \,+\, \frac{e^2}{4\pi\epsilon_0}\int \frac{\rho(\vec{r}\,')}{|\,\vec{r}-\vec{r}\,'|}\,d\vec{r}\,' \,+\, V_{xc} \,+\, V_{ext} \tag{1.17}$$

where the exchange-correlation potential is given by the functional derivative

$$\hat{V}_{xc} \;=\; \frac{\delta V_{xc}[\rho]}{\delta\rho} \tag{1.18}$$

The theorem of Kohn and Sham can now be formulated as follows:

The exact ground-state density $\rho(\vec{r})$ of an N-electron system is

$$\rho(\vec{r}) \;=\; \sum_{i=1}^{N} \phi_i(\vec{r})^*\,\phi_i(\vec{r}) \tag{1.19}$$

where the single-particle wave functions $\phi_i(\vec{r})$ are the N lowest-energy solutions of the Kohn-Sham equation

$$\hat{H}_{KS}\,\phi_i \;=\; \epsilon_i\,\phi_i \tag{1.20}$$

And now we did won a lot. To find the ground-state density, we don't need to use the second Hohenberg-Kohn theorem any more, but we can rely on solving familiar Schrödinger-like non-interacting single-particle equations. The alternative of using the regular Schrödinger equation, would have led to a far more difficult system of coupled differential equations, because of the electron-electron interaction.

Be aware that the single-particle wave functions ϕ_i are not the wave functions of electrons! They describe mathematical quasi-particles, without a direct physical meaning. Only the overall density of these quasi-particles is guaranteed to be equal to the true electron density. Also the single-particle energies ϵ_i are not single-electron energies.

Both the Hartree operator V_H and the exchange-correlation operator V_{xc} depend on the density $\rho(\vec{r})$, which in turn depends on the ϕ_i which are being searched. This means we are dealing with a *self-consistency problem*: the solutions (ϕ_i) determine the original equation (V_H and V_{xc} in H_{KS}), and the equation cannot be written down and solved before its solution is

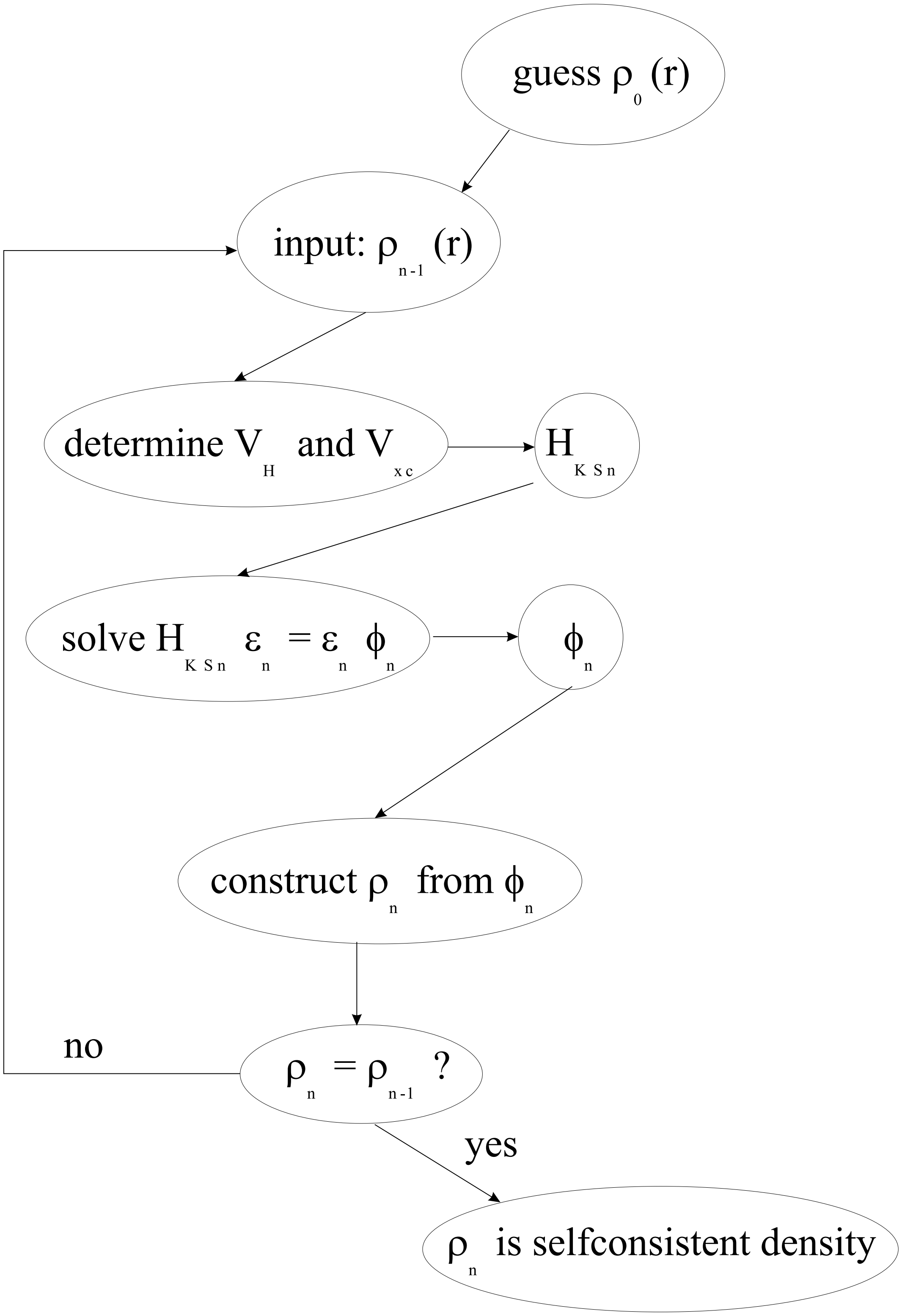
guess ρ_0 (r)
input: ρ_n-1 (r)
determine V_H and V_xc
H_KSn
solve H_KSn ε_n = ε_n ϕ_n
ϕ_n
construct ρ_n from ϕ_n
no
ρ_n = ρ_n-1 ?
yes
ρ_n is selfconsistent density

Figure 1.1: *Flow chart for the* n^{th} *iteration in the selfconsistent procedure to solve Hartree-Fock or Kohn-Sham equations.*

known. An iterative procedure is needed to escape from this paradox (see fig. 1.1). Some starting density ρ_0 is guessed, and a hamiltonian H_{KS1} is constructed with it. The eigenvalue problem is solved, and results in a set of ϕ_1 from which a density ρ_1 can be derived. Most probably ρ_0 will differ from ρ_1. Now ρ_1 is used to construct H_{KS2}, which will yield a ρ_2, etc. The procedure can be set up in such a way that this series will converge to a density ρ_f which generates a H_{KSf} which yields as solution again ρ_f: this final density is then *consistent* with the hamiltonian.

1.2.3 The exchange-correlation functional

The Kohn-Sham scheme described above was exact: apart from the preceding Born-Oppenheimer approximation, no other approximations were made. But we neglected so far the fact that we do not know the exchange-correlation functional. It is here that approximations enter our theory.

A widely used approximation – called the *Local Density Approximation* (LDA) – is to postulate that the exchange-correlation functional has the following form:

$$E_{xc}^{LDA} = \int \rho(\vec{r})\, \epsilon_{xc}(\rho(\vec{r}))\, d\vec{r} \tag{1.21}$$

The *function* (not: *functional*) $\epsilon_{xc}(\rho)$ for the homogeneous electron gas has been defined in Appendix E, and is numerically known. This postulate is somehow reasonable: it means that the exchange-correlation energy due to a particular density $\rho(\vec{r})$ could be found by dividing the material in infinitesimally small volumes with a constant density. Each such volume contributes to the total exchange correlation energy by an amount equal to the exchange correlation energy of an identical volume filled with a homogeneous electron gas, that has the same overall density as the original material has in this volume (see Fig. 1.2). No law of nature guarantees that the true E_{xc} is of this form, it is only a reasonable guess. By construction, LDA is expected to perform well for systems with a slowly varying density. But rather surprisingly, it appears to be very accurate in many other (realistic) cases too.

A next logical step to improve on LDA, is to make the exchange-correlation contribution of every infinitesimal volume not only dependent on the local density in that volume, but also on the density in the neighbouring volumes. In other words, the *gradient* of the density will play a role. This approximation is therefore called the *Generalized*[7] *Gradient Approximation* (GGA). Although GGA performs in general slightly better than LDA, there are a few drawbacks. There is only one LDA exchange-correlation functional, because there is a unique definition for ϵ_{xc}. But there is some freedom to incorporate the density gradient, and therefore several versions of GGA exist (first drawback). Moreover, in practice one often fits a candidate GGA-functional with (hopefully only a few) free parameters to a large set of experimental data on atoms and molecules. The best values for these parameters are fixed then, and the functional is ready to be used routinely in solids. Therefore such a GGA-calculation is strictly spoken not an *ab initio* calculation, as some experimental information is used (second drawback). Nevertheless, there exist GGA's that are parameter free.

---

[7] "Generalized", because a straightforward gradient would result in a functional which violates some relationships which can be proven to be exact for the true functional (and for LDA)

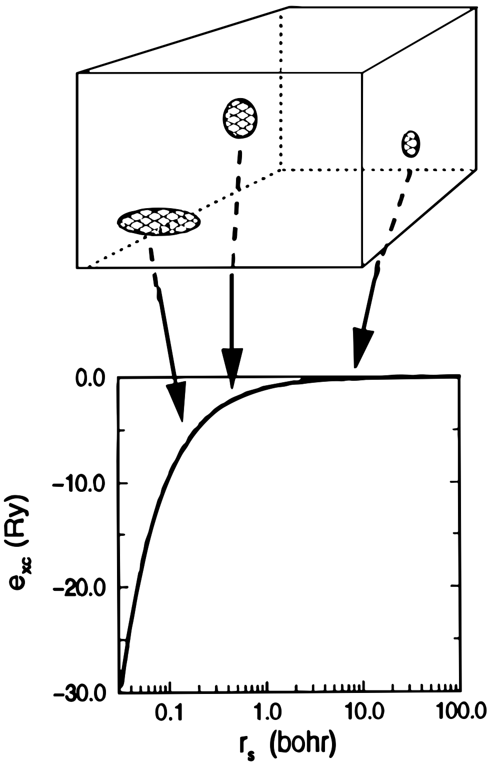

Figure 1.2: *Illustration of the idea behind the LDA postulate. Every infinitesimally small volume of the material contributes to the exchange-correlation energy with an amount equal to the contribution of a homogeneous electron gas that occupies that same infinitesimally small volume, and that has the same (overall) charge density as the charge density of the original material in that volume. The horizontal axis is proportional to the density of the homogeneous electron gas. The vertical axis displays the exchange-correlation energy of the homogeneous electron gas. (picture taken from Martin Fuchs, lecture notes, ICTP Workshop august 1999)*

1.3 Level 3: Solving the equations

Irrespective whether one has used HF or DFT as level 2 approximation, one ends up with an infinite set of one-electron equations of the following type (m is an integer number that counts the members of the set):

$$\underbrace{\left(- \frac{\hbar^2}{2m_e} \vec{\nabla}_m^2 + \frac{e^2}{4\pi\epsilon_0} \int \frac{\rho(\vec{r}\,')}{|\vec{r} - \vec{r}\,'|} d\vec{r}\,' + V_\alpha + V_{ext} \right)}_{\hat{H}_{sp}} \phi_m(\vec{r}) = \epsilon_m \, \phi_m(\vec{r}) \qquad (1.22)$$

We call $\hat{H}_{sp}$ the single-particle hamiltonian. For HF, V_α is the exchange operator. The ϕ_m are true one-electron (or single-particle) orbitals for HF. Exchange is treated exactly, but correlation effects are not included at all. They can be added only in elaborations on the HF-method.

For DFT, V_α is the exchange-correlation operator, in the L(S)DA, GGA or another approximation. Exchange and correlation are both treated, but both approximately. The ϕ_m are mathematical single-particle orbitals.

The similarity between the Hartree-Fock and Kohn-Sham equations means that the same mathematical techniques can be used to solve them. 'Solving' in most methods means that we want to find the coefficients c_p^m needed to express ϕ_m in a given basis set ϕ_p^b:

$$\phi_m = \sum_{p=1}^{P} c_p^m \phi_p^b \tag{1.23}$$

The wave functions ϕ_m belong to a function space which has an infinite dimension, P is therefore in principle infinite. In practice one works with a limited set of basis functions. Such a limited basis will never be able to describe ϕ_m exactly, but one could try to find a basis that can generate a function that is 'close' to ϕ_m.

Having chosen a basis (and hence a finite value for P) we realize that we can tackle the equations 1.22 as an eigenvalue problem, in the same way as described in Appendix C. For a given m, substitute equation 1.23 in 1.22, and left-multiply with $\left\langle \phi_i^b \right|$ $(i = 1, \ldots, P)$. As in equation C.8, this leads to

$$\begin{bmatrix} \cdots & \cdots & \cdots \\ \vdots & \left\langle \phi_i^b \middle| \hat{H}_{sp} \middle| \phi_j^b \right\rangle - \epsilon_m \left\langle \phi_i^b \middle| \phi_j^b \right\rangle & \vdots \\ \cdots & \cdots & \cdots \end{bmatrix} \begin{bmatrix} c_1^m \\ \vdots \\ c_P^m \end{bmatrix} = \begin{bmatrix} 0 \\ \vdots \\ 0 \end{bmatrix} \tag{1.24}$$

We recognize here the matrix elements of the single-particle hamiltonian in the basis states, and the overlap matrix elements S_{ij}. Remember that the overlap matrix is a unit matrix if the basis set is orthonormal. Diagonalization of the hamiltonian matrix will lead to P eigenvalues and P sets of coefficients that express each of the P eigenfunctions in the given basis (if more eigenfunctions are needed, P must be increased). The larger P, the better the approximation of the eigenfunction, but the more time-consuming the diagonalization of the matrix in equation 1.24.

In the following chapters, we will elaborate on this quick sketch. We will see that the label m stands for the quantum numbers $(n, \vec{k})$. It will be possible and advantageous to use a different basis set for each $\vec{k}$, which will enormously reduce the size of P, but will lead to as many different eigenvalue problems to solve as there are $\vec{k}$-vectors needed to be known.

What is a good basis set? If the functions of the basis set are very similar to ϕ_m, one needs only a few of them to accurately describe the wave function, and hence P and the matrix size are small. Such a basis set is called *efficient.* However, this assumes that you know the solution of your problem almost before you start solving it. Such a basis set can therefore never be very general: for some specific problems it will very quickly yield the solution, but for the majority of cases it will poorly describe the eigenfunctions. In the latter cases the required P is much higher than what is affordable, and limiting P would lead to approximate eigenfunctions that are not acceptable. These approximations carry too much properties from the basis function, and such a basis set is therefore called *biased.* The art of theoretical condensed matter physics is to find a basis set that is a simultaneously efficient and unbiased. In the following chapters, two families of basis sets will be described – plane waves and augmented plane waves – that each in their own way try to realize this compromise.

Chapter 2

The pseudopotential method (in brief)

At the end of section 1.3, we formulated two principal requirements for a basis set in which we want to expand the eigenstates of the solid state hamiltonian: the basis set should be *unbiased* (=it should not in a hidden way force the solution into a built-in direction) and *efficient* (= P in equation 1.23 should be as low as possible). Furthermore, it would be nice if the basis functions are mathematically simple. This makes both theory development and programming work easier. A basis set that is certainly unbiased and simple is the plane wave basis set that was introduced in section A.4. We saw in equation A.22 that any eigenfunction $\psi^n_{\vec{k}}$ of a periodic hamiltonian can be expressed exactly in this basis set by means of an infinite set of coefficients $c^{n,\vec{k}}_{\vec{K}}$:

$$\psi^n_{\vec{k}}(\vec{r}) \;=\; \sum_{\vec{K}} c^{n,\vec{k}}_{\vec{K}} e^{i(\vec{k}+\vec{K})\cdot\vec{r}} \tag{2.1}$$

This has to be compared with the general formulation in equation 1.23, where m stands for $(n,\,\vec{k})$ and p for $\vec{k}+\vec{K}$. One basis function for $\psi^n_{\vec{k}}(\vec{r})$ (or $\vec{K}$) is therefore:

$$\phi^{\vec{k}}_{\vec{K}}(\vec{r}) \;=\; \left|\vec{K}\right\rangle \;=\; e^{i(\vec{k}+\vec{K})\cdot\vec{r}} \tag{2.2}$$

Note that this basis-set is $\vec{k}$-dependent: all eigenstates $\psi^n_{\vec{k}}$ that have the same $\vec{k}$ but a different n will be expressed in the basis set with this particular value of $\vec{k}$. For eigenstates with another $\vec{k}$, a new basis set using that other $\vec{k}$ has to be used.

In practice we cannot work with an infinite basis set, and will have to limit it somehow (the P in equation 1.23). For plane waves, this can be easily done by limiting the set to all $\vec{K}$ with $K \leq K_{max}$. This corresponds to a sphere with radius K_{max} centered at the origin of reciprocal space. All reciprocal lattice vectors that are inside this sphere are taken into the basis set. Instead of K_{max}, often the free electron energy corresponding to K_{max} is specified, and is called the *cut-off energy*:

$$E_{cut} \;=\; \frac{\hbar^2 K^2_{max}}{2m_e} \tag{2.3}$$

Plane waves are orthogonal:

$$\begin{aligned}\left\langle \vec{K}_1 \middle| \vec{K}_2 \right\rangle &= \int e^{i\left(\vec{K}_2-\vec{K}_1\right)\cdot\vec{r}} d^3\vec{r} &(2.4)\\ &= \delta(\vec{K}_2-\vec{K}_1) &(2.5)\end{aligned}$$

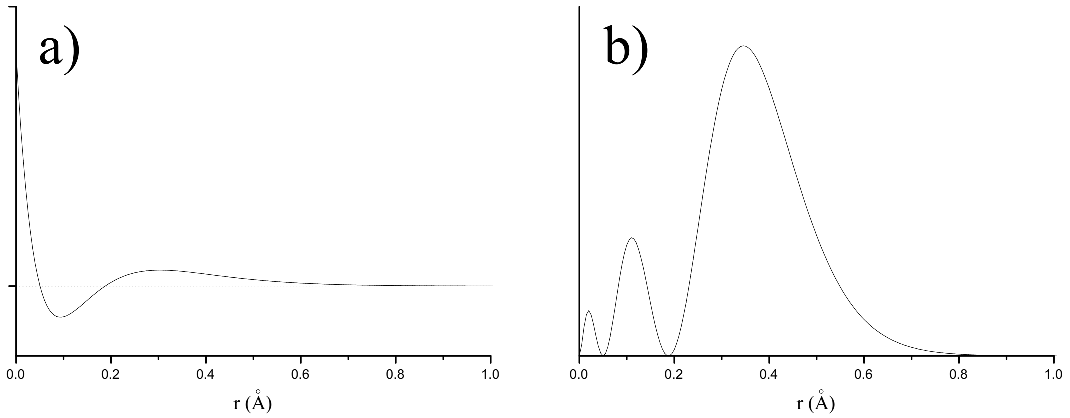

Figure 2.1: *Radial part of wave function (a) and radial probability distribution (b) of a 3s electron in Ca (Y-axis has arbitrary units).*

and therefore the overlap matrix S in equations C.18 or 1.24 is a unit matrix. The matrix elements of the effective single-particle hamiltonian (Kohn-Sham hamiltonian) are calculated in the plane wave basis, and the resulting matrix is diagonalized (see section C.3 for a description of the full procedure). To each eigenvalue $\epsilon^{n,\vec{k}}$ an eigenvector $\left[c^{n,\vec{k}}_{\vec{K}}\right]_{P\times 1}$ of P values for $c^{n,\vec{k}}_{\vec{K}}$ corresponds. Actually, an infinite number of eigenvectors, as multiplying with any real number gives again an eigenvector (see Appendix C). This degree of freedom is used to take that $\left[c^{n,\vec{k}}_{\vec{K}}\right]_{P\times 1}$ for which – by equation 2.1 – $\phi^{n}_{\vec{k}}$ is a normalized function.

So far, so good. But: we were looking for the eigenvalue $\epsilon^{n,\vec{k}}$ of $\psi^{n}_{\vec{k}}$, and instead of one eigenvalue we found P of them, each with their own set of coefficients and hence each leading to another eigenfunction $\psi^{n}_{\vec{k}}$. What has happened? Each eigenfunction is uniquely labeled by n and $\vec{k}$. Eq. 2.1 applies for all our solutions with the same $\vec{k}$. We are therefore forced to conclude that for all these eigenfunctions, n – which does not explicitly appears in the equations – must be different. In one go, we have found P different eigenfunctions with the same $\vec{k}$ but a different band index!

This process has to be repeated for as many $\vec{k}$-points that are needed for a densely enough sampling of the first Brillouin zone.

We did not yet discuss the efficiency of a plane wave basis set. Will we need many or only a few plane waves? The number of plane waves is determined by the smallest length scales that are to be described in real space. Consider the radial part of a 3s wave function in Ca (fig. 2-a). Near the nucleus, the wave function shows steep behaviour. In order to describe the sharp part between 0 and the mimimum at 0.1 Å plane waves with a period as small as roughly an order of magnitude less than this distance are needed (0.01 Å or 10^{-12} m). This sets the order of magnitude for K_{max} to $K_{max} = 2\pi/(10^{-12}\ m) = 6.3 \cdot 10^{12}\ m^{-1}$. Calculate the volume of a sphere with this radius ($10^{39}\ m^{-3}$), divide it by the volume of the first Brillouin zone ($\vec{c}^{*} \cdot (\vec{a}^{*} \times \vec{b}^{*})$, about $9.2 \cdot 10^{30}\ m^{-3}$ for a cubic lattice with a lattice constant of 3 Å), and because the first Brillouin zone contains 1 $\vec{K}$ you find that about 10^{8} plane waves are needed. That would require diagonalization of (many) ($10^{8} \times 10^{8}$)-matrices, which is way beyond the capability of even supercomputers.

Should we conclude now that a plane wave basis set cannot be used? The most oscillating part of the wave functions are the tails that reach out into the region close to the nucleus. But this region of the solid is quite shielded from the more outer regions of the atoms where chemistry happens, and the electrons will not behave very differently from free atom electrons here. One can therefore replace the potential in these inner regions by a *pseudopotential*, that is designed to yield very smooth tails of wavefunctions inside the atom (hence only a few plane waves are needed). More to the outer regions of the atoms, the pseudopotential continuously evolves into the true potential, such that this region of the crystal behaves as if nothing happened. In this way, it is possible to use an (ultrasoft) plane wave basis set for realistic cases with a cut-off energy $E_{cut} = 20\ Ry = 272\ eV$. This corresponds to $K_{max} \approx 4.5\ au^{-1}$, which for our example of a cubic lattice with lattice constant 3 Å calls for about 270 plane waves[1]. A basis set of this size is manageable.

There is no unique recipe to construct a pseudopotential for a particular element. An infinite amount of choices are possible. Two criteria to judge whether a particular pseudopotential is good, are *softness* and *transferability.* A pseudopotential is called *soft* when few plane waves are needed. A specific class of pseudopotentials is even called *ultrasoft*, because of the very small amount of plane waves it calls for. Generally, making a pseudopotential soft means that it gets tailored for an element in a specific environment. What one really wants is a potential that can be used in whatever environment (molecule, cluster, solid, surface, insulator, metal,...) the corresponding element can be. Such a pseudopotential is called *transferable.* The art of creating good pseudopotentials is to find potentials that are both (ultra)soft and transferable.

---

[1] Only *accidentally* similar to E_{cut} in eV.

Chapter 3

The APW method

Although the pseudopotential method is extremely useful, there are reasons why alternatives could be attractive. Is the introduction of the pseudopotential completely innocent? What do you do if you are interested in information that is inherently contained in the region near the nucleus (hyperfine fields for instance, or core level excitations)? Can the basis set be made more efficient? Therefore, we will search for a basis set that uses other functions than plane waves, and that does not require the introduction of a pseudopotential. Such a basis set will have to be more efficient, but of course we do not want it to be biased. Our first example of this will be the *Augmented Plane Wave* (APW) basis set. Right from the beginning it has to be said that the APW-method itself is of no practical use any more today. But for didactical reasons it is advantageous to discuss APW first, before going to its successors, LAPW and APW+lo.

The ideas that lead to the APW basis set are very similar to what made us to introduce the pseudopotential. In the region far away from the nuclei, the electrons are more or less 'free'. Free electrons are described by plane waves[1]. Close to the nuclei, the electrons behave quite as they were in a free atom, and they could be described more efficiently by atomic like functions. Space is therefore divided now in two regions: around each atom[2] a sphere with radius R_α is drawn (call it S_α). Such a sphere is often called a *muffin tin sphere*, the part of space occupied by the spheres is the *muffin tin region*. The remaining space outside the spheres is called the *interstitial region* (call it I). One augmented plane wave (APW) used in the expansion of $\psi^n_{\vec{k}}$ is defined as:

$$\phi^{\vec{k}}_{\vec{K}}(\vec{r},\,E) \;=\; \begin{cases} \frac{1}{\sqrt{V}} e^{i(\vec{k}+\vec{K})\cdot\vec{r}} & \vec{r} \in I \\ \\ \sum_{\ell,\,m} A^{\alpha,\,\vec{k}+\vec{K}}_{\ell m} u^\alpha_\ell(r',E) Y^\ell_m(\hat{r}') & \vec{r} \in S_\alpha \end{cases} \tag{3.1}$$

The symbols $\vec{k}$, $\vec{K}$ and $\vec{r}$ keep their usual meaning, V is the volume of the unit cell. Note that the APW basis set is $\vec{k}$-dependent, as was the plane wave basis set. The position inside the spheres is given with respect to the center of each sphere by $\vec{r}\,' = \vec{r} - \vec{r}_\alpha$ (see fig. 3). The length of $\vec{r}\,'$ is $r\,'$, and the angles θ' and ϕ' specifying the direction of $\vec{r}\,'$ in spherical coordinates, are indicated as $\hat{r}'$. The Y^ℓ_m are spherical harmonics. The $A^{\alpha,\,\vec{k}+\vec{K}}_{\ell m}$ are yet undetermined parameters, as is E. The latter has the dimension of energy. The u^α_ℓ are solutions to the radial part of

---

[1] Plane waves are eigenfunctions of a hamiltonian with zero potential.

[2] Mark the different atoms in the unit cell by a label α. This label is different for all atoms in the unit cell, not just for all inequivalent atoms.

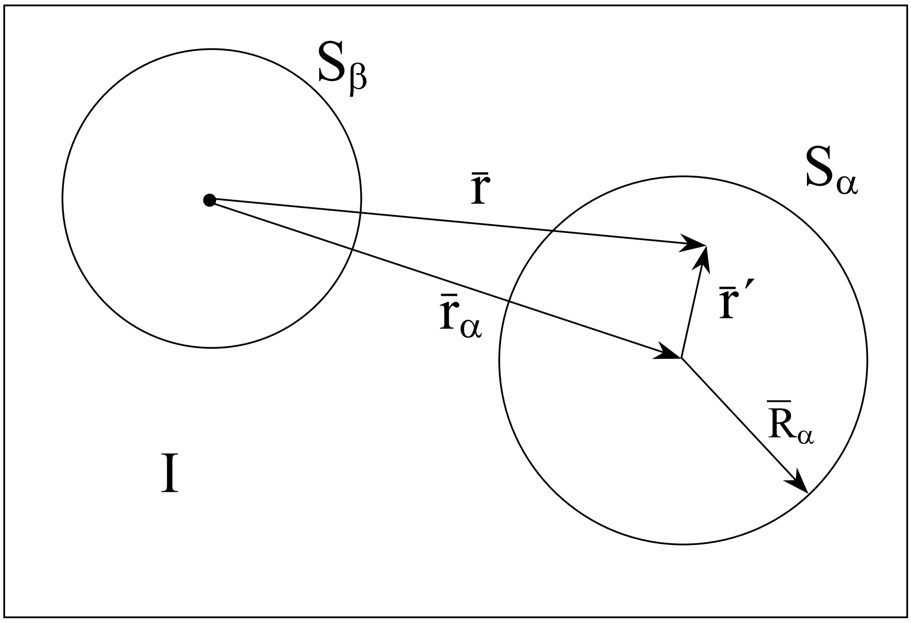

Figure 3.1: *Division of a unit cell in muffin tin regions and the interstitial region, for a case with two atoms. The black dot is the origin of the axis system (which may but need not to coincide with the nucleus of an atom).*

the Schrödinger equation for a free atom α, and this at the energy E. For a true free atom, the boundary condition that $u_\ell^\alpha(r, E)$ should vanish for $r \to \infty$, limits the number of energies E for which a solution u_ℓ^α can be found. But as this boundary condition does not apply here, we can find a numerical solution for any E. Hence, the u_ℓ^α themselves do not correspond to something physical, but that doesn't harm: they are only part of a basis function, not of the searched eigenfunction itself. And because they are close to how the actual eigenfunction will look like in that region of the crystal, they will do their job as basis function very efficiently.

If an eigenfunction would be discontinuous, its kinetic energy would not be well-defined. Such a situation can therefore never happen, and we have to require that the plane wave outside the sphere matches the function inside the sphere over the complete surface of the sphere (in value, not in slope). That seems a weird thing to do: a plane wave is oscillating and has a unique direction built in, how can it match another function based on spherical harmonics over the entire surface of a sphere? To see how this is possible, we expand the plane wave in spherical harmonics about the origin of the sphere of atom α:

$$\frac{1}{\sqrt{V}} e^{i(\vec{k}+\vec{K})\cdot\vec{r}} = \frac{4\pi}{\sqrt{V}} e^{i(\vec{k}+\vec{K})\cdot\vec{r}_\alpha} \sum_{\ell,m} i^\ell j_\ell\left(\left|\vec{k}+\vec{K}\right| \left|\vec{r}'\right|\right) Y_m^{\ell*}\left(\widehat{\vec{k}+\vec{K}}\right) Y_m^\ell\left(\hat{r}'\right) \qquad (3.2)$$

$j_\ell(x)$ is the Bessel function of order ℓ. Requiring this at the sphere boundary (where $\vec{r}' = \vec{R}_\alpha$, which defines $\vec{R}_\alpha$) to be equal to the ℓm-part of Eq. 3.1 easily yields:

$$A_{\ell m}^{\alpha,\vec{k}+\vec{K}} = \frac{4\pi i^\ell e^{i\left(\vec{k}+\vec{K}\right)\cdot\vec{r}_\alpha}}{\sqrt{V} u_\ell^\alpha(\vec{R}_\alpha, E)} j_\ell\left(\left|\vec{k}+\vec{K}\right| R_\alpha\right) Y_m^{\ell*}(\widehat{\vec{k}+\vec{K}}) \qquad (3.3)$$

This uniquely defines the $A_{\ell m}^{\alpha,\vec{k}+\vec{K}}$, apart from the still undetermined E. In principle there are an infinite number of terms in Eq. 3.2, which would force us to use an infinite number of $A_{\ell m}^{\alpha,\vec{k}+\vec{K}}$

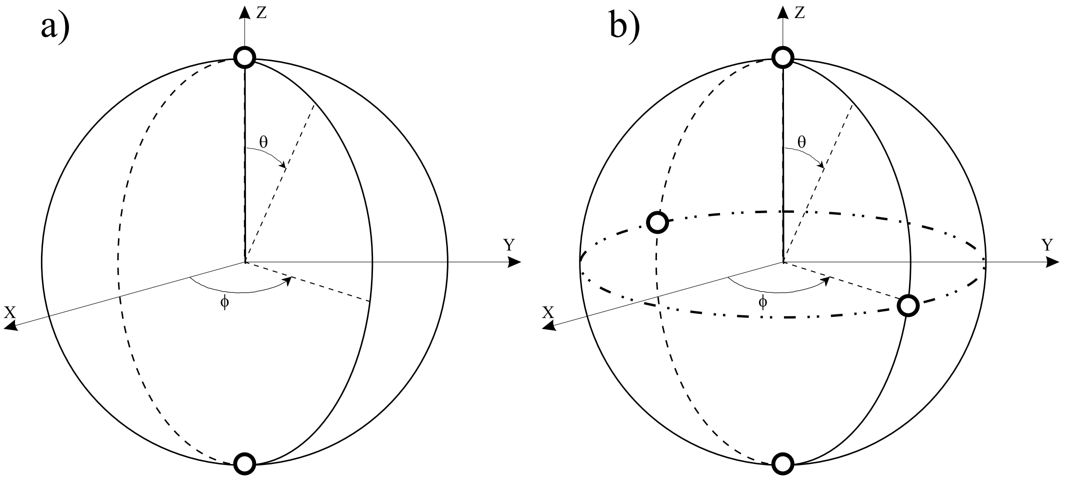

Figure 3.2: *In a) and b), a great circle for some arbitrary value of* ϕ *is drawn. The points on the circle for which a)* $Y^{\ell=2}_{m=2}$ *or b)* $Y^{\ell=2}_{m=1}$ *is zero, are indicated by white points. In a) there are no other points on which* $Y^{\ell=2}_{m=2}$ *is zero. In b) all other such points for* $Y^{\ell=2}_{m=1}$ *fall on a horizontal circle with* $\theta = \pi/2$ *(point-dashed circle). In both cases, the number of zeros along a great circle is at most* $2 \cdot \ell$*. Remember:* $Y^{\ell=2}_{m=2} = \frac{1}{4}\sqrt{\frac{15}{2\pi}}\,\sin^2\theta\, e^{2i\phi}$ *and* $Y^{\ell=2}_{m=1} = -\sqrt{\frac{15}{8\pi}}\sin\theta\,\cos\theta\, e^{i\phi}$*.*

in order to create the matching. In practice we will have to truncate at some value ℓ_{max}. What would be a reasonable choice? For a given ℓ_{max}, $Y^{\ell_{max}}_{m}(\theta,\,\phi)$ can have at most $2\ell_{max}$ nodes along a great circle (i.e. $\theta = 0 \rightarrow 2\pi$ for any fixed ϕ) of the α-sphere (see fig. 3). Converted into nodes per unit of length, this is $2\ell_{max}/(2\pi R_\alpha) = \ell_{max}/(\pi R_\alpha)$. If a plane wave should match with this, there should be plane waves with at least a similar number of nodes per unit of length available. The plane wave with the shortest period $2\pi/K_{max}$ has $2/(2\pi/K_{max}) = K_{max}/\pi$ nodes per unit of length. The cut-off for the plane waves (K_{max}) and for the angular functions (ℓ_{max}) are of comparable quality if the number of nodes per unit of length is identical. This yields the condition $R_\alpha K_{max} = \ell_{max}$. This allows to determine a good ℓ_{max} for a given K_{max}. A finite value for ℓ_{max} means that for each APW the matching at the sphere boundaries will not be exact, but good enough to work with. It is not useful to make ℓ_{max} larger than the condition $R_\alpha K_{max}$ requires, as it would lead to unstable behaviour at the sphere boundary (you can compare this with fitting a polynomial of high order through a limited number of points: the fit will be ‘perfect’, but not very meaningful). Therefore, it is also clear now that the muffin tin radii for the different atoms should not be too different: if they were, a value for ℓ_{max} that is suitable for each atom would not exist.

Now you should be able to visualize the meaning of a single APW $\phi^{\vec{k}}_{\vec{K}}(\vec{r},\,E)$ of equation 3.1: it is an oscillating function that runs through the unit cell. Whenever it encounters an atom on its path, the simple oscillating behaviour is changed into something more complex inside the muffin tin sphere of that atom. Nevertheless, the function values inside and outside the sphere smoothly match, which is taken care of by a set of $\sum_{\ell=1}^{\ell_{max}} 2\ell_{max}+1$ coefficients $A^{\alpha,\vec{k}+\vec{K}}_{\ell m}$ that is different for each atom (the atom determines α, the APW under consideration determines $\vec{k}$ and $\vec{K}$, all ℓ up to ℓ_{max} are present, with the corresponding values of m).

At first sight, it looks like we can now use the APW’s as a basis set, and proceed in the same

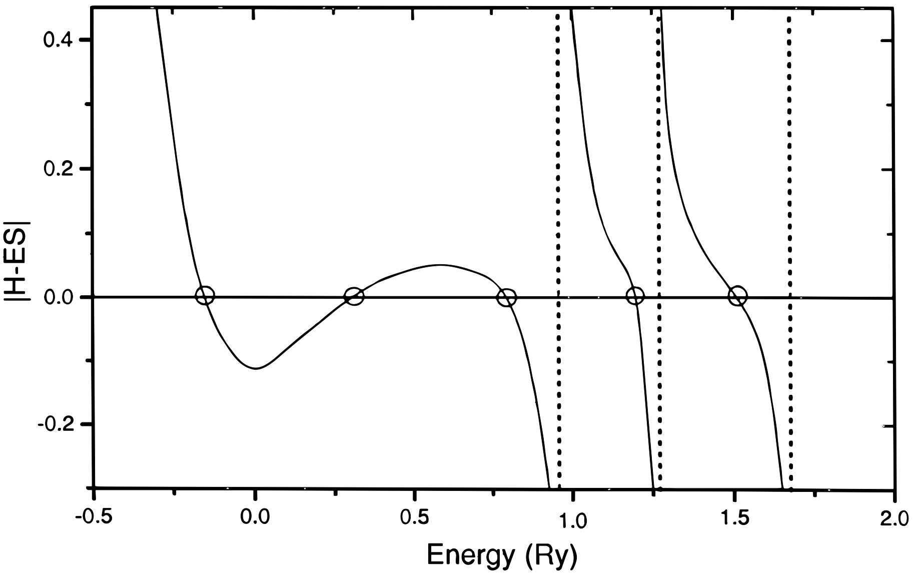

Figure 3.3: *A suggestive visualization of how the roots of the secular equation are searched. (picture copied from K. Schwarz et al., Computer Physics Communications 147 (2002) 71)*

way as for the plane wave basis set in order to determine the coefficients $c_{\vec{K}}^{n,\vec{k}}$ in the expansion of the searched eigenfunction. However, this does not work. We did not settle the parameter E yet. It turns out that in order to describe an eigenstate $\psi_{\vec{k}}^{n}(\vec{r})$ accurately with APW's, one has to set E equal to the eigenvalue (or band energy) $\epsilon_{\vec{k}}^{n}$ of that state. But this is exactly what we are trying to determine! We are hence forced to start with a guessed value for $\epsilon_{\vec{k}}^{n}$ and take this as E. Now we can determine the APW's, and construct the hamiltonian matrix elements and overlap matrix (the APW's are not orthogonal). The secular equation is determined, and our guessed $\epsilon_{\vec{k}}^{n}$ should be a root of it. Usually it is not, hence we have to try a second guess. Due to this new E, the APW's have to be determined again, and similarly for all matrix elements. With the help of root determination algorithms, this guessing continues until a root – say $\epsilon_{\vec{k}}^{(n=1)}$ – is found. And then the whole procedure starts over for $\epsilon_{\vec{k}}^{(n=2)}$, etc. (see fig. 3 for a suggestive visualization of the roots of a secular equation, and fig. 3 for a flow chart of the APW method).

In practice, $K_{max} \approx 3.5\ au^{-1}$ is needed for sufficient accuracy. This is less than the typical value of 5.5 for plane waves and pseudopotentials. As on page 12, the basis set size can be estimated to be about $P = 131$ for APW, compared to roughly $P = 270$ for plane waves. The calculation time (mainly determined by matrix diagonalization) scales with the third power of the basis set size, which would suggest APW to be 10 times faster than pseudopotentials. However, with a plane wave basis set, P eigenvalues are found by a single diagonalization, while with APW one diagonalization is needed for every eigenvalue. This makes the APW method inherently slow, much slower than the pseudopotential method.

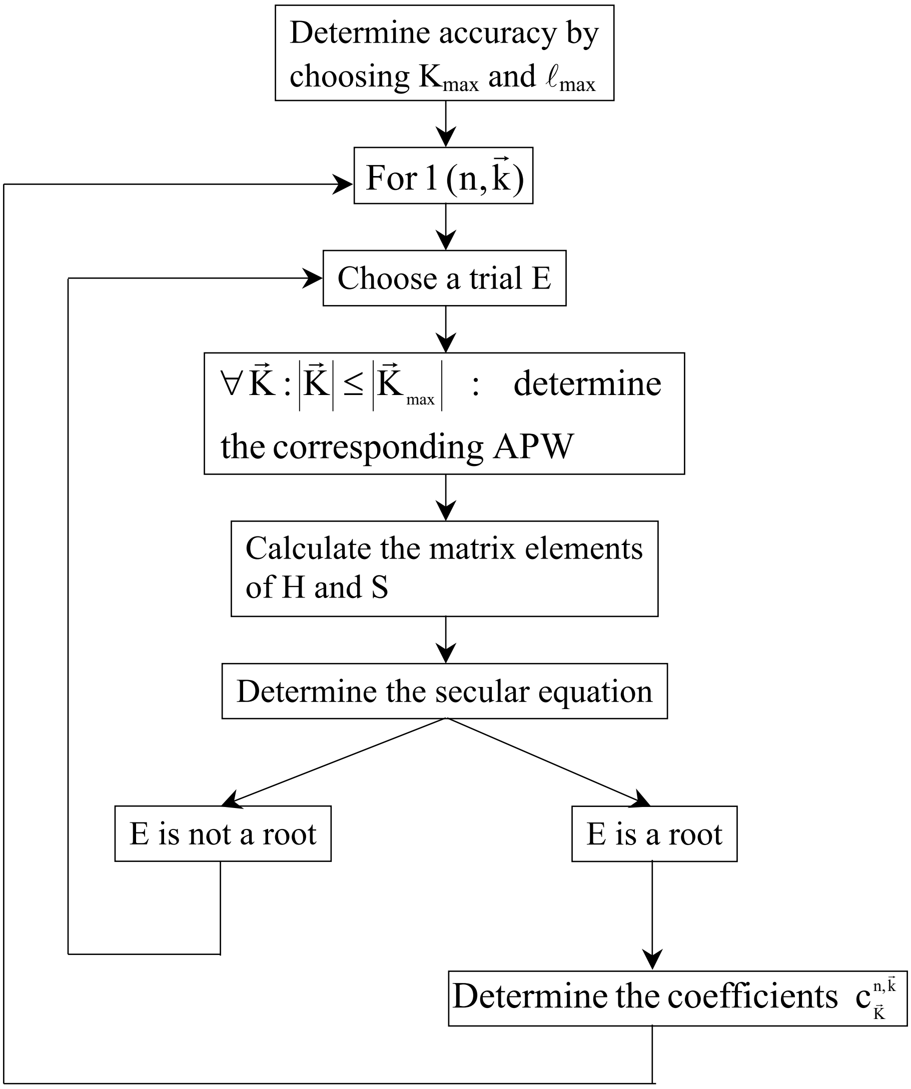
Determine accuracy by choosing K_{max} and ℓ_{max}
For l $(n, \vec{k})$
Choose a trial E
$\forall \vec{K} : |\vec{K}| \leq |\vec{K}_{max}|$: determine the corresponding APW
Calculate the matrix elements of H and S
Determine the secular equation
E is not a root
E is a root
Determine the coefficients $c_{\vec{K}}^{n,\vec{k}}$

Figure 3.4: *Flowchart of the APW method.*

Chapter 4

The LAPW method

4.1 The regular LAPW method

The problem with the APW method was that the $u_\ell^\alpha(r',\, E)$ have to be constructed at the – yet unknown – eigenenergy $E = \epsilon^n_{\vec{k}}$ of the searched eigenstate. It would be helpful if we were able to recover $u_\ell^\alpha(r', \epsilon^n_{\vec{k}})$ on the fly from known quantities. That is exactly what the *Linearized Augmented Plane Wave* method enables us to do. If we have calculated u_ℓ^α at some energy E_0, we could make a Taylor expansion to find it at energies not far away from it:

$$u_\ell^\alpha(r',\, \epsilon^n_{\vec{k}}) \;=\; u_\ell^\alpha(r',\, E_0) + (E_0 - \epsilon^n_{\vec{k}}) \underbrace{\left. \frac{\partial u_\ell^\alpha(r',\, E)}{\partial E} \right|_{E=E_0}}_{\dot{u}_\ell^\alpha(r', E_0)} + O(E_0 - \epsilon^n_{\vec{k}})^2 \tag{4.1}$$

Substituting the first two terms of the expansion in the APW for a fixed E_0 gives the definition of an LAPW. This has a price: the energy difference $(E_0 - \epsilon^n_{\vec{k}})$ is unknown, and hence a yet undetermined $B_{\ell m}^{\alpha,\, \vec{k}+\vec{K}}$ has to be introduced:

$$\phi^{\vec{k}}_{\vec{K}}(\vec{r}) \;=\; \begin{cases} \frac{1}{\sqrt{V}} e^{i(\vec{k}+\vec{K})\cdot\vec{r}} & \vec{r} \in I \\ \\ \sum_{\ell,\, m} \left(A_{\ell m}^{\alpha,\, \vec{k}+\vec{K}} u_\ell^\alpha(r', E_0) + B_{\ell m}^{\alpha,\, \vec{k}+\vec{K}} \dot{u}_\ell^\alpha(r', E_0) \right) Y_m^\ell(\hat{r}') & \vec{r} \in S_\alpha \end{cases} \tag{4.2}$$

In order to determine both $A_{\ell m}^{\alpha,\, \vec{k}+\vec{K}}$ and $B_{\ell m}^{\alpha,\, \vec{k}+\vec{K}}$, we will require that the function in the sphere matches the plane wave both in value and in slope at the sphere boundary. This can be done by using an expression similar to equation 3.2 and its radial derivative. This results in a 2×2 system from which both coefficients can be solved.

Equation 4.2 is not the final definition of an LAPW yet. Imagine we want to describe an eigenstate $\psi^n_{\vec{k}}$ that has predominantly p-character ($\ell = 1$) for atom α. This means that in its expansion in LAPW's, the $A_{(\ell=\mathbf{1})m}^{\alpha,\, \vec{k}+\vec{K}}$ are large. It is therefore advantageous to choose E_0 near the centre of the p-band (how this can be done will be made clear in the examples of Chapter 6). In this way, the $O(E_0 - \epsilon^n_{\vec{k}})^2$-term in equation 4.1 will remain small, and cutting after the linear term is certainly allowed. We can repeat this argument for every physically important ℓ (s-, p-, d- and f-states, i.e. up to $\ell = 3$) and for every atom. As a result, we should not choose one universal E_0, but a set of well-chosen $E_{1,\ell}^\alpha$ up to $\ell = 3$ (the meaning of the index '1' will become clear in section 4.2. For higher ℓ, a fixed value can be kept. The final definition of an LAPW is then:

$$\phi_{\vec{K}}^{\vec{k}}(\vec{r}) = \begin{cases} \frac{1}{\sqrt{V}} e^{i(\vec{k}+\vec{K})\cdot\vec{r}} & \vec{r} \in I \\ \sum_{\ell,m} \left(A_{\ell m}^{\alpha,\vec{k}+\vec{K}} u_\ell^\alpha(r', E_{1,\ell}^\alpha) + B_{\ell m}^{\alpha,\vec{k}+\vec{K}} \dot{u}_\ell^\alpha(r', E_{1,\ell}^\alpha) \right) Y_m^\ell(\hat{r}') & \vec{r} \in S_\alpha \end{cases} \tag{4.3}$$

With the $E_{1,\ell}^\alpha$ being fixed, the basis functions can be calculated once and for all. The same procedure as used for the plane wave basis set can now be applied. One diagonalization will yield P different band energies for this $\vec{k}$.

The accuracy of a plane wave basis set was determined by K_{max}. For the APW or LAPW basis set, it is not incorrect to use the same criterion. However, a better quantity to judge the accuracy here is the product $R_\alpha^{min} K_{max}$ between the smallest muffin tin radius and K_{max}. This can be understood as follows. If the smallest muffin tin radius is increased, the closest point a plane wave can come to a nucleus moves farther away from the nucleus. The part of the wave function that need not to be described with plane waves any more, in general will have displayed the steepest behaviour, steeper than anywhere else in the interstitial region (it was closest to the nucleus). Less plane waves are needed to describe the remaining, smoother parts of the wave function. K_{max} can be reduced, and a good rule of thumb is that the product $R_\alpha^{min} K_{max}$ should remain constant in order to have comparable accuracy. Reducing K_{max} means reducing the size of the matrices, and because matrix diagonalization is very expensive, a larger R_α^{min} can significantly reduce the computation time. R_α^{min} cannot be too large on the other hand, as the spherical harmonics are not suited to describe the wave functions in the region far away from the nuclei.

Compared to a plane wave basis set, the LAPW basis set can be much smaller. The required K_{max} turns out to be $K_{max} = \frac{7.5 \leftrightarrow 9.0}{R_\alpha^{min}} \approx 4\,\mathrm{au}^{-1}$, depending on the desired accuracy. This yields $P \approx 195$ as basis set size, compared to $P \approx 270$ for plane waves. The calculation time (mainly determined by matrix diagonalization) scales with the third power of the basis set size, which makes LAPW in this respect about 2 to 3 times faster than plane waves. There are other aspects however that slow down LAPW[1], such that in the end it is comparable in speed with plane waves.

4.2 LAPW with Local Orbitals (LAPW+LO)

It was not explicitly stated so far which electron states are calculated with the LAPW method. Does it make sense to calculate the 1s orbital of Fe in bcc-Fe? No, because this electron is extremely well bound to the nucleus (-514 Ry), and will behave almost exactly as if it were in a free Fe atom. Such a state is called a *core state*. The criterion for a core state is that it does not participate directly in chemical bonding with other atoms. Therefore, it must be contained entirely in the muffin tin sphere. States that leak out of the muffin tin sphere, are called *valence states*. Valence states participate in chemical bonds, and these states are treated by LAPW. Core states are treated as in free atoms, but subject to the potential due to the valence states.

When applying this definition, it frequently happens that states with the same ℓ but a different principal quantum number n are both valence states. For instance, due to hybridisation,

---

[1] The non-orthogonal basis set for instance.

Fe in bcc-Fe will have a non-negligible amount of 4p-character in its valence states that are about 0.2 Ry below the Fermi level. But the 3p-states that are 4.3 Ry below the Fermi level are not entirely confined in the core too. Such low-lying valence states are called *semi-core states*. It is not clear how $E^{Fe}_{1,(\ell=1)}$ should be chosen: close to 3p, close to 4p, at an intermediate value, . . .? None of the choices is optimal. This dilemma is solved by adding another type of basis function to the LAPW basis set, called a *local orbital* (LO). A local orbital is defined as:

$$\phi^{\ell m}_{\alpha,\,LO}(\vec{r}) = \begin{cases} 0 & \vec{r} \notin S_\alpha \\ \left(A^{\alpha,LO}_{\ell m} u^\alpha_\ell(r', E^\alpha_{1,\ell}) + B^{\alpha,LO}_{\ell m} \dot{u}^\alpha_\ell(r', E^\alpha_{1,\ell}) + C^{\alpha,LO}_{\ell m} u^\alpha_\ell(r', E^\alpha_{2,\ell})\right) Y^\ell_m(\hat{r}') & \vec{r} \in S_\alpha \end{cases} \tag{4.4}$$

A local orbital is defined for a particular ℓ and m, and for a particular atom α. A local orbital is zero in the interstitial region and in the muffin tin spheres of other atoms, hence its name *local* orbital. In the muffin tin sphere of atom α, the same $u^\alpha_\ell(r', E^\alpha_{1,\ell})$ and $\dot{u}^\alpha_\ell(r', E^\alpha_{1,\ell})$ as in the LAPW basis set are used, with as linearization energy $E^\alpha_{1,\ell}$ a value suitable for the highest of the two valence states (4p in our example)[2]. The lower valence state – that is much more free-atom-like – is sharply peaked at an energy $E^\alpha_{2,\ell}$. A single radial function $u^\alpha_\ell(r', E^\alpha_{2,\ell})$ at that same energy will be sufficient to describe it. Local orbitals are not connected to plane waves in the interstitial region, they have hence no $\vec{k}$- or $\vec{K}$-dependence. The three coefficients $A^{\alpha,LO}_{\ell m}$, $B^{\alpha,LO}_{\ell m}$ and $C^{\alpha,LO}_{\ell m}$ are determined by requiring that the LO is normalized, and has zero value and zero slope at the muffin tin boundary (= it does not leak out of the muffin tin sphere).

Adding local orbitals increases the LAPW basis set size. If for each atom local orbitals for p- and d-states are added, the basis set increases with 3+5=8 functions per atom[3] in the unit cell. This number is rather small compared to typical LAPW basis set sizes of a few hundred functions. The slightly increased computational time is a small price to be paid for the much better accuracy that local orbitals offer, and therefore they are always used.

---

[2] Linearization energies for two atoms that are equivalent, are taken to be equal.

[3] Indeed, the more atoms in the unit cell, the more LO's have to be added. In contrast to this, the number of LAPW's does not depend on the number of atoms in the unit cell, but – for a fixed $R^{min}_\alpha K_{max}$ and cell symmetry – on the volume of the unit cell, independent on how many atoms appear in it (more atoms mean just more sets of coefficients $A^{\alpha,\vec{k}+\vec{K}}_{\ell m}$ and $B^{\alpha,\vec{k}+\vec{K}}_{\ell m}$).

Chapter 5

The APW+lo method

5.1 The 'pure' APW+lo basis set

The problem with the APW method was the energy dependence of the basis set. This energy dependence could be removed in the LAPW+LO method, at the cost of a somewhat larger basis set size[1]. In the APW+lo method that will be described now, the basis set will be energy independent and still have the same size as in the APW method. In this sense, APW+lo combines the good features of APW and LAPW+LO.

The APW+lo basis set contains two kinds of functions. The first kind are APW's, with a set of fixed energies $E^{\alpha}_{1,\ell}$:

$$\phi^{\vec{k}}_{\vec{K}}(\vec{r}) = \begin{cases} \frac{1}{\sqrt{V}} e^{i(\vec{k}+\vec{K})\cdot\vec{r}} & \vec{r} \in I \\ \sum_{\ell,m} A^{\alpha,\vec{k}+\vec{K}}_{\ell m} u^{\alpha}_{\ell}(r', E^{\alpha}_{1,\ell}) Y^{\ell}_{m}(\hat{r}') & \vec{r} \in S_{\alpha} \end{cases} \tag{5.1}$$

We know from the discussion in Chapter 3 that with fixed energies this basis set does not give a good description of the eigenfunctions. The basis set is therefore augmented with a second type of functions. These are *local orbitals*, but another type as the one used in connection with the LAPW method. We abbreviate them therefore as 'lo' instead of as 'LO'. Their definition is:

$$\phi^{\ell m}_{\alpha,\,lo}(\vec{r}) = \begin{cases} 0 & \vec{r} \notin S_{\alpha} \\ \left(A^{\alpha,lo}_{\ell m} u^{\alpha}_{\ell}(r', E^{\alpha}_{1,\ell}) + B^{\alpha,lo}_{\ell m} \dot{u}^{\alpha}_{\ell}(r', E^{\alpha}_{1,\ell})\right) Y^{\ell}_{m}(\hat{r}') & \vec{r} \in S_{\alpha} \end{cases} \tag{5.2}$$

The same set of energies $E^{\alpha}_{1,\ell}$ is used as for the corresponding APW's (although this is not strictly needed). The two coefficients $A^{\alpha,lo}_{\ell m}$ and $B^{\alpha,lo}_{\ell m}$ are determined by normalization, and by requiring that the local orbital has zero value at the muffin tin boundary (not zero slope). Hence, both the APW and the local orbital are continuous at the sphere boundary, but for both their first derivative is discontinuous.

For accurate results, the APW+lo basis set appears to require a size that is comparable to the APW method ($K_{max} \approx 3.5$ au^{-1}, $P \approx 130$). This is less than in the LAPW+LO method

---

[1] Due to LAPW itself, and due to adding local orbitals.

($K_{max} \approx 4$ au^{-1}, $P \approx 200$). Nevertheless, P eigenvalues are obtained by a single diagonalization, as in LAPW+LO.

5.2 Mixed LAPW/APW+lo basis sets

The reason why LAPW needs a higher K_{max} than APW+lo can be traced back to a few states that are hard for LAPW. They are:

- Valence d- and f-states.
- States in atoms that have a muffin tin sphere that is much smaller than others spheres in the unit cell.

It is advantageous to treat those states with APW+lo, and keep using LAPW for all other states. Why? Using APW+lo for a state means that per atom $2\ell + 1$ local orbitals are added to the basis set. This makes an APW+lo basis set for the same $R_{\alpha}^{min} K_{max}$ considerably larger than the LAPW basis set. This is compensated by the fact that a lower $R_{\alpha}^{min} K_{max}$ is needed for accurate results, but nevertheless, it is better to use these extra basis functions only there where they are useful. Such an approach leads to a mixed LAPW/APW+lo basis set: for all atoms α and values of ℓ, equation 4.2 is used. But for one or more atoms α_0 ($r \in S_{\alpha_0}$) and one or more ℓ_0, equation 5.1 is used. Corresponding $\phi_{\alpha_0}^{\ell_0 m}$ according to equation 5.2 are then added to the basis set. Such a mixed basis set is the recommended choice in `WIEN2k`.

5.3 APW+lo with Local Orbitals (APW+lo+LO)

With APW+lo basis functions, the same problem with semi-core states appears as described in section 4.2. It will be remediated in the same way: by adding Local Orbitals (LO). The definition of a local orbital for APW+lo is:

$$\phi_{\alpha,LO}^{\ell m}(\vec{r}) = \begin{cases} 0 & \vec{r} \notin S_\alpha \\ \left(A_{\ell m}^{\alpha,LO} u_\ell^\alpha(r', E_{1,\ell}^\alpha) + C_{\ell m}^{\alpha,LO} u_\ell^\alpha(r', E_{2,\ell}^\alpha)\right) Y_m^\ell(\hat{r}') & \vec{r} \in S_\alpha \end{cases} \tag{5.3}$$

In contrast to the LO for LAPW, there is no derivative of u_ℓ^α here. The two coefficients $A_{\ell m}^{\alpha,LO}$ and $C_{\ell m}^{\alpha,LO}$ are determined by the requirement that the LO is normalized, and has zero value (not zero slope) at the sphere boundary.

Chapter 6

Examples for WIEN2k

In this chapter, various "real life" situations that you will encounter while working with WIEN2k are discussed in detail. It is assumed that you know how to set up a simple case. Working through these examples will give you a better understanding of several input and output files, and it will show you strategies for setting up meaningful but economical calculations.

6.1 Linearization energies in LAPW and APW+lo

In this example, mainly the input file `case.in1` for WIEN2k and some output that is related to the basis set will be discussed. The file `case.in1` contains crucial information for the LAPW and APW+lo methods. As our example, we take non-magnetic bcc Fe (B-type lattice with atom at (0, 0, 0), or alternatively space group Im$\bar{3}$m (nr. 229)) at the experimental lattice constant $a = 5.4169$ au, and with a muffin tin radius $R_{mt} = 2.2$ au.

6.1.1 Default initialization for LAPW

In this section, you will produce some starting information without much thinking. Don't try yet to understand what happens at this stage, reasoning will follow later.

Initialize the Fe case, using -6.0 Ry and GGA in `lstart`, 500 k-points[1] in the full first Brillouin zone (20 in the irreducible part) in `kgen`, and the mixing factor at 0.10 in `case.inm`. Use the following `case.in1`:

```
WFFIL        (WFPRI, SUPWF)
  7.00       10    4 (R-MT*K-MAX; MAX L IN WF, V-NMT
  0.30    4  0       (GLOBAL E-PARAMETER WITH n OTHER CHOICES, global APW/LAPW)
 1    0.30      0.000 CONT 0
 1   -3.97      0.005 STOP 0
 2    0.30      0.010 CONT 0
 0    0.30      0.000 CONT 0
K-VECTORS FROM UNIT:4    -7.0      2.5      emin/emax window
```

Compared to the default file, `emax` is changed from 1.5 to 2.5 (in order to see the Density Of States (DOS) over a larger energy window) and the LAPW-method is chosen. The latter is done by changing `NAPWL` at the end of the 4 lines from 1 to 0, but we will soon treat this in more detail. Now iterate this case to selfconsistency[2], and produce the DOS with the following `case.int`:

---

[1] This value is way to low for an accurate calculation, but it provides a fast example – ready while you wait.

```
bcc-Fe
 -0.50 0.001 3.000 0.003   EMIN, DE, EMAX, Gauss-broadening(>de)
   7                       NUMBER OF DOS-CASES specified below
   0    1   Cryst-tot      atom, case=column in qtl-header, label
   1    1   Fe-tot
   1    2   Fe-s
   1    3   Fe-p
   1    4   Fe-d
   1    7   Fe-f
   2    1   Inter
```

This produces the total DOS of the crystal (first line), the total DOS inside the Fe muffin tin sphere, The s-DOS of Fe inside the sphere, similarly for the p-, d- and f-DOS of Fe, and the total DOS in the interstitial region. These results are shown in fig. 6.1.1 [3]. The Fermi energy (:FER) – indicated by a vertical line in fig. 6.1.1 – is close to 0.66 Ry[4] (in your less precise calculation, you will find 0.63 Ry).

6.1.2 Choosing linearization energies for LAPW+LO

We will now scrutinize `case.in1` and interpret the DOS we obtained in fig. 6.1.1. The value '7.00' in the first line of `case.in1` is the product of the minimal muffin-tin radius R_{mt}^{min} (here 2.2 au) and the K_{max} discussed in Chapter 4. By '7.00', we have asked for $K_{max} = 7.00/2.2 = 3.2$ au^{-1}. Following the reasoning on page 12[5], this gives an estimate for the basis set size of 43. Now check the label :RKM in `case.scf`:

```
:RKM  : MATRIX SIZE   49LOs:   3  RKM= 6.85  WEIGHT= 8.00
```

This tells you that the actual number of LAPW's is 46 (the total basis set size (49) minus the number of local orbitals (3))[6]

---

[2] -cc 0.00005 is a sufficient criterion.

[3] The pictures are made with a k-mesh of 165 points. Your results with 20 k-points will look a little bit more shaky, but they should show the same main features.

[4] In a free atom, the zero of the energy scale is chosen such that the potential is zero at infinite distance from the atom. In an infinite lattice, this is not possible. In WIEN2k, the zero corresponds to the average potential in the interstitial region. Whenever an energy is expressed in Ry, it is relative to this zero. The DOS can be chosen to be plot against a Ry-scale (and then this zero is used) or against a scale in eV. In the latter case, the Fermi energy is taken zero by default.

[5] The volume of the first Brillouin zone for Fe can be calculated to be $V_{BZ} = 8\pi^3/V = 3.1212au^{-3}$. Here V is the volume of the unit cell in direct space, which is for the bcc structure one half of the volume of the conventional cube ($V = (5.4169au)^3/2$). This volume can always be found at the `:VOL` label in `case.scf`. Alternatively, V_{BZ} can be found by equations A.14 to A.16, considering that $\vec{a} = (a, 0, 0)$, $\vec{b} = (0, a, 0)$ and $\vec{c} = \left(\frac{a}{2}, \frac{a}{2}, \frac{a}{2}\right)$ with a the lattice constant.

[6] Our value of 43 is an estimate, not an exact value. The unit cell of Fe is quite small, hence the distances between $\vec{K}$-points are large. $K_{max} = 3.2$ au^{-1} is not very large. Our estimate approximates the number of $\vec{K}$-points in a sphere with radius K_{max}, and for a small radius the surface region is rather important relative

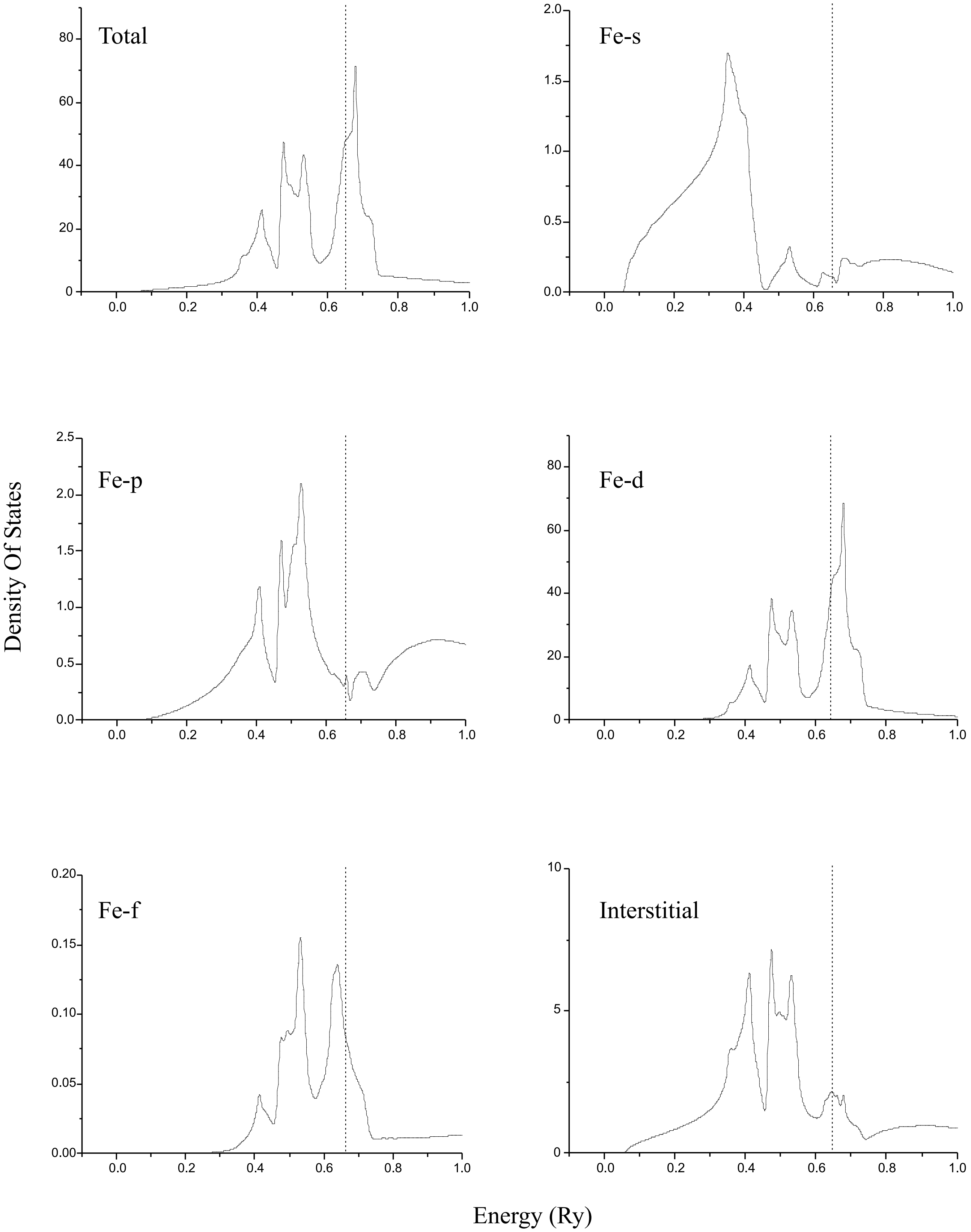

Figure 6.1: *DOS for nonmagnetic bcc-Fe. The dashed line is the Fermi energy.*

The number '10' to the right of $R_{mt}^{min}K_{max}$ is the ℓ_{max} from page 17. The '4' to the right of it, is not discussed in the preceding sections (it is highest ℓ in the expansion of the potential in *lattice harmonics*, a kind of spherical harmonics that is adapted to the symmetry of the lattice).

In the next line, the '0.30' is the E_ℓ^α ($\ell = 1, \ldots, \ell_{max}$, $\alpha =$ Fe) of equation 4.2, expressed in Ry[7]. The '4' to the right means that for 4 values of ℓ an exception will be made. These 4 exceptions will be listed below. The '0' to the right means that for all ℓ an LAPW is used, except if explicitly stated otherwise (with '1', an APW would be used everywhere).

Is the general choice $E_\ell^\alpha = 0.30$ Ry a good choice? Look at the total crystalline DOS of fig. 6.1.1. The valence region starts from about +0.05 Ry. As our calculations are by definition at 0 K (DFT is a ground state theory), we want to describe in the first place the region between 0.05 Ry and E_F=0.66 Ry. If we take a fixed energy at 0.30 Ry, then it falls right in the middle of this region of interest. This is what we want, as in this way linearization errors are minimized. If you would see that your valence band region covers for instance the region 0.35-0.90 Ry, a better choice for E_ℓ^α would be 0.62 Ry (the arithmetic mean of the valence band region). Applying this procedure in the case of bcc Fe would lead to $E_\ell^\alpha = 0.35$ Ry. This is so close to the original 0.30 Ry that it is hardly worth changing it, but we will nevertheless do it as a demonstration.

The following 4 lines list exceptions to this general choice of E_ℓ^α. The first number in each line is ℓ. The first line (with $\ell = 1$) gives E_ℓ^α for p-states (the p-states that are in the valence band region for Fe, are the 4p). A look at fig. 6.1.1 learns that the main weight of the occupied part of the p-band is roughly at 0.50 Ry. $E_{\ell=1}^\alpha = 0.50$ Ry will minimize the linearization error for the p-band. To the right there is a number '0.000'. It means that starting from the chosen value for $E_{\ell=1}^\alpha$, a better linearization energy will be searched by making steps of the size '0.000'. Hence, in this case no search will be done. The `CONT` means that even if no acceptable linearization energy would be found (in the case there would have been done a search), the program will nevertheless continue.

The second line we skip for a while. The third line is for $\ell = 2$ (Fe-3d). From looking at fig. 6.1.1, a good value for $E_{\ell=2}^\alpha$ would be about 0.50 Ry. Here, a search was asked, with a step of 0.010 Ry. What has been the result of that search? Look in the last iteration of `case.scf` for this block:

```
ATOMIC SPHERE DEPENDENT PARAMETERS FOR ATOM  Fe
OVERALL ENERGY PARAMETER IS    0.3000
OVERALL BASIS SET ON ATOM IS LAPW
E( 1)=    0.3000
   LAPW
E( 1)=   -3.3025   E(BOTTOM)=   -3.375   E(TOP)=   -3.230
   LOCAL ORBITAL
E( 2)=    0.5200   E(BOTTOM)=    0.200   E(TOP)=    0.840
   LAPW
E( 0)=    0.3000
   LAPW
```

---

to the volume. Details of the reciprocal lattice will determine whether certain $\vec{K}$ are just inside or just outside this sphere. For larger unit cells (try with the Fe lattice constant increased to 6.5 au) or larger K_{max} (try with $R_{mt}^{min}K_{max} = 9$), the estimate reproduces almost the exact basis set size. A question to test yourself: explain why in both of these cases the basis set size increases (77 and 92, respectively).

[7] Remember: using the same energy scale as for the Fermi energy.

There you see that in the first line no search was done for $E^{\alpha}_{\ell=1}$ and the 0.30 Ry was kept. The second line we skipped. In the third line, the bottom of the d-band was determined as 0.20 Ry. If you blow up your own version of fig. 6.1.1, you will see the d-band starts at about 0.05 Ry and gets appreciable values near 0.14 Ry [8]. The top of the d-band is determined as 0.84 Ry, which is also acceptable from fig. 6.1.1. The arithmetic mean 0.52 Ry is taken as the linearization energy. This is close to our estimate. It does not matter whether we will take 0.50 or 0.52 Ry as $E^{\alpha}_{\ell=2}$. Both are close enough to the center of the band.

The next line deals with $\ell = 0$ (Fe-4s). From fig. 6.1.1, a good linearization energy seems to be $E^{\alpha}_{\ell=0} = 0.30$ Ry (this is accidentally very close to the overall E^{α}_{ℓ}).

There was one line we skipped. It had a label $\ell = 1$, and follows the line that set $E^{\alpha}_{\ell=1}$. Such a second line is used to define a local orbital (LO) according to equation 4.4. In this case, the local orbital will describe Fe-3p states. The energy from the first line with $\ell = 1$ is used as $E^{\alpha}_{1,(\ell=1)}$. The current line is used to specify $E^{\alpha}_{2,(\ell=1)}$. In all cases, the energy suggested by the program can be accepted for a LO. Local orbitals are much tighter bound than valence states, and therefore their energy is sharply defined. You can see from the output in `case.scf` given above that bottom and top of the 3d-band are very close together, almost uniquely defining $E^{\alpha}_{2,(\ell=1)}$. This is also the reason why a label '`STOP`' is used: if the program does not succeed to find such a well-defined energy, something must be wrong. The calculation will then terminate.

There are 3 possible m-values for $\ell = 1$. Therefore 3 LO's will be added to the basis set, one for each m. If there would have been another Fe atom at an inequivalent position, the same would happen there, increasing the number of LO's. In the output from `case.scf` above, indeed '`LOs:    3`' was reported.

It is now also clear why no LO's are needed for the other possible candidate, Fe-3s: these states will be even better bound than -3.7 Ry. Apparently they are at energies lower than -6.0 Ry, which was during the initialization in `lstart` used as the criterion to distinguish between core and valence states. States below -6.0 Ry are core states, and are not described by the the LAPW+LO basis.

The last line in `case.in1` tells that the $\vec{k}$-vectors at which eigenvalues and eigenfunctions are calculated, are read from the default file created by `kgen`. And that eigenvalues are searched in an energy region between -7.0 Ry and +2.5 Ry.

After this analysis, we can construct a better `case.in1` that will lead to a more precise calculation. One possibility is:

```
WFFIL        (WFPRI, SUPWF)
  7.00       10    4 (R-MT*K-MAX; MAX L IN WF, V-NMT
  0.35    4  0       (GLOBAL E-PARAMETER WITH n OTHER CHOICES, global APW/LAPW)
 1    0.50      0.000 CONT 0
 1   -3.97      0.005 STOP 0
 2    0.50      0.010 CONT 0
 0    0.30      0.000 CONT 0
K-VECTORS FROM UNIT:4    -7.0      2.5      emin/emax window
```

---

[8] See the Usersguide, section about `lapw1`, to see how this is determined: the E_b for which $u_{\ell=2}(r, E_b)$ has zero slope at $r = R_{mt}$ is taken as the bottom of the band (bonding state). The E_t for which $u_{\ell=2}(r, E_t)$ has zero value at $r = R_{mt}$ is taken as the top of the band (antibonding state).

The changes in $E^\alpha_{\ell=1}$ will influence (improve) our calculation. The changes in the global E^α_ℓ and in $E^\alpha_{\ell=0}$ were zero or small. If this would not have been the case, they would have improved the calculation as well. The change in $E^\alpha_{\ell=2}$ will have no effect, it only avoids some searching. Alternatively, we could even omit the searching by putting:

```
 2    0.52       0.000 CONT 0
```

In this specific case, where accidentally $E^\alpha_{\ell=0} \approx E^\alpha_\ell$, we could also omit the line with $\ell = 0$ (and change the number of exceptions from 4 to 3).

There is another way in which the accuracy of the calculation can be improved[9]. In the output from the initialization program `lstart`, you can see that the Fe-3s states are close to the energy -6.0 Ry that separates core and valence states:

```
           19          350
   19   3.65E-07     1.918188E+00
     -514.09504     -514.09427
      -60.04156      -59.96008
      -51.73511      -51.67440
      -50.81845      -50.75597
       -6.80000       -6.61035
       -4.38376       -4.19666
       -4.27103       -4.08485
       -0.47568       -0.30733
       -0.46495       -0.29712
       -0.37634       -0.30776
```

The left column is for spin-up, the right for spin-down. Both columns list energies (in Ry, relative to the zero-level for a free atom) for Fe-states in the order that can be found at the beginning of the file. In this case: 1s, 2s, 2p*, 2p, 3s, 3p*, 3p, 3d*, 3d, 4s. The '*' indicates states with a different relativistic quantum number. The Fe-3s (6^{th} line from the bottom, energy of about -6.7 Ry) are considered as core states, but they are close to the separating -6.0 Ry. This is confirmed if we look at the energy of this state in the crystal (in Ry, relative to the zero-level in the crystal) in `case.scf`:

```
1.ATOM              5 CORE STATES
:1S 01:  1S            -513.202669 Ry
:2S 01:  2S             -59.139463 Ry
:2PP01:  2P*            -50.840718 Ry
:2P 01:  2P             -49.923370 Ry
:3S 01:  3S              -5.784221 Ry
```

If you give -7.0 Ry as input for `lstart`, then the Fe-3s will be taken as valence states. In `case.outputst`, you will see these lines (with -6.0 Ry):

---

[9] Whenever we speak here about 'improving accuracy' it means 'as far as the linearization energies and local orbitals are concerned'. Of course you should also consider the size of the basis set itself, by the choice of $R^{min}_{mt} K_{max}$, and the sampling of the first Brillouin zone by the size of the k-mesh.

```
 TOTAL CORE-CHARGE:                    12.00000000005831
 TOTAL CORE-CHARGE INSIDE SPHERE:      11.99785897382221
```

that with -7.0 Ry will change into:

```
 TOTAL CORE-CHARGE:                    10.00000000005758
 TOTAL CORE-CHARGE INSIDE SPHERE:      9.999999992285076
```

Instead of 12 core electrons, there are only 10 now (two Fe-3s are considered as valence now). A charge of 0.00214 electrons was leaking out of the core with the Fe-3s included. If they are valence electrons, the remaining 10 core electrons let only a charge of 0.000000008 leak out of the core.

The program will suggest a new `case.in1` now, including an LO for Fe-3s. After incorporating our previous changes, it will look like:

```
WFFIL        (WFPRI, SUPWF)
  7.00       10    4 (R-MT*K-MAX; MAX L IN WF, V-NMT
  0.35    5  0       (GLOBAL E-PARAMETER WITH n OTHER CHOICES, global APW/LAPW)
 0    0.30      0.000 CONT 0
 0   -6.50      0.005 STOP 0
 1    0.50      0.000 CONT 0
 1   -3.97      0.005 STOP 0
 2    0.52      0.000 CONT 0
K-VECTORS FROM UNIT:4    -8.0      2.5      emin/emax window
```

A last noteworthy point is, that it can sometimes be useful to increase the basis set by extra LO's with an energy in the region way above the Fermi energy. This increases the flexibility of the basis set in that energy region, and allows a better description of the unoccupied states. For Fe, this extra flexibility turns out to be necessary in order to obtain a good total energy (energy-versus-volume curves without this extra LO display scatter, with the extra LO they are smooth). The energy for such an LO has to be set fixed. Check that you will have 9 local orbitals now in your calculation (in the line with `:RKM`, you will read `LOs:    9`): the 3 that where there from the 3p-electrons, 1 which we added for the 3s, and 5 from the high-lying 4d. This would then be our final `case.in1`:

```
WFFIL        (WFPRI, SUPWF)
  7.00       10    4 (R-MT*K-MAX; MAX L IN WF, V-NMT
  0.35    6  0       (GLOBAL E-PARAMETER WITH n OTHER CHOICES, global APW/LAPW)
 0    0.30      0.000 CONT 0
 0   -6.50      0.005 STOP 0
 1    0.50      0.000 CONT 0
 1   -3.97      0.005 STOP 0
 2    0.52      0.000 CONT 0
 2    1.50      0.000 CONT 0
K-VECTORS FROM UNIT:4    -8.0      2.5      emin/emax window
```

Really accurate LAPW+LO calculations for bcc-Fe can be made with this `case.in1`, and with $R_{mt}^{min} K_{max} = 9.0$ and 406 points in the k-mesh.

6.1.3 Choosing linearization energies for APW+lo

The procedure for APW+lo is very similar as for LAPW+LO. According to section 5.2, we can keep using LAPW+LO for most states. Only for Fe-3d we should switch to APW+lo. This leads to the following `case.in1`:

```
WFFIL        (WFPRI, SUPWF)
  7.00       10    4 (R-MT*K-MAX; MAX L IN WF, V-NMT
  0.35    6  0       (GLOBAL E-PARAMETER WITH n OTHER CHOICES, global APW/LAPW)
 0    0.30      0.000 CONT 0
 0   -6.50      0.005 STOP 0
 1    0.50      0.000 CONT 0
 1   -3.97      0.005 STOP 0
 2    0.52      0.000 CONT 1
 2    1.50      0.000 CONT 1
K-VECTORS FROM UNIT:4    -8.0      2.5      emin/emax window
```

It doesn't matter very much whether you take APW+lo or LAPW for the high-lying d local orbital. But for the regular 3d states it has an effect. First of all, you can see in the `:RKM`-line that 5 functions has been added to the basis set, due to 5 additional lo's. Indeed, in the APW+lo method, $2\ell + 1$ local orbitals of the type as defined in Eq. 5.2 are added for each regular state, not for deeper states with the same n only, as was the case in LAPW. APW+lo increases the basis set, which should slow down your calculation. You can check for the accurate calculation ($R_{mt}^{min} K_{max} = 9.0$ and 406 points in the k-mesh) that it needs slightly more time for 1 iteration than LAPW. However, if you do a series of calculations with different $R_{mt}K_{max}$, with APW+lo you will reach converged values at a lower $R_{mt}K_{max}$, which will substantially lower the basis set size (an the time) you need to obtain an accurate result (see also section 6.2).

6.1.4 Automatic determination of linearization energies

There is a way to determine the linearization energies needed in `case.in1` automatically. It requires less thinking than doing it yourself as explained in the previous sections, but it can also be a source of errors if the automatic procedure does not do exactly what you expected it to do. If you start calculating on a new compound, it might be wise to think about the linearization energies yourself first. Later on, when you are familiar with the kind of DOS that can be expected, you can switch to the automatic procedure.

The automatic procedure first needs to know what are true valence states, and what are states that are treated as LO's (semi-core states). The semi-core states are separated by a gap from the continuous region of valence states. By the following line in `case.in2`:

```
      -9.0      14.0 0.50 0.05                EMIN, NE, ESEPERMIN, ESEPER0
```

the program is told that it should start to look at $E^0_{seper} = 0.05$ Ry below the Fermi energy, for a gap between states with a width of at least $E^{min}_{seper} = 0.50$ Ry. In our Fe-example where the Fermi energy is about 0.66 Ry, it will start to search at 0.66-0.05=0.61 Ry. Then it goes downwards, and at about 0.05 Ry it finds a gap that extends down to -3.7 Ry: above this gap are the true valence states, below are the semi-core states that are treated by LO's.

In `case.scf`, the line labeled with `:EPLxx` ('xx' is the number of the atom) keeps track of the average energy for every ℓ for atom 'xx': the arithmetic mean between the bottom and top of the band. The line labeled with `:EPHxx` does the same for the valence states, but unlike for the semi-core states the average energy is now the arithmetic mean between the bottom of the band and the Fermi energy (because higher states are not occupied).

All this is done always, but has no effect unless with the run-command the switch `-in1new N` is specified, with `N` a number. With this switch, after `N` iterations a new `case.in1` is made based on the information in the `:EPLxx` and `:EPHxx` from the previous iteration. The old `case.in1` is copied to `case.in1_orig`. Using the automatic procedure in our case for Fe, would lead for LAPW to

```
WFFIL        (WFPRI, SUPWF)
  7.00       10    4 (R-MT*K-MAX; MAX L IN WF, V-NMT
 .33385   5   0      global e-param with N other choices, napw
 0    0.290     0.000 CONT 0
 0   -5.775     0.000 CONT 0
 1    0.438     0.000 CONT 0
 1   -3.279     0.000 CONT 0
 2    0.505     0.000 CONT 0
 K-VECTORS FROM UNIT:4   -8.0       2.5     emin/emax window
```

which is close to the linearization values we determined ourselves. Be careful: the new `case.in1` will not include the manually added `LO` at 1.5 Ry. If you really need this `LO`, you should after having determined the linearization energies automatically, add the `LO` again:

```
WFFIL        (WFPRI, SUPWF)
  7.00       10    4 (R-MT*K-MAX; MAX L IN WF, V-NMT
 .33385   6   0      global e-param with N other choices, napw
 0    0.290     0.000 CONT 0
 0   -5.775     0.001 CONT 0
 1    0.438     0.000 CONT 0
 1   -3.279     0.001 CONT 0
 2    0.505     0.000 CONT 0
 2    1.500     0.000 CONT 0
 K-VECTORS FROM UNIT:4   -8.0       2.5     emin/emax window
```

Then you run the calculation again, but now without the `-in1new` switch. Furthermore, using `-in1new` will set `0.000` everywhere: no energy searching will be done anymore. This means that this file becomes possibly invalid – and should be recreated with `-in1new` – if you change the lattice constant. If you add highlying `LO`'s and run again without `-in1new`, don't forget to change the `0.000` to a small non-zero value for the deep-lying valence states (see above, the lines with `0.001`). These sharply peaked states could move a tiny bit when adding the high-lying `LO`, and if you would not allow to research the linearization energy, the old value could easily fall entirely outside the peak.

6.2 Finding the best $R_{mt}K_{max}$ and k-mesh

We want to use a basis set that is large enough to be accurate, but only just large enough in order not to waste computer time. The same holds for the density of k-mesh. Therefore, you need to do some testing before you go to real calculations for your problem. Often it is possible to find a small analog to your problem, where you can relatively quickly do the tests. For instance, if you want to calculate properties of an impurity in a solid, you need a large supercell. But the pure solid is described by a much smaller cell. You can determine a good basis set and k-mesh size for this small cell, and extrapolate it to the supercell. *It is very important to determine a good basis set size and k-mesh before doing lengthy calculations and/or interpreting physical results.* If you don't, you might be wasting lots of computer time (if you use higher precision than you need) or you are producing nonsense instead of physics (if you use less precision than you need).

These tests can be done either before or after determination of the linearization energies, as $R_{mt}K_{max}$ and the k-mesh size are only very weakly dependent on the linearization energies[10] As an example, we take hcp-Cd, with the following `case.struct`:

```
hcp-Cd
H                            1
             RELA
  5.630252  5.630252 10.617619 90.000000 90.000000120.000000
ATOM= -1: X=0.33333333 Y=0.66666667 Z=0.25000000
          MULT= 2          ISPLIT= 4
      -1: X=0.66666667 Y=0.33333333 Z=0.75000000
Cd         NPT=  781  R0=0.00001000 RMT=    2.4000   Z: 48.0
LOCAL ROT MATRIX:    1.0000000 0.0000000 0.0000000
                     0.0000000 1.0000000 0.0000000
                     0.0000000 0.0000000 1.0000000
   0      NUMBER OF SYMMETRY OPERATIONS
```

First we will use LAPW. Testing as described in the preceding sections, leads to this `case.in1`:

```
WFFIL        (WFPRI, SUPWF)
  7.00       10    4 (R-MT*K-MAX; MAX L IN WF, V-NMT
 .19836   5   0      global e-param with N other choices, napw
 0    0.168     0.000 CONT 0
 0   -7.144     0.000 CONT 0
 1    0.274     0.000 CONT 0
 1   -4.275     0.000 CONT 0
 2   -0.191     0.000 CONT 0
K-VECTORS FROM UNIT:4   -9.0       2.0      emin/emax window
```

As a start, we take a low but safe value of $R_{mt}K_{max} = 7.0$. Calculate the total energy (`:ENE`) and a sensitive quantity as the main component of the electric-field gradient tensor (`:EFG`) as a function of the size of the k-mesh, where this size is roughly doubled at each step. This gives fig. 6.2, which shows that 912 k-points are enough for a total energy that is accurate up to about

---

[10] For the same $R_{mt}K_{max}$ and k-mesh, using the default linearization energies can lead to slightly different values when compared to the optimized ones (the latter values being more accurate), but they will stabilize at the same $R_{mt}K_{max}$ and k-mesh if you increase the latter.

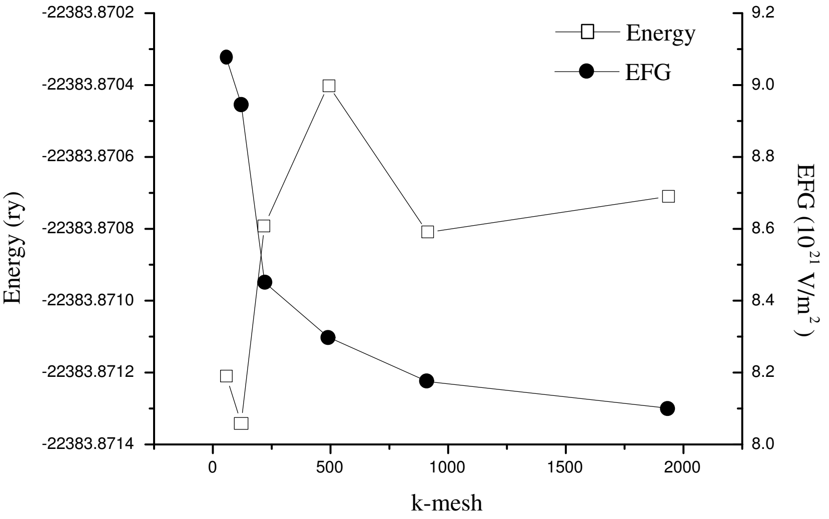

Figure 6.2: *Total energy and electric-field gradient as a function of k-mesh size for hcp-Cd (*$R_{mt}K_{max} = 7.0$*).*

1 mRy or better, and a field gradient accurate up to a few percent. Note that the behaviour of the total energy as a function of the number of k-points is not smooth: the number of k-points is not a variational parameter, and therefore it yields total energies that go in an erratic way towards the exact value.

Having determined a good k-mesh, we now change $R_{mt}K_{max}$ from 6.0 to 9.5 in steps of 0.5 (fig. 6.2). The number of plane waves is a variational quantity, and therefore the total energy drops monotonicly. After $R_{mt}K_{max} = 8.5$ it starts going up, however. This is a signal that the basis set became *overcomplete*: there are much more basis functions available than needed to describe this situation. Consequently, there are many possibilities how to expand the density into this basis set: the coefficients become linearly dependent on each other. This leads to numerical inaccuracies, and the final result is *less* accurate, in spite of the larger basis set. For this case, an optimal choice seems to be $R_{mt}K_{max} = 8.0$

Now we examine the influence of using APW+lo for some orbitals (fig. 6.2). We use APW+lo either for the valence d-orbitals only:

```
WFFIL        (WFPRI, SUPWF)
  7.00       10    4 (R-MT*K-MAX; MAX L IN WF, V-NMT
 .19836   5   0      global e-param with N other choices, napw
 0    0.168     0.000 CONT 0
 0   -7.144     0.000 CONT 0
 1    0.274     0.000 CONT 0
 1   -4.275     0.000 CONT 0
 2   -0.191     0.000 CONT 1
K-VECTORS FROM UNIT:4   -9.0       2.0      emin/emax window
```

or for all valence orbitals:

```
WFFIL        (WFPRI, SUPWF)
```

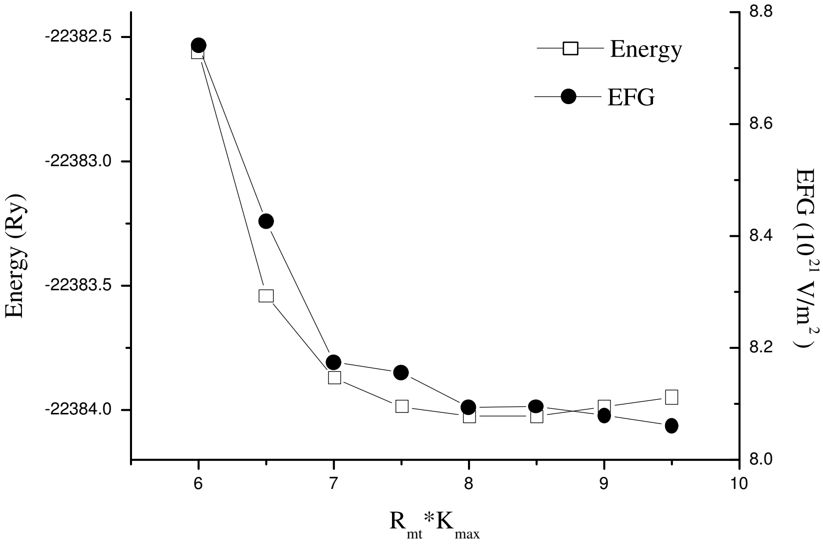

Figure 6.3: *Total energy and electric-field gradient as a function of* $R_{mt}K_{max}$ *for hcp-Cd (912 k-points).*

```
  7.00       10    4 (R-MT*K-MAX; MAX L IN WF, V-NMT
 .19836   5   0      global e-param with N other choices, napw
 0    0.168     0.000 CONT 1
 0   -7.144     0.000 CONT 1
 1    0.274     0.000 CONT 1
 1   -4.275     0.000 CONT 1
 2   -0.191     0.000 CONT 1
K-VECTORS FROM UNIT:4   -9.0       2.0      emin/emax window
```

From fig. 6.2, it is seen that in all three ways the same value is reached for the electric field gradient at the largest basis set size. For LAPW, one comes close enough to this value for $R_{mt}K_{max} \geq 8.0$. If APW+lo is used for the d-orbitals, $R_{mt}K_{max} \geq 7.0$ is sufficient. And if we use APW+lo for all valence orbitals, even $R_{mt}K_{max} \geq 6.0$ is acceptable. For $R_{mt}K_{max} = 8.0$ and LAPW, the basis set has a size of 181 (argue for yourself why there are 8 local orbitals here):

```
RKM  : MATRIX SIZE  181LOs:   8  RKM= 7.81  WEIGHT= 1.00  PGR:
```

For $R_{mt}K_{max} = 6.0$ and APW+lo for all states, there are only 95 basis functions (argue for yourself why there are 26 local orbitals):

```
RKM  : MATRIX SIZE   95LOs:  26  RKM= 5.68  WEIGHT= 1.00  PGR:
```

This reduction of the basis set size by a factor of 1.9, means a reduction in computation time by a factor of $1.9^3 \approx 7$, almost one order of magnitude.

The tests in this example have been done for the electric-field gradient (which is never bad: it is a sensitive quantity, and if this one is accurate, other quantities will usually be accurate

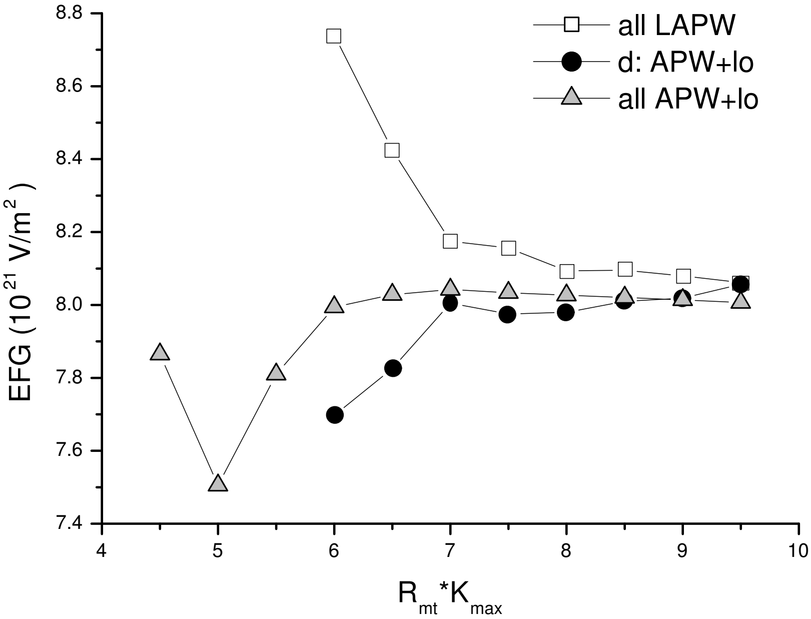

Figure 6.4: *Electric-field gradient as a function of* $R_{mt}K_{max}$ *for hcp-Cd (912 k-points), using LAPW for all orbitals, APW+lo for the valence d-orbitals, and APW+lo for the valence s-, p- and d-orbitals.*

too). For your own problems, you should check the quantities that are of interest for you. It is always wise to check not only one quantity, but a few of them. It could happen that a quantity converges by accident.

Appendix A

Fourier transforms, plane waves, the reciprocal lattice and Bloch's theorem

A.1 Fourier transforms

Most often, one's first encounter with Fourier transforms is in the context of time dependent functions. The Fourier transform $\mathcal{F}$ of a function $f(t)$ is a function $F(\omega)$ in frequency domain, and is defined as:

$$F(\omega) \;=\; \mathcal{F}\{f\} \;=\; \int_{-\infty}^{\infty} f(t)e^{-i\omega t}dt \tag{A.1}$$

$$f(t) \;=\; \mathcal{F}^{-1}\{F\} \;=\; \frac{-1}{2\pi}\int_{-\infty}^{\infty} F(\omega)e^{i\omega t}d\omega \tag{A.2}$$

If one would like to construct $f(t)$ by adding different functions $e^{i\omega t}$, then $F(\omega)$ tells what the weight of each function is. For instance, for $f(t) = \cos(\omega_0 t)$ it can be shown that:

$$F(\omega) \;=\; \frac{1}{2}\delta(\omega_0 + \omega) + \frac{1}{2}\delta(\omega_0 - \omega) \tag{A.3}$$

(see fig. A.1) Therefore, $\cos(\omega_0 t)$ must be equal to a sum of two $e^{i\omega t}$-functions, each with a weight given above:

$$\cos(\omega_0 t) \;=\; \frac{1}{2}e^{i(-\omega_0)t} + \frac{1}{2}e^{i\omega_0 t} \tag{A.4}$$

which indeed is true, considering the definition:

$$e^{i\omega t} \;=\; \cos(\omega t) + i\sin(\omega t) \tag{A.5}$$

In this example, there is a discrete set of ω-values where $F(\omega)$ is non-zero. In general, this is true for any *periodic* function: its Fourier transform is nonzero at a discrete set of frequencies (although there might be an infinite number of frequencies in the set). A periodic function can therefore be written as a *sum* of functions $e^{i\omega t}$. If $f(t)$ is aperiodic, $F(\omega)$ is non-zero over a continuous and often infinite range, and $f(t)$ can therefore be written only as an *integral* over $e^{i\omega t}$.

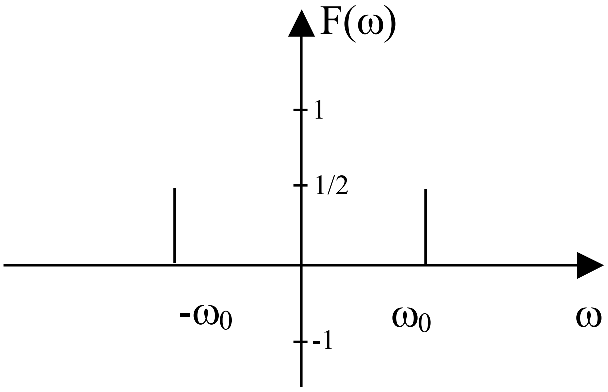

Figure A.1: *Fourier transform of* $\cos(\omega_0 t)$.

A.2 Plane waves

The concept of Fourier transforms can be translated for use with functions $f(\vec{r})$ in real space. The role of ω will be taken over by so-called *reciprocal vectors* $\vec{g}$ that have as dimension 1/length. In the same way as the infinite amount of position vectors continuously fills real space, so the infinite amount of reciprocal vectors continuously fills *reciprocal space*. Fourier and inverse Fourier transforms between real and reciprocal space are now defined as:

$$F(\vec{g}) = \mathcal{F}\{f\} = \int f(\vec{r})\, e^{-i\vec{g}\cdot\vec{r}}\, d^3\vec{r} \tag{A.6}$$

$$f(\vec{r}) = \mathcal{F}^{-1}\{F\} = \frac{-1}{2\pi}\int F(\vec{g})\, e^{i\vec{g}\cdot\vec{r}}\, d^3\vec{g} \tag{A.7}$$

A particular kind of function in real space is a *plane wave*, defined as:

$$f(\vec{r}) = e^{i\vec{g}_0\cdot\vec{r}} \tag{A.8}$$

with $\vec{g}_0$ any vector in reciprocal space. Convince yourself that a plane wave is periodic in space with wave vector $\vec{g}_0$ (= it takes the same value if you go along a distance $2\pi/|\vec{g}_0|$ parallel with $\vec{g}_0$), and that for any plane perpendicular to $\vec{g}_0$ it has the same value everywhere in that plane. The Fourier transform of a plane wave is non-zero at a single point $\vec{g}_0$ in reciprocal space only:

$$F(\vec{g}) = \int e^{i(\vec{g}_0-\vec{g})\cdot\vec{r}} d^3\vec{g} \tag{A.9}$$

$$= \delta(\vec{g}_0 - \vec{g}) \tag{A.10}$$

which is very logical, as we need only one function $e^{i\vec{g}\cdot\vec{r}}$ at $\vec{g} = \vec{g}_0$ with weight 1 to 'build' $e^{i\vec{g}_0\cdot\vec{r}}$. The shorter the period of the plane wave in real space, the further away the point indicated by $\vec{g}_0$ lies from the origin of reciprocal space.

Functions that are periodic in real space, will have a Fourier transform that is non-zero only at discrete points in reciprocal space (possibly an infinite number of points). For aperiodic functions in real space, the Fourier transform will be non-zero over a continuous volume of reciprocal space.

A.3 The reciprocal lattice

So far, we did not make any distinction between different points in real or reciprocal space[1], none of them were 'special'. Now, we introduce an infinite and regular (Bravais) lattice of special points in real space, which we call the *real lattice* or *direct lattice.* As an example, take the following basis vectors:

$$\vec{a} = (5,\, 0,\, 0) \qquad |\vec{a}| = 5 \tag{A.11}$$

$$\vec{b} = \left(5\sqrt{2},\, 5\sqrt{2},\, 0\right) \qquad |\vec{b}| = 10 \tag{A.12}$$

$$\vec{c} = (0,\, 0,\, 1) \qquad |\vec{c}| = 1 \tag{A.13}$$

A cut along the XY-plane of this lattice is shown in fig. A.2-a. Define now three basis vectors to construct a lattice of special points in reciprocal space in the following way:

$$a^* = 2\pi\frac{\vec{b}\times\vec{c}}{\vec{a}\cdot\left(\vec{b}\times\vec{c}\right)} \qquad a^* = \frac{2\pi}{5/\sqrt{2}}\left(\frac{\sqrt{2}}{2},\, -\frac{\sqrt{2}}{2},\, 0\right) \qquad \frac{2\pi}{|a^*|} = \frac{5}{\sqrt{2}} \approx 3.536 \tag{A.14}$$

$$b^* = 2\pi\frac{\vec{c}\times\vec{a}}{\vec{b}\cdot\left(\vec{c}\times\vec{a}\right)} \qquad b^* = \frac{2\pi}{10/\sqrt{2}}(0,\, 1,\, 0) \qquad \frac{2\pi}{|b^*|} = \frac{10}{\sqrt{2}} \approx 7.071 \tag{A.15}$$

$$c^* = 2\pi\frac{\vec{a}\times\vec{b}}{\vec{c}\cdot\left(\vec{a}\times\vec{b}\right)} \qquad c^* = 2\pi\,(0,\, 0,\, 1) \qquad \frac{2\pi}{|a^*|} = 1 \tag{A.16}$$

This definition depends on the details of the real lattice. The lattice of special points in reciprocal space defined by these basis vectors is called the *reciprocal lattice* corresponding to this particular real lattice. We will indicate these special points by $\vec{K}$, reciprocal lattice vectors. In Fig. A.2-a a cut through the XY-plane of our real and reciprocal lattices are shown superimposed on each other[2], Fig. A.2-b shows a detail of the reciprocal lattice close to the origin. As for any point in reciprocal space, also the points of the reciprocal lattice correspond to plane waves in real space. Fig. A.2-c shows wave fronts of the three waves corresponding to $\vec{a}^*$, $3\vec{a}^*$ and $\vec{a}^* + \vec{b}^*$. Apparently, the plane waves corresponding to the reciprocal lattice have a special property: they are commensurate with the real lattice. This means: if you are on a particular wave front and you make a displacement along any $n\vec{a} + m\vec{b} + p\vec{c}$, then you end up on a wavefront with exactly the same value. An alternative definition of the reciprocal lattice could therefore be: the collection of all points in reciprocal space that correspond to plane waves that are commensurate with the real lattice.

---

[1] Be careful in the remainder of this section not to confuse the words *space* and *lattice*: they have clearly distinct meanings!

[2] Note that such a plot depends on the length unit chosen: if $a = 5$ mm in real space, $a^* = (2\pi)/(5/\sqrt{2}) \approx 1.78$ mm^{-1} in reciprocal space. In Fig. A.2-a, this would be represented as 1.78 *mm*. But if a would be measured in meter, $a = 0.05$ m would correspond to 177.7 m. This at first sight strange behaviour is due to the fact that we want to display the reciprocal lattice with inverse length as dimension into our real world where we have to draw lines with real lengths. Fig. A.2 is made with an arbitrary length unit in which $\vec{a}$ has length 5 and $\vec{b}$ has length 10. Anyway, this affects only the appearance of the plot and not the physical results: whatever units are used, the wavelengths (which have dimension of length) are found to be the same.

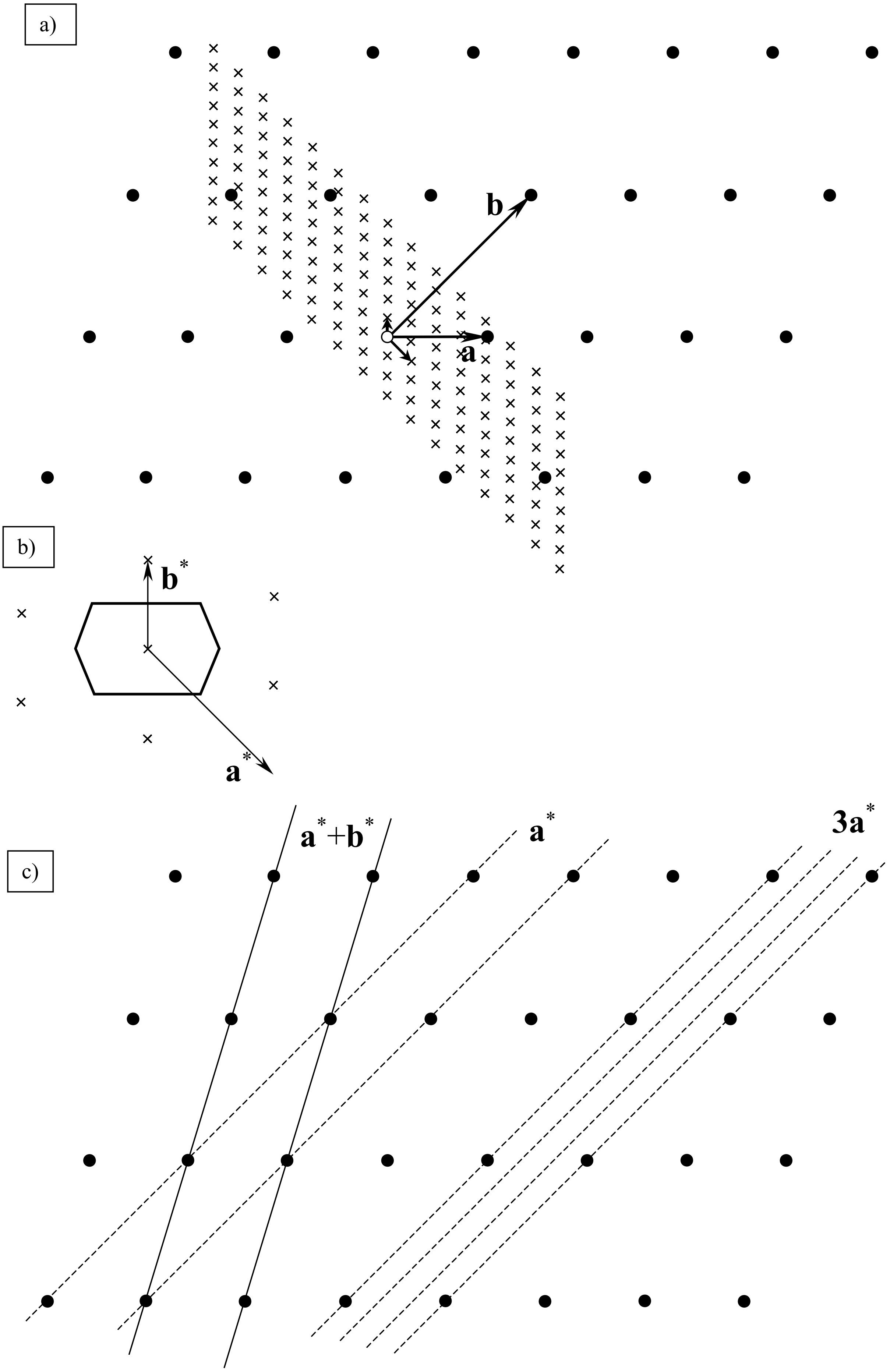

Figure A.2: *a) Lattice and reciprocal lattice. b) The first Brillouin zone. c) Wave fronts of plane waves that are commensurate with the real lattice.*

As for any lattice, several equivalent primitive unit cells can be defined for the reciprocal lattice. One such choice is a cell that contains all points in reciprocal space that are closer to the origin then to any other point of the reciprocal lattice (see Fig. A.2-b). This particular choice of primitive unit cell is called the *first Brillouin zone.* Analogously, the *second Brillouin zone* is defined as containing all points that have the origin of the reciprocal lattice as their second nearest neighbour. Any point in reciprocal space can uniquely be assigned to some *n^{th} Brillouin zone.* By definition of the concept 'unit cell', any point (in the n^{th} Brillouin zone) can be connected to a point in the first Brillouin zone by a vector of the reciprocal lattice.

A.4 Bloch's theorem

As is explained in B, the (discrete) set of eigenfunctions of a quantummechanical hamiltonian can always be labeled by quantum numbers. Bloch's theorem tells how this can be done for a hamiltonian that has a lattice periodicity. The theorem says: *Any eigenfunction $\psi(\vec{r})$ can be written as a product of a function $u_{\vec{g}}(\vec{r})$ that has the periodicity of the lattice, and a plane wave $e^{i\vec{g}\cdot\vec{r}}$ with $\vec{g}$ any*[3] *vector in reciprocal space.*

$$\psi(\vec{r}) \;=\; u_{\vec{g}}(\vec{r})e^{i\vec{g}\cdot\vec{r}} \tag{A.17}$$

As there are an infinite number of vectors in reciprocal space, there are an infinite number of eigenstates of such a hamiltonian. The reciprocal vectors $\vec{g}$ serve as labels (quantum numbers) of the eigenstates, and we could therefore rename $\psi(\vec{r})$ into $\psi_{\vec{g}}(\vec{r})$.

Usually this is not done exactly in this way however. Every $\vec{g}$ can be written as the sum of a vector in the first Brillouin zone (we call this $\vec{k}$ [4]) and a reciprocal lattice vector $\vec{K}$:

$$\vec{g} \;=\; \vec{k}+\vec{K} \tag{A.18}$$

Bloch's theorem can now be rewritten as follows:

$$\psi_{\vec{g}}(\vec{r}) \;=\; \left\{u_{\vec{g}}(\vec{r})e^{i\vec{K}\cdot\vec{r}}\right\} e^{i\vec{k}\cdot\vec{r}} \tag{A.19}$$

Because $e^{i\vec{K}\cdot\vec{r}}$ is commensurate with the lattice, the function between brackets still has the periodicity of the lattice. We could rename it into $u^n_{\vec{k}}(\vec{r})$, where n indicates the number of the Brillouin zone where $\vec{g}$ was in. Indeed, n and $\vec{k}$ contain the same information as $\vec{g}$, and can therefore be used as an alternative way of labeling. For $n = 1$, $\vec{k}$ and $\vec{g}$ are identical. For $\vec{g}$ in the second Brillouin zone, we reuse the same set of $\vec{k}$, but n is increased to 2, etc. For each $\vec{k}$, an infinite number of n is possible. The parameter n is called the *band index.*

---

[3] Strictly spoken only for an infinitely large lattice every $\vec{g}$ is allowed. If a finite crystal and periodic boundary conditions are used, the allowed $\vec{g}$ form a dense mesh in reciprocal space: in the first Brillouin zone, there are as many allowed $\vec{g}$ as there are unit cells contained in the finite (real) lattice. For a crystal of macroscopic size, this is a huge number. For any $\vec{g}$ there will be an allowed $\vec{g}$ very nearby, and therefore we can for practical purposes we assume that all $\vec{g}$ are allowed. The exact property that distinguishes the allowed $\vec{g}$ from the not-allowed ones, is that they correspond to plane waves that are commensurate with the boundaries of the finite crystal.

[4] Pay attention to the consistent names that are use: $\vec{g}$ for any reciprocal vector, $\vec{K}$ for a reciprocal lattice vector and $\vec{k}$ for a vector in the first Brillouin zone. Other texts may use other notations.

Bloch's theorem can now be restated in its most frequently used form: *Any eigenfunction* $\psi^n_{\vec{k}}(\vec{r})$ *can be written as a product of a function* $u^n_{\vec{k}}(\vec{r})$ *that has the periodicity of the lattice, and a plane wave* $e^{i\vec{k}\cdot\vec{r}}$ *with* $\vec{k}$ *any vector in the first Brillouin zone*:

$$\psi^n_{\vec{k}}(\vec{r}) \;\; = \;\; u^n_{\vec{k}}(\vec{r})e^{i\vec{k}\cdot\vec{r}} \tag{A.20}$$

Actually we have now split off a known part ($e^{i\vec{k}\cdot\vec{r}}$) from the eigenstate, such that only the unknown remainder $u^n_{\vec{k}}(\vec{r})$ has to be determined. An important advantage is that this part is now known to have the periodicity of the lattice (which is not true for $\psi^n_{\vec{k}}(\vec{r})$ itself). If a plane wave basis set is used, it can therefore be written as a sum (no integral needed !) over plane waves that have this same periodicity, and these precisely are the plane waves corresponding to reciprocal lattice vectors:

$$u^n_{\vec{k}}(\vec{r}) \;\; = \;\; \sum_{\vec{K}} c^{n,\vec{k}}_{\vec{K}} e^{i\vec{K}\cdot\vec{r}} \tag{A.21}$$

The expansion of $\psi^n_{\vec{k}}(\vec{r})$ in the same basis is then

$$\psi^n_{\vec{k}}(\vec{r}) \;\; = \;\; \sum_{\vec{K}} c^{n,\vec{k}}_{\vec{K}} e^{i(\vec{k}+\vec{K})\cdot\vec{r}} \tag{A.22}$$

and what have to be searched are the coefficients $c^{n,\vec{k}}_{\vec{K}}$.

Appendix B

Quantum numbers and the Density Of States

B.1 Familiar examples

In quantum mechanics, every physical situation that one wants to study is completely defined by its Hamiltonian $\hat{H}$. Every stationary solution of such a problem is described by a state ψ_κ that is an eigenstate of that Hamiltonian. The energy of that solution is the corresponding eigenvalue E_κ:

$$\hat{H}\psi_\kappa = E_\kappa\psi_\kappa \tag{B.1}$$

This is called the (time-independent) Schrödinger equation. In a physical situation, there are always boundary conditions. This limits the possible eigenvalues, and only a discrete but infinite set E_κ remains. The symbols κ stands for one or more *quantum numbers* that are used to label the different eigenfunctions and eigenvalues that satisfy the boundary conditions.

Some well-known examples:

- *The one-dimensional harmonic quantum oscillator*
  A (light) particle with mass M moves in a one-dimensional harmonic potential $V(x) = Cx^2/2$. The Schrödinger equation is:

  $$\underbrace{\left(-\frac{\hbar^2}{2M}\frac{d^2}{dx^2} + \frac{Cx^2}{2}\right)}_{\hat{H}}\psi_n(x) = E_n\psi_n(x) \tag{B.2}$$

  The boundary condition is that the particle is bound: its probability to appear at $x \to \infty$ is zero. The discrete set of eigenvalues E_n and eigenfunctions $\psi_n(x)$ can be labeled by a single quantum number n ($n = 0, 1, 2, \ldots$) and are

  $$E_n = (n + \frac{1}{2})\, h\nu \tag{B.3}$$

  $$\psi_n(x) = \sqrt{\frac{1}{\sqrt{\pi}\, 2^n\, n!}}\, e^{-\,\xi^2/2}\, H_n(\xi) \tag{B.4}$$

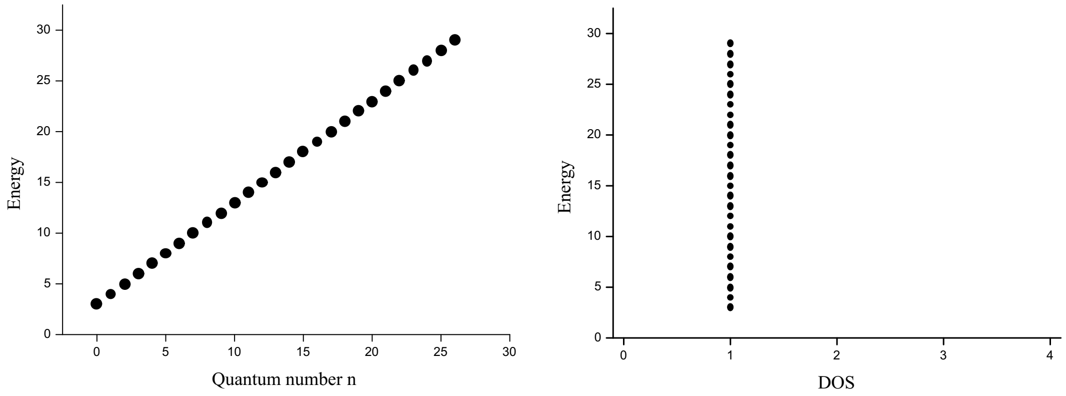

Figure B.1: *Left: Eigenvalues as a function of the label (quantum number) n for the harmonic oscillator. Right: DOS as a function of energy for the harmonic oscillator (the picture is rotated over 90° in order to make comparison with the energy axis of the left picture easy).*

The vibration frequency ν is fully determined by C and M. The variable $\xi = \sqrt{\alpha}\, x$, with α fully determined by C and M. $H_n(\xi)$ is the Hermite polynomial of order n.

Fig. B.1 shows the eigenvalues as a function of the quantum number n, which is called the *eigenvalue spectrum.* In the same picture, the following function of the energy is shown:

$$g(E) \;=\; \sum_n \delta\left(E - E_n\right) \tag{B.5}$$

Here δ is the Dirac *delta function.* The function $g(E)$ is called the *Density Of States* (DOS). Whenever E equals an eigenvalue, the DOS is non-zero. If two eigenvalues would be degenerate (which does not happen in this example), the DOS would be twice as high for that energy. If in some energy region many eigenvalues would exist, the DOS would be non-zero at many places in that region. Obviously, the eigenvalue spectrum contains the full physical information about the problem. The DOS contains less information than the eigenvalue spectrum[1], but carries still some of the characteristics of the problem. The DOS will serve as an easy 'fingerprint' of the problem in cases where the eigenvalue spectrum is difficult to visualize (see soon).

- *A single electron orbiting a given nucleus with infinite mass*
  This case is treated in Appendix D. The Hamiltonian is given in equations D.1 and D.2. Boundary conditions for the 3 separate differential equations are that the solutions are single-valued (for Φ), everywhere finite (for Θ) and vanishing for $r \to \infty$ (for u). Eigenfunctions and eigenvalues are labeled by the quantum numbers n, ℓ and m.

  Fig. B.2 shows the eigenvalue spectrum, that is clearly different from the case of the harmonic oscillator. For a given n, there are n values of ℓ ranging from 0 to $n-1$ that are degenerate. Each of these ℓ itself is $2\ell+1$-fold degenerate due to m. That makes for each n a $\sum_{\ell=0}^{n-1}(2\ell+1)$-fold degeneracy, which is reflected in the increasing value of the DOS. The energies where the DOS is non-zero, are also not evenly spaced any more.

---

[1] Try to construct an imaginary eigenvalue spectrum that would yield exactly the same DOS as in fig. B.1.

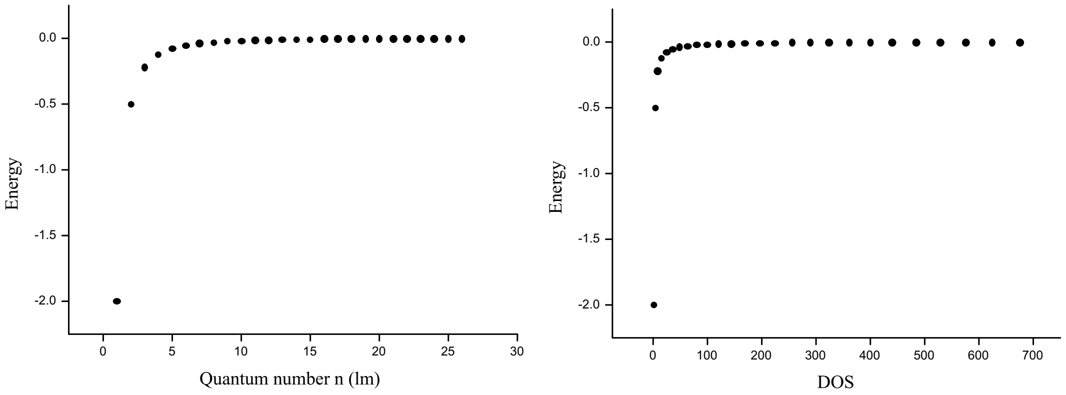

Figure B.2: *Left: Eigenvalues as a function of the label (quantum number) n for a free one-electron atom (states with the same n but a different ℓ and m are degenerate). Right: DOS as a function of energy for a free one-electron atom (the picture is rotated over 90° in order to make comparison with the energy axis of the left picture easy).*

- *A free particle in a one-dimensional box with length L*
  The Schrödinger equation is:

$$\underbrace{\left(-\frac{\hbar^2}{2M}\frac{d^2}{dx^2}\right)}_{\hat{H}}\psi_p(x) \quad = \quad E_p\psi_p(x) \tag{B.6}$$

  with M the mass of the particle. The boundary condition is that the parcticle cannot leave the box. Eigenvalues and eigenstates are labeled by a quantum number p that can take the values

$$p \quad = \quad \frac{2\pi n}{L} \tag{B.7}$$

  with n any integer (negative, zero or positive). Eigenvalues and eigenstates are:

$$E_p \quad = \quad \frac{\hbar^2 p^2}{2M} \tag{B.8}$$

$$\psi_p(x) \quad = \quad \frac{1}{\sqrt{L}}e^{ipx} \tag{B.9}$$

The eigenvalue spectrum and the DOS are shown in fig. B.3. Apart from $\psi_{p=0}$, all states are doubly degenerate. When going to higher energies, the intervals between eigenvalues grow larger.

B.2 Crystalline solids

A general property of crystalline solids is that the potential due to the nuclei is periodic:

$$V(\vec{r}+\vec{R}) \quad = \quad V(\vec{r}) \tag{B.10}$$

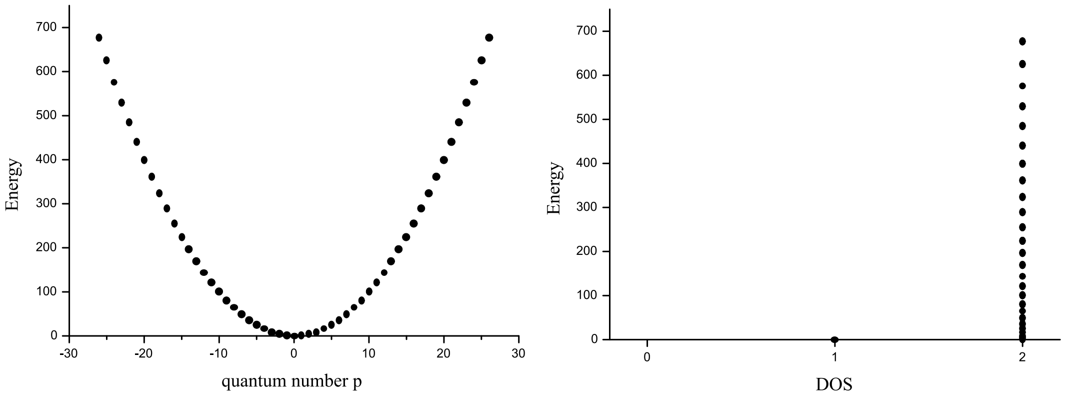

Figure B.3: *Left: Eigenvalues as a function of the label (quantum number) p for a particle in a one-dimensional box. Right: DOS as a function of energy for a particle in one-dimensional box (the picture is rotated over 90° in order to make comparison with the energy axis of the left picture easy).*

with $\vec{R}$ any vector of the Bravais lattice[2]. It can be shown that for the kinetic part of the Hamiltonian, this condition is always fulfilled. Therefore, if the potential is periodic, the total hamiltonian is. We look now at the case where *periodic boundary conditions* are used for the macroscopic piece of crystal: an infinite solid is built by filling space periodically with copies of the crystal (fig. B.4). Under these conditions, eigenvalues and eigenfunctions of a periodic Hamiltonian can be labeled with the quantum 'numbers' n and $\vec{k}$. $n = 1, 2, 3, \ldots$ and $\vec{k}$ is any vector in the first Brillouin zone that corresponds to a plane wave that is commensurate with the macroscopic crystal (fig. B.4). For every valid $\vec{k}$, all values of n occur. The number of valid $\vec{k}$-vectors is equal to the number of unit cells in the macroscopic crystal. For real solids, this number is huge (order of magnitude 10^{23}). Therefore, the $\vec{k}$-vectors are very close to each other, and we can assume that they continuously fill the first Brillouin zone. Note the difference between the $\vec{K}$-vectors from section A.3 that build the reciprocal lattice: they correspond to plane waves that are commensurate with the unit cell of the crystal. The special $\vec{k}$-vectors that are selected inside the first Brillouin zone here, are commensurate with the entire macroscopic piece of crystal. Eigenfunctions are written as $\psi^n_{\vec{k}}$, eigenvalues as $\epsilon^n_{\vec{k}}$ (or $E^n_{\vec{k}}$).

When we want to visualize the spectrum of eigenvalues, a problem is met: the eigenvalues are labeled by 4 independent numbers (n, k_x, k_y and k_z), such that 5 dimensions would be needed to make a plot as in the preceding examples. Obviously, this cannot be done. An alternative procedure is to select a path through the first Brillouin zone, and plot for every n the energies $\epsilon^n_{\vec{k}}$ of the $\vec{k}$ that happen to fall onto that path (see fig. B.5 for an example). The way the path is selected, is largely conventional. Straight intervals of the path are given conventional capital letters (often Greek) as names, and special points in the first Brillouin zone are given conventional capital letters (often Latin). Figures as fig. B.5 (in `WIEN2K` they are called 'spaghetti plots') represent most but not all essential information of the eigenvalue spectrum.

As in the previous cases, it is possible to plot the DOS here. The DOS is a function of the energy only, and hence not suffers from too high dimensions. Due to the presence of the

---

[2] The Bravais lattice contains all essential information on the underlying symmetry of the crystal lattice.

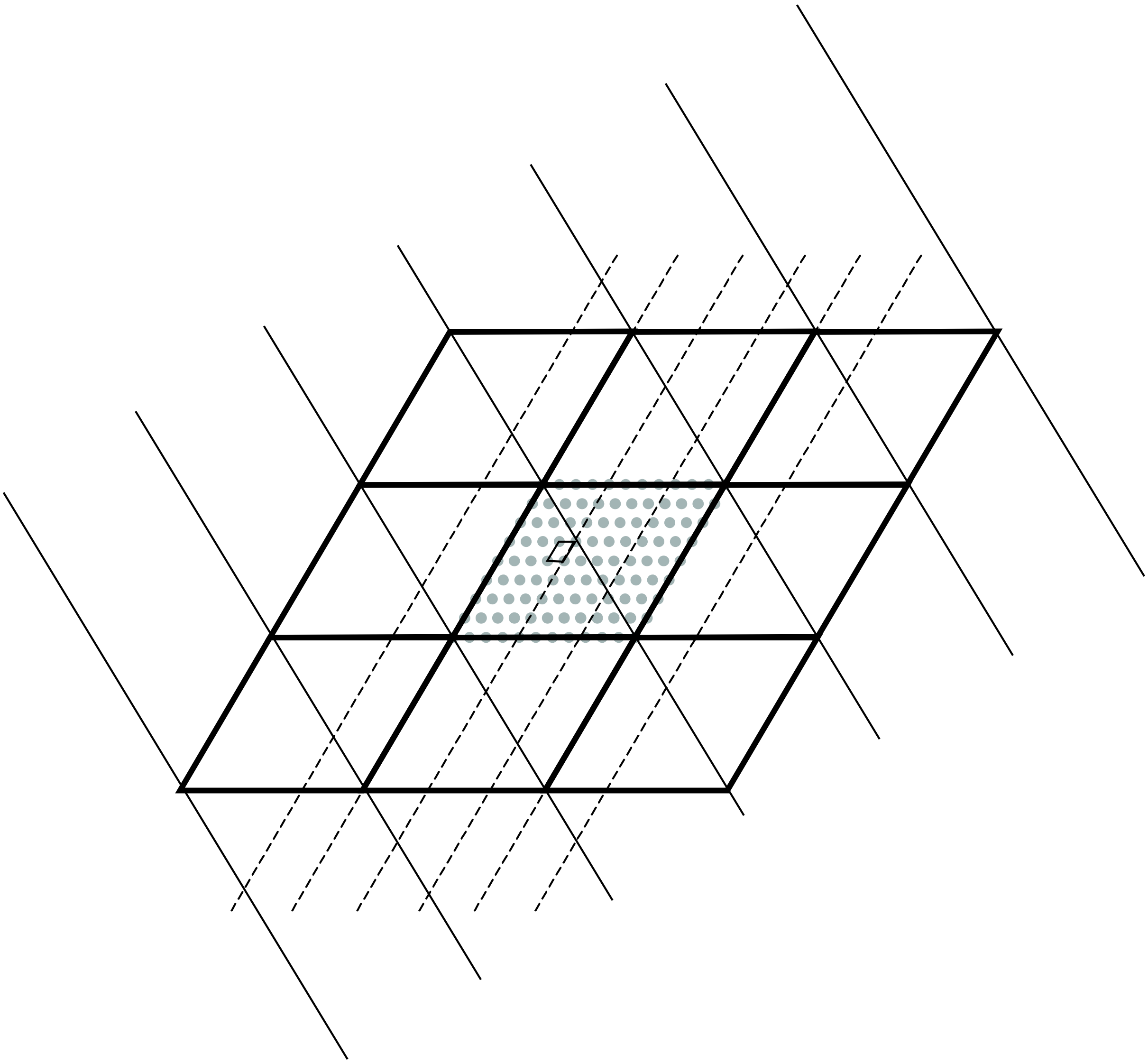

Figure B.4: *Periodic boundary conditions and wave fronts of plane waves that are commensurate with the entire crystal.*

(almost) continuous extra quantum number $\vec{k}$, the definition of the DOS has to be extended:

$$g(E) \quad = \quad \frac{2}{V_{BZ}} \sum_n \int \delta(\epsilon - \epsilon^n_{\vec{k}})\, d\vec{k} \tag{B.11}$$

Here V_{BZ} is the volume of the first Brillouin zone, and the integral is over the first Brillouin zone. The factor 2 is included to account explicitly for (degenerate) spin. An example of a DOS – which is a continuous function now due to $\vec{k}$ – is shown in fig. B.5. When there are flat regions in the spaghetti plot, many $\vec{k}$ have the same energy. This will show up as a peak in the DOS.

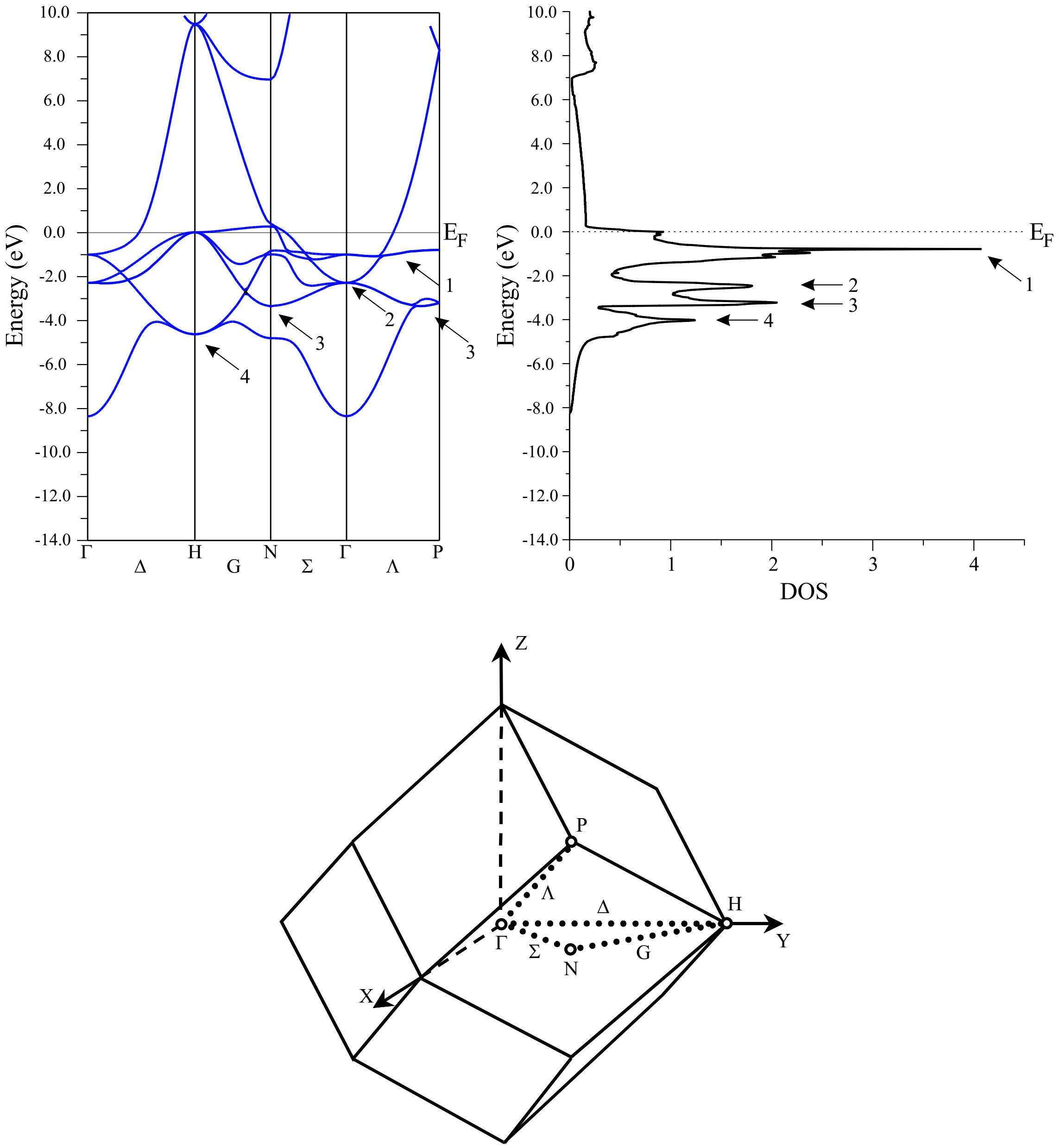

Figure B.5: *Bottom: The first Brillouin zone of a bcc lattice, with 4 high symmetry points indicated: Γ, H, N and P. These points are connected by lines, labeled Δ, G, Σ and Λ. These 4 lines define a path through the Brillouin zone. Top left: the eigenvalues of all the k-vectors along the path are plotted, for some bcc compound (bcc-Fe, spin up). This so-called* band structure plot *contains less information than the full eigenvalues-versus-*$\vec{k}$ *plot, but the latter would require 5 dimensions. The fact that for every* $\vec{k}$ *multiple eigenvalues are possible, is due to different n. Top right: the DOS for the same material, plotted in such a way that the energy scale matches with the energy scale of the band structure plot. The DOS too contains less information than the full eigenvalues-versus-*$\vec{k}$ *plot, but compared to the band structure other information is omitted or included: the DOS is an integral over the full Brillouin zone (and not limited to a onedimensional path only), but on the other hand does not give information for a specific* $\vec{k}$*. Peaks in the DOS mean that many* $\vec{k}$*-vectors have the same eigenvalue. If the path goes through the region of the Brillouin zone that is responsible for this, we see relatively flat (horizontal) lines in the left picture, at the same energies where a peak is in the DOS. Four peaks are labeled 1-4 in the DOS, and are related to features of the band structure plot.*

Appendix C

The eigenvalue problem

C.1 Eigenvalues and eigenvectors

Consider a vector space ($I\!R^n$, take $n = 2$ for simplicity) with a basis set ($\vec{e}_1$, $\vec{e}_2$) that need not to be orthogonal. With respect to this basis, every vector $\vec{x}$ can be uniquely characterized by two numbers (x_1, x_2). With $\hat{H}$, we denote an operator that can operate on $\vec{x}$, resulting in a new vector $\vec{y} = (y_1,\, y_2)$:

$$\hat{H}\vec{x} \;=\; \vec{y} \tag{C.1}$$

Given $\hat{H}$, we want to find all vectors $\vec{x}$ that under $\hat{H}$ transform into a vector that is parallel with itself:

$$\hat{H}\vec{x} \;=\; \lambda\vec{x} \tag{C.2}$$

We rewrite this condition using the basis vectors, and will search for the (x_1, x_2) that satisfy it:

$$\hat{H}\left(x_1\vec{e}_1 + x_2\vec{e}_2\right) \;=\; \lambda\left(x_1\vec{e}_1 + x_2\vec{e}_2\right) \tag{C.3}$$

Left-multiply this equation with $\vec{e}_1$:

$$x_1\,\vec{e}_1\cdot\left(\hat{H}\vec{e}_1\right) + x_2\,\vec{e}_1\cdot\left(\hat{H}\vec{e}_2\right) \;=\; \lambda\left(x_1\,\vec{e}_1\cdot\vec{e}_1 + x_2\,\vec{e}_1\cdot\vec{e}_2\right) \tag{C.4}$$

Do the same with $\vec{e}_2$:

$$x_1\,\vec{e}_2\cdot\left(\hat{H}\vec{e}_1\right) + x_2\,\vec{e}_2\cdot\left(\hat{H}\vec{e}_2\right) \;=\; \lambda\left(x_1\,\vec{e}_2\cdot\vec{e}_1 + x_2\,\vec{e}_2\cdot\vec{e}_2\right) \tag{C.5}$$

Equations C.4 and C.5 can be summarized in matrix notation as

$$\begin{bmatrix} \vec{e}_1\cdot\left(\hat{H}\vec{e}_1\right) & \vec{e}_1\cdot\left(\hat{H}\vec{e}_2\right) \\ \vec{e}_2\cdot\left(\hat{H}\vec{e}_1\right) & \vec{e}_2\cdot\left(\hat{H}\vec{e}_2\right) \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} - \lambda \begin{bmatrix} \vec{e}_1\cdot\vec{e}_1 & \vec{e}_1\cdot\vec{e}_2 \\ \vec{e}_2\cdot\vec{e}_1 & \vec{e}_2\cdot\vec{e}_2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \;=\; \begin{bmatrix} 0 \\ 0 \end{bmatrix} \tag{C.6}$$

The ij^{th} element (call it H_{ij}) of the first matrix is a number, as it is the result of a dot product. This matrix is completely determined if the action of $\hat{H}$ on the basis vectors is known[1]. The

---

[1] It is straightforward to show that with this matrix $\hat{H}\vec{x} = \vec{y}$ can be expressed as

$$\begin{bmatrix} H_{11} & H_{12} \\ H_{21} & H_{22} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \;=\; \begin{bmatrix} y_1 \\ y_2 \end{bmatrix} \tag{C.7}$$

If the H_{ij} are known, the effect of $\hat{H}$ on any vector $\vec{x}$ can be found by a simple matrix multiplication.

elements of the second matrix (call them S_{ij}) are determined by the basis only. This matrix is called the *overlap matrix.* With this notation, our 2 equations to find suitable $(x_1,\, x_2)$ become:

$$\begin{bmatrix} H_{11} - \lambda S_{11} & H_{12} - \lambda S_{12} \\ H_{21} - \lambda S_{21} & H_{22} - \lambda S_{22} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \tag{C.8}$$

The left matrix is fully known, apart from λ that is a parameter. For every value of λ, equation C.8 can be solved for x_1 and x_2. For most values of λ, the determinant of the matrix will be different from zero. Equation C.8 will have then one unique solution, which is obviously (0, 0). This vector is parallel with the original $\vec{x}$, but in a trivial way. Other, more interesting solutions can occur only for those values of λ that give a vanishing determinant:

$$\begin{vmatrix} H_{11} - \lambda S_{11} & H_{12} - \lambda S_{12} \\ H_{21} - \lambda S_{21} & H_{22} - \lambda S_{22} \end{vmatrix} = 0 \tag{C.9}$$

The above equation is called the *secular equation* of $\hat{H}$. It is a polynomial equation in λ, with as highest power the dimension of the space n (2 in this case). The roots[2] of the secular equation are called the *eigenvalues* of $\hat{H}$. If $\lambda = \lambda_1$ is an eigenvalue, $[H - \lambda_1 S]\,[x] = [0]$ has an infinite number of solutions: indeed, if $(x_1 = a,\, x_2 = b)$ is a solution, $(\beta a,\, \beta b)$ is one too, for β any real number. This conforms with our initial requirement: if $\hat{H}$ transforms $\vec{x}$ into a vector that is parallel with $\vec{x}$, it will do the same with any vector $\beta\vec{x}$. These vectors are called the *eigenvectors* of $\hat{H}$, belonging to the eigenvalue λ_1. Usually a unit vector among them is chosen to represent this set of vectors. Eigenvectors belonging to different eigenvalues, can be shown to be perpendicular[3]. Therefore, unit vectors belonging to n different eigenvalues can be taken as an orthonormal basis $\left(\vec{e}_1^{\,\lambda_1},\, \vec{e}_2^{\,\lambda_2}\right)$ for the vector space.

What will be the matrix representation of the operator $\hat{H}$, written in this new basis? According to equation C.6, the matrix elements are:

$$\begin{aligned} H_{ij} &= \vec{e}_i^{\,\lambda_i} \cdot \left(\hat{H}\vec{e}_j^{\,\lambda_j}\right) && \text{(C.10)} \\ &= \vec{e}_i^{\,\lambda_i} \cdot \left(\lambda_j \vec{e}_j^{\,\lambda_j}\right) && \text{(C.11)} \\ &= \lambda_j \delta_{ij} && \text{(C.12)} \end{aligned}$$

This is a diagonal matrix with the former eigenvalues on the diagonal:

$$[H] = \begin{bmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{bmatrix} \tag{C.13}$$

We could now try to find the eigenvalues and eigenvectors of this operator in the new basis, using the same procedure as described above. Obviously we will find the same λ_1 and λ_2 as before, with as eigenvectors (1, 0) and (0, 1) (i.e. the old eigenvectors which are now the basis vectors). This illustrates that eigenvalues and eigenvectors are intrinsic properties of the operator, and do not depend on the choice of the basis.

If the basis we originally started with was already orthonormal, then the overlap matrix will be a unit matrix. This simplifies equation C.6.

---

[2] There will be *at most* n roots. In all cases that will be of interest for us, there will be *exactly* n roots.

[3] At least if the matrix representing the operator is hermitian, which is the case for 'well-behaved' operators.

It could happen that 2 of the roots of the secular equation coincide. In that case, that eigenvalue leads to a plane instead of a line of eigenvectors. There is freedom then to choose two mutually perpendicular basis vectors in that plane, but still they are perpendicular to eigenvectors of the other – different – eigenvalues.

This idea can be extended to vector spaces of infinite dimension (e.g. function spaces, in which single particle wave functions are described).

C.2 Basis transformations

Consider an operator $\hat{A}$ with matrix representation A [4], in a vector space $I\!R^n$ with a normalized but not necessarily orthogonal basis $(\vec{e}_1^{\,\alpha}, \ldots, \vec{e}_n^{\,\alpha})$. This operator transforms every vector into a new vector. Also the basis vectors are transformed, and in this way you can understand that the j^{th} column of A contains the coefficients that express the transformed j^{th} basis vector in the original basis:

$$\begin{bmatrix} \cdots & a_{1j} & \cdots \\ \cdots & \vdots & \cdots \\ \cdots & a_{jj} & \cdots \\ \cdots & \vdots & \cdots \\ \cdots & a_{nj} & \cdots \end{bmatrix}_{n\times n} \begin{bmatrix} e^{\alpha}_{1j} = 0 \\ \vdots \\ e^{\alpha}_{jj} = 1 \\ \vdots \\ e^{\alpha}_{nj} = 0 \end{bmatrix}_{n\times 1} = \begin{bmatrix} a_{1j} \\ \vdots \\ a_{jj} \\ \vdots \\ a_{nj} \end{bmatrix}_{n\times 1} \tag{C.14}$$

This provides a practical recipe to find the matrix representation of an operator about which we want that it transforms the basis vectors $\vec{e}_j^{\,\alpha}$ in a new set $\vec{e}_j^{\,\beta}$: express each $\vec{e}_j^{\,\beta}$ in the old basis $\vec{e}_j^{\,\alpha}$, and these expansion coefficients form the columns of A. Note that in this way a non-orthogonal basis can be transformed into an orthogonal basis. If both bases are orthonormal, than the matrix A has special properties (it is a unitary matrix).

Conversely, the j^{th} column of the inverse matrix A^{-1} contains the coefficients that express the vector $\vec{e}_j^{\,\alpha}$ of the old basis in the new basis of the $\vec{e}_j^{\,\beta}$. This is illustrated in fig. C.1 for an operator

$$A = \begin{bmatrix} \frac{\sqrt{2}}{2} & -\frac{\sqrt{2}}{2} \\ \frac{\sqrt{2}}{2} & \frac{\sqrt{2}}{2} \end{bmatrix} \qquad A^{-1} = \begin{bmatrix} \frac{\sqrt{2}}{2} & \frac{\sqrt{2}}{2} \\ -\frac{\sqrt{2}}{2} & \frac{\sqrt{2}}{2} \end{bmatrix} \tag{C.15}$$

that rotates every vector of $I\!R^2$ counterclockwise over 45°. The old basis $(\vec{e}_1^{\,\alpha}, \vec{e}_2^{\,\alpha})$ is transformed in the new basis $\left(\vec{e}_1^{\,\beta}, \vec{e}_2^{\,\beta}\right)$. The coordinates of these 4 vectors in both bases are:

	old	new
$\vec{e}_1^{\,\alpha}$	(1, 0)	$\left(\frac{\sqrt{2}}{2}, -\frac{\sqrt{2}}{2}\right)$
$\vec{e}_2^{\,\alpha}$	(0, 1)	$\left(\frac{\sqrt{2}}{2}, \frac{\sqrt{2}}{2}\right)$
$\vec{e}_1^{\,\beta}$	$\left(\frac{\sqrt{2}}{2}, \frac{\sqrt{2}}{2}\right)$	(1, 0)
$\vec{e}_2^{\,\beta}$	$\left(-\frac{\sqrt{2}}{2}, \frac{\sqrt{2}}{2}\right)$	(0, 1)

---

[4] $\hat{A}$ is a general operator, not necessarily identical to the $\hat{H}$ we were discussing before. But everything that is said about $\hat{A}$, applies for $\hat{H}$ too.

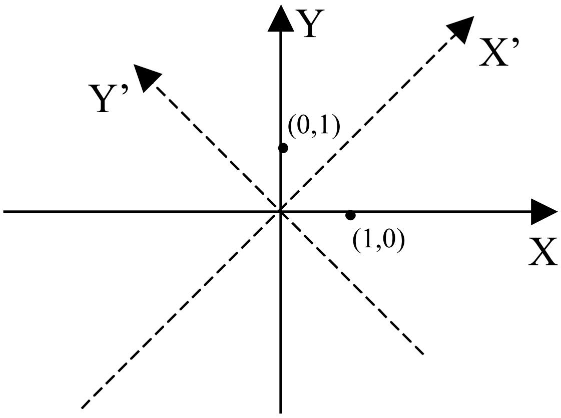

Figure C.1: *An operator $\hat{A}$ rotates every vector counterclockwise over 45°. The new axis system X'Y' formed by the rotated original basis vectors is shown.*

Check with this table and with fig. C.1 that the following expression

$$AX = Y \tag{C.16}$$

(X and Y are $(n \times 1)$-matrices) can be interpreted in these two different ways:

- First interpretation: X contains the coefficients of any vector expressed in the old basis. Then Y will contain the coefficients of the transformation of X under $\hat{A}$, expressed in the old basis.
- Second interpretation: X contains the coefficients of any vector expressed in the new basis. Then Y will contain the coefficients of this same vector, expressed in the old basis.

Similarly, this expression

$$A^{-1}X = Y \tag{C.17}$$

can be interpreted in two ways too:

- First interpretation: X contains the coefficients of the transformation under $\hat{A}$ of any vector, expressed in the old basis. Then Y will contain the coefficients of this original vector (=before transformation) expressed in the old basis.
- Second interpretation: X contains the coefficients of any vector expressed in the old basis. Then Y will contain the coefficients of this same vector, expressed in the new basis.

C.3 A practical procedure to find eigenvalues and eigenvectors

Given a vector space $\mathbb{R}^n$ with a normalized but not necessarily orthogonal basis, we want to find the n eigenvalues and the n eigenvectors of a given operator $\hat{H}$. The matrix elements H_{ij}

and S_{ij} of the matrix representation H of the operator and of the overlap matrix S of the basis set, can be obtained as described before. Both are $n \times n$ matrices. What we want to determine are the coefficients c_{ij} that express in the given basis the j^{th} normalized eigenvector. In total, $n \times n$ coefficients c_{ij} are needed to specify all eigenvectors, and they can be arranged in an $n \times n$ matrix C[5]. Equation C.8 can then be extended to:

$$HC = SCE \tag{C.18}$$

where E is an $n \times n$ diagonal matrix that contains as its j^{th} diagonal element the eigenvalue corresponding to the j^{th} eigenvector[6]. In equation C.18, the n^2 coefficients c_{ij} are unknown, as are the n eigenvalues E_{jj}. Equation C.18 represents n^2 independent equations, and $n^2 + n$ unknowns have to be determined. The remaining n equations come from the requirement that in each class of eigenvectors, we search a normalized one[7]:

$$\left[c_{1j} \; c_{2j} \; \ldots \; c_{nj}\right]_{1\times n} \; \left[S_{ij}\right]_{n\times n} \begin{bmatrix} c_{1j} \\ c_{2j} \\ \vdots \\ c_{nj} \end{bmatrix}_{n\times 1} = 1 \tag{C.19}$$

Equations C.18 and C.19 define the problem to be solved.

Now, we left-multiply equation C.18 by an $n \times n$ matrix A^{-1}. At two other places, we insert a multiplication by the unit matrix $\mathbb{I} = AA^{-1}$:

$$\underbrace{A^{-1}HA}_{H_0} \; \underbrace{A^{-1}C}_{C_0} = \underbrace{A^{-1}SA}_{S_0} \; \underbrace{A^{-1}C}_{C_0} \; E \tag{C.20}$$

It can be proven that for 'well-behaved' operators $\hat{H}$, always a matrix A can be found such that $H_0 = A^{-1}HA$ is a diagonal matrix. We know that in such a case the elements on the diagonal must be the eigenvalues of $\hat{H}$. Consequently, $H_0 = E$. But if H_0 is diagonal, we know that it must be expressed in a basis that consists of its eigenvectors. Therefore, the operator $\hat{A}$ must transform the original basis into the new basis of eigenvectors. Because C contained the (unknown) coefficients of the eigenvectors in the old basis, $C_0 = A^{-1}C$ contains the coefficients of the eigenvectors in the basis of eigenvectors: C_0 must be diagonal. (We cannot say that C_0 should be the unit matrix: according to equation C.18 it should contain coefficients of eigenvectors, not of *normalized* eigenvectors.) Equation C.20 is identical to equation C.18, but written in the new basis of eigenvectors:

$$H_0C_0 = S_0C_0E \tag{C.21}$$

with $H_0 = E$ and C_0 diagonal. Because the eigenvectors are mutually orthogonal, this new basis is orthonormal and hence S_0 is a unit matrix. The final expression

$$EC_0 = C_0E \tag{C.22}$$

---

[5] Note that the coefficients that describe the j^{th} normalized eigenvector form the j^{th} column of C.

[6] Check for $n = 2$ that equation C.18 is indeed equivalent to twice equation C.8, once for both eigenvalues/eigenvectors.

[7] In an orthonormal basis (S is a unit matrix), this reduces to the more familiar formula for the square of the length of a vector.

is true, because multiplication among diagonal matrices is commutative.

We conclude that if we can find the matrix A that diagonalizes H, the eigenvalues are read from its diagonal, and the eigenvectors can be calculated from

$$C = AC_0 \tag{C.23}$$

where the n undetermined values on the diagonal of C_0 are fixed by the still unused normalization requirements C.19. Diagonalizing the Hamiltonian matrix is the key to solve the eigenvalue problem. A lot of standard numerical procedures are available to perform a diagonalization. An important common feature is that the time it takes to execute them on a computer scales with n^3.

Appendix D

Solutions of the radial part of the Schrödinger equation

Many text books on quantum mechanics explicitly solve the Schrödinger equation for a single electron with charge $-e$ and mass m_e in the Coulomb potential of a nucleus with charge Ze and infinite mass[1]. In spherical coordinates $\vec{r} = (r,\, \theta,\, \phi)$, the Schrödinger equation is:

$$-\frac{\hbar^2}{2m_e}\,\nabla^2\,\psi(r,\,\theta,\,\phi) \,+\, V(r) \;=\; E\,\psi(r,\,\theta,\,\phi) \qquad \text{(D.1)}$$

with

$$\begin{aligned} V(r) &= -\frac{Z\,e^2}{4\pi\epsilon_0\,r} && \text{(D.2)} \\ \nabla^2 &= \frac{1}{r^2}\,\frac{\partial}{\partial r}\left(r^2\,\frac{\partial}{\partial r}\right) \,+\, \frac{1}{r^2\,\sin^2\theta}\,\frac{\partial^2}{\partial\phi^2} \,+\, \frac{1}{r^2\,\sin\theta}\,\frac{\partial}{\partial\theta}\left(\sin\theta\,\frac{\partial}{\partial\theta}\right) && \text{(D.3)} \end{aligned}$$

This differential equation can be shown to be separable in 3 independent differential equations, each of them involving only r, θ or ϕ respectively. General solutions of equation D.1 will therefore be of the form

$$\psi(r,\,\theta,\,\phi) \;=\; u(r)\,\Theta(\theta)\,\Phi(\phi) \qquad \text{(D.4)}$$

The three differential equations of which $\Phi(\phi)$, $\Theta(\theta)$ and $u(r)$ are solutions, are:

$$\begin{aligned} \frac{d^2\Phi}{d\phi^2} &= -m^2\Phi && \text{(D.5)} \\ -\frac{1}{\sin\theta}\,\frac{d}{d\theta}\left(\sin\theta\,\frac{d\Theta}{d\theta}\right) \,+\, \frac{m^2\Theta}{\sin^2\theta} &= \ell(\ell+1)\,\Theta && \text{(D.6)} \\ \frac{1}{r^2}\,\frac{d}{dr}\left(r^2\,\frac{du}{dr}\right) \,+\, \frac{2m_e}{\hbar^2}\,(E \,-\, V(r))\;u &= \ell(\ell+1)\,\frac{u}{r^2} && \text{(D.7)} \end{aligned}$$

Here ℓ and m are dimensionless parameters. In order to have well-behaved solutions, it can be shown that $\ell = 0,\, 1,\, 2,\, 3,\, \ldots$ and $m = -\ell,\, -\ell+1,\, \ldots,\, \ell-1,\, \ell$. The product of $\Theta(\theta)$ and $\Phi(\phi)$ is called a *spherical harmonic* $Y^\ell_m(\theta,\phi)$:

$$Y^\ell_m(\theta,\phi) \;=\; \Theta(\theta)\,\Phi(\phi) \qquad \text{(D.8)}$$

---

[1] An infinite mass avoids the rotation of nucleus and electron around their common center of mass.

Spherical harmonics can be found tabulated for many (ℓ, m) at various places. The full solution of equation D.1 can now be written as

$$\psi(r,\,\theta,\,\phi) \;=\; u_\ell(r)\,Y^\ell_m(\theta,\,\phi) \tag{D.9}$$

with $u_\ell(r)$ a solution of equation D.7 with the proper value of ℓ. Equation D.7 is the *radial Schrödinger equation*.

What will be the solutions of equation D.7? Using these transformations of variables:

$$\rho \;=\; 2\beta r \tag{D.10}$$

$$\beta^2 \;=\; -\,\frac{2m_e E}{\hbar^2} \tag{D.11}$$

$$\gamma \;=\; \frac{m_e Z e^2}{\hbar^2 \beta} \tag{D.12}$$

it can be shown that the solutions have the form

$$u_\ell(r) \;=\; e^{-\rho/2}\,\rho^\ell\,G_\gamma(\rho) \tag{D.13}$$

where $G_\gamma(\rho)$ is defined by the above relation. For the case of a free atom which we are considering here, the boundary condition for the radial equation is that $u_\ell(r) \to 0$ as $r \to \infty$ (or equivalently, as $\rho \to \infty$). This can be shown to be possible only if $\gamma = n$, with n any of the integers

$$n \;=\; \ell+1,\, \ell+2,\, \ell+3,\, \ldots \tag{D.14}$$

For these specific values of γ, the $G_\gamma(\rho)$ – now written $G_n(\rho)$ – are called the *associated Laguerre functions*. The radial solutions $u(r)$ carry a double index $n\ell$ now ($u_{n\ell}(r)$), and the solutions of the entire Schrödinger equation D.1 are labeled by n, ℓ and m:

$$\psi_{n\ell m}(r,\theta,\phi) \;=\; u_{n\ell}(r)\,Y^\ell_m(\theta,\,\phi) \tag{D.15}$$

The corresponding discrete set of energies (eigenvalues) is given by

$$E_{n\ell m} \;=\; E_n \;=\; -\,\frac{m_e Z^2 e^4}{2\hbar^2 n^2} \qquad n \;=\; 1,\,2,\,3,\,4,\,\ldots \tag{D.16}$$

which is degenerate in ℓ and m.

In the APW, LAPW+LO and APW+lo methods, we will use radial functions $u_\ell(r)$ that are solutions of equation D.7. But we will need these functions only inside the muffin tin spheres. What their behaviour for $r \to \infty$ is, is not important for us: we will never use these functions in that region. The boundary condition that $u_\ell(r)$ should vanish at infinity therefore disappears. Hence, the requirement that γ must be integer (leading to a discrete set of energies) disappears too. *If no boundary condition is used, solutions $u_\ell(r)$ can be found for every energy E.* Note that the subscript n is dropped now, as it has no meaning without a boundary condition.

Appendix E

The homogeneous electron gas

The *homogeneous electron gas*, *uniform electron gas* or *jellium model* is an imaginary solid where all nuclear charge is homogeneously smeared out over space. This material is completely isotropic, and identical on every length scale. Therefore the electron density is constant:

$$\rho(\vec{r}) \;\equiv\; \rho \;\equiv\; \frac{N}{V} \tag{E.1}$$

with N the number of electrons in the material, and V its volume. The parameter ρ is the only thing we need to specify a particular homogeneous electron gas completely.

If the electrons do not interact, we are in the case of the free electron gas, which can be solved analytically in a straightforward way. The problem is much more difficult for an interacting electron gas. Here numerical calculations for the total energy are possible by quantum Monte Carlo. Subtracting the non-interacting kinetic energy and the Hartree energy gives a numerical result for the *exchange-correlation energy* ϵ_{xc}. If this is done for several densities ρ, the function $\epsilon_{xc}(\rho)$ is obtained. Note that ϵ_{xc} is a *function* of ρ, not a *functional*.

Appendix F

Functionals

F.1 Definition and Examples

To say that F is a *functional* of the function $f(\tau)$ means that F is a number whose value depends on the *form* of the function $f(\tau)$. In other words: for every function f (defined on the same domain, given by τ) the functional $F[f]$ returns a unique number:

$$F : \mathcal{F} \mapsto I\!R : f \mapsto F[f] \tag{F.1}$$

($\mathcal{F}$ is a space of functions) A functional is therefore a 'function of functions'. Some examples of functionals are:

- $F[f] = \int_{-\infty}^{+\infty} f(\tau)\, e^{-\tau^2}\, d\tau$

  Clearly, if f is changed the value of the integral and therefore $F[f]$ is changed. Note that F does not depend on the coordinates τ of f, only the form of f matters.

- $F[f] = f(0)$

  This very simple functional depends only on the value of f at $\tau = 0$. Another f gives another value for $F[f]$.

- The integral of a function is a functional. Another function gives another integral, and hence another value for the functional.

- The total energy of a quantum system is a functional of the density, as stated by the first theorem of Hohenberg and Kohn. The density is a function defined on $I\!R^3$.

F.2 Functional derivatives

The generalization of the derivative $\frac{df(x)}{dx}$ of a function $f(x)$, is the functional derivative $\frac{\delta F[\rho]}{\delta \rho}$ of the functional $F[\rho]$. As an example, we calculate here the functional derivative of the nucleus-electron part of the total-energy functional used in Chapter 1. The same method can be applied to functionals of other form too.

Consider a functional

$$E^{n-e}_{V_{ext}}[\rho] \;=\; \int_a^b \rho(x) V_{ext}(x)\, dx \tag{F.2}$$

where $V_{ext}(x)$ is a given function. The integral runs over the 1-dimensional volume $[a,\, b]$ (it is straightforward to extend the following reasoning to functions $\rho(\vec{r})$ and $V_{ext}(\vec{r})$ and integrals over 3-dimensional volumes, as is needed in Chapter 1).

The variation of $E^{n-e}_{V_{ext}}[\rho]$ with respect to its argument $\rho(x)$ is given by

$$\begin{aligned} \delta E^{n-e}_{V_{ext}}[\rho] &= E^{n-e}_{V_{ext}}[\rho + \delta\rho] - E^{n-e}_{V_{ext}}[\rho] && \text{(F.3)} \\ &= \left(\int_a^b \rho(x)\, V_{ext}(x)\, dx \;+\; \int_a^b \delta\rho(x)\, V_{ext}(x)\, dx \right) \;-\; \int_a^b \rho(x)\, V_{ext}(x)\, dx && \text{(F.4)} \\ &= \int_a^b \delta\rho(x)\, V_{ext}(x)\, dx && \text{(F.5)} \end{aligned}$$

In standard calculus, the differential df of a function $f(\vec{x})$ defined on an n-dimensional domain, is given by the chain rule:

$$df \;=\; \sum_{i=1}^{n} \frac{\partial f}{\partial x_i}\, dx_i \tag{F.6}$$

In the calculus of functionals, the dimension of the domain $(\mathcal{F})$ is infinite. If one divides the interval $[a,\, b]$ into small enough sub-intervals defined by the set $\{x_1 = a,\, x_2,\, \ldots,\, x_{n-1},\, x_p = b\}$ so that the variation $\delta\rho(x)$ is approximately constant in each interval, then equation F.5 is given approximately by

$$\delta E^{n-e}_{V_{ext}}[\rho] \;\approx\; \sum_{i=1}^{p} V_{ext}(x_i)\; [\delta\rho(x_i)\, \Delta x] \tag{F.7}$$

In the limit of $p \to \infty$, equation F.5 is recovered. Comparing equations F.5 and F.6 one can identify

$$\begin{aligned} dx_i &\;\to\; \delta\rho(x)\, dx && \text{(F.8)} \\ \frac{\partial f}{\partial x_i} &\;\to\; V_{ext}(x) \;=\; \frac{\delta E^{n-e}_{V_{ext}}[\rho]}{\delta\rho} && \text{(F.9)} \end{aligned}$$

Generalization to functionals

$$E^{n-e}_{V_{ext}}[\rho] \;=\; \int_{vol} \rho(\vec{r})\, V_{ext}(\vec{r})\, d\vec{r} \tag{F.10}$$

defined with functions $\rho(\vec{r}) : \mathbb{R}^3 \mapsto \mathbb{R}$ leads to the following formula for the functional derivative:

$$\frac{\delta E^{n-e}_{V_{ext}}[\rho]}{\delta\rho} \;=\; V_{ext}(\vec{r}) \tag{F.11}$$

ISBN 90-807215-1-4