Lecture Notes in Physics

Editorial Board

H. Araki, Kyoto, Japan
R. Beig, Vienna, Austria
J. Ehlers, Potsdam, Germany
U. Frisch, Nice, France
K. Hepp, Zürich, Switzerland
R. L. Jaffe, Cambridge, MA, USA
R. Kippenhahn, Göttingen, Germany
H. A. Weidenmüller, Heidelberg, Germany
J. Wess, München, Germany
J. Zittartz, Köln, Germany

Managing Editor

W. Beiglböck
Assisted by Mrs. Sabine Lehr
c/o Springer-Verlag, Physics Editorial Department II
Tiergartenstrasse 17, D-69121 Heidelberg, Germany

Springer
Berlin
Heidelberg
New York
Barcelona
Budapest
Hong Kong
London
Milan
Paris
Santa Clara
Singapore
Tokyo

The Editorial Policy for Proceedings

The series Lecture Notes in Physics reports new developments in physical research and teaching – quickly, *informally, and at a high level.* The proceedings to be considered for publication in this series should be limited to only a few areas of research, and these should be closely related to each other. The contributions should be of a high standard and should avoid lengthy redraftings of papers already published or about to be published elsewhere. As a whole, the proceedings should aim for a balanced presentation of the theme of the conference including a description of the techniques used and enough motivation for a broad readership. It should not be assumed that the published proceedings must reflect the conference in its entirety. (A listing or abstracts of papers presented at the meeting but not included in the proceedings could be added as an appendix.)
When applying for publication in the series Lecture Notes in Physics the volume's editor(s) should submit sufficient material to enable the series editors and their referees to make a fairly accurate evaluation (e.g. a complete list of speakers and titles of papers to be presented and abstracts). If, based on this information, the proceedings are (tentatively) accepted, the volume's editor(s), whose name(s) will appear on the title pages, should select the papers suitable for publication and have them refereed (as for a journal) when appropriate. As a rule discussions will not be accepted. The series editors and Springer-Verlag will normally not interfere with the detailed editing except in fairly obvious cases or on technical matters.
Final acceptance is expressed by the series editor in charge, in consultation with Springer-Verlag only after receiving the complete manuscript. It might help to send a copy of the authors' manuscripts in advance to the editor in charge to discuss possible revisions with him. As a general rule, the series editor will confirm his tentative acceptance if the final manuscript corresponds to the original concept discussed, if the quality of the contribution meets the requirements of the series, and if the final size of the manuscript does not greatly exceed the number of pages originally agreed upon. The manuscript should be forwarded to Springer-Verlag shortly after the meeting. In cases of extreme delay (more than six months after the conference) the series editors will check once more the timeliness of the papers. Therefore, the volume's editor(s) should establish strict deadlines, or collect the articles during the conference and have them revised on the spot. If a delay is unavoidable, one should encourage the authors to update their contributions if appropriate. The editors of proceedings are strongly advised to inform contributors about these points at an early stage.
The final manuscript should contain a table of contents and an informative introduction accessible also to readers not particularly familiar with the topic of the conference. The contributions should be in English. The volume's editor(s) should check the contributions for the correct use of language. At Springer-Verlag only the prefaces will be checked by a copy-editor for language and style. Grave linguistic or technical shortcomings may lead to the rejection of contributions by the series editors. A conference report should not exceed a total of 500 pages. Keeping the size within this bound should be achieved by a stricter selection of articles and not by imposing an upper limit to the length of the individual papers. Editors receive jointly 30 complimentary copies of their book. They are entitled to purchase further copies of their book at a reduced rate. As a rule no reprints of individual contributions can be supplied. No royalty is paid on Lecture Notes in Physics volumes. Commitment to publish is made by letter of interest rather than by signing a formal contract. Springer-Verlag secures the copyright for each volume.

The Production Process

The books are hardbound, and the publisher will select quality paper appropriate to the needs of the author(s). Publication time is about ten weeks. More than twenty years of experience guarantee authors the best possible service. To reach the goal of rapid publication at a low price the technique of photographic reproduction from a camera-ready manuscript was chosen. This process shifts the main responsibility for the technical quality considerably from the publisher to the authors. We therefore urge all authors and editors of proceedings to observe very carefully the essentials for the preparation of camera-ready manuscripts, which we will supply on request. This applies especially to the quality of figures and halftones submitted for publication. In addition, it might be useful to look at some of the volumes already published. As a special service, we offer free of charge LaTeX and TeX macro packages to format the text according to Springer-Verlag's quality requirements. We strongly recommend that you make use of this offer, since the result will be a book of considerably improved technical quality. To avoid mistakes and time-consuming correspondence during the production period the conference editors should request special instructions from the publisher well before the beginning of the conference. Manuscripts not meeting the technical standard of the series will have to be returned for improvement.

For further information please contact Springer-Verlag, Physics Editorial Department II, Tiergartenstrasse 17,
D-69121 Heidelberg, Germany

Daniel Joubert (Ed.)

Density Functionals: Theory and Applications

Proceedings of the Tenth Chris Engelbrecht
Summer School in Theoretical Physics
Held at Meerensee, near Cape Town
South Africa, 19–29 January 1997

Springer

Editor

Daniel Joubert
Physics Department
University of the Witwatersrand
P.O. Wits 2050
Johannesburg, South Africa

Cataloging-in-Publication Data applied for.

Die Deutsche Bibliothek - CIP-Einheitsaufnahme

Density functionals : theory and applications ; proceedings of the Tenth Chris Engelbrecht Summer School in Theoretical Physics, held at Meerensee, near Cape Town, South Africa, 19 - 29 January 1997 / Daniel Joubert (ed.). - Berlin ; Heidelberg ; New York ; Barcelona ; Budapest ; Hong Kong ; London ; Milan ; Paris ; Santa Clara ; Singapore ; Tokyo : Springer, 1998
(Lecture notes in physics ; Vol. 500)
ISBN 3-540-63937-3

ISSN 0075-8450
ISBN 3-540-63937-3 Springer-Verlag Berlin Heidelberg New York

This work is subject to copyright. All rights are reserved, whether the whole or part of the material is concerned, specifically the rights of translation, reprinting, re-use of illustrations, recitation, broadcasting, reproduction on microfilms or in any other way, and storage in data banks. Duplication of this publication or parts thereof is permitted only under the provisions of the German Copyright Law of September 9, 1965, in its current version, and permission for use must always be obtained from Springer-Verlag. Violations are liable for prosecution under the German Copyright Law.

© Springer-Verlag Berlin Heidelberg 1998

Printed in Germany

The use of general descriptive names, registered names, trademarks, etc. in this publication does not imply, even in the absence of a specific statement, that such names are exempt from the relevant protective laws and regulations and therefore free for general use.

Typesetting: Camera-ready by the authors and editor
Cover design: *design & production* GmbH, Heidelberg
SPIN: 10644026 55/3144-543210 - Printed on acid-free paper

Preface

Density functional theory for electrons in atoms, molecules and solids has a long and increasingly rich and successful history that has its roots in the early development of quantum mechanics. During the past decades, application of density functional theory has become the most effective method for the calculation of ground-state structural and electronic properties of molecules and solids. Exchange-and-correlation functionals of increasing sophistication now make it possible to perform calculations to chemical accuracy ($\sim$ 1 kcal/mol) for many systems. Active research is leading to successful methods dealing with excited states and there is significant progress in the development of a relativistic density functional formalism.

The 45 lectures delivered at the 10th Chris Engelbrecht Summer School in Theoretical Physics covered a wide range of aspects of density functional theory for electrons, including non-relativistic as well as relativistic approaches. The lecturers are leading experts in density functional theory and their knowledge and enthusiasm made the lectures exciting and the school a great success. The background of the participants was mixed, ranging from graduate students, who were setting out on research programmers, to established practicants of density functional theory. They included experimentalists as well as theoreticians. The material presented at the School and reflected in this volume consequently offers an introduction to and an overview of the subject as well as reviews of some of the most recent developments in the field. The material contained in this volume was contributed by the lecturers and collaborators who did not participate in the school. Not all the material discussed at the school is included here, but the volume covers selected topics in more depth and with extensive references.

The Organising Committee is indebted to the Foundation for Research Development for its continuing financial support, without which it would not be possible to organise this series of summer schools, which is of great value to the South African theoretical physics community. We would like to thank Erna Pheiffer, without whose administrative expertise and unstinting dedication the school would not have been a success.

I would like to thank the members of the Organising Committee for their support in organising the school. Finally I wish to express my thanks to the editors of *Lecture Notes in Physics* for assistance and advice during the preparation of the manuscript.

Exeter, United Kingdom
November 1997 Daniel Joubert

Contents

Lecturers

Walter Kohn, University of California, Santa Barbara
John Perdew, Tulane University
Matthias Ernzerhof, Tulane University
Hardy Gross, Würzburg University
Reiner Dreizler, Frankfurt University

Collaborators

Kieron Burke, Rutgers University
Eberhard Engel, Frankfurt University
Stefan Kurth, Tulane University

Organising Committee

Nithaya Chetty, University of Natal, Pietermaritzburg
Fritz Hahne, University of Stellenbosch
Daniel Joubert, (Chair) University of the Witwatersrand
Frikkie Scholtz, University of Stellenbosch Erna Pheiffer (Secretariat), FRD

List of Participants

Anderson, Ylva `ylvaa@fy.chalmers.se`
Department of Applied Physics, Chalmers University of Technology, Fysikgrand S-412-96, Götenburg, Sweden,

Carter, Rachel `rcarter@scientia.up.ac.za`
Department of Physics, University of Pretoria, Pretoria 0002, South Africa,

Chetty, Nithaya `chetty@paul.phy.unp.ac.za`
Department of Physics, University of Natal, Private Bag x01, Scotsville, 3209, South Africa,

De Lange, Owen `delange@phsc.unp.ac.za`
Department of Physics, University of Natal, Private Bag x01, Scotsville, 3209, South Africa,

De Wet, Antony
P O Box 514, Plettenberg Bay, 6600, South Africa,

Dewhurst, John `dewhurst@physnet.phys.wits.ac.za`
Department of Physics, University of the Witwatersrand, P O Wits, Johannesburg 2050, South Africa,

Dreizler, Reiner `dreizler@th.physik.uni-frankfurt.de`
Institut fü Theoretische Physik, J W Goethe Universität, Frankfurt am Main, Robert Mayer Str. 8-10, D-60054 Frankfurt am Main, Germany,

Engelbrecht, Chris `engelca@karee.unisa.ac.za`
Department of Physics, UNISA, P O Box 392, Pretoria 0001, South Africa,

Ernzerhof, Matthias `matthias@merlin.phy.tulane.edu`
Department of Physics, Tulane University, New Orleans, LA 70118, USA,

Geyer, Hendrik `hbg@sunvax.sun.ac.za`
Department of Physics, University of Stellenbosch, Private Bag x1, Matieland 7602, South Africa,

Gross, Hardy `gross@physik.uni-wuerzburg.de`
Institute for Theoretical Physics, University of Würzburg, Am Hubland, D-97074 Würzburg, Germany,

Hahne, Fritz `fjwh@maties.sun.ac.za`
Dean of Natural Sciences, University of Stellenbosch, Private Bag x1, Matieland 7602, South Africa,

Hlungwani, Ndhuma phala_mf@unin1.unorth.ac.za
Department of Physics, University of the North, Private Bag x1106, Sovenga 0727, South Africa,

Hoch, Michael hoch@physnet.phys.wits.ac.za
Department of Physics, University of the Witwatersrand, P O Wits, Johannesburg 2050, South Africa,

Hoffman, Matie nwfsmjh@med.uovs.ac.za
Department of Physics, University of the Free State, P O Box 339, Bloemfontein 9300, South Africa,

Hult, Erika tfyhult@fy.chalmers.se
Department of Applied Physics, Chalmers University of Technology, Fysikgrand S-412-96, Götenburg, Sweden,

Joubert, Daniel joubert@physnet.phys.wits.ac.za
Department of Physics, University of the Witwatersrand, P O Wits, Johannesburg 2050, South Africa,

Kganyago, Khomotso reggy@unin1.unorth.ac.za
Materials Modelling Centre, University of the North, Private Bag x1106, Sovenga 0727, South Africa,

Kohn, Walter kohn@physics.ucsb.edu
Department of Physics, University of California, Santa Barbara, CA 93106-9530, USA,

Lynch, Richard lynch@physnet.phys.wits.ac.za
Department of Physics, University of the Witwatersrand, P O Wits, Johannesburg 2050, South Africa,

Machi, Isaac machi@schonlan.src.wits.ac.za
Schnland Research Centre, University of the Witwatersrand, Private Bag x3, Johannesburg 2050, South Africa,

Malherbe, Johan malherbe@scientia.up.ac.za
Department of Physics, University of Pretoria, Pretoria 0002, South Africa,

Martin, Richard rmartin@uiuc.edu
Department of Physics, University of Illinois, 1110 W. Green Street, Urbana, IL 61801, USA,

Moncada, Antonie
Department of Physics, University of Stellenbosch, Private Bag x1, Matieland 7602, South Africa,

Mpholle, Lekopa ncoville@aurum.chem.wits.ac.za
Department of Chemistry, University of the Witwatersrand, P O Wits 2050, Johannesburg, South Africa,

Müller, Martin muller@physnet.phys.wits.ac.za
Department of Physics, University of the Witwatersrand, P O Wits 2050, South Africa,

Ngoepe, Phuti ngoepe_pe@unin1.unorth.ac.za
Materials Modelling Centre, University of the North, Private Bag x1106, Sovenga 0727, South Africa,

Nkosibomvu, Zola `nkosi@physnet.phys.wits.ac.za`
Department of Physics, University of the Witwatersrand, P O Wits 2050,
South AFrica,

Ntoahae, Peter
Department of Physics, University of the North, Private Bag x1106,
Sovenga 0727, South Africa,

Paulin-Campbell, Alastair `paulin@physnet.phys.wits.ac.za`
Department of Physics, University of the Witwatersrand, P O Wits 2050,
South Africa,

Perdew, John `perdew@mailhost.tcs.tulane.edu`
Department of Physics, Quantum Theory Group, Tulane University, New
Orleans, LA 70118, USA,

Pettifor, David `david.pettifor@materials.oxford.ac.uk`
Department of Materials, University of Oxford, Parks Road, Oxford OX1
3PH, UK,

Phala, Feredi `phala_mf@unin1.unorth.ac.za`
Materials Modelling Centre, University of the North, Private Bag x1106,
Sovenga 0727, South Africa,

Rydberg, Henrik `rydberg@fy.chalmers.se`
Department of Applied Physics, Chalmers University of Technology,
Fysikgrand S-412-96, Götenborg, Sweden,

Scholtz, Frikkie `fgs@sunvax.sun.ac.za`
Department of Physics, University of Stellenbosch, Private Bag x1,
Matieland 7602, South Africa,

Stoop, Paulina `stooppm@karee.unisa.ac.za`
Department of Physics, UNISA, P O Box 392, Pretoria 0001, South
Africa,

Tlali, Spirit `spirit@physics.nul.ls`
Department of Physics, National University of Lesotho, P O ROMA 180,
Lesotho,

Tseki, P F `tseki@chem.nul.ls`
Department of Chemistry, National University of Lesotho, P O ROMA
180, Lesotho,

Van Biljon, Andrew `s9209301@itu3.sun.ac.za`
Department of Physics, University of Stellenbosch, Private Bag x1,
Matieland 7602, South Africa,

Van der Ventel, Brandon `9020381@sunvax.sun.ac.za`
Department of Physics, University of Stellenbosch, Private Bag x1,
Matieland 7602, South Africa,

Verwoerd,Wynand `verwows@karee.unisa.ac.za`
Department of Physics, UNISA, P O Box 392, Pretoria 0001, South
Africa,

Wyngaardt, Shaun `shaun@sun014.sun.ac.za`
Department of Physics, University of Stellenbosch, Private Bag x1, Matieland 7602, South Africa

Fundamentals of Density Functional Theory

Walter Kohn

University of California, Department of Physics, Santa Barbara, CA 93106

1 Introduction

Density functional theory (DFT) is primarily a theory of the electronic structure of atoms, molecules and solids in their ground states, in which the electronic density distribution $n(\mathbf{r})$ plays the central role.

Traditionally electronic structure problems were attacked by finding approximate solutions of the Schrödinger equation

$$\left\{-\frac{h^2}{2m}\Sigma_j\nabla_j^2+\sum_j v(\mathbf{r}_j)+\frac{1}{2}\sum_{j\neq i}\frac{e^2}{|\mathbf{r}_j-\mathbf{r}_i|}-E\right\}\Phi(\mathbf{r}_1,\cdots\mathbf{r}_N)=0\ , \qquad (1)$$

where

$$v(\mathbf{r}_j)=-\sum_i\frac{Z_ic^2}{|\mathbf{r}_j-\mathbf{R}_i|}\ , \qquad (2)$$

N is the number of electrons, i and j run from 1 to N, $\mathbf{R}_i$ and Z_i are the coordinates and atomic numbers of the nuclei. The spin variables have been suppressed. The boundary condition is that Φ vanishes when any $|\mathbf{r}_i| \to \infty$. The most important characteristics of the ground state electronic structure are the electronic density distribution $n(\mathbf{r})$ and the total energy E, for given locations $\mathbf{R}_i$ of the nuclei.

Eq. (1) is a non-trivial many body problem even for N is as small as 2 or 3 (He, Li, H_2 etc.). The traditional configuration interaction methods of quantum chemistry can provide the ground state energy of molecules with the chemically required accuracy of $\sim 0.2\,eV$ only up to a maximum number N_0, which is in the vicinity of 10. For large values of N, the computing time T rises very rapidly so that, even with major improvements in computing power, one runs into an almost impenetrable "wall", and traditional methods generally cannot handle chemical systems with $N \gtrsim 10$.

DFT has attracted the interests of solid state physicists, and in recent years also of chemists and others, primarily for the following three reasons: (1) the 3-dimensional density, $n(\mathbf{r})$, is much easier to think about than the 3N-dimensional wavefunction, Φ; (2) computational simplicity; and (3) its capability to handle infinite periodic systems (*no* N-dependence) and non-periodic systems of very many atoms, currently on the order of 10^3.

In section 2 I sketch the elements of DFT. In section 3, I list, and briefly comment on, the most important generalizations, Section 4 presents concluding remarks. Many quantitative applications will be given by other speakers.

The aim of these lectures is to provide a basis for the following more specialized presentations. There exist several excellent monographs and reviews such as [1], [6].

2 Basic Density Functional Theory

DFT in its modern form originated from the observation [2] that, for an interacting N-electron system in a general external potential $v(\mathbf{r})$, a knowledge of the groundstate density $n(\mathbf{r})$ **uniquely** determines $v(\mathbf{r})$ (to within a physically irrelevant additive constant):

$$n(\mathbf{r}) \rightarrow v(\mathbf{r}) \ . \tag{3}$$

This is true whether the ground state is non-degenerate or degenerate (In the latter case any of the possible ground state densities uniquely determines the potential $v(\mathbf{r})$).

Because of the key importance of this observation, the simple proof, for non-degenerate ground states, will now be presented:

Let $v(\mathbf{r})$ be the external potential of the system, with associated ground state density $n(\mathbf{r})$, total number of particles $N = \int n(\mathbf{r})d\mathbf{r}$, Hamiltonian H and ground state and energy, Ψ and E,

$$v : H, N, n(\mathbf{r}), \Phi, E \ . \tag{4}$$

Similarly consider a second system of N particles with

$$v' : H', n'(\mathbf{r}), N, \Phi', E' \ , \tag{5}$$

where $v' \neq v + C$, and hence $\Psi' \neq \Psi$. Then, by the Rayleigh Ritz variational principle

$$E = (\Phi, H\Phi) < (\Phi', H\Phi') = (\Phi', H'\Phi') + \int (v(\mathbf{r}) - v'(\mathbf{r}))n'(\mathbf{r})d\mathbf{r} \ , \tag{6}$$

or

$$E < E' + \int (v(\mathbf{r}) - v'(\ bfr))n'(\mathbf{r})d\mathbf{r} \ . \tag{7a}$$

The inequality, <, follows from the fact that $\Psi' \neq \Psi$. Similarly,

$$E' < E + \int (v'(\mathbf{r}) - v(\mathbf{r}))n(\mathbf{r})d\mathbf{r} \ . \tag{7b}$$

Adding the inequalities (7a) and (7b) gives

$$(E + E') < (E + E') + \int (v(\mathbf{r}) - v'(\mathbf{r}))(n'(\mathbf{r}) - n(\mathbf{r}))d\mathbf{r} \ . \tag{8}$$

The possibility $n'(\mathbf{r}) \equiv n(\mathbf{r})$ is excluded since (8) would result in $0 < 0$. Thus any potential $v'(\mathbf{r})$ except $v(\mathbf{r}) + c$, leads to an $n'(\mathbf{r}) \not\equiv n(\mathbf{r})$. Q.E.D.

The simple observation (3) has far-reaching consequences. Since $n(\mathbf{r})$ determines $v(\mathbf{r})$ and, trivially, also N, it determines H; hence, implicitly, also all properties derivable from H such as the many electron ground state wave function $\Phi(\mathbf{r}_1 \cdots \mathbf{r}_N)$ and energy E, excited state wave functions and energies, Green's functions etc.

The total ground state energy of a system can be written as

$$E = (\Phi, V\Phi) + (\Phi, (T+U)\Phi) \ , \tag{9}$$

where the terms on the right hand side are the expectation values of the external potential-, kinetic energy- and interaction energy operators. Clearly

$$(\Phi, V\Phi) = \int v(\mathbf{r}) n(\mathbf{r}) d\mathbf{r} \ , \tag{10}$$

while the quantity

$$F[n] \equiv (\Phi, (T+U)\Phi) \tag{11}$$

is a functional of $n(\mathbf{r})$ (through $\Psi[n]$, defined for any physical $n(\mathbf{r})$ corresponding to **some** $v(\mathbf{r})$. (Such densities are called v-representable.) Thus,

$$E \equiv E_{v(\mathbf{r})}[n(\mathbf{r})] \equiv \int v(\mathbf{r})\, n(\mathbf{r}) d\mathbf{r} + F[n(\mathbf{r})] \ . \tag{12}$$

Using the Rayleigh Ritz principle for the ground state energy leads rather simply to the conclusion that, for a given $v(\mathbf{r})$, the expression (12) is a **minimum** for the correct ground state density $n(\mathbf{r})$ [2]:

Here is the original proof, given for a non-degenerate groundstate Φ' with energy E_0. Let Φ' be a trial state. Then, by the conventional Rayleigh-Ritz principle,

$$E[\Phi'] \equiv (\Phi', (T+V+U)\Phi') \geq E_0 \ , \tag{13}$$

or

$$E_v[n'(\mathbf{r})] \equiv \int v(\mathbf{r}) n'(\mathbf{r}) d\mathbf{r} + F[n'(\mathbf{r})] \geq E_0 \ . \tag{14}$$

The equality sign holds only if $\Phi' = \Phi$. This is the Hohenberg-Kohn energy variational principle. Note that, so far, $F[n'(\mathbf{r})]$ is defined only abstractly through Eq. (11) but not yet constructively.

The proof can be extended to degenerate ground states, leading again to Eq. (14). The equality is now obtained for *any* $n'(\mathbf{r})$ of one of the many groundstates. An improved derivation was later given independently by [4], and [5], by the so-called constrained search process:

By the conventional Rayleigh Ritz principle we have

$$E_0 = \min_{\Phi'}(\Phi', H\Phi') \tag{15}$$

We now sort all trial functions into classes according to the densities $n'(\mathbf{r})$ to which they give rise. We then minimize in two stages

$$E_0 \equiv \min_{n'(\mathbf{r})} \min_{\substack{\Phi' \\ (n'[\Phi']=n'(\mathbf{r}))}} (\Phi', H\Phi') = \min_{n'(\mathbf{r})}\left[\int v(\mathbf{r})n'(\mathbf{r})\, d\mathbf{r} + F[n'(\mathbf{r})]\right] , \tag{16}$$

where

$$F\left[n'(\mathbf{r})\right] \equiv \min_{\substack{\Phi' \\ (n'[\Phi']=n'(\mathbf{r}))}} (\Phi', (T+U)\Phi' , \tag{17}$$

i.e. $F[n'(\mathbf{r})]$ is the minimum subject to the constraint $n[\Phi'] = n'(\mathbf{r})$. (Note that the Hohenberg-Kohn definition of $F[n'(\mathbf{r})]$ applies only to densities which are ground state densities, while the definition (17) pertains to a broader class of densities. Further, degenerate states are automatically covered.)

Thus if the functional $F[n(\mathbf{r})]$ is known with sufficient accuracy, the ground state energy and density for any electronic system, no matter the number of electrons, can be determined by minimizing (14) with respect to the 3-dimensional $n'(\mathbf{r})$. (The crudest Ansatz for $F[n(\mathbf{r})]$ gives the familiar Thomas Fermi approximation).

The functional $F[n]$ represents the kinetic and interaction energies. It is advantageous to write it as

$$F[n] \equiv T_s[n] + \frac{1}{2}\int \frac{n(\mathbf{r})n(\mathbf{r}')}{|\mathbf{r}-\mathbf{r}'|}\, d\mathbf{r}d\mathbf{r}' + E_{xc}[n] , \tag{18}$$

where $T_s[n]$ is the kinetic energy of non-interacting electrons of density $n(\mathbf{r})$, the next term is the classical (or mean value) electron-electron interaction energy, and $E_{xc}[n]$ is the remainder, normally a small fraction of $F[n]$. The first two terms are chosen so that, when E_{xc} is completely neglected, minimization of $E_{v(\mathbf{r})}[n]$, Eq. (8) leads to the Hartree equations. **With** inclusion of $E_{xc}[n]$ one is led to the so-called Kohn Sham (KS) self-consistent equations, [3],

$$\left[-\frac{\hbar^2}{2m}\nabla^2 + v(\mathbf{r}) + \int \frac{n(\mathbf{r}')}{|\mathbf{r}-\mathbf{r}'|}\, d\mathbf{r}' + v_{xc}(\mathbf{r})\right]\varphi_j(\mathbf{r}) = \epsilon_j\varphi_j(\mathbf{r}) , \tag{19}$$

where

$$v_{xc}(\mathbf{r}) \equiv \frac{\delta E_{xc}[n(\mathbf{r})]}{\delta n(\mathbf{r})} \tag{20}$$

and

$$n(\mathbf{r}) = \sum_1^N |\varphi_j(\mathbf{r})|^2 . \tag{21}$$

The total energy is

$$E = \sum_1^N \epsilon_j - \frac{1}{2}\int \frac{n(\mathbf{r})\, n(\mathbf{r}')}{|\mathbf{r}-\mathbf{r}'|} d\mathbf{r}d\mathbf{r}' - \int v_{xc}(\mathbf{r})\, n(\mathbf{r}) d\mathbf{r} + E_{xc}[n(\mathbf{r})] . \tag{22}$$

In principle, if the exact $E_{xc}[n]$ were used in these equations, the resulting self-consistent density $n(\mathbf{r})$ and energy, E, would be **exact**, including all many body effects. Of course, in real life one must content oneself with approximation for E_{xc}. The simplest of these–which has become a benchmark for all others–is the local density approximation, LDA:

$$E_{xc}^{LDA}[n(\mathbf{r})] \equiv \int \epsilon_{xc}(n(\mathbf{r}))n(\mathbf{r})d\mathbf{r} \ , \tag{23}$$

where $\epsilon_{xc}(n)$ is the exchange-correlation energy per electron of a **uniform** electron gas of density n, which is known to great accuracy (~0.1%) from independent studies of the uniform electron gas. The corresponding v_{xc} is, by Eq. (20), given by

$$v_{xc}(\mathbf{r}) = n(\mathbf{r})(\partial\epsilon_{xc}(n)/\partial n)_{n=n(\mathbf{r})} + \epsilon_{xc}(n(\mathbf{r})) \ . \tag{24}$$

Refinements of the LDA include the so-called generalized gradient approximations, GGA, ([7]),

$$E_{xc}^{GGA} = \int f(n(\mathbf{r}), |\nabla n(\mathbf{r})|)d\mathbf{r} \ . \tag{25}$$

Judicious choice of the form of $f(n, |\nabla n|)$, a function of **two** variables, typically reduces the error of the LDA by a factor of 2 to 10. This is an area of active research. (The most obvious gradient correction to the LDA is a systematic gradient expansion

$$E_{xc} = \int \epsilon_{xc}(n(\mathbf{r}))n(\mathbf{r})d\mathbf{r} + \int g(n(\mathbf{r}))|\nabla n(\mathbf{r})|^2 d\mathbf{r} + \cdots \tag{26}$$

but this is generally of little quantitative value.)

3 Generalizations

(See Parr and Yang (1989) [6] and Dreizler and Gross (1990) [1].)
The principle of DFT, considering the density $n(\mathbf{r})$ as the basic variable and representing many body effects in a density functional such as $E_{xc}[n(\mathbf{r})]$, has been generalized in many directions. Here we list the most important ones, with indications of the relevant densities.

a) **Degenerate Ground States.**

b) **Spin Density Functional Theory.**
Separate spin up and spin down densities $n_\uparrow(\mathbf{r}), n_\downarrow(\mathbf{r})$ are introduced. Paramagnetism, cooperative magnetism.

c) **Multicomponent Systems.**
$n_1(\mathbf{r}), \cdots n_M(\mathbf{r})$. Nuclei (= protons + neutrons); electron-hole droplets in semiconductors.

d) **Free Energy for Finite Temperature.**
$n(\mathbf{r}; T)$. Plasmas

e) **Excited States.**
$n(\mathbf{r}) = M^{-1}\sum_1^M n_j(\mathbf{r})$, where $n_j(\mathbf{r})$ is the density of the j' the excited state
$E = \frac{1}{M}\sum_1^M E_j$.
(See also **i** below.)

f) **Orbital Magnetism.**
$n(\mathbf{r})$ (normal density) and $\tilde{n}(\mathbf{r})$ (superconducting density)
Diamagnetism, quantum Hall effect.

g) **Superconductors.**
$n(\mathbf{r})$ (normal density) and $\tilde{n}(\mathbf{r})$ (superconducting density)

h) **Relativistic Electrons.**
$n(\mathbf{r})$ $(E_{xc}^{rel}(n) \neq E_{xc}^{non}(n))$

i) **Time-Dependent Phenomena.**
$v(\mathbf{r}, t), n(\mathbf{r}, t)$
Applications to first order response theory for given frequency ω. Poles of response function give excited state energies.

4 Concluding Remarks

More than 30 years after its beginning, DFT has become one of the standard methods for studying ground state energies and density distributions of systems containing many electrons. It is in the writer's opinion, limited to a modest accuracy, typically $0.1 - 0.5\, eV$, for physically or chemically relevant energy differences. For determination of structures (corresponding to energy minima) it is generally surprisingly accurate, giving interatomic distances with a typical accuracy of $\pm 0.02 \mathring{A}$. In the last decade, thoughtful and energetic analyses and improved approximations, which will be described in subsequent lectures, have been improving the accuracy of DFT until, today, it is, for many types of systems, close to or within the so-called chemical accuracy of $0.1 - 0.2\, eV$. Although further progress is becoming more difficult it is continuing.

Two other important developments, covered in later lectures of this course have been taking place in recent years: DFT algorithms which scale *linearly* in the number of atoms; and inclusion of polarization (or Van der Waals–) energies, which are not included in traditional approximations.

An important concept in DFT is the exchange-correlation hole $n_{xc}(\mathbf{r}, \mathbf{r}')$ which describes the depletion of the density at point $\mathbf{r}$, given that there is an electron at point $\mathbf{r}'$. A knowledge of n_{xc} (as function of the "coupling constant" e) allows one to calculate the all-important E_{xc} of Eq. (13). Detailed studies of different classes of chemical and physical systems are providing increasing understanding of n_{xc} and throw new light on electronic energies.

It would be too long to try and summarize all the other ongoing directions of research and applications. Let me conclude by remarking that the electronic density distribution, $n(\mathbf{r})$, has now joined the traditional single electron orbitals, $\varphi_j(\mathbf{r})$, as an important characterization, both qualitative and quantitative, of physical and chemical systems with many electrons.

Acknowledgment

I would like to express my appreciation to the organizers of this school for giving me an opportunity to contribute to the scientific life of the new South Africa. Support of NSF grant No. DMR9630452 is gratefully acknowledged.

References

1. R.M. Dreizler and E.K.U. Gross, *Density Functional Theory* (Springer, Berlin, 1990). (Primarily for physicists).
2. P. Hohenberg and W. Kohn, Inhomogeneous Electron Gas, Phys. Rev. **136**, B864 (1964).
3. W. Kohn, and L. J. Sham, Self-Consistent Equations including Exchange and Correlation Effects, Phys. Rev. **140A**, 1133 (1965).
4. M. Levy, Electron Densities in Search of Hamiltonians, Phys. Rev. A **26**, 1200 (1982).
5. E. H. Lieb, *Physics as Natural Philosophy*, ed. by A. Skimory and H. Feshback (MIT Press, Cambridge, MA, 1982), 111.
6. R.G. Parr and W. Yang, *Density-Functional Theory of Atoms and Molecules* (Oxford University Press, New York, 1989). (Primarily for chemists).
7. J. P. Perdew, K. Burke, M. Ernzerhof, Generalized Gradient Approximation Made Simple, Phys. Rev. Lett. **77**, 3865 (1996).

Density Functionals for Non-Relativistic Coulomb Systems

John P. Perdew and Stefan Kurth

Department of Physics and Quantum Theory Group
Tulane University, New Orleans, LA 70118, USA

Outline

1 Introduction

1.1 Quantum Mechanical Many-Electron Problem

The material world of everyday experience, as studied by chemistry and condensed-matter physics, is built up from electrons and a few (or at most a few hundred) kinds of nuclei. The basic interaction is electrostatic or Coulombic: An electron at position $\mathbf{r}$ is attracted to a nucleus of charge Z at $\mathbf{R}$ by the potential energy $-Z/|\mathbf{r}-\mathbf{R}|$, a pair of electrons at $\mathbf{r}$ and $\mathbf{r}'$ repel one another by the potential energy $1/|\mathbf{r}-\mathbf{r}'|$, and two nuclei at $\mathbf{R}$ and $\mathbf{R}'$ repel one another as $Z'Z/|\mathbf{R}-\mathbf{R}'|$. The electrons must be described by quantum mechanics, while the more massive nuclei can sometimes be regarded as classical particles. All of the electrons in the lighter elements, and the chemically important valence electrons in most elements, move at speeds much less than the speed of light, and so are non-relativistic.

In essence, that is the simple story of practically everything. But there is still a long path from these general principles to theoretical prediction of the structures and properties of atoms, molecules, and solids, and eventually to the design of new chemicals or materials. If we restrict our focus to the important class of ground-state properties, we can take a shortcut through density functional theory.

These lectures present an introduction to density functionals for non-relativistic Coulomb systems. The reader is assumed to have a working knowledge of quantum mechanics at the level of one-particle wavefunctions $\psi(\mathbf{r})$ [1]. The many-electron wavefunction $\Psi(\mathbf{r}_1, \mathbf{r}_2, \ldots, \mathbf{r}_N)$ [2] is briefly introduced here, and then replaced as basic variable by the electron density $n(\mathbf{r})$. Various terms of the total energy are defined as functionals of the electron density, and some formal properties of these functionals are discussed. The most widely-used density functionals - the local spin density and generalized gradient approximations - are then introduced and discussed. At the end, the reader should be prepared to approach the broad literature of quantum chemistry and condensed-matter physics in which these density functionals are applied to predict diverse properties: the shapes and sizes of molecules, the crystal structures of solids, binding or atomization energies, ionization energies and electron affinities, the heights of energy barriers to various processes, static response functions, vibrational frequencies of nuclei, etc. Moreover, the reader's approach will be an informed and discerning one, based upon an understanding of where these functionals come from, why they work, and how they work.

These lectures are intended to teach at the introductory level, and not to serve as a comprehensive treatise. The reader who wants more can go to several excellent general sources [3–5] or to the original literature. Atomic units (in which all electromagnetic equations are written in cgs form, and the fundamental constants $\hbar$, e^2, and m are set to unity) have been used throughout.

1.2 Summary of Kohn-Sham Spin-Density Functional Theory

This introduction closes with a brief presentation of the Kohn-Sham [6] spin-density functional method, the most widely-used method of electronic-structure calculation in condensed-matter physics and one of the most widely-used methods in quantum chemistry. We seek the ground-state total energy E and spin densities $n_\uparrow(\mathbf{r})$, $n_\downarrow(\mathbf{r})$ for a collection of N electrons interacting with one another and with an external potential $v(\mathbf{r})$ (due to the nuclei in most practical cases). These are found by the self-consistent solution of an auxiliary (fictitious) one-electron Schrödinger equation:

$$\left(-\frac{1}{2}\nabla^2 + v(\mathbf{r}) + u([n];\mathbf{r}) + v_{\mathrm{xc}}^{\sigma}([n_\uparrow, n_\downarrow];\mathbf{r})\right)\psi_{\alpha\sigma}(\mathbf{r}) = \varepsilon_{\alpha\sigma}\psi_{\alpha\sigma}(\mathbf{r}) \ , \tag{1}$$

$$n_\sigma(\mathbf{r}) = \sum_{\alpha}\theta(\mu - \varepsilon_{\alpha\sigma})|\psi_{\alpha\sigma}(\mathbf{r})|^2 \ . \tag{2}$$

Here $\sigma = \uparrow$ or $\downarrow$ is the z-component of spin, and α stands for the set of remaining one-electron quantum numbers. The effective potential includes a classical Hartree potential

$$u([n];\mathbf{r}) = \int \mathrm{d}^3 r' \frac{n(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|} \ , \tag{3}$$

$$n(\mathbf{r}) = n_\uparrow(\mathbf{r}) + n_\downarrow(\mathbf{r}) \ , \tag{4}$$

and a multiplicative spin-dependent exchange-correlation potential $v_{\mathrm{xc}}^{\sigma}([n_\uparrow, n_\downarrow];\mathbf{r})$, which is a functional of the spin densities. The step function $\theta(\mu - \varepsilon_{\alpha\sigma})$ in Eq. (2) ensures that all Kohn-Sham spin orbitals with $\varepsilon_{\alpha\sigma} < \mu$ are singly occupied, and those with $\varepsilon_{\alpha\sigma} > \mu$ are empty. The chemical potential μ is chosen to satisfy

$$\int \mathrm{d}^3 r\, n(\mathbf{r}) = N \ . \tag{5}$$

Because Eqs. (1) and (2) are interlinked, they can only be solved by iteration to self-consistency.

The total energy is

$$E = T_{\mathrm{s}}[n_\uparrow, n_\downarrow] + \int \mathrm{d}^3 r\, n(\mathbf{r}) v(\mathbf{r}) + U[n] + E_{\mathrm{xc}}[n_\uparrow, n_\downarrow] \ , \tag{6}$$

where

$$T_{\mathrm{s}}[n_\uparrow, n_\downarrow] = \sum_{\sigma}\sum_{\alpha}\theta(\mu - \varepsilon_{\alpha\sigma})\langle\psi_{\alpha\sigma}| - \frac{1}{2}\nabla^2|\psi_{\alpha\sigma}\rangle \tag{7}$$

is the non-interacting kinetic energy, a functional of the spin densities because (as we shall see) the external potential $v(\mathbf{r})$ and hence the Kohn-Sham orbitals are functionals of the spin densities. In our notation,

$$\langle \psi_{\alpha\sigma} | \hat{O} | \psi_{\alpha\sigma} \rangle = \int \mathrm{d}^3 r \psi^*_{\alpha\sigma}(\mathbf{r}) \hat{O} \psi_{\alpha\sigma}(\mathbf{r}) \ . \tag{8}$$

The second term of Eq. (6) is the interaction of the electrons with the external potential. The third term of Eq. (6) is the Hartree electrostatic self-repulsion of the electron density

$$U[n] = \frac{1}{2} \int \mathrm{d}^3 r \int \mathrm{d}^3 r' \frac{n(\mathbf{r}) n(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|} \ . \tag{9}$$

The last term of Eq. (6) is the exchange-correlation energy, whose functional derivative (as explained later) yields the exchange-correlation potential

$$v^\sigma_{\mathrm{xc}}([n_\uparrow, n_\downarrow]; \mathbf{r}) = \frac{\delta E_{\mathrm{xc}}}{\delta n_\sigma(\mathbf{r})} \ . \tag{10}$$

Not displayed in Eq. (6), but needed for a system of electrons and nuclei, is the electrostatic repulsion among the nuclei. E_{xc} is defined to include everything else omitted from the first three terms of Eq. (6).

If the exact dependence of E_{xc} upon $n_\uparrow$ and $n_\downarrow$ were known, these equations would predict the exact ground-state energy and spin-densities of a many-electron system. The forces on the nuclei , and their equilibrium positions, could then be found from $-\frac{\partial E}{\partial \mathbf{R}}$.

In practice, the exchange-correlation energy functional must be approximated. The local spin density [6,7] (LSD) approximation has long been popular in solid state physics:

$$E^{\mathrm{LSD}}_{\mathrm{xc}}[n_\uparrow, n_\downarrow] = \int \mathrm{d}^3 r n(\mathbf{r}) e_{\mathrm{xc}}(n_\uparrow(\mathbf{r}), n_\downarrow(\mathbf{r})) \ , \tag{11}$$

where $e_{\mathrm{xc}}(n_\uparrow, n_\downarrow)$ is the known [8–10] exchange-correlation energy per particle for an electron gas of uniform spin densities $n_\uparrow$, $n_\downarrow$. More recently, generalized gradient approximations (GGA's) [11–21] have become popular in quantum chemistry:

$$E^{\mathrm{GGA}}_{\mathrm{xc}}[n_\uparrow, n_\downarrow] = \int \mathrm{d}^3 r f(n_\uparrow, n_\downarrow, \nabla n_\uparrow, \nabla n_\downarrow) \ . \tag{12}$$

The input $e_{\mathrm{xc}}(n_\uparrow, n_\downarrow)$ to LSD is in principle unique, since there is a possible system in which $n_\uparrow$ and $n_\downarrow$ are constant and for which LSD is exact. At least in this sense, there is no unique input $f(n_\uparrow, n_\downarrow, \nabla n_\uparrow, \nabla n_\downarrow)$ to GGA. These lectures will stress a conservative "philosophy of approximation" [20, 21], in which we construct a nearly-unique GGA with all the known correct formal features of LSD, plus others.

The equations presented here are really all that we need to do a practical calculation for a many-electron system. They allow us to draw upon the intuition and experience we have developed for one-particle systems. The many-body effects are in $U[n]$ (trivially) and $E_{\mathrm{xc}}[n_\uparrow, n_\downarrow]$ (less trivially), but we shall also develop an intuitive appreciation for E_{xc}.

While E_{xc} is often a relatively small fraction of the total energy of an atom, molecule, or solid (minus the work needed to break up the system into separated electrons and nuclei), the contribution from E_{xc} is typically about 100 % or more of the chemical bonding or atomization energy (the work needed to break up the system into separated neutral atoms). E_{xc} is a kind of "glue", without which atoms would bond weakly if at all. Thus, accurate approximations to E_{xc} are essential to the whole enterprise of density functional theory. Table 1 shows the typical relative errors we find from self-consistent calculations within the LSD or GGA approximations of Eqs. (11) and (12). Table 2 shows the mean absolute errors in the atomization energies of 20 molecules when calculated by LSD, by GGA, and in the Hartree-Fock approximation . Hartree-Fock treats exchange exactly, but neglects correlation completely. While the Hartree-Fock total energy is an upper bound to the true ground-state total energy, the LSD and GGA energies are not.

Table 1. Typical errors for atoms, molecules, and solids from self-consistent Kohn-Sham calculations within the LSD and GGA approximations of Eqs. (11) and (12). Note that there is typically some cancellation of errors between the exchange (E_x) and correlation (E_c) contributions to E_{xc}. The "energy barrier" is the barrier to a chemical reaction that arises at a highly-bonded intermediate state.

Property	LSD	GGA
E_x	5 % (not negative enough)	0.5 %
E_c	100 % (too negative)	5 %
bond length	1 % (too short)	1 % (too long)
structure	overly favors close packing	more correct
energy barrier	100 % (too low)	30 % (too low)

Table 2. Mean absolute error of the atomization energies for 20 molecules, evaluated by various approximations. (1 hartree = 27.21 eV) (From Ref. [20].)

Approximation	Mean absolute error (eV)
Unrestricted Hartree-Fock	3.1 (underbinding)
LSD	1.3 (overbinding)
GGA	0.3 (mostly overbinding)
Desired "chemical accuracy"	0.05

In most cases we are only interested in small total-energy changes associated with re-arrangements of the outer or valence electrons, to which the inner or core electrons of the atoms do not contribute. In these cases, we can replace each core by the pseudopotential [22] it presents to the valence electrons, and then expand the valence-electron orbitals in an economical and convenient basis of plane waves. Pseudopotentials are routinely combined with density functionals. Although the most realistic pseudopotentials are nonlocal operators and not simply local or multiplication operators, and although density functional theory in principle requires a local external potential, this inconsistency does not seem to cause any practical difficulties.

There are empirical versions of LSD and GGA, but these lectures will only discuss non-empirical versions. If every electronic-structure calculation were done at least twice, once with nonempirical LSD and once with nonempirical GGA, the results would be useful not only to those interested in the systems under consideration but also to those interested in the development and understanding of density functionals.

2 Wavefunction Theory

2.1 Wavefunctions and Their Interpretation

We begin with a brief review of one-particle quantum mechanics [1]. An electron has spin $s = \frac{1}{2}$ and z-component of spin $\sigma = +\frac{1}{2}$ ($\uparrow$) or $-\frac{1}{2}$ ($\downarrow$). The Hamiltonian or energy operator for one electron in the presence of an external potential $v(\mathbf{r})$ is

$$\hat{h} = -\frac{1}{2}\nabla^2 + v(\mathbf{r}) \ . \tag{13}$$

The energy eigenstates $\psi_\alpha(\mathbf{r},\sigma)$ and eigenvalues ε_α are solutions of the time-independent Schrödinger equation

$$\hat{h}\psi_\alpha(\mathbf{r},\sigma) = \varepsilon_\alpha \psi_\alpha(\mathbf{r},\sigma) \ , \tag{14}$$

and $|\psi_\alpha(\mathbf{r},\sigma)|^2 \mathrm{d}^3r$ is the probability to find the electron with spin σ in volume element d^3r at $\mathbf{r}$, given that it is in energy eigenstate ψ_α. Thus

$$\sum_\sigma \int \mathrm{d}^3r |\psi_\alpha(\mathbf{r},\sigma)|^2 = \langle\psi|\psi\rangle = 1 \ . \tag{15}$$

Since $\hat{h}$ commutes with $\hat{s}_z$, we can choose the ψ_α to be eigenstates of $\hat{s}_z$, i.e., we can choose $\sigma =\uparrow$ or $\downarrow$ as a one-electron quantum number.

The Hamiltonian for N electrons in the presence of an external potential $v(\mathbf{r})$ is [2]

$$\begin{aligned} \hat{H} &= \sum_{i=1}^{N} -\frac{1}{2}\nabla_i^2 + \sum_{i=1}^{N} v(\mathbf{r}_i) + \frac{1}{2}\sum_i \sum_{j\neq i} \frac{1}{|\mathbf{r}_i - \mathbf{r}_j|} \\ &= \hat{T} + \hat{V}_{\mathrm{ext}} + \hat{V}_{\mathrm{ee}} \ . \end{aligned} \tag{16}$$

The electron-electron repulsion $\hat{V}_{\mathrm{ee}}$ sums over distinct pairs of different electrons. The states of well-defined energy are the eigenstates of $\hat{H}$:

$$\hat{H}\Psi_k(\mathbf{r}_1\sigma_1,\ldots,\mathbf{r}_N\sigma_N) = E_k\Psi_k(\mathbf{r}_1\sigma_1,\ldots,\mathbf{r}_N\sigma_N) \ , \tag{17}$$

where k is a complete set of many-electron quantum numbers; we shall be interested mainly in the ground state or state of lowest energy, the zero-temperature equilibrium state for the electrons.

Because electrons are fermions, the only physical solutions of Eq. (17) are those wavefunctions that are antisymmetric [2] under exchange of two electron labels i and j:

$$\begin{aligned}&\Psi(\mathbf{r}_1\sigma_1,\ldots,\mathbf{r}_i\sigma_i,\ldots,\mathbf{r}_j\sigma_j,\ldots,\mathbf{r}_N\sigma_N) = \\ &-\Psi(\mathbf{r}_1\sigma_1,\ldots,\mathbf{r}_j\sigma_j,\ldots,\mathbf{r}_i\sigma_i,\ldots,\mathbf{r}_N\sigma_N) \ .\end{aligned} \tag{18}$$

There are $N!$ distinct permutations of the labels $1, 2, \ldots, N$, which by Eq. (18) all have the same $|\Psi|^2$. Thus $N!|\Psi(\mathbf{r}_1\sigma_1,\ldots,\mathbf{r}_N\sigma_N)|^2 \mathrm{d}^3r_1 \ldots \mathrm{d}^3r_N$ is the probability to find *any* electron with spin σ_1 in volume element d^3r_1, etc., and

$$\frac{1}{N!}\sum_{\sigma_1\ldots\sigma_N}\int \mathrm{d}^3r_1\ldots\int \mathrm{d}^3r_N N!|\Psi(\mathbf{r}_1\sigma_1,\ldots,\mathbf{r}_N\sigma_N)|^2 = \int|\Psi|^2 = \langle\Psi|\Psi\rangle = 1 \ . \tag{19}$$

We define the electron spin density $n_\sigma(\mathbf{r})$ so that $n_\sigma(\mathbf{r})\mathrm{d}^3r$ is the probability to find an electron with spin σ in volume element d^3r at $\mathbf{r}$. We find $n_\sigma(\mathbf{r})$ by integrating over the coordinates and spins of the $(N-1)$ other electrons, i.e.,

$$\begin{aligned}n_\sigma(\mathbf{r}) &= \frac{1}{(N-1)!}\sum_{\sigma_2\ldots\sigma_N}\int \mathrm{d}^3r_2\ldots\int \mathrm{d}^3r_N N!|\Psi(\mathbf{r}\sigma,\mathbf{r}_2\sigma_2,\ldots,\mathbf{r}_N\sigma_N)|^2 \\ &= N\sum_{\sigma_2\ldots\sigma_N}\int \mathrm{d}^3r_2\ldots\int \mathrm{d}^3r_N|\Psi(\mathbf{r}\sigma,\mathbf{r}_2\sigma_2,\ldots,\mathbf{r}_N\sigma_N)|^2 \ .\end{aligned} \tag{20}$$

Equations (19) and (20) yield

$$\sum_\sigma\int \mathrm{d}^3r n_\sigma(\mathbf{r}) = N \ . \tag{21}$$

Based on the probability interpretation of $n_\sigma(\mathbf{r})$, we might have expected the right-hand side of Eq. (21) to be 1, but that is wrong; the sum of probabilities of all mutually-exclusive events equals 1, but finding an electron at $\mathbf{r}$ does *not* exclude the possibility of finding one at $\mathbf{r}'$, except in a one-electron system. Eq. (21) shows that $n_\sigma(\mathbf{r})\mathrm{d}^3r$ is the average number of electrons of spin σ in volume element d^3r. Moreover, the expectation value of the external potential is

$$\langle\hat{V}_{\mathrm{ext}}\rangle = \langle\Psi|\sum_{i=1}^N v(\mathbf{r}_i)|\Psi\rangle = \int \mathrm{d}^3r n(\mathbf{r})v(\mathbf{r}) \ , \tag{22}$$

with the electron density $n(\mathbf{r})$ given by Eq. (4).

2.2 Wavefunctions for Non-Interacting Electrons

As an important special case, consider the Hamiltonian for N non-interacting electrons:

$$\hat{H}_{\text{non}} = \sum_{i=1}^{N} \left(-\frac{1}{2}\nabla_i^2 + v(\mathbf{r}_i) \right) \quad . \tag{23}$$

The eigenfunctions of the one-electron problem of Eqs. (13) and (14) are spin orbitals which can be used to construct the antisymmetric eigenfunctions Φ of $\hat{H}_{non}$:

$$\hat{H}_{\text{non}}\Phi = E_{\text{non}}\Phi \quad . \tag{24}$$

Let i stand for $\mathbf{r}_i, \sigma_i$ and construct the Slater determinant or antisymmetrized product [2]

$$\Phi = \frac{1}{\sqrt{N!}} \sum_P (-1)^P \psi_{\alpha_1}(P1)\psi_{\alpha_2}(P2)\ldots\psi_{\alpha_N}(PN) \quad , \tag{25}$$

where the quantum label α_i now includes the spin quantum number σ. Here P is any permutation of the labels $1, 2, \ldots, N$, and $(-1)^P$ equals $+1$ for an even permutation and -1 for an odd permutation. The total energy is

$$E_{\text{non}} = \varepsilon_{\alpha_1} + \varepsilon_{\alpha_2} + \ldots + \varepsilon_{\alpha_N} \quad , \tag{26}$$

and the density is given by the sum of $|\psi_{\alpha_i}(\mathbf{r})|^2$. If any α_i equals any α_j in Eq. (25), we find $\Phi = 0$, which is not a normalizable wavefunction. This is the Pauli exclusion principle : two or more non-interacting electrons may not occupy the same spin orbital.

As an example, consider the ground state for the non-interacting helium atom ($N = 2$). The occupied spin orbitals are

$$\psi_1(\mathbf{r}, \sigma) = \psi_{1s}(\mathbf{r})\delta_{\sigma,\uparrow} \quad , \tag{27}$$

$$\psi_2(\mathbf{r}, \sigma) = \psi_{1s}(\mathbf{r})\delta_{\sigma,\downarrow} \quad , \tag{28}$$

and the 2-electron Slater determinant is

$$\begin{aligned} \Phi(1,2) &= \frac{1}{\sqrt{2}} \begin{vmatrix} \psi_1(\mathbf{r}_1, \sigma_1) & \psi_2(\mathbf{r}_1, \sigma_1) \\ \psi_1(\mathbf{r}_2, \sigma_2) & \psi_2(\mathbf{r}_2, \sigma_2) \end{vmatrix} \\ &= \psi_{1s}(\mathbf{r}_1)\psi_{1s}(\mathbf{r}_2)\frac{1}{\sqrt{2}}\left(\delta_{\sigma_1,\uparrow}\delta_{\sigma_2,\downarrow} - \delta_{\sigma_2,\uparrow}\delta_{\sigma_1,\downarrow}\right) \quad , \end{aligned} \tag{29}$$

which is symmetric in space but antisymmetric in spin (whence the total spin is $S = 0$).

If several different Slater determinants yield the same non-interacting energy E_{non}, then a linear combination of them will be another antisymmetric eigenstate of $\hat{H}_{\text{non}}$. More generally, the Slater-determinant eigenstates of $\hat{H}_{\text{non}}$ define a complete orthonormal basis for expansion of the antisymmetric eigenstates of $\hat{H}$, the interacting Hamiltonian of Eq. (16).

2.3 Wavefunction Variational Principle

The Schrödinger equation (17) is equivalent to a wavefunction variational principle [2]: Extremize $\langle\Psi|\hat{H}|\Psi\rangle$ subject to the constraint $\langle\Psi|\Psi\rangle = 1$, i.e., set the following first variation to zero:

$$\delta\left\{\langle\Psi|\hat{H}|\Psi\rangle/\langle\Psi|\Psi\rangle\right\} = 0 \ . \tag{30}$$

The ground state energy and wavefunction are found by minimizing the expression in curly brackets.

The Rayleigh-Ritz method finds the extrema or the minimum in a *restricted* space of wavefunctions. For example, the Hartree-Fock approximation to the ground-state wavefunction is the single Slater determinant Φ that minimizes $\langle\Phi|\hat{H}|\Phi\rangle/\langle\Phi|\Phi\rangle$. The configuration-interaction ground-state wavefunction [23] is an energy-minimizing linear combination of Slater determinants, restricted to certain kinds of excitations out of a reference determinant. The Quantum Monte Carlo method typically employs a trial wavefunction which is a single Slater determinant times a Jastrow pair-correlation factor [24]. Those widely-used many-electron wavefunction methods are both approximate and computationally demanding, especially for large systems where density functional methods are distinctly more efficient.

The unrestricted solution of Eq. (30) is equivalent by the method of Lagrange multipliers to the unconstrained solution of

$$\delta\left\{\langle\Psi|\hat{H}|\Psi\rangle - E\langle\Psi|\Psi\rangle\right\} = 0 \ , \tag{31}$$

i.e.,

$$\langle\delta\Psi|(\hat{H} - E)|\Psi\rangle = 0 \ . \tag{32}$$

Since $\delta\Psi$ is an arbitrary variation, we recover the Schrödinger equation (17). Every eigenstate of $\hat{H}$ is an extremum of $\langle\Psi|\hat{H}|\Psi\rangle/\langle\Psi|\Psi\rangle$ and vice versa.

The wavefunction variational principle implies the Hellmann-Feynman and virial theorems below and also implies the Hohenberg-Kohn [25] density functional variational principle to be presented later.

2.4 Hellmann-Feynman Theorem

Often the Hamiltonian $\hat{H}_\lambda$ depends upon a parameter λ, and we want to know how the energy E_λ depends upon this parameter. For any normalized variational solution Ψ_λ (including in particular any eigenstate of $\hat{H}_\lambda$), we define

$$E_\lambda = \langle\Psi_\lambda|\hat{H}_\lambda|\Psi_\lambda\rangle \ . \tag{33}$$

Then

$$\frac{\mathrm{d}E_\lambda}{\mathrm{d}\lambda} = \frac{\mathrm{d}}{\mathrm{d}\lambda'}\langle\Psi_{\lambda'}|\hat{H}_\lambda|\Psi_{\lambda'}\rangle\bigg|_{\lambda'=\lambda} + \langle\Psi_\lambda|\frac{\partial\hat{H}_\lambda}{\partial\lambda}|\Psi_\lambda\rangle \ . \tag{34}$$

The first term of Eq. (34) vanishes by the variational principle, and we find the Hellmann-Feynman theorem [26]

$$\frac{\mathrm{d}E_\lambda}{\mathrm{d}\lambda} = \langle \Psi_\lambda | \frac{\partial \hat{H}_\lambda}{\partial \lambda} | \Psi_\lambda \rangle \ . \tag{35}$$

Eq. (35) will be useful later for our understanding of E_{xc}. For now, we shall use Eq. (35) to derive the electrostatic force theorem[26]. Let $\mathbf{r}_i$ be the position of the i-th electron, and $\mathbf{R}_I$ the position of the (static) nucleus I with atomic number Z_I. The Hamiltonian

$$\hat{H} = \sum_{i=1}^{N} -\frac{1}{2}\nabla_i^2 + \sum_i \sum_I \frac{-Z_I}{|\mathbf{r}_i - \mathbf{R}_I|} + \frac{1}{2}\sum_i \sum_{j \neq i} \frac{1}{|\mathbf{r}_i - \mathbf{r}_j|} + \frac{1}{2}\sum_I \sum_{J \neq I} \frac{Z_I Z_J}{|\mathbf{R}_I - \mathbf{R}_J|} \tag{36}$$

depends parametrically upon the position $\mathbf{R}_I$, so the force on nucleus I is

$$\begin{aligned} -\frac{\partial E}{\partial \mathbf{R}_I} &= \left\langle \Psi \middle| -\frac{\partial \hat{H}}{\partial \mathbf{R}_I} \middle| \Psi \right\rangle \\ &= \int \mathrm{d}^3 r n(\mathbf{r}) \frac{Z_I(\mathbf{r} - \mathbf{R}_I)}{|\mathbf{r} - \mathbf{R}_I|^3} + \sum_{J \neq I} \frac{Z_I Z_J (\mathbf{R}_I - \mathbf{R}_J)}{|\mathbf{R}_I - \mathbf{R}_J|^3} \ , \end{aligned} \tag{37}$$

just as classical electrostatics would predict. Eq. (37) can be used to find the equilibrium geometries of a molecule or solid by varying all the $\mathbf{R}_I$ until the energy is a minimum and $-\partial E/\partial \mathbf{R}_I = 0$. Eq. (37) also forms the basis for a possible density functional molecular dynamics , in which the nuclei move under these forces by Newton's second law. In principle, all we need for either application is an accurate electron density for each set of nuclear positions.

2.5 Virial Theorem

The density scaling relations to be presented in section 4, which constitute important constraints on the density functionals, are rooted in the same wavefunction scaling that will be used here to derive the virial theorem [26].

Let $\Psi(\mathbf{r}_1, \ldots, \mathbf{r}_N)$ be any extremum of $\langle \Psi | \hat{H} | \Psi \rangle$ over normalized wavefunctions, i.e., any eigenstate or optimized restricted trial wavefunction (where irrelevant spin variables have been suppressed). For any scale parameter $\gamma > 0$, define the uniformly-scaled wavefunction

$$\Psi_\gamma(\mathbf{r}_1, \ldots, \mathbf{r}_N) = \gamma^{3N/2} \Psi(\gamma \mathbf{r}_1, \ldots, \gamma \mathbf{r}_N) \tag{38}$$

and observe that

$$\langle \Psi_\gamma | \Psi_\gamma \rangle = \langle \Psi | \Psi \rangle = 1 \ . \tag{39}$$

The density corresponding to the scaled wavefunction is the scaled density

$$n_\gamma(\mathbf{r}) = \gamma^3 n(\gamma \mathbf{r}) \ , \tag{40}$$

which clearly conserves the electron number:

$$\int d^3r n_\gamma(\mathbf{r}) = \int d^3r n(\mathbf{r}) = N \ . \tag{41}$$

$\gamma > 1$ leads to densities $n_\gamma(\mathbf{r})$ that are higher (on average) and more contracted than $n(\mathbf{r})$, while $\gamma < 1$ produces densities that are lower and more expanded.

Now consider what happens to $\langle \hat{H} \rangle = \langle \hat{T} + \hat{V} \rangle$ under scaling. By definition of Ψ,

$$\frac{d}{d\gamma} \langle \Psi_\gamma | \hat{T} + \hat{V} | \Psi_\gamma \rangle \bigg|_{\gamma=1} = 0 \ . \tag{42}$$

But $\hat{T}$ is homogeneous of degree -2 in $\mathbf{r}$, so

$$\langle \Psi_\gamma | \hat{T} | \Psi_\gamma \rangle = \gamma^2 \langle \Psi | \hat{T} | \Psi \rangle \ , \tag{43}$$

and Eq. (42) becomes

$$2 \langle \Psi | \hat{T} | \Psi \rangle + \frac{d}{d\gamma} \langle \Psi_\gamma | \hat{V} | \Psi_\gamma \rangle \bigg|_{\gamma=1} = 0 \ , \tag{44}$$

or

$$2 \langle \hat{T} \rangle - \langle \sum_{i=1}^{N} \mathbf{r}_i \cdot \frac{\partial \hat{V}}{\partial \mathbf{r}_i} \rangle = 0 \ . \tag{45}$$

If the potential energy $\hat{V}$ is homogeneous of degree n, i.e., if

$$V(\gamma \mathbf{r}_i, \ldots, \gamma \mathbf{r}_N) = \gamma^n V(\mathbf{r}_i, \ldots, \mathbf{r}_N) \ , \tag{46}$$

then

$$\langle \Psi_\gamma | \hat{V} | \Psi_\gamma \rangle = \gamma^{-n} \langle \Psi | \hat{V} | \Psi \rangle \ , \tag{47}$$

and Eq. (44) becomes simply

$$2 \langle \Psi | \hat{T} | \Psi \rangle - n \langle \Psi | \hat{V} | \Psi \rangle = 0 \ . \tag{48}$$

For example, $n = -1$ for the Hamiltonian of Eq. (36) in the presence of a single nucleus, or more generally when the Hellmann-Feynman forces of Eq. (37) vanish for the state Ψ.

3 Definitions of Density Functionals

3.1 Introduction to Density Functionals

The many-electron wavefunction $\Psi(\mathbf{r}_1\sigma_1,\ldots,\mathbf{r}_N\sigma_N)$ contains a great deal of information - all we could ever have, but more than we usually want. Because it is a function of many variables, it is not easy to calculate, store, apply or even think about. Often we want no more than the total energy E (and its changes), or perhaps also the spin densities $n_\uparrow(\mathbf{r})$ and $n_\downarrow(\mathbf{r})$, for the ground state. As we shall see, we can formally replace Ψ by the observables $n_\uparrow$ and $n_\downarrow$ as the basic variational objects.

While a *function* is a rule which assigns a number $f(x)$ to a number x, a *functional* is a rule which assigns a number $F[f]$ to a function f. For example, $h[\Psi] = \langle\Psi|\hat{H}|\Psi\rangle$ is a functional of the trial wavefunction Ψ, given the Hamiltonian $\hat{H}$. $U[n]$ of Eq. (9) is a functional of the density $n(\mathbf{r})$, as is the local density approximation for the exchange energy:

$$E_{\mathrm{x}}^{\mathrm{LDA}}[n] = A_{\mathrm{x}}\int \mathrm{d}^3r\, n(\mathbf{r})^{4/3} \ . \tag{49}$$

The *functional derivative* $\delta F/\delta n(\mathbf{r})$ tells us how the functional $F[n]$ changes under a small variation $\delta n(\mathbf{r})$:

$$\delta F = \int \mathrm{d}^3r \left(\frac{\delta F}{\delta n(\mathbf{r})}\right)\delta n(\mathbf{r}) \ . \tag{50}$$

For example,

$$\begin{aligned}\delta E_{\mathrm{x}}^{\mathrm{LDA}} &= A_{\mathrm{x}}\int \mathrm{d}^3r\left\{[n(\mathbf{r})+\delta n(\mathbf{r})]^{4/3} - n(\mathbf{r})^{4/3}\right\}\\ &= A_{\mathrm{x}}\int \mathrm{d}^3r\,\frac{4}{3}n(\mathbf{r})^{1/3}\delta n(\mathbf{r}) \ ,\end{aligned}$$

so

$$\frac{\delta E_{\mathrm{x}}^{\mathrm{LDA}}}{\delta n(\mathbf{r})} = A_{\mathrm{x}}\frac{4}{3}n(\mathbf{r})^{1/3} \ . \tag{51}$$

Similarly,

$$\frac{\delta U[n]}{\delta n(\mathbf{r})} = u([n];\mathbf{r}) \ , \tag{52}$$

where the right hand side is given by Eq. (3). Functional derivatives of various orders can be linked through the translational and rotational symmetries of empty space [27].

3.2 Density Variational Principle

We seek a density functional analog of Eq. (30). Instead of the original derivation of Hohenberg, Kohn and Sham [25,6], which was based upon "reductio ad absurdum", we follow the "constrained search" approach of Levy [28], which is in some respects simpler and more constructive.

Eq. (30) tells us that the ground state energy can be found by minimizing $\langle\Psi|\hat{H}|\Psi\rangle$ over all normalized, antisymmetric N-particle wavefunctions:

$$E = \min_{\Psi}\langle\Psi|\hat{H}|\Psi\rangle \ . \tag{53}$$

We now separate the minimization of Eq. (53) into two steps. First we consider all wavefunctions Ψ which yield a given density $n(\mathbf{r})$, and minimize over those wavefunctions:

$$\min_{\Psi\to n}\langle\Psi|\hat{H}|\Psi\rangle = \min_{\Psi\to n}\langle\Psi|\hat{T}+\hat{V}_{\mathrm{ee}}|\Psi\rangle + \int \mathrm{d}^3r v(\mathbf{r})n(\mathbf{r}) \ , \tag{54}$$

where we have exploited the fact that all wavefunctions that yield the same $n(\mathbf{r})$ also yield the same $\langle\Psi|\hat{V}_{\mathrm{ext}}|\Psi\rangle$. Then we define the universal functional

$$F[n] = \min_{\Psi\to n}\langle\Psi|\hat{T}+\hat{V}_{\mathrm{ee}}|\Psi\rangle = \langle\Psi_n^{min}|\hat{T}+\hat{V}_{\mathrm{ee}}|\Psi_n^{\mathrm{min}}\rangle \tag{55}$$

where Ψ_n^{min} is that wavefunction which delivers the minimum for a given n. Finally we minimize over all N-electron densities $n(\mathbf{r})$:

$$\begin{aligned} E &= \min_n E_v[n] \\ &= \min_n\left\{F[n] + \int \mathrm{d}^3r v(\mathbf{r})n(\mathbf{r})\right\} \ , \end{aligned} \tag{56}$$

where of course $v(\mathbf{r})$ is held fixed during the minimization. The minimizing density is then the ground-state density.

The constraint of fixed N can be handled formally through introduction of a Lagrange multiplier μ:

$$\delta\left\{F[n] + \int \mathrm{d}^3r v(\mathbf{r})n(\mathbf{r}) - \mu\int \mathrm{d}^3r n(\mathbf{r})\right\} = 0 \ , \tag{57}$$

which is equivalent to the Euler equation

$$\frac{\delta F}{\delta n(\mathbf{r})} + v(\mathbf{r}) = \mu \ . \tag{58}$$

μ is to be adjusted until Eq. (5) is satisfied. Eq. (58) shows that the external potential $v(\mathbf{r})$ is uniquely determined by the ground state density (or by any one of them, if the ground state is degenerate).

The functional $F[n]$ is defined via Eq. (55) for all densities $n(\mathbf{r})$ which are "N-representable", i.e., come from an antisymmetric N-electron wavefunction. We shall discuss the extension from wavefunctions to ensembles in section 4.5. The functional derivative $\delta F/\delta n(\mathbf{r})$ is defined via Eq. (58) for all densities which are "v-representable", i.e., come from antisymmetric N-electron ground-state wavefunctions for some choice of external potential $v(\mathbf{r})$.

This formal development requires only the total density of Eq. (4), and not the separate spin densities $n_\uparrow(\mathbf{r})$ and $n_\downarrow(\mathbf{r})$. However, it is clear how to get to a spin-density functional theory: just replace the constraint of fixed n in Eq. (54) and subsequent equations by that of fixed $n_\uparrow$ and $n_\downarrow$. There are two practical reasons to do so: (1) This extension is required when the external potential is spin-dependent, i.e., $v(\mathbf{r}) \to v_\sigma(\mathbf{r})$, as when an external magnetic field couples to the z-component of electron spin. (If this field also couples to the current density $\mathbf{j}(\mathbf{r})$, then we must resort to a current-density functional theory.) (2) Even when $v(\mathbf{r})$ is spin-independent, we may be interested in the physical spin magnetization (e.g., in magnetic materials). (3) Even when neither (1) nor (2) applies, our local and semi-local approximations (Eqs. (11) and (12)) typically work better when we use $n_\uparrow$ and $n_\downarrow$ instead of n.

3.3 Kohn-Sham Non-Interacting System

For a system of non-interacting electrons, $\hat{V}_{\mathrm{ee}}$ of Eq. (16) vanishes so $F[n]$ of Eq. (55) reduces to

$$T_{\mathrm{s}}[n] = \min_{\Psi \to n} \langle \Psi | \hat{T} | \Psi \rangle = \langle \Phi_n^{\mathrm{min}} | \hat{T} | \Phi_n^{\mathrm{min}} \rangle \ . \tag{59}$$

Although we can search over all antisymmetric N-electron wavefunctions in Eq. (59), the minimizing wavefunction Φ_n^{min} for a given density will be a non-interacting wavefunction (a single Slater determinant or a linear combination of a few) for some external potential $\hat{V}_{\mathrm{s}}$ such that

$$\frac{\delta T_{\mathrm{s}}}{\delta n(\mathbf{r})} + v_{\mathrm{s}}(\mathbf{r}) = \mu \ , \tag{60}$$

as in Eq. (58). In Eq. (60), the Kohn-Sham potential $v_{\mathrm{s}}(\mathbf{r})$ is a functional of $n(\mathbf{r})$. If there were any difference between μ and μ_{s}, the chemical potentials for interacting and non-interacting systems of the same density, it could be absorbed into $v_{\mathrm{s}}(\mathbf{r})$. We have assumed that $n(\mathbf{r})$ is both interacting and non-interacting v-representable.

Now we *define* the exchange-correlation energy $E_{\mathrm{xc}}[n]$ by

$$F[n] = T_{\mathrm{s}}[n] + U[n] + E_{\mathrm{xc}}[n] \ , \tag{61}$$

where $U[n]$ is given by Eq. (9). The Euler equations (58) and (60) are consistent with one another if and only if

$$v_{\mathrm{s}}(\mathbf{r}) = v(\mathbf{r}) + \frac{\delta U[n]}{\delta n(\mathbf{r})} + \frac{\delta E_{\mathrm{xc}}}{\delta n(\mathbf{r})} \ . \tag{62}$$

Thus we have derived the Kohn-Sham method [6] of section 1.2.

The Kohn-Sham method treats $T_s[n]$ exactly, leaving only $E_{xc}[n]$ to be approximated. This makes good sense, for several reasons: (1) T_s is typically a very large part of the energy, while E_{xc} is a smaller part. (2) T_s is largely responsible for density oscillations of the shell structure and Friedel types, which are accurately described by the Kohn-Sham method. (3) E is somewhat better suited to the local and semi-local approximations than is $T_s[n]$, for reasons to be discussed later. The price to be paid for these benefits is the appearance of orbitals. If we had a very accurate approximation for T_s directly in terms of n, we could dispense with the orbitals and solve the Euler equation (60) directly for $n(\mathbf{r})$.

The total energy of Eq. (6) may also be written as

$$E = \sum_{\alpha\sigma} \theta(\mu - \varepsilon_{\alpha\sigma})\varepsilon_{\alpha\sigma} - U[n] - \int \mathrm{d}^3 r\, n(\mathbf{r}) v_{xc}([n];\mathbf{r}) + E_{xc}[n] \ , \tag{63}$$

where the second and third terms on the right-hand-side simply remove contributions to the first term which do not belong in the total energy. The first term on the right of Eq. (63), the non-interacting energy E_{non}, is the only term that appears in the semi-empirical Hückel theory [26]. This first term includes most of the electronic shell structure effects which arise when $T_s[n]$ is treated exactly (but not when $T_s[n]$ is treated in a continuum model like the Thomas-Fermi approximation or the gradient expansion).

3.4 Exchange Energy and Correlation Energy

$E_{xc}[n]$ is the sum of distinct exchange and correlation terms:

$$E_{xc}[n] = E_x[n] + E_c[n] \ , \tag{64}$$

where [29]

$$E_x[n] = \langle \Phi_n^{\min} | \hat{V}_{ee} | \Phi_n^{\min} \rangle - U[n] \ . \tag{65}$$

When $\Phi_n^{\min}$ is a single Slater determinant, Eq. (65) is just the usual Fock integral applied to the Kohn-Sham orbitals, i.e., it differs from the Hartree-Fock exchange energy only to the extent that the Kohn-Sham orbitals differ from the Hartree-Fock orbitals for a given system or density (in the same way that $T_s[n]$ differs from the Hartree-Fock kinetic energy). We note that

$$\langle \Phi_n^{\min} | \hat{T} + \hat{V}_{ee} | \Phi_n^{\min} \rangle = T_s[n] + U[n] + E_x[n] \ , \tag{66}$$

and that, in the one-electron ($\hat{V}_{ee} = 0$) limit [9],

$$E_x[n] = -U[n] \qquad (N = 1) \ . \tag{67}$$

The correlation energy is

$$E_{\mathrm{c}}[n] = F[n] - \{T_{\mathrm{s}}[n] + U[n] + E_{\mathrm{x}}[n]\}$$
$$= \langle \Psi_n^{\mathrm{min}} | \hat{T} + \hat{V}_{\mathrm{ee}} | \Psi_n^{\mathrm{min}} \rangle - \langle \Phi_n^{\mathrm{min}} | \hat{T} + \hat{V}_{\mathrm{ee}} | \Phi_n^{\mathrm{min}} \rangle \ . \qquad (68)$$

Since Ψ_n^{min} is that wavefunction which yields density n and minimizes $\langle \hat{T} + \hat{V}_{\mathrm{ee}} \rangle$, Eq. (68) shows that

$$E_{\mathrm{c}}[n] \leq 0 \ . \qquad (69)$$

Since Φ_n^{min} is that wavefunction which yields density n and minimizes $\langle \hat{T} \rangle$, Eq. (68) shows that $E_{\mathrm{c}}[n]$ is the sum of a positive kinetic energy piece and a negative potential energy piece. These pieces of E_c contribute respectively to the first and second terms of the virial theorem, Eq. (45). Clearly for any one-electron system [9]

$$E_{\mathrm{c}}[n] = 0 \qquad (N = 1) \ . \qquad (70)$$

Eqs. (67) and (70) show that the exchange-correlation energy of a one-electron system simply cancels the spurious self-interaction $U[n]$. In the same way, the exchange-correlation potential cancels the spurious self-interaction in the Kohn-Sham potential [9]

$$\frac{\delta E_{\mathrm{x}}}{\delta n(\mathbf{r})} = -u([n];\mathbf{r}) \qquad (N = 1) \ , \qquad (71)$$

$$\frac{\delta E_{\mathrm{c}}}{\delta n(\mathbf{r})} = 0 \qquad (N = 1) \ . \qquad (72)$$

Thus

$$\lim_{r \to \infty} \frac{\delta E_{\mathrm{xc}}}{\delta n(\mathbf{r})} = -\frac{1}{r} \qquad (N = 1) \ . \qquad (73)$$

The extension of these one-electron results to spin-density functional theory is straightforward, since a one-electron system is fully spin-polarized.

3.5 Coupling-Constant Integration

The definitions (65) and (68) are formal ones, and do not provide much intuitive or physical insight into the exchange and correlation energies, or much guidance for the approximation of their density functionals. These insights are provided by the coupling-constant integration [30–33] to be derived below.

Let us define $\Psi_n^{\mathrm{min},\lambda}$ as that normalized, antisymmetric wavefunction which yields density $n(\mathbf{r})$ and minimizes the expectation value of $\hat{T} + \lambda \hat{V}_{\mathrm{ee}}$, where we have introduced a non-negative coupling constant λ. When $\lambda = 1$, $\Psi_n^{\mathrm{min},\lambda}$ is Ψ_n^{min}, the interacting ground-state wavefunction for density n. When $\lambda = 0$, $\Psi_n^{\mathrm{min},\lambda}$ is Φ_n^{min}, the non-interacting or Kohn-Sham wavefunction for density n. Varying λ at fixed $n(\mathbf{r})$ amounts to varying the external potential $v_\lambda(\mathbf{r})$: At $\lambda = 1$, $v_\lambda(\mathbf{r})$ is the true external potential, while at $\lambda = 0$ it is the Kohn-Sham effective potential $v_{\mathrm{s}}(\mathbf{r})$. We normally assume a smooth, "adiabatic connection" between the interacting and non-interacting ground states as λ is reduced from 1 to 0.

Now we write Eqs. (64), (65) and (68) as

$$\begin{aligned}E_{\mathrm{xc}}[n] &\\ &= \langle\Psi_n^{\mathrm{min},\lambda}|\hat{T}+\lambda\hat{V}_{\mathrm{ee}}|\Psi_n^{\mathrm{min},\lambda}\rangle\Big|_{\lambda=1} - \langle\Psi_n^{\mathrm{min},\lambda}|\hat{T}+\lambda\hat{V}_{\mathrm{ee}}|\Psi_n^{\mathrm{min},\lambda}\rangle\Big|_{\lambda=0} - U[n]\\ &= \int_0^1 \mathrm{d}\lambda\frac{\mathrm{d}}{\mathrm{d}\lambda}\langle\Psi_n^{\mathrm{min},\lambda}|\hat{T}+\lambda\hat{V}_{\mathrm{ee}}|\Psi_n^{\mathrm{min},\lambda}\rangle - U[n]\ . \qquad (74)\end{aligned}$$

The Hellmann-Feynman theorem of section 2.4 allows us to simplify Eq. (74) to

$$E_{\mathrm{xc}}[n] = \int_0^1 \mathrm{d}\lambda\langle\Psi_n^{\mathrm{min},\lambda}|\hat{V}_{\mathrm{ee}}|\Psi_n^{\mathrm{min},\lambda}\rangle - U[n]\ . \qquad (75)$$

Eq. (75) "looks like" a potential energy; the kinetic energy contribution to E_{xc} has been subsumed by the coupling-constant integration. We should remember, of course, that only $\lambda = 1$ is real or physical. The Kohn-Sham system at $\lambda = 0$, and all the intermediate values of λ, are convenient mathematical fictions.

To make further progress, we need to know how to evaluate the N-electron expectation value of a sum of one-body operators like $\hat{T}$, or a sum of two-body operators like $\hat{V}_{\mathrm{ee}}$. For this purpose, we introduce one-electron (ρ_1) and two-electron (ρ_2) reduced density matrices [34] :

$$\begin{aligned}&\rho_1(\mathbf{r}'\sigma,\mathbf{r}\sigma) \equiv N\sum_{\sigma_2\ldots\sigma_N}\\ &\int\mathrm{d}^3r_2\ldots\int\mathrm{d}^3r_N\Psi^*(\mathbf{r}'\sigma,\mathbf{r}_2\sigma_2,\ldots,\mathbf{r}_N\sigma_N)\Psi(\mathbf{r}\sigma,\mathbf{r}_2\sigma_2,\ldots,\mathbf{r}_N\sigma_N)\ , \qquad (76)\end{aligned}$$

$$\rho_2(\mathbf{r}',\mathbf{r}) \equiv N(N-1)\sum_{\sigma_1\ldots\sigma_N}\int\mathrm{d}^3r_3\ldots\int\mathrm{d}^3r_N|\Psi(\mathbf{r}'\sigma_1,\mathbf{r}\sigma_2,\ldots,\mathbf{r}_N\sigma_N)|^2\ . \qquad (77)$$

From Eq. (20),

$$n_\sigma(\mathbf{r}) = \rho_1(\mathbf{r}\sigma,\mathbf{r}\sigma)\ . \qquad (78)$$

Clearly also

$$\langle\hat{T}\rangle = \sum_\sigma\int\mathrm{d}^3r - \frac{1}{2}\frac{\partial}{\partial\mathbf{r}}\cdot\frac{\partial}{\partial\mathbf{r}}\rho_1(\mathbf{r}'\sigma,\mathbf{r}\sigma)\Big|_{\mathbf{r}'=\mathbf{r}}\ , \qquad (79)$$

$$\langle\hat{V}_{\mathrm{ee}}\rangle = \frac{1}{2}\int\mathrm{d}^3r\int\mathrm{d}^3r'\frac{\rho_2(\mathbf{r}',\mathbf{r})}{|\mathbf{r}-\mathbf{r}'|}\ . \qquad (80)$$

We interpret the positive number $\rho_2(\mathbf{r}',\mathbf{r})\mathrm{d}^3r'\mathrm{d}^3r$ as the joint probability of finding an electron in volume element d^3r' at $\mathbf{r}'$, *and* an electron in d^3r at $\mathbf{r}$. By standard probability theory, this is the product of the probability of finding an electron in d^3r ($n(\mathbf{r})\mathrm{d}^3r$) and the conditional probability of finding an electron in d^3r', given that there is one at $\mathbf{r}$ ($n_2(\mathbf{r},\mathbf{r}')\mathrm{d}^3r'$):

$$\rho_2(\mathbf{r}',\mathbf{r}) = n(\mathbf{r})n_2(\mathbf{r},\mathbf{r}') \ . \tag{81}$$

By arguments similar to those used in section 2.1, we interpret $n_2(\mathbf{r},\mathbf{r}')$ as the average density of electrons at $\mathbf{r}'$, given that there is an electron at $\mathbf{r}$. Clearly then

$$\int \mathrm{d}^3r' n_2(\mathbf{r},\mathbf{r}') = N - 1 \ . \tag{82}$$

For the wavefunction $\Psi_n^{\min,\lambda}$, we write

$$n_2(\mathbf{r},\mathbf{r}') = n(\mathbf{r}') + n_{\mathrm{xc}}^{\lambda}(\mathbf{r},\mathbf{r}') \ , \tag{83}$$

an equation which defines $n_{\mathrm{xc}}^{\lambda}(\mathbf{r},\mathbf{r}')$, the density at $\mathbf{r}'$ of the exchange-correlation hole [33] about an electron at $\mathbf{r}$. Eqs. (5) and (83) imply that

$$\int \mathrm{d}^3r' n_{\mathrm{xc}}^{\lambda}(\mathbf{r},\mathbf{r}') = -1 \ , \tag{84}$$

which says that, if an electron is definitely at $\mathbf{r}$, it is missing from the rest of the system.

Because the Coulomb interaction $1/u$ is singular as $u = |\mathbf{r} - \mathbf{r}'| \to 0$, the exchange-correlation hole density has a cusp [35,34] around $u = 0$:

$$\frac{\partial}{\partial u} \int \frac{\mathrm{d}\Omega_{\mathbf{u}}}{4\pi} n_{\mathrm{xc}}^{\lambda}(\mathbf{r},\mathbf{r}+\mathbf{u}) \bigg|_{u=0} = \lambda \left[n(\mathbf{r}) + n_{\mathrm{xc}}^{\lambda}(\mathbf{r},\mathbf{r}) \right] \ , \tag{85}$$

where $\int \mathrm{d}\Omega_{\mathbf{u}}/(4\pi)$ is an angular average. This cusp vanishes when $\lambda = 0$, and also in the fully-spin-polarized and low-density limits, in which all other electrons are excluded from the position of a given electron: $n_{\mathrm{xc}}^{\lambda}(\mathbf{r},\mathbf{r}) = -n(\mathbf{r})$.

We can now rewrite Eq. (75) as [33]

$$E_{\mathrm{xc}}[n] = \frac{1}{2} \int \mathrm{d}^3r \int \mathrm{d}^3r' \frac{n(\mathbf{r})\bar{n}_{\mathrm{xc}}(\mathbf{r},\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|} \ , \tag{86}$$

where

$$\bar{n}_{\mathrm{xc}}(\mathbf{r},\mathbf{r}') = \int_0^1 \mathrm{d}\lambda n_{\mathrm{xc}}^{\lambda}(\mathbf{r},\mathbf{r}') \tag{87}$$

is the coupling-constant averaged hole density. The exchange-correlation energy is just the electrostatic interaction between each electron and the coupling-constant averaged exchange-correlation hole which surrounds it. The hole is created by three effects: (1) self-interaction correction, a classical effect which guarantees that an electron cannot interact with itself, (2) the Pauli exclusion principle, which tends to keep two electrons with parallel spins apart in space, and (3) the Coulomb repulsion, which tends to keep any two electrons apart in space. Effects (1) and (2) are responsible for the exchange energy, which is present even at $\lambda = 0$, while effect (3) is responsible for the correlation energy, and arises only for $\lambda \neq 0$.

If $\Psi_n^{\min,\lambda=0}$ is a single Slater determinant, as it typically is, then the one- and two-electron density matrices at $\lambda = 0$ can be constructed explicitly from the Kohn-Sham spin orbitals $\psi_{\alpha\sigma}(\mathbf{r})$:

$$\rho_1^{\lambda=0}(\mathbf{r}'\sigma, \mathbf{r}\sigma) = \sum_\alpha \theta(\mu - \varepsilon_{\alpha\sigma})\psi^*_{\alpha\sigma}(\mathbf{r}')\psi_{\alpha\sigma}(\mathbf{r}) \ , \tag{88}$$

$$\rho_2^{\lambda=0}(\mathbf{r}', \mathbf{r}) = n(\mathbf{r})n(\mathbf{r}') + n(\mathbf{r})n_{\mathrm{x}}(\mathbf{r}, \mathbf{r}') \ , \tag{89}$$

where

$$n_{\mathrm{x}}(\mathbf{r}, \mathbf{r}') = n_{\mathrm{xc}}^{\lambda=0}(\mathbf{r}, \mathbf{r}') = -\sum_\sigma \frac{|\rho_1^{\lambda=0}(\mathbf{r}'\sigma, \mathbf{r}\sigma)|^2}{n(\mathbf{r})} \tag{90}$$

is the exact exchange-hole density. Eq. (90) shows that

$$n_{\mathrm{x}}(\mathbf{r}, \mathbf{r}') \leq 0 \ , \tag{91}$$

so the exact exchange energy

$$E_{\mathrm{x}}[n] = \frac{1}{2}\int \mathrm{d}^3r \int \mathrm{d}^3r' \frac{n(\mathbf{r})n_{\mathrm{x}}(\mathbf{r}, \mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|} \tag{92}$$

is also negative, and can be written as the sum of up-spin and down-spin contributions:

$$E_{\mathrm{x}} = E_{\mathrm{x}}^{\uparrow} + E_{\mathrm{x}}^{\downarrow} < 0 \ . \tag{93}$$

Eq. (84) provides a sum rule for the exchange hole:

$$\int \mathrm{d}^3r' n_{\mathrm{x}}(\mathbf{r}, \mathbf{r}') = -1 \ . \tag{94}$$

Eqs. (90) and (78) show that the "on-top" exchange hole density is [36]

$$n_{\mathrm{x}}(\mathbf{r}, \mathbf{r}) = -\frac{n_\uparrow^2(\mathbf{r}) + n_\downarrow^2(\mathbf{r})}{n(\mathbf{r})} \ , \tag{95}$$

which is determined by just the local spin densities at position $\mathbf{r}$ - suggesting a reason why local spin density approximations work better than local density approximations.

The correlation hole density is defined by

$$\bar{n}_{\mathrm{xc}}(\mathbf{r}, \mathbf{r}') = n_{\mathrm{x}}(\mathbf{r}, \mathbf{r}') + \bar{n}_{\mathrm{c}}(\mathbf{r}, \mathbf{r}') \ , \tag{96}$$

and satisfies the sum rule

$$\int \mathrm{d}^3r' \bar{n}_{\mathrm{c}}(\mathbf{r}, \mathbf{r}') = 0 \ , \tag{97}$$

which says that Coulomb repulsion changes the shape of the hole but not its integral. In fact, this repulsion typically makes the hole deeper but more short-ranged, with a negative on-top correlation hole density:

$$\bar{n}_c(\mathbf{r},\mathbf{r}) \leq 0 \ . \tag{98}$$

The positivity of Eq. (77) is equivalent via Eqs. (81) and (83) to the inequality

$$\bar{n}_{\mathrm{xc}}(\mathbf{r},\mathbf{r}') \geq -n(\mathbf{r}') \ , \tag{99}$$

which asserts that the hole cannot take away electrons that weren't there initially. By the sum rule (97), the correlation hole density $\bar{n}_c(\mathbf{r},\mathbf{r}')$ must have positive as well as negative contributions. Moreover, unlike the exchange hole density $n_x(\mathbf{r},\mathbf{r}')$, the exchange-correlation hole density $\bar{n}_{xc}(\mathbf{r},\mathbf{r}')$ can be positive.

To better understand E_{xc}, we can simplify Eq. (86) to the "real-space analysis" [37]

$$E_{\mathrm{xc}}[n] = \frac{N}{2}\int_0^\infty \mathrm{d}u\, 4\pi u^2 \frac{\langle \bar{n}_{\mathrm{xc}}(u)\rangle}{u} \ , \tag{100}$$

where

$$\langle \bar{n}_{\mathrm{xc}}(u)\rangle = \frac{1}{N}\int \mathrm{d}^3 r\, n(\mathbf{r}) \int \frac{\mathrm{d}\Omega_{\mathbf{u}}}{4\pi} \bar{n}_{\mathrm{xc}}(\mathbf{r},\mathbf{r}+\mathbf{u}) \tag{101}$$

is the system- and spherical-average of the coupling-constant-averaged hole density. The sum rule of Eq. (84) becomes

$$\int_0^\infty \mathrm{d}u\, 4\pi u^2 \langle \bar{n}_{\mathrm{xc}}(u)\rangle = -1 \ . \tag{102}$$

As u increases from 0, $\langle n_{\mathrm{x}}(u)\rangle$ rises analytically like $\langle n_{\mathrm{x}}(0)\rangle + \mathcal{O}(u^2)$, while $\langle \bar{n}_{\mathrm{c}}(u)\rangle$ rises like $\langle \bar{n}_{\mathrm{c}}(0)\rangle + \mathcal{O}(|u|)$ as a consequence of the cusp of Eq. (85). Because of the constraint of Eq. (102) and because of the factor $1/u$ in Eq. (100), E_{xc} typically becomes more negative as the on-top hole density $\langle \bar{n}_{\mathrm{xc}}(u)\rangle$ gets more negative.

4 Formal Properties of Functionals

4.1 Uniform Coordinate Scaling

The more we know of the exact properties of the density functionals $E_{xc}[n]$ and $T_{\mathrm{s}}[n]$, the better we shall understand and be able to approximate these functionals. We start with the behavior of the functionals under a uniform coordinate scaling of the density, Eq. (40).

The Hartree electrostatic self-repulsion of the electrons is known exactly (Eq. (9)), and has a simple coordinate scaling:

$$\begin{aligned} U[n_\gamma] &= \frac{1}{2}\int \mathrm{d}^3(\gamma r)\int \mathrm{d}^3(\gamma r')\frac{n(\gamma\mathbf{r})n(\gamma\mathbf{r}')}{|\mathbf{r}-\mathbf{r}'|} \\ &= \gamma\frac{1}{2}\int \mathrm{d}^3 r_1 \int \mathrm{d}^3 r_1' \frac{n(\mathbf{r}_1)n(\mathbf{r}_1')}{|\mathbf{r}_1-\mathbf{r}_1'|} = \gamma U[n] \ , \end{aligned} \tag{103}$$

where $\mathbf{r}_1 = \gamma\mathbf{r}$ and $\mathbf{r}'_1 = \gamma\mathbf{r}'$.

Next consider the non-interacting kinetic energy of Eq. (59). Scaling all the wavefunctions Ψ in the constrained search as in Eq. (38) will scale the density as in Eq. (40) and scale each kinetic energy expectation value as in Eq. (43). Thus the constrained search for the unscaled density maps into the constrained search for the scaled density, and [38]

$$T_s[n_\gamma] = \gamma^2 \, T_s[n] \ . \tag{104}$$

We turn now to the exchange energy of Eq. (65). By the argument of the last paragraph, $\Phi^{\min}_{n_\gamma}$ is the scaled version of $\Phi^{\min}_{n}$. Since also

$$\hat{V}_{ee}(\gamma\mathbf{r}_1, \ldots, \gamma\mathbf{r}_N) = \gamma^{-1} \, \hat{V}_{ee}(\mathbf{r}_1, \ldots, \mathbf{r}_N) \ , \tag{105}$$

and with the help of Eq. (103), we find [38]

$$E_x[n_\gamma] = \gamma E_x[n] \ . \tag{106}$$

In the high-density ($\gamma \to \infty$) limit, $T_s[n_\gamma]$ dominates $U[n_\gamma]$ and $E_x[n_\gamma]$. An example would be an ion with a fixed number of electrons N and a nuclear charge Z which tends to infinity; in this limit, the density and energy become essentially hydrogenic, and the effects of U and E_x become relatively negligible. In the low-density ($\gamma \to 0$) limit, $U[n_\gamma]$ and $E_x[n_\gamma]$ dominate $T_s[n_\gamma]$.

We can use coordinate scaling relations to fix the form of a local density approximation

$$F[n] = \int \mathrm{d}^3 r f(n(\mathbf{r})) \ . \tag{107}$$

If $F[n_\lambda] = \lambda^p F[n]$, then

$$\lambda^{-3} \int \mathrm{d}^3(\lambda\mathbf{r}) f\left(\lambda^3 n(\lambda\mathbf{r})\right) = \lambda^p \int \mathrm{d}^3 r f(n(\mathbf{r})) \ , \tag{108}$$

or $f(\lambda^3 n) = \lambda^{p+3} f(n)$, whence

$$f(n) = n^{1+p/3} \ . \tag{109}$$

For the exchange energy of Eq. (106), $p = 1$ so Eqs. (107) and (109) imply Eq. (49). For the non-interacting kinetic energy of Eq. (104), $p = 2$ so Eqs. (107) and (109) imply the Thomas-Fermi approximation

$$T_0[n] = A_s \int \mathrm{d}^3 r n^{5/3}(\mathbf{r}) \ . \tag{110}$$

$U[n]$ of Eq. (9) is too strongly nonlocal for any local approximation.

While $T_s[n]$, $U[n]$ and $E_x[n]$ have simple scalings, $E_c[n]$ of Eq. (68) does not. This is because $\Psi^{\min}_{n_\gamma}$, the wavefunction which via Eq. (55) yields the scaled density $n_\gamma(\mathbf{r})$ and minimizes the expectation value of $\hat{T} + \hat{V}_{ee}$, is *not* the scaled wavefunction $\gamma^{3N/2}\Psi_n^{\min}(\gamma\mathbf{r}_1,\ldots,\gamma\mathbf{r}_N)$. The scaled wavefunction yields $n_\gamma(\mathbf{r})$ but minimizes the expectation value of $\hat{T}+\gamma\hat{V}_{\mathrm{ee}}$, and it is this latter expectation value which scales like γ^2 under wavefunction scaling. Thus [39]

$$E_{\mathrm{c}}[n_\gamma]=\gamma^2 E_{\mathrm{c}}^{1/\gamma}[n] \ , \tag{111}$$

where $E_{\mathrm{c}}^{1/\gamma}[n]$ is the density functional for the correlation energy in a system for which the electron-electron interaction is not $\hat{V}_{\mathrm{ee}}$ but $\gamma^{-1}\hat{V}_{\mathrm{ee}}$.

To understand these results, let us assume that the Kohn-Sham non-interacting Hamiltonian has a non-degenerate ground state. In the high-density limit ($\gamma\to\infty$), $\Psi_{n_\gamma}^{\min}$ minimizes just $\langle\hat{T}\rangle$ and reduces to $\Phi_{n_\gamma}^{\min}$. Now we treat

$$\Delta\equiv\hat{V}_{\mathrm{ee}}-\sum_{i=1}^{N}\left(\frac{\delta U}{\delta n(\mathbf{r}_i)}+\frac{\delta E_x}{\delta n(\mathbf{r}_i)}\right) \tag{112}$$

as a weak perturbation [40,41] on the Kohn-Sham non-interacting Hamiltonian, and find

$$E_{\mathrm{c}}=\sum_{n\neq 0}\frac{|\langle n|\Delta|0\rangle|^2}{E_0-E_n} \ , \tag{113}$$

where the $|n\rangle$ are the eigenfunctions of the Kohn-Sham non-interacting Hamiltonian, and $|0\rangle$ is its ground state. Both the numerator and the denominator of Eq. (113) scale like γ^2, so [42]

$$\lim_{\gamma\to\infty} E_{\mathrm{c}}[n_\gamma]=\mathit{constant} \ . \tag{114}$$

In the low-density limit, $\Psi_{n_\gamma}^{\min}$ minimizes just $\langle\hat{V}_{\mathrm{ee}}\rangle$, and Eq. (68) then shows that [43]

$$E_{\mathrm{c}}[n_\gamma]\approx\gamma D[n] \qquad (\gamma\to 0) \ . \tag{115}$$

Generally, we have a scaling inequality [38]

$$E_{\mathrm{c}}[n_\gamma]>\gamma E_{\mathrm{c}}[n] \qquad (\gamma>1) \ , \tag{116}$$

$$E_{\mathrm{c}}[n_\gamma]<\gamma E_{\mathrm{c}}[n] \qquad (\gamma<1) \ . \tag{117}$$

If we choose a density n, we can plot $E_{\mathrm{c}}[n_\gamma]$ versus γ, and compare the result to the straight line $\gamma E_{\mathrm{c}}[n]$. These two curves will drop away from zero as γ increases from zero (with different initial slopes), then cross at $\gamma=1$. The convex $E_{\mathrm{c}}[n_\gamma]$ will then approach a negative constant as $\gamma\to\infty$.

4.2 Local Lower Bounds

Because of the importance of local and semilocal approximations like Eqs. (11) and (12), bounds on the exact functionals are especially useful when the bounds are themselves local functionals.

Lieb and Thirring [44] have conjectured that $T_s[n]$ is bounded from below by the Thomas-Fermi functional

$$T_s[n] \geq T_0[n] \ , \tag{118}$$

where $T_0[n]$ is given by Eqs. (110) with

$$A_s = \frac{3}{10}\,(3\pi^2)^{2/3} \ . \tag{119}$$

We have already established that

$$E_x[n] \geq E_{xc}[n] \geq E_{xc}^{\lambda=1}[n] \ , \tag{120}$$

where the final term of Eq. (120) is the integrand $E_{xc}^{\lambda}[n]$ of the coupling-constant integration of Eq. (75) ,

$$E_{xc}^{\lambda}[n] = \langle \Psi_n^{\min,\lambda} | \hat{V}_{ee} | \Psi_n^{\min,\lambda} \rangle - U[n] \ , \tag{121}$$

evaluated at the upper limit $\lambda = 1$. Lieb and Oxford [45] have proved that

$$E_{xc}^{\lambda=1}[n] \geq 2.273 E_x^{\mathrm{LDA}}[n] \ , \tag{122}$$

where $E_x^{\mathrm{LDA}}[n]$ is the local density approximation for the exchange energy, Eq. (49), with

$$A_x = -\frac{3}{4\pi}\,(3\pi^2)^{1/3} \ . \tag{123}$$

4.3 Spin Scaling Relations

Spin scaling relations can be used to convert density functionals into spin-density functionals.

For example, the non-interacting kinetic energy is the sum of the separate kinetic energies of the spin-up and spin-down electrons:

$$T_s[n_\uparrow, n_\downarrow] = T_s[n_\uparrow, 0] + T_s[0, n_\downarrow] \ . \tag{124}$$

The corresponding density functional, appropriate to a spin-unpolarized system, is [46]

$$T_s[n] = T_s[n/2, n/2] = 2T_s[n/2, 0] \ , \tag{125}$$

whence $T_s[n/2, 0] = \frac{1}{2}T_s[n]$ and Eq. (124) becomes

$$T_s[n_\uparrow, n_\downarrow] = \frac{1}{2}T_s[2n_\uparrow] + \frac{1}{2}T_s[2n_\downarrow] \ . \tag{126}$$

Similarly, Eq. (93) implies [46]

$$E_{\mathrm{x}}[n_\uparrow, n_\downarrow] = \frac{1}{2}E_{\mathrm{x}}[2n_\uparrow] + \frac{1}{2}E_{\mathrm{x}}[2n_\downarrow] \ . \tag{127}$$

For example, we can start with the local density approximations (110) and (49), then apply (126) and (127) to generate the corresponding local spin density approximations.

Because two electrons of anti-parallel spin repel one another Coulombically, making an important contribution to the correlation energy, there is no simple spin scaling relation for E_{c}.

4.4 Size Consistency

Common sense tells us that the total energy E and density $n(\mathbf{r})$ for a system, comprised of two well-separated subsystems with energies E_1 and E_2 and densities $n_1(\mathbf{r})$ and $n_2(\mathbf{r})$, must be $E = E_1 + E_2$ and $n(\mathbf{r}) = n_1(\mathbf{r}) + n_2(\mathbf{r})$. Approximations which satisfy this expectation, such as the LSD of Eq. (11) or the GGA of Eq. (12), are properly size consistent [47]. Size consistency is not only a principle of physics, it is almost a principle of epistemology: How could we analyze or understand complex systems, if they could not be separated into simpler components?

Density functionals which are *not* size consistent are to be avoided. An example is the Fermi-Amaldi [48] approximation for the exchange energy,

$$E_{\mathrm{x}}^{\mathrm{FA}}[n] = -U[n/N] \ , \tag{128}$$

where N is given by Eq. (5), which was constructed to satisfy Eq. (67).

4.5 Derivative Discontinuity

In section 3, our density functionals were defined as constrained searches over wavefunctions. Because all wavefunctions searched have the same electron number, there is no way to make a number-nonconserving density variation $\delta n(\mathbf{r})$. The functional derivatives are defined only up to an arbitrary constant, which has no effect on Eq. (50) when $\int d^3r \delta n(\mathbf{r}) = 0$.

To complete the definition of the functional derivatives and of the chemical potential μ, we extend the constrained search from wavefunctions to ensembles [49,50]. An ensemble or mixed state is a set of wavefunctions or pure states and their respective probabilities. By including wavefunctions with different electron numbers in the same ensemble, we can develop a density functional theory for non-integer particle number. Fractional particle numbers can arise in an open system that shares electrons with its environment, and in which the electron number fluctuates between integers.

The upshot is that the ground-state energy $E(N)$ varies linearly between two adjacent integers, and has a derivative discontinuity at each integer. This discontinuity arises in part from the exchange-correlation energy (and entirely so in cases for which the integer does not fall on the boundary of an electronic shell or subshell, e.g., for $N = 6$ in the carbon atom but not for $N = 10$ in the neon atom).

By Janak's theorem [51], the highest partly-occupied Kohn-Sham eigenvalue $\varepsilon_{\mathrm{HO}}$ equals $\partial E/\partial N = \mu$, and so changes discontinuously [49,50] at an integer Z:

$$\varepsilon_{\mathrm{HO}} = \begin{cases} -I_Z & (Z-1 < N < Z) \\ -A_Z & (Z < N < Z+1) \end{cases} , \tag{129}$$

where I_Z is the first ionization energy of the Z-electron system (i.e., the least energy needed to remove an electron from this system), and A_Z is the electron affinity of the Z-electron system (i.e., $A_Z = I_{Z+1}$). If Z does not fall on the boundary of an electronic shell or subshell, all of the difference between $-I_Z$ and $-A_Z$ must arise from a discontinuous jump in the exchange-correlation potential $\delta E_{\mathrm{xc}}/\delta n(\mathbf{r})$ as the electron number N crosses the integer Z.

Since the asymptotic decay of the density of a finite system with Z electrons is controlled by I_Z, we can show that the exchange-correlation potential tends to zero as $|\mathbf{r}| \to \infty$ [52]:

$$\lim_{|\mathbf{r}|\to\infty} \frac{\delta E_{\mathrm{xc}}}{\delta n(\mathbf{r})} = 0 \quad (Z-1 < N < Z) \ , \tag{130}$$

or more precisely

$$\lim_{|\mathbf{r}|\to\infty} \frac{\delta E_{\mathrm{xc}}}{\delta n(\mathbf{r})} = -\frac{1}{r} \quad (Z-1 < N < Z) \ . \tag{131}$$

As N increases through the integer Z, $\delta E_{\mathrm{xc}}/\delta n(\mathbf{r})$ jumps up by a positive additive constant. With further increases in N above Z, this "constant" vanishes, first at very large $|\mathbf{r}|$ and then at smaller and smaller $|\mathbf{r}|$, until it is all gone in the limit where N approaches the integer $Z+1$ from below.

Simple continuum approximations to $E_{\mathrm{xc}}[n_\uparrow, n_\downarrow]$, such as the LSD of Eq. (11) or the GGA of Eq. (12), miss much or all the derivative discontinuity, and can at best average over it. For example, the highest occupied orbital energy for a neutral atom becomes approximately $-\frac{1}{2}(I_Z + A_Z)$, the average of Eq. (129) from the electron-deficient and electron-rich sides of neutrality. We must never forget, when we make these approximations, that we are fitting a round peg into a square hole. The areas (integrated properties) of a circle and a square can be matched, but their perimeters (differential properties) will remain stubbornly different.

5 Uniform Electron Gas

5.1 Kinetic Energy

Simple systems play an important paradigmatic role in science. For example, the hydrogen atom is a paradigm for all of atomic physics. In the same way, the uniform electron gas [24] is a paradigm for solid-state physics, and also for density functional theory. In this system, the electron density $n(\mathbf{r})$ is uniform or constant over space, and thus the electron number is infinite. The negative charge of the electrons is neutralized by a rigid uniform positive background. We could imagine creating such a system by starting with a simple metal, regarded as a perfect crystal of valence electrons and ions, and then smearing out the ions to make the uniform background of positive charge. In fact, the simple metal sodium is physically very much like a uniform electron gas.

We begin by evaluating the non-interacting kinetic energy (this section) and exchange energy (next section) per electron for a spin-unpolarized electron gas of uniform density n. The corresponding energies for the spin-polarized case can then be found from Eqs. (126) and (127).

By symmetry, the Kohn-Sham potential $v_s(\mathbf{r})$ must be uniform or constant, and we take it to be zero. We impose boundary conditions within a cube of volume $\mathcal{V} \to \infty$, i.e., we require that the orbitals repeat from one face of the cube to its opposite face. (Presumably any choice of boundary conditions would give the same answer as $\mathcal{V} \to \infty$.) The Kohn-Sham orbitals are then plane waves $\exp(i\mathbf{k}\cdot\mathbf{r})/\sqrt{\mathcal{V}}$, with momenta or wavevectors $\mathbf{k}$ and energies $k^2/2$. The number of orbitals of both spins in a volume d^3k of wavevector space is $2[\mathcal{V}/(2\pi)^3]\mathrm{d}^3k$, by an elementary geometrical argument [53].

Let $N = n\mathcal{V}$ be the number of electrons in volume $\mathcal{V}$. These electrons occupy the N lowest Kohn-Sham spin orbitals, i.e., those with $k < k_\mathrm{F}$:

$$N = 2\sum_{\mathbf{k}} \theta(k_\mathrm{F} - k) = 2\frac{\mathcal{V}}{(2\pi)^3}\int_0^{k_\mathrm{F}} \mathrm{d}k\ 4\pi k^2 = \mathcal{V}\frac{k_\mathrm{F}^3}{3\pi^2}\ , \tag{132}$$

where k_F is called the Fermi wavevector. The Fermi wavelength $2\pi/k_\mathrm{F}$ is the shortest de Broglie wavelength for the non-interacting electrons. Clearly

$$n = \frac{k_\mathrm{F}^3}{3\pi^2} = \frac{3}{4\pi r_\mathrm{s}^3}\ , \tag{133}$$

where we have introduced the Seitz radius r_s - the radius of a sphere which on average contains one electron.

The kinetic energy of an orbital is $k^2/2$, and the average kinetic energy per electron is

$$t_\mathrm{s}(n) = \frac{2}{N}\sum_{\mathbf{k}} \theta(k_\mathrm{F} - k)\frac{k^2}{2} = \frac{2\mathcal{V}}{N(2\pi)^3}\int_0^{k_\mathrm{F}} \mathrm{d}k\ 4\pi k^2\frac{k^2}{2} = \frac{3}{5}\frac{k_\mathrm{F}^2}{2}\ , \tag{134}$$

or 3/5 of the Fermi energy. In other notation,

$$t_\mathrm{s}(n) = \frac{3}{10}(3\pi^2 n)^{2/3} = \frac{3}{10}\frac{(9\pi/4)^{2/3}}{r_\mathrm{s}^2}\ . \tag{135}$$

All of this kinetic energy follows from the Pauli exclusion principle, i.e., from the fermion character of the electron.

5.2 Exchange Energy

To evaluate the exchange energy, we need the Kohn-Sham one-matrix for electrons of spin σ, as defined in Eq. (88):

$$\begin{aligned}\rho_1^{\lambda=0}(\mathbf{r}+\mathbf{u}\sigma,\mathbf{r}\sigma) &= \sum_{\mathbf{k}} \theta(k_\mathrm{F}-k)\frac{\exp(-i\mathbf{k}\cdot(\mathbf{r}+\mathbf{u})}{\sqrt{\mathcal{V}}}\frac{\exp(i\mathbf{k}\cdot\mathbf{r})}{\sqrt{\mathcal{V}}} \\ &= \frac{1}{(2\pi)^3}\int_0^{k_\mathrm{F}} \mathrm{d}k\; 4\pi k^2 \int \frac{\mathrm{d}\Omega_\mathbf{k}}{4\pi}\exp(-i\mathbf{k}\cdot\mathbf{u}) \\ &= \frac{1}{2\pi^2}\int_0^{k_\mathrm{F}} \mathrm{d}k\; k^2\frac{\sin(ku)}{ku} \\ &= \frac{k_\mathrm{F}^3}{2\pi^2}\frac{\sin(k_\mathrm{F}u)-k_\mathrm{F}u\cos(k_\mathrm{F}u)}{(k_\mathrm{F}u)^3} \quad . \end{aligned} \tag{136}$$

The exchange hole density at distance u from an electron is, by Eq. (90),

$$n_\mathrm{x}(u) = -2\frac{|\rho_1^{\lambda=0}(\mathbf{r}+\mathbf{u}\sigma,\mathbf{r}\sigma)|^2}{n} \ , \tag{137}$$

which ranges from $-n/2$ at $u=0$ (where all other electrons of the same spin are excluded by the Pauli principle) to 0 (like $1/u^4$) as $u \to \infty$. The exchange energy per electron is

$$e_\mathrm{x}(n) = \int_0^\infty \mathrm{d}u\; 2\pi u n_\mathrm{x}(u) = -\frac{3}{4\pi}k_\mathrm{F} \ . \tag{138}$$

In other notation,

$$e_\mathrm{x}(n) = -\frac{3}{4\pi}(3\pi^2 n)^{1/3} = -\frac{3}{4\pi}\frac{(9\pi/4)^{1/3}}{r_\mathrm{s}} \ . \tag{139}$$

Since the self-interaction correction vanishes for the diffuse orbitals of the uniform gas, all of this exchange energy is due to the Pauli exclusion principle.

5.3 Correlation Energy

Exact analytic expressions for $e_\mathrm{c}(n)$, the correlation energy per electron of the uniform gas, are known only in extreme limits. The high-density ($r_\mathrm{s} \to 0$) limit is also the weak-coupling limit, in which

$$e_\mathrm{c}(n) = c_0 \ln r_s - c_1 + c_2 r_s \ln r_s - c_3 r_s + \dots \quad (r_s \to 0) \tag{140}$$

from many-body perturbation theory [54]. The positive constants $c_0 = 0.031091$ [54] and $c_1 = 0.046644$ [55] are known. Eq. (140) does not quite tend to a constant when $r_\mathrm{s} \to 0$, as Eq. (114) would suggest, because the excited states of the non-interacting system lie arbitrarily close in energy to the ground state.

The low-density ($r_s \to \infty$) limit is also the strong coupling limit in which the uniform fluid phase is unstable against the formation of a close-packed Wigner lattice of localized electrons. Because the energies of these two phases remain nearly degenerate as $r_s \to \infty$, they have the same kind of dependence upon r_s [56]:

$$e_c(n) \to -\frac{d_0}{r_s} + \frac{d_1}{r_s^{3/2}} + \dots \qquad (r_s \to \infty) \ . \tag{141}$$

The constants d_0 and d_1 in Eq. (141) can be estimated from the Madelung electrostatic and zero-point vibrational energies of the Wigner crystal, respectively. The estimate

$$d_0 \approx -\frac{9}{10} \tag{142}$$

can be found from the electrostatic energy of a neutral spherical cell: just add the electrostatic self-repulsion $3/5r_s$ of a sphere of uniform positive background (with radius r_s) to the interaction $-3/2r_s$ between this background and the electron at its center. The origin of the $r_s^{-3/2}$ term in Eq. (141) is also simple: Think of the potential energy of the electron at small distance u from the center of the sphere as $-3/2r_s + \frac{1}{2}ku^2$, where k is a spring constant. Since this potential energy must vanish for $u \approx r_s$, we find that $k \sim r_s^{-3}$ and thus the zero-point vibrational energy is $3\omega/2 = 1.5\sqrt{k/m} \sim r_s^{-3/2}$.

An expression which encompasses both limits (140)and (141) is [8]

$$e_c(n) = -2c_0(1+\alpha_1 r_s)\ln\left[1 + \frac{1}{2c_0(\beta_1 r_s^{1/2} + \beta_2 r_s + \beta_3 r_s^{3/2} + \beta_4 r_s^2)}\right] \ , \tag{143}$$

where

$$\beta_1 = \frac{1}{2c_0}\exp\left(-\frac{c_1}{2c_0}\right) \ , \tag{144}$$

$$\beta_2 = 2c_0\beta_1^2 \ . \tag{145}$$

The coefficients $\alpha_1 = 0.21370$, $\beta_3 = 1.6382$, and $\beta_4 = 0.49294$ are found by fitting to accurate Quantum Monte Carlo correlation energies [57] for r_s =2, 5, 10, 20, 50, and 100.

The uniform electron gas is in equilibrium when the density n minimizes the total energy per electron, i.e., when

$$\frac{\partial}{\partial n}\left[t_s(n) + e_x(n) + e_c(n)\right] = 0 \ . \tag{146}$$

This condition is met at $r_s \simeq 4.1$, close to the observed valence electron density of sodium. At any r_s, we have

$$\frac{\delta T_s}{\delta n(\mathbf{r})} = \frac{\partial}{\partial n}\left[n t_s(n)\right] = \frac{1}{2}k_F^2 \ , \tag{147}$$

$$\frac{\delta E_x}{\delta n(\mathbf{r})} = \frac{\partial}{\partial n}\left[n e_x(n)\right] = -\frac{1}{\pi}k_F \ . \tag{148}$$

Eq. (143) with the parameters listed above provides a representation of $e_c(n_\uparrow, n_\downarrow)$ for $n_\uparrow = n_\downarrow = n/2$; other accurate representations are also available [9,10]. Eq. (143) with different parameters ($c_0 = 0.015545$, $c_1 = 0.025599$, $\alpha_1 = 0.20548$, $\beta_3 = 3.3662$, $\beta_4 = 0.62517$) can represent $e_c(n_\uparrow, n_\downarrow)$ for $n_\uparrow = n$ and $n_\downarrow = 0$, the correlation energy per electron for a fully spin-polarized uniform gas. But we shall need $e_c(n_\uparrow, n_\downarrow)$ for arbitrary relative spin polarization

$$\zeta = \frac{(n_\uparrow - n_\downarrow)}{(n_\uparrow + n_\downarrow)} , \tag{149}$$

which ranges from 0 for an unpolarized system to ±1 for a fully-spin-polarized system. A useful interpolation formula, based upon a study of the random phase approximation, is [10]

$$\begin{aligned} e_c(n_\uparrow, n_\downarrow) &= e_c(n) + \alpha_c(n) \frac{f(\zeta)}{f''(0)}(1 - \zeta^4) + [e_c(n, 0) - e_c(n)] f(\zeta) \zeta^4 \\ &= e_c(n) + \alpha_c(n) \zeta^2 + \mathcal{O}(\zeta^4) , \end{aligned} \tag{150}$$

where

$$f(\zeta) = \frac{[(1 + \zeta)^{4/3} + (1 - \zeta)^{4/3} - 2]}{(2^{4/3} - 2)} . \tag{151}$$

In Eq. (150), $\alpha_c(n)$ is the correlation contribution to the spin stiffness. Roughly $\alpha_c(n) \approx e_c(n, 0) - e_c(n)$, but more precisely $-\alpha_c(n)$ can be parametrized in the form of Eq. (143) (with $c_0 = 0.016887$, $c_1 = 0.035475$, $\alpha_1 = 0.11125$, $\beta_3 = 0.88026$, $\beta_4 = 0.49671$).

For completeness, we note that the spin-scaling relations (126) and (127) imply that

$$e_x(n_\uparrow, n_\downarrow) = e_x(n) \frac{[(1 + \zeta)^{4/3} + (1 - \zeta)^{4/3}]}{2} , \tag{152}$$

$$t_s(n_\uparrow, n_\downarrow) = t_s(n) \frac{[(1 + \zeta)^{5/3} + (1 - \zeta)^{5/3}]}{2} . \tag{153}$$

The exchange-hole density of Eq. (137) can also be spin scaled. Expressions for the exchange and correlation holes for arbitrary r_s and ζ are given in Ref. [58].

5.4 Linear Response

We now discuss the linear response of the spin-unpolarized uniform electron gas to a weak, static, external potential $\delta v(\mathbf{r})$. This is a well-studied problem [59], and a practical one for the local-pseudopotential description of a simple metal [60].

Because the unperturbed system is homogeneous, we find that, to first order in $\delta v(\mathbf{r})$, the electron density response is

$$\delta n(\mathbf{r}) = \int d^3 r' \chi(|\mathbf{r} - \mathbf{r}'|) \delta v(\mathbf{r}') \tag{154}$$

where χ is a linear response function. If

$$\delta v(\mathbf{r}) = \delta v(\mathbf{q}) \exp(i\mathbf{q} \cdot \mathbf{r}) \tag{155}$$

is a wave of wavevector $\mathbf{q}$ and small amplitude $\delta v(\mathbf{q})$, then Eq. (154) becomes $\delta n(\mathbf{r}) = \delta n(\mathbf{q}) \exp(i\mathbf{q} \cdot \mathbf{r})$, where

$$\delta n(\mathbf{q}) = \chi(\mathbf{q}) \delta v(\mathbf{q}) \ , \tag{156}$$

and

$$\chi(\mathbf{q}) = \int \mathrm{d}^3 x \exp(-i\mathbf{q} \cdot \mathbf{x}) \chi(|\mathbf{x}|) \tag{157}$$

is the Fourier transform of $\chi(|\mathbf{r} - \mathbf{r}'|)$ with respect to $\mathbf{x} = \mathbf{r} - \mathbf{r}'$. (In Eq. (155), the real part of the complex exponential $\exp(i\alpha) = \cos(\alpha) + i \sin(\alpha)$ is understood.)

By the Kohn-Sham theorem, we also have

$$\delta n(\mathbf{q}) = \chi_{\mathrm{s}}(q) \delta v_{\mathrm{s}}(\mathbf{q}) \ , \tag{158}$$

where $\delta v_{\mathrm{s}}(\mathbf{q})$ is the change in the Kohn-Sham effective one-electron potential of Eq. (62), and

$$\chi_{\mathrm{s}}(q) = -\frac{k_{\mathrm{F}}}{\pi^2} F(q/2k_{\mathrm{F}}) \tag{159}$$

is the density response function for the non-interacting uniform electron gas. The Lindhard function

$$F(x) = \frac{1}{2} + \frac{1 - x^2}{4x} \ln \left| \frac{1 + x}{1 - x} \right| \tag{160}$$

equals $1 - x^2/3 - x^4/15$ as $x \to 0$, $1/2$ at $x = 1$, and $1/(3x^2) + 1/(15x^4)$ as $x \to \infty$. $\mathrm{d}F/\mathrm{d}x$ diverges logarithmically as $x \to 1$.

Besides $\delta v(\mathbf{r})$, the other contributions to $\delta v_{\mathrm{s}}(\mathbf{r})$ of Eq. (62) are

$$\delta \left(\frac{\delta U}{\delta n(\mathbf{r})} \right) = \int \mathrm{d}^3 r' \frac{\delta n(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|} \ , \tag{161}$$

$$\delta \left(\frac{\delta E_{\mathrm{xc}}}{\delta n(\mathbf{r})} \right) = \int \mathrm{d}^3 r' \frac{\delta^2 E_{\mathrm{xc}}}{\delta n(\mathbf{r}) \delta n(\mathbf{r}')} \delta n(\mathbf{r}') \ . \tag{162}$$

In other words,

$$\delta v_{\mathrm{s}}(\mathbf{q}) = \delta v(\mathbf{q}) + \frac{4\pi}{q^2} \delta n(\mathbf{q}) - \frac{\pi}{k_F^2} \gamma_{\mathrm{xc}}(q) \delta n(\mathbf{q}) \ , \tag{163}$$

where the coefficient of the first $\delta n(\mathbf{q})$ is the Fourier transform of the Coulomb interaction $1/|\mathbf{r} - \mathbf{r}'|$, and the coefficient of the second $\delta n(\mathbf{q})$ is the Fourier transform of $\delta^2 E_{\mathrm{xc}} / \delta n(\mathbf{r}) \delta n(\mathbf{r}')$.

We re-write Eq. (163) as

$$\delta v_{\mathrm{s}}(\mathbf{q}) = \delta v(\mathbf{q}) + \frac{4\pi}{q^2}\left[1 - G_{\mathrm{xc}}(q)\right]\delta n(\mathbf{q}) \ , \tag{164}$$

where

$$G_{\mathrm{xc}}(q) = \gamma_{\mathrm{xc}}(q)\left(\frac{q}{2k_{\mathrm{F}}}\right)^2 \tag{165}$$

is the so-called local-field factor. Then we insert Eq. (158) into Eq. (164) and find

$$\delta v_{\mathrm{s}}(\mathbf{q}) = \frac{\delta v(\mathbf{q})}{\epsilon_{\mathrm{s}}(q)} \tag{166}$$

where

$$\epsilon_{\mathrm{s}}(q) = 1 - \frac{4\pi}{q^2}\left[1 - G_{\mathrm{xc}}(q)\right]\chi_{\mathrm{s}}(q) \ . \tag{167}$$

In other words, the density response function of the interacting uniform electron gas is

$$\chi(q) = \frac{\chi_{\mathrm{s}}(q)}{\epsilon_{\mathrm{s}}(q)} \ . \tag{168}$$

These results are particularly simple in the long-wavelength ($q \to 0$) limit, in which $\gamma_{\mathrm{xc}}(q)$ tends to a constant and

$$\epsilon_{\mathrm{s}}(q) \to 1 - \frac{\gamma_{\mathrm{xc}}(q=0)}{\pi k_{\mathrm{F}}} + \frac{k_{\mathrm{s}}^2}{q^2} \quad (q \to 0) \ , \tag{169}$$

where

$$k_{\mathrm{s}} = \left(\frac{4k_{\mathrm{F}}}{\pi}\right)^{1/2} = \left(\frac{4}{\pi}\right)^{1/2}\left(\frac{9\pi}{4}\right)^{1/6}\frac{1}{r_{\mathrm{s}}^{1/2}} \tag{170}$$

is the inverse of the Thomas-Fermi screening length - the characteristic distance over which an external perturbation is screened out. Eqs. (166) and (167) show that a slowly-varying external perturbation $\delta v(\mathbf{q})$ is strongly "screened out" by the uniform electron gas, leaving only a very weak Kohn-Sham potential $\delta v_{\mathrm{s}}(\mathbf{q})$. Eq. (168) shows that the response function $\chi(q)$ is weaker than $\chi_{\mathrm{s}}(q)$ by a factor $(q/k_{\mathrm{s}})^2$ in the limit $q \to 0$.

In Eq. (166), $\epsilon_{\mathrm{s}}(q)$ is a kind of dielectric function, but it is not the standard dielectric function $\epsilon(q)$ which predicts the response of the electrostatic potential alone:

$$\delta v(\mathbf{q}) + \frac{4\pi}{q^2}\delta n(\mathbf{q}) = \frac{\delta v(\mathbf{q})}{\epsilon(q)} \ . \tag{171}$$

By inserting Eq. (156) into Eq. (171), we find

$$\frac{1}{\epsilon(q)} = 1 + \frac{4\pi}{q^2}\chi(q) \ . \tag{172}$$

It is only when we neglect exchange and correlation that we find the simple Lindhard result

$$\epsilon(q) \to \epsilon_{\mathrm{s}}(q) \to \epsilon_{\mathrm{L}}(q) = 1 - \frac{4\pi}{q^2}\chi_{\mathrm{s}}(q) \quad (\gamma_{\mathrm{xc}} \to 0) \ . \tag{173}$$

Neglecting correlation, γ_{x} is a numerically-tabulated function of $(q/2k_{\mathrm{F}})$ with the small-q expansion [61]

$$\gamma_{\mathrm{x}}(q) = 1 + \frac{5}{9}\left(\frac{q}{2k_{\mathrm{F}}}\right)^2 + \frac{73}{225}\left(\frac{q}{2k_{\mathrm{F}}}\right)^4 \quad (q \to 0) \ . \tag{174}$$

When correlation is included, $\gamma_{\mathrm{xc}}(q)$ depends upon r_{s} as well as $(q/2k_{\mathrm{F}})$, in a way that is known from Quantum Monte Carlo studies [62] of the weakly-perturbed uniform gas.

The second-order change δE in the total energy may be found from the Hellmann-Feynman theorem of Sect. 2.4. Replace $\delta v(\mathbf{r})$ by $v_\lambda(\mathbf{r}) = \lambda\delta v(\mathbf{r})$ and $\delta n(\mathbf{r})$ by $\lambda\delta n(\mathbf{r})$, to find

$$\begin{aligned}\delta E &= \int_0^1 \mathrm{d}\lambda \int \mathrm{d}^3 r n_\lambda(\mathbf{r}) \frac{\mathrm{d}}{\mathrm{d}\lambda} v_\lambda(\mathbf{r}) \\ &= \int_0^1 \mathrm{d}\lambda \int \mathrm{d}^3 r [n + \lambda\delta n(\mathbf{r})]\delta v(\mathbf{r}) \\ &= \frac{1}{2}\int \mathrm{d}^3 r \delta n(\mathbf{r})\delta v(\mathbf{r}) \\ &= \frac{1}{2}\delta n(-\mathbf{q})\delta v(\mathbf{q}) \ .\end{aligned} \tag{175}$$

5.5 Clumping and Adiabatic Connection

The uniform electron gas for $r_{\mathrm{s}} \lesssim 30$ provides a nice example of the adiabatic connection discussed in Sect. 3.5. As the coupling constant λ turns on from 0 to 1, the ground state wavefunction evolves continuously from the Kohn-Sham determinant of plane waves to the ground state of interacting electrons in the presence of the external potential, while the density remains fixed. (One should of course regard the infinite system as the infinite-volume limit of a finite chunk of uniform background neutralized by electrons.)

The adiabatic connection between non-interacting and interacting uniform-density ground states could be destroyed by any tendency of the density to clump. A fictitious attractive interaction between electrons would yield such a tendency. Even in the absence of attractive interactions, clumping appears in the very-low-density electron gas as a charge density wave or Wigner crystallization [56,59]. Then there is probably no external potential which will hold the interacting system in a uniform-density ground state, but one can still find the energy of the uniform state by imposing density uniformity as a constraint on a trial interacting wavefunction.

The uniform phase becomes unstable against a charge density wave of wavevector $\mathbf{q}$ and infinitesimal amplitude when $\epsilon_{\mathrm{s}}(q)$ of Eq. (167) vanishes [59]. This instability for $q \approx 2k_{\mathrm{F}}$ arises at low density as a consequence of exchange and correlation.

Table 3. Exchange-correlation energies of atoms, in hartree.

Atom	LSD	GGA	Exact
H	-0.29	-0.31	-0.31
He	-1.00	-1.06	-1.09
Li	-1.69	-1.81	-1.83
Be	-2.54	-2.72	-2.76
N	-6.32	-6.73	-6.78
Ne	-11.78	-12.42	-12.50

Table 4. Atomization energies of molecules, in eV. (1 hartree = 27.21 eV). From Ref. [20].

Molecule	LSD	GGA	Exact
H_2	4.9	4.6	4.7
CH_4	20.0	18.2	18.2
NH_3	14.6	13.1	12.9
H_2O	11.6	10.1	10.1
CO	13.0	11.7	11.2
O_2	7.6	6.2	5.2

6 Local and Semi-Local Approximations

6.1 Local Spin Density Approximation

The local spin density approximation (LSD) for the exchange-correlation energy, Eq. (11), was proposed in the original work of Kohn and Sham [6], and has proved to be remarkably accurate, useful, and hard to improve upon. The generalized gradient approximation (GGA) of Eq. (12), a kind of simple extension of LSD, is now more widely used in quantum chemistry, but LSD remains the most popular way to do electronic-structure calculations in solid state physics. Tables 1 and 2 provide a summary of typical errors for LSD and GGA, while Tables 3 and 4 make this comparison for a few specific atoms and molecules. The LSD is parametrized as in Sect. 5, while the GGA is the non-empirical one of Perdew, Burke, and Ernzerhof [20], to be presented later.

The LSD approximation to any energy component G is

$$G^{\mathrm{LSD}}[n_\uparrow, n_\downarrow] = \int \mathrm{d}^3 r n(\mathbf{r}) g(n_\uparrow(\mathbf{r}), n_\downarrow(\mathbf{r})) \quad , \tag{176}$$

where $g(n_\uparrow, n_\downarrow)$ is that energy component per particle in an electron gas with uniform spin densities $n_\uparrow$ and $n_\downarrow$, and $n(\mathbf{r})\mathrm{d}^3r$ is the average number of electrons in volume element d^3r. Sections 5.1-5.3 provide the ingredients for $T_s^{LSD} = T_0$, E_x^{LSD}, and E_c^{LSD}. The functional derivative of Eq. (176) is

$$\frac{\delta G^{LSD}}{\delta n_\sigma(\mathbf{r})} = \frac{\partial}{\partial n_\sigma}\left[(n_\uparrow + n_\downarrow)g(n_\uparrow, n_\downarrow)\right] . \tag{177}$$

By construction, LSD is exact for a uniform density, or more generally for a density that varies slowly over space [6]. More precisely, LSD should be valid when the length scale of the density variation is large in comparison with length scales set by the local density, such as the Fermi wavelength $2\pi/k_F$ or the screening length $1/k_s$. This condition is rarely satisfied in real electronic systems, so we must look elsewhere to understand why LSD works.

We need to understand why LSD works, for three reasons: to justify LSD calculations, to understand the physics, and to develop improved density functional approximations. Thus we will start with the good news about LSD, proceed to the mixed good/bad news, and close with the bad news.

LSD has *many* correct formal features. It is exact for uniform densities and nearly-exact for slowly-varying ones, a feature that makes LSD well suited at least to the description of the crystalline simple metals. It satisfies the inequalities $E_x < 0$ (Eq. (93)) and $E_c < 0$ (Eq. (69)), the correct uniform coordinate scaling of E_x (Eq. (106)), the correct spin scaling of E_x (Eq. (127)), the correct coordinate scaling for E_c (Eqs. (111), (116), (117)), the correct low-density behavior of E_c (Eq. (115)), and the correct Lieb-Oxford bound on E_{xc} (Eqs. (120) and (122)). LSD is properly size-consistent (Sect. 4.4).

LSD provides a surprisingly good account of the linear response of the spin-unpolarized uniform electron gas (Sect. 5.4). Since

$$\frac{\delta^2 E_{xc}^{LSD}}{\delta n(\mathbf{r})\delta n(\mathbf{r}')} = \delta(\mathbf{r} - \mathbf{r}')\frac{\partial^2[n e_{xc}(n)]}{\partial n^2} , \tag{178}$$

where $\delta(\mathbf{r} - \mathbf{r}')$ is the Dirac delta function, we find

$$\gamma_{xc}^{LSD}(q) = 1 - \frac{k_F^2}{\pi}\frac{\partial^2}{\partial n^2}[n e_c(n)] , \tag{179}$$

a constant independent of q, which must be the exact $q \to 0$ or slowly-varying limit of $\gamma_{xc}(q)$. Figure 1 of Ref. [20] shows that the "exact" $\gamma_{xc}(q)$ from a Quantum Monte Carlo calculation [62] for $r_s = 4$ is remarkably close to the LSD prediction for $q \lesssim 2k_F$. The same is true over the whole valence-electron density range $2 \lesssim r_s \lesssim 5$, and results from a strong cancellation between the nonlocalities of exchange and correlation. Indeed the exact result for exchange (neglecting correlation), Eq. (174), is strongly q-dependent or nonlocal. The displayed terms of Eq. (174) suffice for $q \lesssim 2k_F$.

Powerful reasons for the success of LSD are provided by the coupling constant integration of Sect. 3.5. Comparison of Eqs. (86) and (11) reveals that the LSD approximations for the exchange and correlation holes of an inhomogeneous system are

$$n_{\rm x}^{\rm LSD}(\mathbf{r},\mathbf{r}') = n_{\rm x}^{\rm unif}(n_\uparrow(\mathbf{r}), n_\downarrow(\mathbf{r}); |\mathbf{r}-\mathbf{r}'|) \ , \tag{180}$$

$$n_{\rm c}^{\rm LSD}(\mathbf{r},\mathbf{r}') = n_{\rm c}^{\rm unif}(n_\uparrow(\mathbf{r}), n_\downarrow(\mathbf{r}); |\mathbf{r}-\mathbf{r}'|) \ , \tag{181}$$

where $n_{\rm xc}^{\rm unif}(n_\uparrow, n_\downarrow; u)$ is the hole in an electron gas with uniform spin densities $n_\uparrow$ and $n_\downarrow$. Since the uniform gas is a possible physical system, Eqs. (180) and (181) obey the exact constraints of Eqs. (91) (negativity of $n_{\rm x}$), (94) (sum rule on $n_{\rm x}$), (95), (97) (sum rule on $\bar{n}_{\rm c}$), (98), and (85) (cusp condition).

By Eq. (95), the LSD on-top exchange hole $n_{\rm x}^{\rm LSD}(\mathbf{r},\mathbf{r})$ is exact, at least when the Kohn-Sham wavefunction is a single Slater determinant. The LSD on-top correlation hole $\bar{n}_{\rm c}^{\rm LSD}(\mathbf{r},\mathbf{r})$ is *not* exact [63] (except in the high-density, low-density, fully spin-polarized, or slowly-varying limit), but it is often quite realistic [64]. By Eq. (85), its cusp is then also realistic.

Because it satisfies all these constraints, the LSD model for the system-, spherically-, and coupling-constant-averaged hole of Eq. (101),

$$\langle \bar{n}_{\rm xc}^{\rm LSD}(u)\rangle = \frac{1}{N}\int {\rm d}^3r\, n(\mathbf{r})\bar{n}_{\rm xc}^{\rm unif}(n_\uparrow(\mathbf{r}), n_\downarrow(\mathbf{r}); u) \ , \tag{182}$$

can be very physical. Moreover, the system average in Eq. (182) "unweights" regions of space where LSD is expected to be least reliable, such as near a nucleus or in the evanescent tail of the electron density [65,64].

Since correlation makes $\langle \bar{n}_{\rm xc}(u=0)\rangle$ deeper, and thus by Eq. (102) makes $\langle \bar{n}_{\rm xc}(u)\rangle$ more short-ranged, $E_{\rm xc}$ can be "more local" than either $E_{\rm x}$ or $E_{\rm c}$. In other words, LSD often benefits from a cancellation of errors between exchange and correlation.

Mixed good and bad news about LSD is the fact that self-consistent LSD calculations can break exact spin symmetries. As an example, consider "stretched H_2", the hydrogen molecule ($N=2$) with a very large separation between the two nuclei. The exact ground state is a spin singlet ($S=0$), with $n_\uparrow(\mathbf{r}) = n_\downarrow(\mathbf{r}) = n(\mathbf{r})/2$. But the LSD ground state localizes all of the spin-up density on one of the nuclei, and all of the spin-down density on the other. Although (or rather *because*) the LSD spin densities are wrong, the LSD total energy is correctly the sum of the energies of two isolated hydrogen atoms, so this symmetry breaking is by no means entirely a bad thing [66,67]. The self-consistent LSD on-top hole density $\langle \bar{n}_{\rm xc}(0)\rangle = -\langle n\rangle$ is also right: Heitler-London correlation ensures that two electrons are never found near one another, or on the same nucleus at the same time.

Finally, we present the bad news about LSD: (1) LSD does not incorporate known inhomogeneity or gradient corrections to the exchange-correlation hole near the electron (Sect. 6.2) (2) It does not satisfy the high-density correlation scaling requirement of Eq. (114), but shows a $\ln\gamma$ divergence associated with the $\ln r_{\rm s}$ term of Eq. (140). (3) LSD is not exact in the one-electron limit, i.e., does not satisfy Eqs. (67), and (70)- (73). Although the "self-interaction error" is small for the exchange-correlation energy, it is more substantial for the exchange-correlation potential and orbital eigenvalues. (4) As a "continuum approximation", based as it is on the uniform electron gas and its continuous one-electron energy spectrum, LSD misses the derivative discontinuity of Sect. 4.5. Effectively, LSD averages over the discontinuity, so its highest occupied orbital energy for a Z-electron system is not Eq. (129) but $\varepsilon^{\mathrm{HO}} \approx -(I_Z + A_Z)/2$. A second consequence is that LSD predicts an incorrect dissociation of a hetero-nuclear molecule or solid to fractionally charged fragments. (In LSD calculations of atomization energies, the dissociation products are constrained to be neutral atoms, and not these unphysical fragments.) (5) LSD does not guarantee satisfaction of Eq. (99), an inherently nonlocal constraint.

The GGA to be derived in Sect. 6.4 will preserve all the good or mixed features of LSD listed above, while eliminating bad features (1) and (2) but not (3) - (5). Elimination of (3) - (5) will probably require the construction of $E_{\mathrm{xc}}[n_\uparrow, n_\downarrow]$ from the Kohn-Sham orbitals (which are themselves highly-nonlocal functionals of the density). For example, the self-interaction correction [9,68] to LSD eliminates most of the bad features (3) and (4), but not in an entirely satisfactory way.

6.2 Gradient Expansion

Gradient expansions [6,69], which offer systematic corrections to LSD for electron densities that vary slowly over space, might appear to be the natural next step beyond LSD. As we shall see, they are not; understanding why not will light the path to the generalized gradient approximations of Sect. 6.3.

As a first measure of inhomogeneity, we define the reduced density gradient

$$s = \frac{|\nabla n|}{2k_{\mathrm{F}}n} = \frac{|\nabla n|}{2(3\pi)^{1/3}n^{4/3}} = \frac{3}{2}\left(\frac{4}{9\pi}\right)^{1/3}|\nabla r_{\mathrm{s}}| \ , \tag{183}$$

which measures how fast and how much the density varies on the scale of the local Fermi wavelength $2\pi/k_{\mathrm{F}}$. For the energy of an atom, molecule, or solid, the range $0 \lesssim s \lesssim 1$ is very important. The range $1 \lesssim s \lesssim 3$ is somewhat important, more so in atoms than in solids, while $s > 3$ (as in the exponential tail of the density) is unimportant [70,71].

Other measures of density inhomogeneity, such as $p = \nabla^2 n/(2k_{\mathrm{F}})^2 n$, are also possible. Note that s and p are small not only for a slow density variation but also for a density variation of small amplitude (as in Sect. 5.4). The slowly-varying limit is one in which p/s is also small [6].

Under the uniform density scaling of Eq. (40), $s(\mathbf{r}) \to s_\gamma(\mathbf{r}) = s(\gamma\mathbf{r})$. The functionals $T_{\mathrm{s}}[n]$ and $E_{\mathrm{x}}[n]$ must scale as in Eqs. (104) and (106), so their gradient expansions are

$$T_{\mathrm{s}}[n] = A_{\mathrm{s}}\int \mathrm{d}^3 r\, n^{5/3}[1 + \alpha s^2 + \ldots] \ , \tag{184}$$

$$E_{\mathrm{x}}[n] = A_{\mathrm{x}}\int \mathrm{d}^3 r\, n^{4/3}[1 + \mu s^2 + \ldots] \ , \tag{185}$$

Because there is no special direction in the uniform electron gas, there can be no term linear in ∇n. Moreover, terms linear in $\nabla^2 n$ can be recast as s^2 terms, since

$$\int d^3r f(n) \nabla^2 n = -\int d^3r \left(\frac{\partial f}{\partial n}\right) |\nabla n|^2 \tag{186}$$

via integration by parts. Neglecting the dotted terms in Eqs. (184) and (185), which are fourth or higher-order in ∇, amounts to the second-order gradient expansion, which we call the gradient expansion approximation (GEA).

Correlation introduces a second length scale, the screening length $1/k_s$, and thus another reduced density gradient

$$t = \frac{|\nabla n|}{2k_s n} = \left(\frac{\pi}{4}\right)^{1/2} \left(\frac{9\pi}{4}\right)^{1/6} \frac{s}{r_s^{1/2}} \ . \tag{187}$$

In the high-density ($r_s \to 0$) limit, the screening length ($1/k_s \sim r_s^{1/2}$) is the only important length scale for the correlation hole.

The gradient expansion of the correlation energy is

$$E_c[n] = \int d^3r n \left[e_c(n) + \beta(n) t^2 + \ldots\right] \ . \tag{188}$$

While $e_c(n)$ does not quite approach a constant as $n \to \infty$, $\beta(n)$ does [69].

While the form of the gradient expansion is easy to guess, the coefficients can only be calculated by hard work. Start with the uniform electron gas, in either its non-interacting (T_s,E_x) or interacting (E_c) ground state, and apply a weak external perturbation $\delta v_s(\mathbf{q}) \exp(i\mathbf{q}\cdot\mathbf{r})$ or $\delta v(\mathbf{q}) \exp(i\mathbf{q}\cdot\mathbf{r})$, respectively. Find the linear response $\delta n(\mathbf{q})$ of the density, and the second-order response δG of the energy component G of interest. Use the linear response of the density (as in Eqs. (157) or (156)) to express δG entirely in terms of $\delta n(\mathbf{q})$. Finally, expand δG in powers of q^2, observing that $|\nabla n|^2 \sim q^2 |\delta n(\mathbf{q})|^2$, and extract the gradient coefficient.

In this way, Kirzhnits [72] found the gradient coefficient for T_s,

$$\alpha = \frac{5}{27} \tag{189}$$

(which respects the conjectured bound of Eq. (118)), Sham [73] found the coefficient of E_x,

$$\mu_{\mathrm{Sham}} = \frac{7}{81} \ , \tag{190}$$

and Ma and Brueckner [69] found the high-density limit of $\beta(n)$:

$$\beta_{\mathrm{MB}} = 0.066725 \ . \tag{191}$$

The weak density dependence of $\beta(n)$ is also known [74], as is its spin-dependence [75]. Neglecting small $\nabla\zeta$ contributions, the gradient coefficients (coefficients of $|\nabla n|^2/n^{4/3}$) for both exchange and correlation at arbitrary relative spin polarization ζ are found from those for $\zeta = 0$ through multiplication by [76]

$$\phi(\zeta) = \frac{1}{2}\left[(1+\zeta)^{2/3} + (1-\zeta)^{2/3}\right] . \tag{192}$$

For exchange, this is easily verified by applying the spin-scaling relation of Eq. (127) to Eqs. (185) and (183).

There is another interesting similarity between the gradient coefficients for exchange and correlation. Generalize the definition of t (Eq. (187)) to

$$t = \frac{|\nabla n|}{2\phi k_s n} = \left(\frac{\pi}{4}\right)^{1/2}\left(\frac{9\pi}{4}\right)^{1/6}\frac{s}{\phi r_s^{1/2}} . \tag{193}$$

Then

$$\beta_{\mathrm{MB}}\phi^3 n t^2 = \mu C_{\mathrm{x}} \phi n^{4/3} s^2 , \tag{194}$$

where

$$\mu = \beta_{\mathrm{MB}}\frac{\pi^2}{3} = 0.21951 . \tag{195}$$

Sham's derivation [73] of Eq. (190) starts with a screened Coulomb interaction $(1/u)\exp(-\kappa u)$, and takes the limit $\kappa \to 0$ at the end of the calculation. Antoniewicz and Kleinman [77] showed that the correct gradient coefficient for the unscreened Coulomb interaction is not μ_{Sham} but

$$\mu_{\mathrm{AK}} = \frac{10}{81} . \tag{196}$$

It is believed [78] that a similar order-of limits problem exists for β, in such a way that the combination of Sham's exchange coefficient with the Ma-Brueckner [69] correlation coefficient yields the correct gradient expansion of E_{xc} in the slowly-varying high-density limit.

Numerical tests of these gradient expansions for atoms show that the second-order gradient term provides a useful correction to the Thomas-Fermi or local density approximation for T_{s}, and a modestly useful correction to the local density approximation for E_{x}, but seriously worsens the local spin density results for E_{c} and E_{xc}. In fact, the GEA correlation energies are positive! The latter fact was pointed out in the original work of Ma and Brueckner [69], who suggested the first generalized gradient approximation as a remedy.

The local spin density approximation to E_{xc}, which is the leading term of the gradient expansion, provides rather realistic results for atoms, molecules, and solids. But the second-order term, which is the next systematic correction for slowly-varying densities, makes E_{xc} worse.

There are two answers to the seeming paradox of the previous paragraph. The first is that realistic electron densities are *not* very close to the slowly-varying limit ($s \ll 1$, $p/s \ll 1$, $t \ll 1$, etc.). The second is this: The LSD approximation to the exchange-correlation hole is the hole of a possible physical system, the uniform electron gas, and so satisfies many exact constraints, as discussed in Sect. 6.1. The second-order gradient expansion or GEA approximation to the hole is not, and does not.

The second-order gradient expansion or GEA models are known for both the exchange hole [12,13] $n_x(\mathbf{r}, \mathbf{r}+\mathbf{u})$ and the correlation hole $\bar{n}_c(\mathbf{r}, \mathbf{r}+\mathbf{u})$ [79]. They appear to be more realistic than the corresponding LSD models at small u, but far less realistic at large u, where several spurious features appear: $n_x(\mathbf{r}, \mathbf{r}+\mathbf{u})_{\mathrm{GEA}}$ has an undamped $\cos(2k_F u)$ oscillation which violates the negativity constraint of Eq. (91), and integrates to -1 (Eq. (94)) only with the help of a convergence factor $\exp(-\kappa u)$ $(\kappa \to 0)$. $\bar{n}_c(\mathbf{r}, \mathbf{r}+\mathbf{u})_{\mathrm{GEA}}$ has a positive u^{-4} tail, and integrates not to zero (Eq. (97)) but to a positive number $\sim s^2$. These spurious large-u behaviors are sampled by the long range of the Coulomb interaction $1/u$, leading to unsatisfactory energies for real systems.

The gradient expansion for the exchange hole density is known [80] to third order in ∇, and suggests the following interpretation of the gradient expansion: When the density does not vary too rapidly over space (e.g., in the weak-pseudopotential description of a simple metal), the addition of each successive order of the gradient expansion improves the description of the hole at small u while worsening it at large u. The bad large-u behavior thwarts our expectation that the hole will remain normalized to each order in ∇.

The non-interacting kinetic energy T_s does not sample the spurious large-u part of the gradient expansion, so its gradient expansion (Eqs. (184) and (189)) works reasonably well even for realistic electron densities. In fact, we can use Eq. (79) to show that

$$T_s[n] = \sum_\sigma \int \mathrm{d}^3r \, -\frac{1}{2}\frac{\partial}{\partial \mathbf{r}} \cdot \frac{\partial}{\partial \mathbf{r}} \rho_1^{\lambda=0}(\mathbf{r}'\sigma, \mathbf{r}\sigma)\Big|_{\mathbf{r}'=\mathbf{r}} \tag{197}$$

samples only the small-u part of the gradient expansion of the Kohn-Sham one-electron reduced density matrix, while $E_x[n]$ of Eqs. (90) and (92) also samples large values of u. The GEA for $T_s[n]$ is, in a sense, its own GGA [81]. Moreover, the sixth-order gradient expansion of T_s is also known: it diverges for finite systems, but provides accurate monovacancy formation energies for jellium [82].

The GEA form of Eqs. (184), (185), and (188) is a special case of the GGA form of Eq. (12). To find the functional derivative, note that

$$\begin{aligned}
\delta F &= \int \mathrm{d}^3r \, \delta f(n_\uparrow, n_\downarrow, \nabla n_\uparrow, \nabla n_\downarrow) \\
&= \sum_\sigma \int \mathrm{d}^3r \left[\frac{\partial f}{\partial n_\sigma(\mathbf{r})} \delta n_\sigma(\mathbf{r}) + \frac{\partial f}{\partial \nabla n_\sigma(\mathbf{r})} \cdot \nabla \delta n_\sigma(\mathbf{r}) \right] \\
&= \sum_\sigma \int \mathrm{d}^3r \frac{\delta F}{\delta n_\sigma(\mathbf{r})} \delta n_\sigma(\mathbf{r}) \ .
\end{aligned} \tag{198}$$

Integration by parts gives

$$\frac{\delta F}{\delta n_\sigma(\mathbf{r})} = \frac{\partial f}{\partial n_\sigma(\mathbf{r})} - \nabla \cdot \frac{\partial f}{\partial \nabla n_\sigma(\mathbf{r})} \ . \tag{199}$$

For example, the functional derivative of the gradient term in the spin-unpolarized high-density limit is

$$\frac{\delta}{\delta n(\mathbf{r})} \int \mathrm{d}^3 r C_{\mathrm{xc}} \frac{|\nabla n(\mathbf{r})|^2}{n^{4/3}} = C_{\mathrm{xc}} \left[\frac{4}{3} \frac{|\nabla n(\mathbf{r})|^2}{n^{7/3}} - 2 \frac{\nabla^2 n}{n^{4/3}} \right] \ , \tag{200}$$

which involves second as well as first derivatives of the density.

The GEA for the linear response function $\gamma_{\mathrm{xc}}(q)$ of Eq. (163) is found by inserting $n(\mathbf{r}) = n + \delta n(\mathbf{q}) \exp(i\mathbf{q} \cdot \mathbf{r})$ into Eq. (199) and linearizing in $\delta n(\mathbf{q})$:

$$\gamma_{\mathrm{xc}}^{\mathrm{GEA}}(q) = \gamma_{\mathrm{xc}}^{\mathrm{LSD}} - 24\pi(3\pi^2)^{1/3} C_{\mathrm{xc}} \left(\frac{q}{2k_{\mathrm{F}}} \right)^2 \ . \tag{201}$$

For example, the Antoniewicz-Kleinman gradient coefficient [77] for exchange of Eq. (196), inserted into Eqs. (200) and (201), yields the q^2 term of Eq. (174).

6.3 History of Several Generalized Gradient Approximations

In 1968, Ma and Brueckner [69] derived the second-order gradient expansion for the correlation energy in the high-density limit, Eqs. (188) and (191). In numerical tests, they found that it led to improperly positive correlation energies for atoms, because of the large size of the positive gradient term. As a remedy, they proposed the first GGA,

$$E_{\mathrm{c}}^{\mathrm{MB}}[n] = \int \mathrm{d}^3 r n e_{\mathrm{c}}(n) \left[1 - \frac{\beta_{\mathrm{MB}} t^2}{\nu n e_{\mathrm{c}}(n)} \right]^{-\nu} \ , \tag{202}$$

where $\nu \approx 0.32$ was fitted to known correlation energies. Eq. (202) reduces to Eqs. (188) and (191) in the slowly-varying ($t \to 0$) limit, but provides a strictly negative "energy density" which tends to zero as $t \to \infty$. In this respect, it is strikingly like the nonempirical GGA's that were developed in 1991 or later, differing from them mainly in the presence of an empirical parameter, the absence of a spin-density generalization, and a less satisfactory high-density limit.

Under the uniform scaling of Eq. (40), $n(\mathbf{r}) \to n_\gamma(\mathbf{r})$, we find $r_{\mathrm{s}}(\mathbf{r}) \to \gamma^{-1} r_{\mathrm{s}}(\gamma\mathbf{r})$, $\zeta(\mathbf{r}) \to \zeta(\gamma\mathbf{r})$, $s(\mathbf{r}) \to s(\gamma\mathbf{r})$, and $t(\mathbf{r}) \to \gamma^{1/2} t(\gamma\mathbf{r})$. Thus $E_{\mathrm{c}}^{\mathrm{MB}}[n_\gamma]$ tends to $E_{\mathrm{c}}^{\mathrm{LSD}}[n_\gamma]$ as $\gamma \to \infty$, and not to a negative constant as required by Eq. (114).

In 1980, Langreth and Perdew [83] explained the failure of the second-order gradient expansion (GEA) for E_{c}. They made a complete wavevector analysis of E_{xc}, i.e., they replaced the Coulomb interaction $1/u$ in Eq. (100) by its Fourier transform and found

$$E_{\mathrm{xc}}[n] = \frac{N}{2} \int_0^\infty \mathrm{d}k \frac{4\pi k^2}{(2\pi)^3} \langle \bar{n}_{\mathrm{xc}}(k) \rangle \frac{4\pi}{k^2} \ , \tag{203}$$

where

$$\langle \bar{n}_{\mathrm{xc}}(k) \rangle = \int_0^\infty du\ 4\pi u^2 \langle \bar{n}_{\mathrm{xc}}(u) \rangle \frac{\sin(ku)}{ku} \tag{204}$$

is the Fourier transform of the system- and spherically-averaged exchange-correlation hole. In Eq. (203), E_{xc} is decomposed into contributions from dynamic density fluctuations of various wavevectors k.

The sum rule of Eq. (102) should emerge from Eq. (204) in the $k \to 0$ limit (since $\sin(x)/x \to 1$ as $x \to 0$), and does so for the exchange energy at the GEA level. But the $k \to 0$ limit of $\bar{n}_{\mathrm{c}}^{\mathrm{GEA}}(k)$ turns out to be a positive number proportional to t^2, and not zero. The reason seems to be that the GEA correlation hole is only a truncated expansion, and not the exact hole for any physical system, so it can and does violate the sum rule.

Langreth and Mehl [11] (1983) proposed a GGA based upon the wavevector analysis of Eq. (203). They introduced a sharp cutoff of the spurious small-k contributions to $E_{\mathrm{c}}^{\mathrm{GEA}}$: all contributions were set to zero for $k < k_{\mathrm{c}} = f|\nabla n/n|$, where $f \approx 0.15$ is only semi-empirical since $f \approx 1/6$ was estimated theoretically. Extension of the Langreth-Mehl $E_{\mathrm{c}}^{\mathrm{GGA}}$ beyond the random phase approximation was made by Perdew [14] in 1986.

The errors of the GEA for the exchange energy are best revealed in real space (Eq. (100)), not in wavevector space (Eq. (203)). In 1985, Perdew [12] showed that the GEA for the exchange hole density $n_{\mathrm{x}}(\mathbf{r}, \mathbf{r}+\mathbf{u})$ contains a spurious undamped $\cos(2k_{\mathrm{F}}u)$ oscillation as $u \to \infty$, which violates the negativity constraint of Eq. (91) and respects the sum rule of Eq. (94) only with the help of a convergence factor (e.g., $\exp(-\kappa u)$ as $\kappa \to 0$). This suggested that the required cutoffs should be done in real space, not in wavevector space. The GEA hole density $n_{\mathrm{x}}^{\mathrm{GEA}}(\mathbf{r}, \mathbf{r}+\mathbf{u})$ was replaced by zero for all $\mathbf{u}$ where $n_{\mathrm{x}}^{\mathrm{GEA}}$ was positive, and for all $u > u_{\mathrm{x}}(\mathbf{r})$ where the cutoff radius $u_{\mathrm{x}}(\mathbf{r})$ was chosen to recover Eq. (94). Eq. (92) then provided a numerically-defined GGA for E_{x}, which turned out to be more accurate than either LSD or GEA. In 1986, Perdew and Wang [13] simplified this GGA in two ways: (1) They replaced $n_{\mathrm{x}}^{\mathrm{GEA}}(\mathbf{r}, \mathbf{r}+\mathbf{u})$, which depends upon both first and second derivatives of $n(\mathbf{r})$, by $\tilde{n}_{\mathrm{x}}^{\mathrm{GEA}}(\mathbf{r}, \mathbf{r}+\mathbf{u})$, an equivalent expression found through integration by parts, which depends only upon $\nabla n(\mathbf{r})$. (2) The resulting numerical GGA has the form

$$E_{\mathrm{x}}^{\mathrm{GGA}}[n] = A_{\mathrm{x}} \int d^3r n^{4/3} F_{\mathrm{x}}(s) \ , \tag{205}$$

which scales properly as in Eq. (106). The function $F_{\mathrm{x}}(s)$ was plotted and fitted by an analytic form. The spin-scaling relation (127) was used to generate a spin-density generalization. Perdew and Wang [13] also coined the term "generalized gradient approximation".

A parallel but more empirical line of GGA development arose in quantum chemistry around 1986. Becke [15,16] showed that a GGA for E_{x} could be constructed with the help of one or two parameters fitted to exchange energies of atoms, and demonstrated numerically that these functionals could greatly reduce the LSD overestimate of atomization energies of molecules. Lee, Yang, and Parr [17] transformed the Colle-Salvetti [84] expression for the correlation energy from a functional of the Kohn-Sham one-particle density matrix into a functional of the density. This functional contains one empirical parameter and works well in conjunction with Becke [16] exchange for many atoms and molecules, although it underestimates the correlation energy of the uniform electron gas by about a factor of two at valence-electron densities.

The real-space cutoff of the GEA hole provides a powerful nonempirical way to construct GGA's. Since exchange and correlation should be treated in a balanced way, there was a need to extend the 1986 real-space cutoff construction [13] from exchange to correlation with the help of a second cutoff radius $u_c(\mathbf{r})$ chosen to satisfy Eq. (97). Without accurate formulas for the correlation hole of the uniform electron gas, this extension had to wait until 1991, when it led to the Perdew-Wang 1991 (PW91) [18,79] GGA for E_{xc}. For most practical purposes, PW91 is equivalent to the Perdew-Burke-Ernzerhof [20,21] (PBE) "GGA made simple", which will be derived, presented, and discussed in the next two sections.

6.4 Construction of a "GGA Made Simple"

The PW91 GGA and its construction [18,79] are simple in principle, but complicated in practice by a mass of detail. In 1996, Perdew, Burke and Ernzerhof [20,21] (PBE) showed how to construct essentially the same GGA in a much simpler form and with a much simpler derivation.

Ideally, an approximate density functional $E_{xc}[n_\uparrow, n_\downarrow]$ should have all of the following features: (1) a non-empirical derivation, since the principles of quantum mechanics are well-known and sufficient; (2) universality, since in principle one functional should work for diverse systems (atoms, molecules, solids) with different bonding characters (covalent, ionic, metallic, hydrogen, and van der Waals); (3) simplicity, since this is our only hope for intuitive understanding and our best hope for practical calculation; and (4) accuracy enough to be useful in calculations for real systems.

The LSD of Eq. (11) and the non-empirical GGA of Eq. (12) nicely balance these desiderata. Both are exact only for the electron gas of uniform density, and represent controlled extrapolations away from the slowly-varying limit (unlike the GEA of Sect. 6.2, which is an uncontrolled extrapolation). LSD is a controlled extrapolation because, even when applied to a density that varies rapidly over space, it preserves many features of the exact E_{xc}, as discussed in Sect. 6.1. LSD has worked well in solid state applications for thirty years.

Our conservative philosophy of GGA construction is to try to retain all the correct features of LSD, while adding others. In particular, we retain the correct uniform-gas limit, for two reasons: (1) This is the only limit in which the restricted GGA form *can* be exact. (2) Nature's data set includes the crystalline simple metals like Na and Al. The success of the stabilized jellium model [85] reaffirms that the valence electrons in these systems are correlated very much as in a uniform gas. Among the welter of possible conditions which could be imposed to construct a GGA, the most natural and important are those respected by LSD or by the real-space cutoff construction of PW91, and these are the conditions chosen in the PBE derivation [20] below. The resulting GGA is one in which all parameters (other than those in LSD) are fundamental constants.

We start by writing the correlation energy in the form

$$E_{\mathrm{c}}^{\mathrm{GGA}}[n_\uparrow, n_\downarrow] = \int \mathrm{d}^3 r n [e_{\mathrm{c}}(r_{\mathrm{s}}, \zeta) + H(r_{\mathrm{s}}, \zeta, t)] \ , \tag{206}$$

where the local density parameters r_{s} and ζ are defined in Eqs. (133) and (149), and the reduced density gradient t in Eq. (193). The small-t behavior of nH should be given by the left-hand side of Eq. (194), which emerges naturally from the real-space cutoff construction of PW91 [79]. In the opposite or $t \to \infty$ limit, we expect that $H \to -e_{\mathrm{c}}(r_{\mathrm{s}}, \zeta)$, the correlation energy per electron of the uniform gas, as it does in the PW91 construction or in the Ma-Brueckner GGA of Eq. (202). Finally, under the uniform scaling of Eq. (40) to the high-density ($\gamma \to \infty$) limit, Eq. (206) should tend to a negative constant, as in Eq. (114) or in the numerically-constructed PW91. This means that H must cancel the logarithmic singularity of e_{c} (Eq. (140)) in this limit.

A simple function which meets these expectations is

$$H = c_0 \phi^3 \ln \left\{ 1 + \frac{\beta_{\mathrm{MB}}}{c_0} t^2 \left[\frac{1 + At^2}{1 + At^2 + A^2 t^4} \right] \right\} \ , \tag{207}$$

where ϕ is given by Eq. (192) and

$$A = \frac{\beta_{\mathrm{MB}}}{c_0} \frac{1}{\exp[-e_{\mathrm{c}}(r_{\mathrm{s}}, \zeta)/c_0 \phi^3] - 1} \ . \tag{208}$$

We now check the required limits:

$$\begin{aligned} t \to 0: \quad H &\to c_0 \phi^3 \ln \left\{ 1 + \frac{\beta_{\mathrm{MB}}}{c_0} t^2 \right\} \\ &\to \beta_{\mathrm{MB}} \phi^3 t^2 \ . \end{aligned} \tag{209}$$

$$\begin{aligned} t \to \infty: \quad H &\to c_0 \phi^3 \ln \left\{ 1 + \frac{\beta_{\mathrm{MB}}}{c_0 A} \right\} \\ &\to c_0 \phi^3 \ln \left\{ \exp \left[-\frac{e_{\mathrm{c}}(r_{\mathrm{s}}, \zeta)}{c_0 \phi^3} \right] \right\} \\ &\to -e_{\mathrm{c}}(r_{\mathrm{s}}, \zeta) \ . \end{aligned} \tag{210}$$

$$r_{\mathrm{s}} \to 0 \text{ at fixed } s: \quad H \to c_0 \phi^3 \ln t^2 \to -c_0 \phi^3 \ln r_{\mathrm{s}} \ . \tag{211}$$

To a good approximation, Eq. (140) can be generalized to

$$e_c(r_s,\zeta) = \phi^3[c_0 \ln r_s - c_1 + \ldots] \ , \tag{212}$$

which cancels the log singularity of Eq. (211).

Under uniform density scaling to the high-density limit, we find

$$\gamma \to \infty : \ E_c^{\mathrm{GGA}}[n_\gamma] \to -c_0 \int \mathrm{d}^3 r n \phi^3 \ln \left[1 + \frac{1}{\chi s^2/\phi^2 + (\chi s^2/\phi^2)^2} \right] \tag{213}$$

(where s is defined by Eq. (183)), a negative constant as required by Eq. (114), with

$$\chi = \left(\frac{3\pi^2}{16} \right)^{2/3} \frac{\beta_{\mathrm{MB}}}{c_0} \exp(-c_1/c_0) \ . \tag{214}$$

For a two-electron ion of nuclear charge Z in the limit $Z \to \infty$, Eq. (213) is -0.0479 hartree and the exact value is -0.0467. Realistic results from Eq. (213) in the $Z \to \infty$ limit have also be found [86] for ions with 3, 9, 10, and 11 electrons.

Now we turn to the construction of a GGA for the exchange energy. Because of the spin-scaling relation (127), we only need to construct $E_x^{\mathrm{GGA}}[n]$, which must be of the form of Eq. (205). To recover the good LSD description of the linear response of the uniform gas (Sect. 5.4), we choose the gradient coefficient for exchange to cancel that for correlation, i.e., we take advantage of Eq. (194) to write

$$s \to 0 : \ F_x(s) = 1 + \mu s^2 \ . \tag{215}$$

Then the gradient coefficients for exchange and correlation will cancel for *all* r_s and ζ, apart from small $\nabla\zeta$ contributions to E_x^{GGA}, as discussed in the next section.

The value of μ of Eq. (195) is 1.78 times bigger than μ_{AK} of Eq. (196), the proper gradient coefficient for exchange in the slowly-varying limit. But this choice can be justified in two other ways as well: (a) It provides a decent fit to the results of the real-space cutoff construction [79] of the PW91 exchange energy, which does *not* recover μ_{AK} in the slowly-varying limit. (b) It provides a reasonable emulation of the exact-exchange linear response function of Eq. (174) over the important range of $0 < q/2k_{\mathrm{F}} \lesssim 1$ (but not of course in the limit $q \to 0$, where μ_{AK} is needed).

Finally, we want to satisfy the Lieb-Oxford bound of Eqs. (120) and (122), which LSD respects. We can achieve this, and also recover the limit of Eq. (215), with the simple form

$$F_x(s) = 1 + \kappa - \frac{\kappa}{(1 + \mu s^2/\kappa)} \ , \tag{216}$$

where κ is a constant less than or equal to 0.804. Taking $\kappa = 0.804$ gives a GGA which is virtually identical to PW91 over the range of densities and reduced density gradients important in most real systems. We shall complete the discussion of this paragraph in the next section.

6.5 GGA Nonlocality: Its Character, Origins, and Effects

A useful way to visualize and think about gradient-corrected nonlocality, or to compare one GGA with another, is to write [19,87]

$$E_{\mathrm{xc}}^{\mathrm{GGA}}[n_\uparrow, n_\downarrow] \approx \int \mathrm{d}^3r\, n \left(-\frac{c}{r_{\mathrm{s}}}\right) F_{\mathrm{xc}}(r_{\mathrm{s}}, \zeta, s) \ , \tag{217}$$

where $c = (3/4\pi)(9\pi/4)^{1/3}$ and $-c/r_{\mathrm{s}} = e_{\mathrm{x}}(r_{\mathrm{s}}, \zeta = 0)$ is the exchange energy per electron of a spin-unpolarized uniform electron gas. The enhancement factor $F_{\mathrm{xc}}(r_{\mathrm{s}}, \zeta, s)$ shows the effects of correlation (through its r_{s} dependence), spin polarization (ζ), and inhomogeneity or nonlocality (s). F_{xc} is the analog of $3\alpha/2$ in Slater's $X\alpha$ method [88], so its variation is bounded and plotable. Figure 1 shows $F_{\mathrm{xc}}(r_{\mathrm{s}}, \zeta = 0, s)$, the enhancement factor for a spin-unpolarized system. Figure 2 shows $F_{\mathrm{xc}}(r_{\mathrm{s}}, \zeta = 1, s) - F_{\mathrm{xc}}(r_{\mathrm{s}}, \zeta = 0, s)$, the enhancement factor for the spin polarization energy. (Roughly, $F_{\mathrm{xc}}(r_{\mathrm{s}}, \zeta, s) \approx F_{\mathrm{xc}}(r_{\mathrm{s}}, \zeta = 0, s) + \zeta^2[F_{\mathrm{xc}}(r_{\mathrm{s}}, \zeta = 1, s) - F_{\mathrm{xc}}(r_{\mathrm{s}}, \zeta = 0, s)]$). The nonlocality is the s-dependence, and

$$F_{\mathrm{xc}}^{\mathrm{LSD}}(r_{\mathrm{s}}, \zeta, s) = F_{\mathrm{xc}}(r_{\mathrm{s}}, \zeta, s = 0) \tag{218}$$

is visualized as a set of horizontal straight lines coinciding with the GGA curves in the limit $s \to 0$.

Clearly, the correlation energy of Eq. (206) can be written in the form of Eq. (217). To get the exchange energy into this form, apply the spin-scaling relation (127) to Eq. (205), then drop small ∇s contributions to find

$$\begin{aligned} F_{\mathrm{x}}(\zeta, s) &= \frac{1}{2}(1+\zeta)^{4/3} F_{\mathrm{x}}\left(s/(1+\zeta)^{1/3}\right) + \frac{1}{2}(1-\zeta)^{4/3} F_{\mathrm{x}}\left(s/(1-\zeta)^{1/3}\right) \\ &= \frac{1}{2}\left[(1+\zeta)^{4/3} + (1-\zeta)^{4/3}\right] + \mu\phi s^2 + \ldots \end{aligned} \tag{219}$$

Now

$$F_{\mathrm{xc}}(r_{\mathrm{s}}, \zeta, s) = F_{\mathrm{x}}(\zeta, s) + F_{\mathrm{c}}(r_{\mathrm{s}}, \zeta, s) \ , \tag{220}$$

where

$$\lim_{r_{\mathrm{s}} \to 0} F_{\mathrm{c}}(r_{\mathrm{s}}, \zeta, s) = 0 \tag{221}$$

by Eqs. (106) and (114). Thus the $r_{\mathrm{s}} = 0$ or high-density-limit curve in each figure is the exchange-only enhancement factor. Clearly $F_{\mathrm{x}} > 0$, $F_{\mathrm{c}} > 0$, and $F_{\mathrm{x}}(\zeta = 0, s = 0) = 1$ by definition.

The Lieb-Oxford bound of Eq. (122) will be satisfied for all densities $n(\mathbf{r})$ if and only if

$$F_{\mathrm{xc}}(r_{\mathrm{s}}, \zeta, s) \geq 2.273 \ . \tag{222}$$

For the PBE GGA of Eqs. (206) and (216), this requires that

$$2^{1/3} F_{\mathrm{x}}(s/2^{1/3}) \leq 2.273 \ , \tag{223}$$

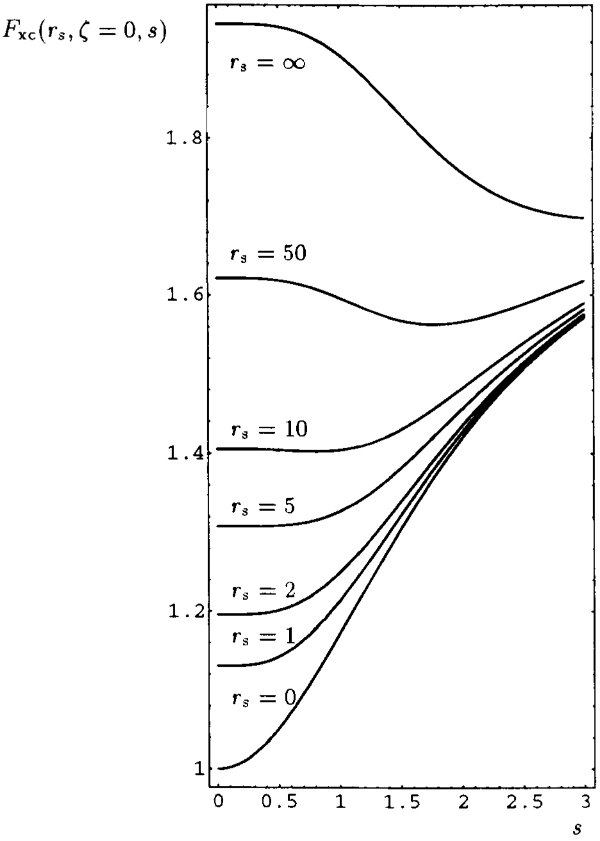


Fig. 1. The enhancment factor F_{xc} of Eq. (217) for the GGA of Perdew, Burke, and Ernzerhof [20], as a function of the reduced density gradient s of Eq. (183), for $\zeta = 0$. The local density parameter r_s and the relative spin polarization ζ are defined in Eqs. (133) and (149), respectively.

or

$$\kappa \leq 2.273/2^{1/3} - 1 = 0.804 \ , \tag{224}$$

as stated in Sect. 6.4.

There is much to be seen and explained [21] in Eq. (217) and Figs. 1 and 2. However, the main qualitative features are simply stated: When we make a density variation in which r_s decreases, ζ increases, or s increases everywhere, we find that $|E_x|$ increases and $|E_c/E_x|$ decreases.

To understand this pattern [21], we note that the second-order gradient expansion for the non-interacting kinetic energy $T_s[n_\uparrow, n_\downarrow]$, which is arguably

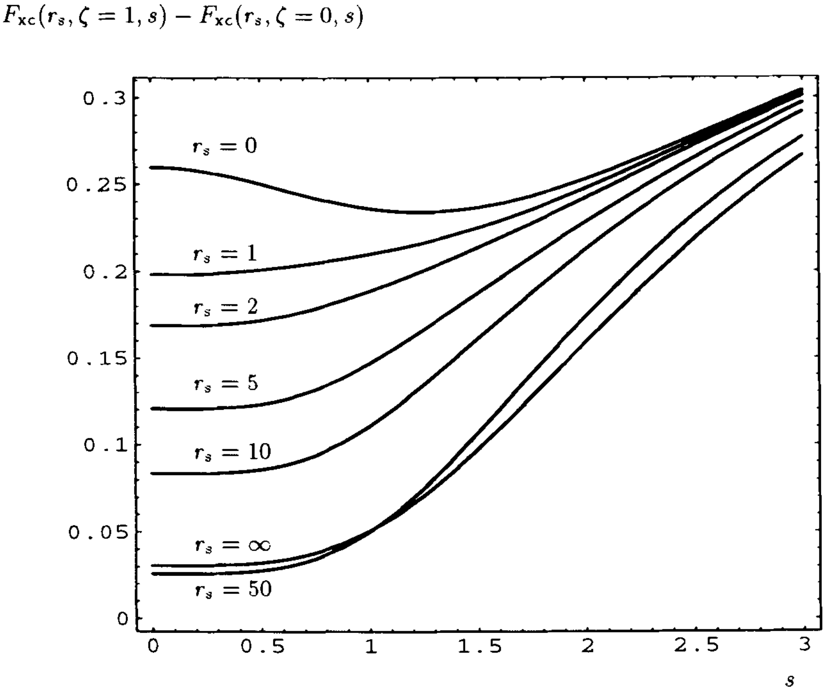


Fig. 2. Same as Fig. 1, but for the difference between the fully spin-polarized ($\zeta = 1$) and unpolarized ($\zeta = 0$) enhancement factors.

its own GGA [81], can be written as

$$T_s^{\mathrm{GGA}}[n_\uparrow, n_\downarrow] = \int d^3r n \frac{3}{10} \frac{\left(\frac{9\pi}{4}\right)^{2/3}}{r_s^2} G(\zeta, s) \ , \tag{225}$$

$$G(\zeta, s) = \frac{1}{2}\left[(1+\zeta)^{5/3} + (1-\zeta)^{5/3}\right] + \frac{5}{27}s^2 \ , \tag{226}$$

using approximate spin scaling (Eq. (126) plus neglect of $\nabla\zeta$ contributions). Eqs. (225) and (226) respect Eq. (104) and confirm our intuition based upon the Pauli exclusion and uncertainty principles: Under a density variation in which r_s decreases, ζ increases, or s increases everywhere, we find that $T_s[n_\uparrow, n_\downarrow]$ increases.

The first effect of such an increase in T_s is an increase in $|E_x|$. T_s and $|E_x|$ are "conjoint" [89], in the sense that both can be constructed from the occupied Kohn-Sham orbitals (Eqs. (7), (88), (90) and (92)). With more kinetic energy, these occupied orbitals will have shorter de Broglie wavelengths. By the uncertainty principle, they can then dig a more short-ranged and deeper exchange hole with a more negative exchange energy. Thus exchange turns on when we decrease r_s, increase ζ, or increase s.

The second effect of such an increase in T_s is to strengthen the Kohn-Sham Hamiltonian which holds non-interacting electrons at the spin densities $n_\uparrow(\mathbf{r})$ and $n_\downarrow(\mathbf{r})$. This makes the electron-electron repulsion of Eq. (112) a relatively weaker perturbation on the Kohn-Sham problem, and so reduces the ratio $|E_c/E_x|$. Thus correlation turns off relative to exchange when we decrease r_s, increase ζ, or increase s.

We note in particular that $F_x(r_s, \zeta, s)$ increases while $F_c(r_s, \zeta, s)$ decreases with increasing s. The nonlocalities of exchange and correlation are opposite, and tend to cancel for valence-electron densities ($1 \lesssim r_s \lesssim 10$) in the range $0 \lesssim s \lesssim 1$. The same remarkable cancellation occurs [62,21] in the linear response function for the uniform gas of Eq. (163), i.e., $\gamma_{xc}(q) \approx \gamma_{xc}^{LSD}(q) = \gamma_{xc}(q=0)$ for $0 \leq q/2k_F \lesssim 1$.

The core electrons in any system, and the valence electrons in solids, sample primarily the range $0 \lesssim s \lesssim 1$. The high-density core electrons see a strong, exchange-like nonlocality of E_{xc} which provides an important correction to the LSD total energy. But the valence electrons in solids see an almost-complete cancellation between the nonlocalities of exchange and correlation. This helps to explain why LSD has been so successful in solid state physics, and why the small residue of GGA nonlocality in solids does not provide a universally-better description than LSD.

The valence electrons in atoms and molecules see $0 \lesssim s \lesssim \infty$, when s diverges in the exponential tail of the density, but the energetically-important range is $0 \lesssim s \lesssim 3$ [70,71]. Figs. 1 and 2 show that GGA nonlocality *is* important in this range, so GGA is almost-always better than LSD for atoms and molecules.

For $r_s \lesssim 10$, the residual GGA nonlocality is exchange-like, i.e., exchange and correlation together turn on stronger with increasing inhomogeneity. It can then be seen from Eq. (217) that gradient corrections will favor greater density inhomogeneity and higher density [70]. Defining average density parameters $\langle r_s \rangle$, $\langle \zeta \rangle$, and $\langle s \rangle$ as in Ref. [70], we find that gradient corrections favor changes $\mathrm{d}\langle s \rangle > 0$ and $\mathrm{d}\langle r_s \rangle < 0$. Gradient corrections tend to drive a process forward when [70]

$$\frac{\mathrm{d}\langle s \rangle}{\langle s \rangle} \geq \frac{\mathrm{d}\langle r_s \rangle}{\langle r_s \rangle} . \tag{227}$$

In a typical process (bond stretching, transition to a more open structure, fragmentation, or atomization), one has $\mathrm{d}\langle s \rangle > 0$ and $\mathrm{d}\langle r_s \rangle > 0$. Thus, by Eq. (227), these effects compete - another reason why LSD has met with some success. In most such cases, the left-hand side of Eq. (227) is bigger than the right, so typically gradient corrections favor larger bond lengths or lattice constants (and thus softer vibration frequencies), more open structures, fragmentation of a highly-bonded transition state, or atomization of a molecule. In the case of bond stretching in H_2, however, the right hand side of Eq. (227) exceeds the left, so gradient corrections actually and correctly shrink the equilibrium bond length relative to LSD.

There have been many interesting tests and applications of GGA to a wide range of atoms, molecules, and solids. Some references will be found in Refs. [19,90,79,21].

We close by discussing those situations in which LSD or GGA can fail badly. They seem to be of two types: (1) When the Kohn-Sham non-interacting wavefunction is not a single Slater determinant, or when the non-interacting energies are nearly degenerate, the LSD and GGA exchange-correlation holes can be unrealistic even very close to or on top of the electron [36,91,66]. (2) In an extended system, the exact hole may display a diffuse long-range tail which is not properly captured by either LSD or GGA. To a limited extent, this effect could be mimicked by reducing the parameter κ in Eq. (216). An example of a diffuse hole arises in the calculation of the surface energy of a metal [19,32]: When an electron wanders out into the vacuum region, the exchange-correlation hole around it can extend significantly backward into the interior of the metal. A more extreme example is "stretched H_2^+", the ground state of one electron in the presence of two protons at very large separation: Half of the exact hole is localized on each proton, a situation which has no analog in the electron gas of uniform or slowly-varying density, and for which LSD and GGA make large self-interaction errors [9,92,68].

"Stretched H_2^+" and related systems are of course unusual. In most systems, the exact exchange-correlation hole is reasonably localized around its electron, as it is in LSD or GGA - and that fact is one of the reasons [93] why LSD and GGA work as well as they do.

Acknowledgements: Work supported in part by the U.S. National Science Foundation under Grant No. DMR95-21353, and in part by the Deutsche Forschungsgemeinschaft. We thank Matthias Ernzerhof for help with the figures.

References

1. D. Park, *Introduction to the Quantum Theory* (McGraw-Hill, New York, 1974).
2. H.A. Bethe, *Intermediate Quantum Mechanics* (Benjamin, New York, 1964).
3. R.O. Jones and O. Gunnarsson, Rev. Mod. Phys. **61**, 689 (1989).
4. R.M. Dreizler and E.K.U. Gross, *Density Functional Theory* (Springer, Berlin, 1990).
5. R.G. Parr and W. Yang, *Density-Functional Theory of Atoms and Molecules* (Oxford University Press, New York, 1989).
6. W. Kohn and L.J. Sham, Phys. Rev. **140**, A1133 (1965).
7. U. von Barth and L. Hedin, J. Phys. C **5**, 1629 (1972).
8. J.P. Perdew and Y. Wang, Phys. Rev. B **45**, 13244 (1992).

9. J.P. Perdew and A. Zunger, Phys. Rev. B **23**, 5048 (1981).
10. S.H. Vosko, L. Wilk, and M. Nusair, Can. J. Phys. **58**, 1200 (1980).
11. D.C. Langreth and M.J. Mehl, Phys. Rev. B **28**, 1809 (1983).
12. J.P. Perdew, Phys. Rev. Lett. **55**, 1665 (1985).
13. J.P. Perdew and Y. Wang, Phys. Rev. B **33**, 8800 (1986).
14. J.P. Perdew, Phys. Rev. B **33**, 8822 (1986).
15. A.D. Becke, J. Chem. Phys. **84**, 4524 (1986).
16. A.D. Becke, Phys. Rev. A **38**, 3098 (1988).
17. C. Lee, W. Yang, and R.G. Parr, Phys. Rev. B **37**, 785 (1988).
18. J.P. Perdew, in *Electronic Structure of Solids '91*, edited by P. Ziesche and H. Eschrig (Akademie Verlag, Berlin, 1991), p. 11.
19. J.P. Perdew, J.A. Chevary, S.H. Vosko, K.A. Jackson, M.R. Pederson, D.J. Singh, and C. Fiolhais , Phys. Rev. B **46**, 6671 (1992), ibid., **48**, 4978 (1993)(E).
20. J.P. Perdew, K. Burke, and M. Ernzerhof, Phys. Rev. Lett. **77**, 3865 (1996), ibid, **78**, 1396 (1997)(E).
21. J.P. Perdew, M. Ernzerhof, A. Zupan, and K. Burke, submitted to J. Chem. Phys.
22. J. Hafner, *From Hamiltonians to Phase Diagrams* (Dover, New York, 1988).
23. A. Szabo and N.S. Ostlund, *Modern Quantum Chemistry* (Macmillan Publishing, New York, 1982).
24. P. Fulde, *Electron Correlations in Molecules and Solids* (Springer, Berlin, 1993).
25. P. Hohenberg and W. Kohn, Phys. Rev. **136**, B864 (1964).
26. I.N. Levine, *Quantum Chemistry* (Allyn and Bacon, Boston, 1974).
27. D.P. Joubert, Int. J. Quantum Chem. **61**, 355 (1997).
28. M. Levy, Proc. Natl. Acad. Sci. USA **76**, 6062 (1979).
29. M. Levy, in *Recent Developments and Applications of Modern Density Functional Theory*, edited by J.M. Seminario (Elsevier, Amsterdam, 1996), p. 3.
30. J. Harris and R.O. Jones, J. Phys. F **4**, 1170 (1974).
31. D.C. Langreth and J.P. Perdew, Solid State Commun. **17**, 1425 (1975).
32. D.C. Langreth and J.P. Perdew, Phys. Rev. B **15**, 2884 (1977).
33. O. Gunnarsson and B.I. Lundqvist, Phys. Rev. B **13**, 4274 (1976).
34. E.R. Davidson, *Reduced Density Matrices in Quantum Chemistry* (Academic Press, New York, 1976).
35. J.C. Kimball, Phys. Rev. A **7**, 1648 (1973).
36. T. Ziegler, A. Rauk, and E.J. Baerends, Theoret. Chim. Acta **43**, 261 (1977).
37. K. Burke and J.P. Perdew, Int. J. Quantum Chem. **56**, 199 (1995).
38. M. Levy and J.P. Perdew, Phys. Rev. A **32**, 2010 (1985).
39. M. Levy, in *Single-Particle Density in Physics and Chemistry*, edited by N.H.March and B.M. Deb (Academic, London, 1987), p. 45.
40. A. Görling and M. Levy, Phys. Rev. B **47**, 13105 (1993).
41. A. Görling and M. Levy, Phys. Rev. A **50**, 196 (1994).
42. M. Levy, Phys. Rev. A **43**, 4637 (1991).
43. M. Levy and J.P. Perdew, Phys. Rev. B **48**, 11638 (1993).
44. E.H. Lieb, Int. J. Quantum Chem. **24**, 243 (1983).
45. E.H. Lieb and S. Oxford, Int. J. Quantum Chem. **19**, 427 (1981).
46. G.L. Oliver and J.P. Perdew, Phys. Rev. A **20**, 397 (1979).

47. J.P. Perdew, in *Density Functional Theory of Many-Fermion Systems*, Vol. 21 of *Advances in Quantum Chemistry*, edited by S.B. Trickey (Academic, New York, 1990), p. 113.
48. E. Fermi and E. Amaldi, Accad. Ital. Rome **6**, 119 (1934).
49. J.P. Perdew, R.G. Parr, M. Levy, and J.L. Balduz, Phys. Rev. Lett. **49**, 1691 (1982).
50. J.P. Perdew, in *Density Functional Methods in Physics*, Vol. B123 of *NATO ASI Series*, edited by R.M. Dreizler and J. da Providencia (Plenum Press, New York, 1985), p. 265.
51. J.F. Janak, Phys. Rev. B **18**, 7165 (1978).
52. M. Levy, J.P. Perdew, and V. Sahni, Phys. Rev. A **30**, 2745 (1984).
53. N.W. Ashcroft and N.D. Mermin, *Solid State Physics* (Holt, Rinehart, and Winston, New York, 1976).
54. M. Gell-Mann and K.A. Brueckner, Phys. Rev. **106**, 364 (1957).
55. L. Onsager, L. Mittag, and M.J. Stephen, Ann. Phys. (Leipzig) **18**, 71 (1966).
56. R.A. Coldwell-Horsfall and A.A. Maradudin, J. Math. Phys. **1**, 395 (1960).
57. D.M. Ceperley and B.J. Alder, Phys. Rev. Lett. **45**, 566 (1980).
58. J.P. Perdew and Y. Wang, Phys. Rev. B **46**, 12947 (1992).
59. J.P. Perdew and T. Datta, Phys. Stat. Sol. (b) **102**, 283 (1980).
60. C. Fiolhais, J.P. Perdew, S.Q. Armster, J.M. MacLaren, and M. Brajczewska, Phys. Rev. B **51**, 14001 (1995), ibid., **53**, 13193 (1996)(E).
61. E. Engel and S.H. Vosko, Phys. Rev. B **42**, 4940 (1990).
62. S. Moroni, D.M. Ceperley, and G. Senatore, Phys. Rev. Lett. **75**, 689 (1995).
63. K. Burke, J.P. Perdew, and D.C. Langreth, Phys. Rev. Lett. **73**, 1283 (1994).
64. K. Burke, J.P. Perdew, and M. Ernzerhof, in preparation.
65. M. Ernzerhof, in *Macmillan Encyclopedia of Physics* (Macmillan Publishing Company, New York, 1996), Vol. 2, p. 733.
66. J.P. Perdew, A. Savin, and K. Burke, Phys. Rev. A **51**, 4531 (1995).
67. J.P. Perdew, M. Ernzerhof, K. Burke, and A. Savin, Int. J. Quantum Chem. Symp. **61**, 197 (1997).
68. J.P. Perdew and M. Ernzerhof, in *Electronic Density Functional Theory: Recent Progress and New Directions*, edited by J.F. Dobson, G. Vignale, and M. P. Das (Plenum Press, New York, 1997).
69. S.-K. Ma and K.A. Brueckner, Phys. Rev. **165**, 18 (1968).
70. A. Zupan, K. Burke, M. Ernzerhof, and J.P. Perdew, J. Chem. Phys. **106**, 10184 (1997).
71. A. Zupan, J.P. Perdew, K. Burke, and M. Causá, Int. J. Quantum Chem. **61**, 287 (1997).
72. D.A. Kirzhnits, Sov. Phys. JETP **5**, 64 (1957).
73. L.J. Sham, in *Computational Methods in Band Theory*, edited by P.M. Marcus, J.F. Janak, and A.R. Williams (Plenum Press, New York, 1971), p. 458.
74. M. Rasolt and D.J.W. Geldart, Phys. Rev. B **34**, 1325 (1986).
75. M. Rasolt and H.L. Davis, Phys. Lett. A **86**, 45 (1981).
76. Y. Wang and J.P. Perdew, Phys. Rev. B **43**, 8911 (1991).
77. P.A. Antoniewicz and L. Kleinman, Phys. Rev. B **31**, 6779 (1985).
78. D.C. Langreth and S.H. Vosko, Phys. Rev. Lett. **59**, 497 (1987), ibid. **60**, 1984 (1988).
79. J.P. Perdew, K. Burke, and Y. Wang, Phys. Rev. B **54**, 16533 (1996).

80. Y. Wang, J.P. Perdew, J.A. Chevary, L.D. MacDonald, and S.H. Vosko, Phys. Rev. A **41**, 78 (1990).
81. J.P. Perdew, Phys. Lett. A **165**, 79 (1992).
82. Z. Yan, J.P. Perdew, T. Korhonen, and P. Ziesche, Phys. Rev. A **55**, 4601 (1997).
83. D.C. Langreth and J.P. Perdew, Phys. Rev. B **21**, 5469 (1980).
84. R. Colle and O. Salvetti, Theoret. Chim. Acta **37**, 329 (1975).
85. J.P. Perdew, H.Q. Tran, and E.D. Smith, Phys. Rev. B **42**, 11627 (1990).
86. S. Ivanov and M. Levy, submitted to J. Chem. Phys.
87. J.P. Perdew and K. Burke, Int. J. Quantum Chem. Symp. **27**, 309 (1996).
88. J.C. Slater, *The Self-Consistent Field for Molecules and Solids* (McGraw-Hill, New York, 1974).
89. H. Lee, C. Lee, and R.G. Parr, Phys. Rev. A **44**, 768 (1991).
90. K. Burke, J.P. Perdew, and M. Levy, in *Modern Density Functional Theory: A Tool for Chemistry*, Vol. 2 of *Theoretical and Computational Chemistry*, edited by J.M. Seminario and P. Politzer (Elsevier, Amsterdam, 1995), p. 29.
91. A.D. Becke, A. Savin, and H. Stoll, Theoret. Chim. Acta **91**, 147 (1995).
92. R. Merkle, A. Savin, and H. Preuss, J. Chem. Phys. **97**, 9216 (1992).
93. V. Tschinke and T. Ziegler, J. Chem. Phys. **93**, 8051 (1990).

Hybrid Methods: Combining Density Functional and Wavefunction Theory

Matthias Ernzerhof

Department of Physics, Tulane University, New Orleans, LA 70118, USA

Contents

Abstract. Hybrid methods of density-functional and wavefunction theory have proven to be powerful tools in electronic structure theory. The Kohn-Sham scheme is already a hybrid scheme, since part of the kinetic energy of the system is calculated from the Kohn-Sham wavefunction. In the empirical hybrid methods a fraction of exact exchange, also calculated from the Kohn-Sham wavefunction, is mixed with density functional approximations to exchange and correlation. These methods give a significant improvement in accuracy over local- and gradient-corrected approximations to the exchange-correlation energy. We show how insight into the nature of exchange and correlation in molecules and atoms can be used to construct nonempirical hybrid schemes. We review density-functional perturbation theory and incorporate second-order density-functional perturbation theory into the nonempirical hybrid scheme. The relation between the nonempirical- and the empirical hybrid methods is analyzed.

1 Introduction

Almost all electronic structure calculations employ the Born-Oppenheimer approximation [1,2]. In this approximation the atomic nuclei appear as space fixed point charges in the Schrödinger equation for the electrons. In atomic units ($\hbar = m_e = e^2 = 1$) the electronic Hamiltonian $\hat{H}_{el}$ is given by

$$\hat{H}_{el} = \sum_{i}^{N_{el}} -\frac{1}{2}\Delta_i - \sum_{i}^{N_{el}} \sum_{j}^{N_{nuc}} \frac{Z_j}{|\mathbf{r}_i - \mathbf{R}_j|} + \frac{1}{2}\sum_{i \neq j}^{N_{el}} \frac{1}{|\mathbf{r}_i - \mathbf{r}_j|} + \frac{1}{2}\sum_{i \neq j}^{N_{nuc}} \frac{Z_i Z_j}{|\mathbf{R}_i - \mathbf{R}_j|} . \tag{1}$$

N_{el} is the number of electrons and N_{nuc} the number of nuclei. The equilibrium geometry is the set of positions of the nuclei $\{\mathbf{R}_i\}_{i=1,\ldots,N_{nuc}}$ which minimizes the electronic energy $E(\{\mathbf{R}_i\}_{i=1,\ldots,N_{nuc}})$. An (electronic) bond dissociation energy is obtained as a change in the electronic energy upon stretching a certain bond length from its equilibrium value to infinity. Dissociation energies of molecules in their ground states will be the main focus in this article. Dissociation energies provide a demanding test for the performance of any method to calculate the electronic energy. Until recently, most standard approaches to calculate dissociation energies (and many other properties) would have started with a Hartree-Fock (HF) [3] calculation. Density functional theory (DFT) was not very popular for the calculation of dissociation energies. A simple approximate density functional scheme, the Kohn-Sham [4] approach with the local spin-density approximation (LSD) to the exchange-correlation energy E_{XC}, leads to a significant overestimation of dissociation energies. In Table 1 (details of the calculations reported here are compiled in the Appendix), a few examples of atomization energies in HF and LSD are listed. Atomization means that the molecule is separated into the atoms, i.e., all bonds are broken. Although the LSD approximation gives better results than the HF method, this advantage is overshadowed by the fact that LSD is quite far away from the goal of chemical accuracy ($\sim$ 1 kcal/mol) and that no significantly improvement over LSD was provided by DFT. HF results can be systematically improved, although at the cost of expensive computations. Popular post-Hartree-Fock methods, such as configuration interaction- and coupled-cluster techniques [3], are able to achieve chemical accuracy, however the computational effort prohibits routine application to large systems.

The development of generalized gradient approximations (GGA) [5,6,54, 8,9] has shifted the attention of chemists to DFT methods. The GGA's combine useful accuracy with computational efficiency. In Table 1 we see that GGA's [1] significantly improves upon LSD. GGA's have proven to be very useful tools in quantum chemistry [13–15], but they still leave room for improvement, especially for the congested systems in the lower part of Table

---

[1] The GGA calculations reported here are performed with the Perdew-Wang 91 (PW91) [10,11] GGA or the Perdew-Burke-Ernzerhof (PBE) [5,12] GGA, which are both derived non-empirically in different ways. Both are exact in the uniform-density limit. The numerical differences between these two functionals are insignificant for the calculations reported here.

Table 1. Electronic atomization energies ΔE ($= E^{\text{atoms}} - E^{\text{molecule}}$) of molecules. UHF is unrestricted Hartree-Fock [3], hyb denotes the exact-exchange mixing scheme defined in Eq. 2, and exp are the experimental results. The GGA employed here is PW91. The mean absolute errors (mae) are also listed. (Energies are in kcal/mol, 1eV=23.06 kcal/mol).

Molecule	ΔE^{UHF}	ΔE^{LSD}	ΔE^{GGA}	ΔE^{hyb}	ΔE^{exp}
CH_4	328	462	422	419	419
NH_3	201	337	301	294	297
H_2O	155	267	233	226	232
HF	97	162	143	137	141
CO	174	299	269	255	259
N_2	115	267	242	224	229
NO	53	199	171	152	153
F_2	-20	78	54	35	39
mae	83	38	8	3	–

1. We use "congested systems" to refer to systems where occupied binding- and/or nonbinding orbitals are in the same region of space and overlap with each other. The electron-localization function defined in [16] typically indicates delocalized electrons in regions of space, where occupied binding- and/or nonbinding overlap.

The next major step towards affordable chemical accuracy was made by Becke [17–19]. Becke suggested an empirical combination of exact exchange E_X and density functional approximations to exchange and correlation E_{XC}. Exact exchange means that the exchange energy is evaluated from the Kohn-Sham single-particle orbitals, as in the HF approximation. A formally simple empirical hybrid scheme is given by [19]

$$E_{XC}^{BeckeII} = a(E_X - E_X^{GGA}) + E_{XC}^{GGA} \quad , \tag{2}$$

where a is an empirical parameter. Table 1 shows that exact exchange mixing with $a = 0.25$ leads to a significant improvement over GGA and brings us close to the goal of chemical accuracy.

In this lecture we develop a theory for the systematic construction of hybrid schemes. The popular empirical hybrid schemes are obtained as approximations to this theory. In particular we obtain an estimate of the parameter a ($a = 0.25$) in Eq. 2, which is close to the empirical values obtained for a [19].

After the initial publications by Becke, the empirical hybrid schemes have undergone modifications [20] and the potential of this method has been documented in a number of studies, among them [17–19,21]. The empirical hybrid schemes have been applied to a vast number of problems and it is impossible to give an overview of the literature on the subject. Here we focus mainly on dissociation energies and explain how the hybrid methods work.

2 Preliminaries

In this section we study a formal decomposition of the ground-state energy. This decomposition is the starting point for modern approximations in DFT.

We define a λ-dependent Hamiltonian $\hat{H}_\lambda$ by

$$\hat{H}_\lambda = \sum_i^{N_{el}} -\frac{1}{2}\Delta_i + \sum_i^{N_{el}} v_\lambda(\mathbf{r}_i) + \frac{1}{2}\sum_{i \neq j}^{N_{el}} \frac{\lambda}{|\mathbf{r}_i - \mathbf{r}'_j|} \; . \tag{3}$$

$v_{\lambda=1}$ is the physical external potential of the system, i.e., it is a potential caused by atomic nuclei, and $\rho_{\lambda=1}(\mathbf{r})$ is the corresponding electron density. Note that we have dropped the term $\frac{1}{2}\sum_{i \neq j}^{N_{nuc}} \frac{Z_i Z_j}{|\mathbf{R}_i - \mathbf{R}_j|}$, contributing to the electronic Hamiltonian Eq. 1. This term gives a geometry-dependent shift in the energy but leaves the wavefunction and the electron density unchanged. The potential $v_\lambda(\mathbf{r})$ for $\lambda \neq 1$ is defined by the condition that $\rho_\lambda = \rho_{\lambda=1}$, i.e., the potential is adjusted to keep the electron density fixed. The ground-state energy E^λ of $\hat{H}_\lambda$ is given by

$$E^\lambda = T^\lambda + \int d^3r \; v_\lambda(\mathbf{r})\rho(\mathbf{r}) + V_{ee}^\lambda \; , \tag{4}$$

where

$$T^\lambda = \langle \psi_\lambda | \hat{T} | \psi_\lambda \rangle \tag{5}$$

and

$$V_{ee}^\lambda = \langle \psi_\lambda | \lambda \hat{V}_{ee} | \psi_\lambda \rangle \; . \tag{6}$$

ψ_λ is the ground-state wavefunction of $\hat{H}_\lambda$ and $\hat{V}_{ee} = 1/2 \sum_{i \neq j}^{N_{el}} 1/|\mathbf{r}_i - \mathbf{r}_j|$. The total energy will be rewritten as

$$E^\lambda = \langle \psi_\lambda | \hat{H}_\lambda | \psi_\lambda \rangle \tag{7}$$

$$= \langle \psi_{\lambda=0} | \hat{H}_{\lambda=0} | \psi_{\lambda=0} \rangle + \int_0^\lambda d\tilde{\lambda} \; \frac{d\langle \psi_{\tilde{\lambda}} | \hat{H}_{\tilde{\lambda}} | \psi_{\tilde{\lambda}} \rangle}{d\tilde{\lambda}} \; . \tag{8}$$

The variational principle ensures that $\langle d\psi_\lambda/d\lambda | \hat{H}_\lambda | \psi_\lambda \rangle = \langle \psi_\lambda | \hat{H}_\lambda | d\psi_\lambda/d\lambda \rangle = 0$ [22,23], and therefore

$$\frac{d\langle \psi_\lambda | \hat{H}_\lambda | \psi_\lambda \rangle}{d\lambda} = \frac{1}{2}\int d^3r d^3r' \; \frac{P_\lambda(\mathbf{r}, \mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|} + \int d^3r \; \rho(\mathbf{r}) \frac{dv_\lambda(\mathbf{r})}{d\lambda} \; . \tag{9}$$

$P_\lambda(\mathbf{r}, \mathbf{r}')$ is the pair density calculated from ψ_λ,

$$P_\lambda(\mathbf{r},\mathbf{r}') = N(N-1)\int d^3r_3\ldots d^3r_N d\sigma_1\ldots d\sigma_N \\ \times\ \psi_\lambda^*(\mathbf{r},\sigma_1,\mathbf{r}',\sigma_2,\mathbf{r}_3,\sigma_3,\ldots\mathbf{r}_N,\sigma_N)\psi_\lambda(\mathbf{r},\sigma_1,\mathbf{r}',\sigma_2,\mathbf{r}_3,\sigma_3,\ldots\mathbf{r}_N,\sigma_N) . \tag{10}$$

The pair density $P_\lambda(\mathbf{r},\mathbf{r}')$ gives the probability density to find an electron at point $\mathbf{r}$ and a second electron at point $\mathbf{r}'$.

The first term in Eq. 8 is given by

$$\langle\psi_{\lambda=0}|H_{\lambda=0}|\psi_{\lambda=0}\rangle = T_s + \int d^3r\ \rho(\mathbf{r})v_{\lambda=0}(\mathbf{r}) \ . \tag{11}$$

$T_s = \langle\psi_{\lambda=0}|\hat{T}|\psi_{\lambda=0}\rangle$ is the kinetic energy of a non-interacting ($\lambda = 0$) ground-state wavefunction, usually a Slater determinant, which yields the density $\rho(\mathbf{r})$ of the interacting ($\lambda = 1$) system. Combining Eqs. 8, 9, and 11, we find

$$E^\lambda = T_s + \int d^3r\ \rho(\mathbf{r})v_\lambda(\mathbf{r}) + \frac{1}{2}\int_0^\lambda d\tilde{\lambda}\int d^3rd^3r'\ \frac{P_{\tilde{\lambda}}(\mathbf{r},\mathbf{r}')}{|\mathbf{r}-\mathbf{r}'|} \ . \tag{12}$$

The λ-dependent pair density is split up into a λ-independent term and the λ-dependent exchange-correlation hole $\rho_{\mathrm{XC},\lambda}(\mathbf{r},\mathbf{r}')$

$$P_\lambda(\mathbf{r},\mathbf{r}') = \rho(\mathbf{r})[\rho(\mathbf{r}') + \rho_{\mathrm{XC},\lambda}(\mathbf{r},\mathbf{r}')] \ . \tag{13}$$

The factor $\rho(\mathbf{r})$ in front of the bracket gives the probability density to find an electron at point $\mathbf{r}$, and the first term inside the bracket is an approximation to the probability to find a second electron at point $\mathbf{r}'$ given one at $\mathbf{r}$. The second term in the bracket, the exchange-correlation hole at coupling strength λ, accounts for the fact that the electron at point $\mathbf{r}$ cannot be found anywhere else in the system, so that the probability to find a second electron at $\mathbf{r}'$ is reduced. Furthermore, Fermi- and Coulomb correlation alter the probability to find a second electron at point $\mathbf{r}'$. The electron-electron interaction term in Eq. 12 becomes

$$\frac{1}{2}\int d^3rd^3r'\frac{P_\lambda(\mathbf{r},\mathbf{r}')}{|\mathbf{r}-\mathbf{r}'|} = U[\rho] + \frac{1}{2}\int d^3rd^3r'\ \frac{\rho(\mathbf{r})\rho_{\mathrm{XC},\lambda}(\mathbf{r},\mathbf{r}')}{|\mathbf{r}-\mathbf{r}'|} \ , \tag{14}$$

where

$$U[\rho] = \frac{1}{2}\int d^3rd^3r'\ \frac{\rho(\mathbf{r})\rho(\mathbf{r}')}{|\mathbf{r}-\mathbf{r}'|} \ . \tag{15}$$

Equation 12 becomes

$$E^\lambda = T_s + \int d^3r\ \rho(\mathbf{r})v_\lambda(\mathbf{r}) + \lambda U[\rho] + \int_0^\lambda d\tilde{\lambda}\ V_{\mathrm{XC},\tilde{\lambda}} \ , \tag{16}$$

with

$$V_{\mathrm{XC},\lambda} = \frac{1}{2}\int d^3rd^3r'\ \frac{\rho(\mathbf{r})\rho_{\mathrm{XC},\lambda}(\mathbf{r},\mathbf{r}')}{|\mathbf{r}-\mathbf{r}'|} \ . \tag{17}$$

The exchange-correlation energy $E^{\lambda}_{\mathrm{XC}}$ of DFT at coupling-strength λ is now given by the λ-integral on the right-hand side of Eq. 16

$$E^{\lambda}_{\mathrm{XC}} = \int_0^{\lambda} d\tilde{\lambda}\; V_{\mathrm{XC},\tilde{\lambda}} \quad . \tag{18}$$

Equation 18 is the famous adiabatic connection formula [24–26], which we will use extensively in this article. The complicated many-body contributions to the ground-state energy are now all contained in the exchange-correlation energy $E^{\lambda}_{\mathrm{XC}}$. From Eqs. 4, 16, and 18 we find

$$E^{\lambda}_{\mathrm{XC}} = T^{\lambda} - T_{\mathrm{S}} + V^{\lambda}_{ee} - \lambda U \quad . \tag{19}$$

Note that the correlation contribution to the kinetic energy $T_{\mathrm{C},\lambda}$, $T_{\mathrm{C},\lambda} = T^{\lambda} - T_{\mathrm{S}}$, has also been absorbed in $E^{\lambda}_{\mathrm{XC}}$. The adiabatic connection formula Eq. 18 accounts for $T_{\mathrm{C},\lambda}$ by averaging the potential contribution to the exchange-correlation energy $V_{\mathrm{XC},\lambda}$ over the coupling constant λ.

$E^{\lambda}_{\mathrm{XC}}$ is further decomposed into an exchange E_{X} and a correlation contribution $E_{\mathrm{C},\lambda}$,

$$E^{\lambda}_{\mathrm{XC}} = \lambda E_{\mathrm{X}} + E^{\lambda}_{\mathrm{C}} \quad , \tag{20}$$

where

$$\begin{aligned} E_{\mathrm{X}} &= V_{\mathrm{XC},\lambda=0} \\ &= \langle \psi_{\lambda=0} | \hat{V}_{ee} | \psi_{\lambda=0} \rangle - U[\rho] \quad , \end{aligned} \tag{21}$$

and

$$\begin{aligned} E^{\lambda}_{\mathrm{C}} &= E^{\lambda}_{\mathrm{XC}} - \lambda E_{\mathrm{X}} \\ &= \int_0^{\lambda} d\tilde{\lambda}\; V_{\mathrm{C},\tilde{\lambda}} \quad , \end{aligned} \tag{22}$$

with $V_{\mathrm{C},\lambda} = V_{\mathrm{XC},\lambda} - E_{\mathrm{X}}$. Now we split up E^{λ}_{C} into a potential-energy and a kinetic-energy contribution

$$E^{\lambda}_{\mathrm{C}} = T_{\mathrm{C},\lambda} + \lambda V_{\mathrm{C},\lambda} \quad . \tag{23}$$

Equations 18 and 22 show that

$$T_{\mathrm{C},\lambda} = \int_0^{\lambda} d\tilde{\lambda}\; V_{\mathrm{XC},\tilde{\lambda}} - \lambda V_{\mathrm{XC},\lambda} = \int_0^{\lambda} d\tilde{\lambda}\; V_{\mathrm{C},\tilde{\lambda}} - \lambda V_{\mathrm{C},\lambda} \quad . \tag{24}$$

Utilizing $E^{\lambda}_{\mathrm{C}} = T^{\lambda} - T_{\mathrm{S}} + V^{\lambda}_{ee} - \lambda U - \lambda E_{\mathrm{X}} = T^{\lambda} - T_{\mathrm{S}} + \lambda V_{\mathrm{C},\lambda}$ and Eq. 23, we confirm that

$$T_{\mathrm{C},\lambda} = T^{\lambda} - T_{\mathrm{S}} \quad . \tag{25}$$

To introduce the electron density as variable, we note that T_{S} and $E^{\lambda}_{\mathrm{XC}}$ are functionals of the electron density [4,27,28]. Obviously $\int d^3r\; \rho(\mathbf{r}) v_{\lambda}(\mathbf{r})$ and $\frac{1}{2} \int d^3r d^3r' \; \frac{\rho(\mathbf{r})\rho(\mathbf{r}')}{|\mathbf{r}-\mathbf{r}'|}$ are also functionals of the electron density, and thus the total ground-state energy E^λ is a functional of the ground-state density. Furthermore, the functional

$$E^\lambda[\tilde{\rho}] = T_{\mathrm{S}}[\tilde{\rho}] + \int d^3r \ \tilde{\rho}(\mathbf{r}) v_\lambda(\mathbf{r}) + \lambda U[\tilde{\rho}] + E^\lambda_{\mathrm{XC}}[\tilde{\rho}] \tag{26}$$

minimizes [4,27,28] if the density $\tilde{\rho}$ is equal to the density ρ generated by $v_\lambda(\mathbf{r})$. The value of $E^\lambda[\rho]$ is the ground-state energy. The stationarity condition for Eq. 26 is

$$\begin{aligned} 0 &= \delta E^\lambda[\tilde{\rho}] \\ &= \int d^3r \ \frac{\delta E[\tilde{\rho}]}{\delta\tilde{\rho}(\mathbf{r})} \delta\tilde{\rho}(\mathbf{r}) \ , \end{aligned} \tag{27}$$

where the variations $\delta\tilde{\rho}$ are constrained such that $\int d^3r \ \delta\tilde{\rho}(\mathbf{r}) = 0$. Equation 27 can only be satisfied if $\frac{\delta E[\tilde{\rho}]}{\tilde{\rho}(\mathbf{r})} = const$. Therefore

$$\begin{aligned} const &= \frac{\delta E^\lambda[\tilde{\rho}]}{\delta\tilde{\rho}(\mathbf{r})} \\ &= \frac{\delta T_{\mathrm{S}}[\tilde{\rho}]}{\delta\tilde{\rho}(\mathbf{r})} + v_\lambda(\mathbf{r}) + \lambda u([\tilde{\rho}]; \mathbf{r}) + \frac{\delta E^\lambda_{\mathrm{XC}}[\tilde{\rho}]}{\delta\tilde{\rho}(\mathbf{r})} \ , \end{aligned} \tag{28}$$

with

$$u([\rho]; \mathbf{r}) = \int d^3r' \frac{\rho(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|} \ . \tag{29}$$

Equation 28 can be interpreted as the Euler equation of non-interacting electrons [4] whose energy E_{S} is

$$E_{\mathrm{S}}[\rho] = T_{\mathrm{S}}[\rho] + \int d^3r \ v_{\mathrm{KS}}([\rho]; \mathbf{r}) \rho(\mathbf{r}) \ , \tag{30}$$

with

$$v_{\mathrm{KS}}([\rho]; \mathbf{r}) = v_\lambda(\mathbf{r}) + \lambda u([\rho]; \mathbf{r}) + \frac{\delta E^\lambda_{\mathrm{XC}}[\rho]}{\delta\rho(\mathbf{r})} \ . \tag{31}$$

Note that the density is independent of λ, and therefore the Kohn-Sham potential v_{KS} [4] is also independent of λ. For $\lambda = 1$, Eq. 31 becomes

$$v_{\mathrm{KS}}([\rho]; \mathbf{r}) = -\sum_j^{N_{nuc}} \frac{Z_j}{|\mathbf{r} - \mathbf{R}_j|} + u([\rho]; \mathbf{r}) + \frac{\delta E^{\lambda=1}_{\mathrm{XC}}[\rho]}{\delta\rho(\mathbf{r})} \ . \tag{32}$$

Assuming that the ground-state wavefunction of the non-interacting problem Eq. 30 is a single Slater determinant, the Schrödinger equation for the single-particle orbitals $\phi_i(\mathbf{r})$ is the Kohn-Sham [4] equation

$$\left[-\frac{1}{2}\nabla^2 + v_{\mathrm{KS}}([\rho]; \mathbf{r})\right] \phi_i(\mathbf{r}) = \epsilon_i \ \phi_i(\mathbf{r}) \ , \tag{33}$$

$$\rho(\mathbf{r}) = \sum_i^{occ} |\phi_i(\mathbf{r})|^2 \ . \tag{34}$$

The effective potential v_{KS} depends on the electron density and the Kohn-Sham equation has to be solved iteratively, starting with an initial guess for the electron density ρ. T_{S} is then obtained from the occupied orbitals

$$T_s = \sum_i^{occ} -\frac{1}{2} \int d^3r \ \phi_i^*(\mathbf{r}) \Delta \phi_i(\mathbf{r}) \ , \tag{35}$$

and the electron density $\rho(\mathbf{r}) = \sum_i^{occ} |\phi_i(\mathbf{r})|^2$ is used to calculate $E_{\mathrm{XC}}^{\lambda}[\rho]$ and $U[\rho]$. The calculation of the ground-state energy and the ground-state density has now formally been reduced to a problem of solving the independent-particle equation Eq. 33. However, the functional $E_{\mathrm{XC}}^{\lambda}[\rho]$, which appears in Eq. 26 and whose density derivative contributes to the Kohn-Sham potential, is not known in an exact and practically useful form. In calculations on physical systems we only need $E_{\mathrm{XC}}^{\lambda=1}$, and approximations to $E_{\mathrm{XC}}^{\lambda=1}$, have been developed to transform the above-described scheme into a useful method for electronic structure theory [29–31].

3 Density-functional perturbation theory

In the previous section, we introduced a particular decomposition of the ground-state energy $E = T_{\mathrm{S}} + \int d^3r \ v_\lambda \rho + \lambda U + E_{\mathrm{XC}}^{\lambda}$ and we derived expressions for T_{S} and E_{X} in terms of $\psi_{\lambda=0}$ (Eqs. 21 and 35). In this section we give an explicit expression for E_{C}^{λ} in terms of the eigenvalues and eigenfunctions of the Kohn-Sham Hamiltonian $\hat{T} + \hat{v}_{\lambda=0}$.

Görling and Levy [32] derived a series expansion of E_{C}^{λ} in powers of λ, with the Kohn-Sham problem as zeroth-order reference system. The Hamiltonian $\hat{H}_\lambda$ is written as

$$\hat{H}_\lambda = \hat{H}_0 + \hat{H}(\lambda) \ , \tag{36}$$

where

$$\hat{H}_0 = \hat{T} + \hat{v}_{\lambda=0} \tag{37}$$

and

$$\hat{H}(\lambda) = \lambda \hat{V}_{ee} + \hat{v}_\lambda - \hat{v}_{\lambda=0} \ . \tag{38}$$

Note that $v_{\lambda=0}(\mathbf{r}) = v_{\mathrm{KS}}(\mathbf{r})$. Standard Rayleigh-Schrödinger perturbation theory [3] is used to expand E^λ in powers of λ.

$$\begin{aligned} E^\lambda = E^{\lambda=0} &+ \left(\int d^3r \ \frac{dv_\lambda(\mathbf{r})}{d\lambda} \rho(\mathbf{r}) + U + E_{\mathrm{X}} \right)_{\lambda=0} \lambda \\ &+ \frac{1}{2} \left(\int d^3r \ \frac{d^2 v_\lambda(\mathbf{r})}{d\lambda^2} \rho(\mathbf{r}) + \frac{d^2 E_{\mathrm{C}}^{\lambda}}{d\lambda^2} \right)_{\lambda=0} \lambda^2 \\ &+ \dots \ . \end{aligned} \tag{39}$$

$E_{\mathrm{C}}^{\lambda} = \int_0^{\lambda} d\tilde{\lambda} V_{c,\tilde{\lambda}}$ has no first-order term in λ. The second-order contribution to E_{C}^{λ} is given by [32,33]

$$\begin{aligned} E_{\mathrm{C},2} &= \frac{1}{2}\frac{d^2 E_{\mathrm{C}}^{\lambda}}{d\lambda^2}\Big|_{\lambda=0} \\ &= -\frac{1}{4}\sum_{ij\alpha\beta} \frac{|\langle \phi_i^{KS}\phi_j^{KS}|\hat{V}_{ee}|\phi_\alpha^{KS}\phi_\beta^{KS}\rangle|^2}{\epsilon_\alpha + \epsilon_\beta - \epsilon_i - \epsilon_j} \\ &\quad - \sum_{i\alpha} \frac{|\langle \phi_i^{KS}|\frac{d\hat{v}_\lambda}{d\lambda}|_{\lambda=0} + \hat{u} + \hat{f}|\phi_\alpha^{KS}\rangle|^2}{\epsilon_\alpha - \epsilon_i} , \end{aligned} \tag{40}$$

where the indices i and j run over the occupied, and α and β over the unoccupied, Kohn-Sham orbitals. ϵ_i denotes a Kohn-Sham eigenvalue, and $\hat{f}$ is the non-local exchange potential of the unrestricted Fock operator formed from the Kohn-Sham orbitals.

To obtain expressions for E_{C}^{λ} we still need to find the potential v_λ, which enters into the calculation of the perturbed wavefunction ψ_λ and is therefore needed to obtain E_{C}^{λ}. In order to calculate the correlation energy in nth order, we need the wavefunction ψ_λ in order $n-1$. The wavefunction in order $n-1$ can be constructed from knowledge of v_λ in order $n-1$. Using the fact that $\delta E^{\lambda}[\tilde{\rho}]/\delta\tilde{\rho}|_{\tilde{\rho}=\rho} = const$ we find

$$\begin{aligned} v_\lambda(\mathbf{r}) &= -\frac{\delta}{\delta\tilde{\rho}(\mathbf{r})}\left\{T_{\mathrm{S}}[\tilde{\rho}] + \lambda U[\tilde{\rho}] + \lambda E_{\mathrm{X}}[\tilde{\rho}] + E_{\mathrm{C}}^{\lambda}[\tilde{\rho}]\right\}_{\tilde{\rho}=\rho} + const \\ &= v_{\mathrm{KS}}([\rho];\mathbf{r}) - \lambda u([\rho];\mathbf{r}) - \lambda v_{\mathrm{X}}([\rho];\mathbf{r}) - \frac{\delta E_{\mathrm{C}}^{\lambda}[\tilde{\rho}]}{\delta\tilde{\rho}(\mathbf{r})}\Big|_{\tilde{\rho}=\rho} , \end{aligned} \tag{41}$$

and therefore

$$\begin{aligned} \frac{d^{n-1}v_\lambda(\mathbf{r})}{d\lambda^{n-1}} &= -\frac{\delta}{\delta\tilde{\rho}(\mathbf{r})}\frac{d^{n-1}}{d\lambda^{n-1}}\left\{T_{\mathrm{S}}[\tilde{\rho}] + \lambda U[\tilde{\rho}] + \lambda E_{\mathrm{X}}[\tilde{\rho}] + E_{\mathrm{C}}^{\lambda}[\tilde{\rho}]\right\}_{\tilde{\rho}=\rho} \\ &\quad + \frac{d^{n-1}const}{d\lambda^{n-1}} . \end{aligned} \tag{42}$$

The knowledge of $\left\{T_{\mathrm{S}}[\tilde{\rho}] + \lambda U[\tilde{\rho}] + \lambda E_{\mathrm{X}}[\tilde{\rho}] + E_{\mathrm{C}}^{\lambda}[\tilde{\rho}]\right\}$ in the order (n-1) of λ determines $d^{n-1}v_\lambda/d\lambda^{n-1}|_{\lambda=0}$, which is what is needed to calculate $\frac{d^n}{d\lambda^n}\left\{T_{\mathrm{S}}[\tilde{\rho}] + \lambda U[\tilde{\rho}] + \lambda E_{\mathrm{X}}[\tilde{\rho}] + E_{\mathrm{C}}^{\lambda}[\tilde{\rho}]\right\}$. Once we have evaluated E_{C}^{λ} up to a certain order, it is easy to obtain a series expansion for $V_{\mathrm{XC},\lambda}$, by simply taking the derivative of $E_{\mathrm{XC}}^{\lambda}$ with respect to λ.

In practice, the calculation of the $d^{n-1}v_\lambda/d\lambda^{n-1}|_{\lambda=0}$ for arbitrary n turns out to be very difficult. In zeroth order, v_λ is just v_{KS} and various methods have been developed to construct v_{KS} for a given density, among them [34–38]. The first-order term of v_λ, which appears in Eq. 40, is $-u - v_{\mathrm{X}}$, where $v_{\mathrm{X}} = \frac{\delta E_{\mathrm{X}}[\tilde{\rho}]}{\delta\tilde{\rho}}|_{\tilde{\rho}=\rho}$. A method to calculate the exchange potential v_{X} has been proposed in Ref. [39], and in Ref. [40] the implementation of yet another approach to calculate v_X is described. The higher-order terms of v_λ can in principle be obtained, but so far no exact calculation of these terms has been reported.

In the present work we evaluate the density-functional perturbation expansion up to second order. To solve the zeroth-order reference problem, we construct the Kohn-Sham potential Eq. 31 using the GGA approximation to $\frac{\delta E^\lambda_{XC}[\rho]}{\delta\rho(\mathbf{r})}$. Furthermore, we also use the GGA approximation to v_X, which is needed to calculate Eq. 40.

4 The adiabatic connection for approximate density functionals

In principle we are now able to construct E^λ_{XC} in a systematic way. This is however not useful for numerical calculations on large systems, since the computational effort to evaluate higher than second-order perturbation expansions is too high to perform these calculations for molecules with hundreds of atoms. Below we will see that second-order density-functional perturbation theory itself is not accurate enough to calculate atomization energies.

To obtain properties of real systems, only E^λ_{XC} for $\lambda = 1$ is needed and in practice local- or gradient-corrected approximations to $E^{\lambda=1}_{XC}$ are made. A simple approximation is the local spin-density approximation [4]. More recently LSD has been significantly improved by adding gradient corrections. Gradient corrected approximations (GGA's) to $E^{\lambda=1}_{XC}[\rho]$ [5–9] are of the general form

$$E^{GGA}_{XC} = \int d^3r\ f(\rho(\mathbf{r}), \nabla\rho(\mathbf{r})) \ . \tag{43}$$

Note that omitting the superscript λ implies $\lambda = 1$. An interesting result obtained in Refs. [41,42] and Eq. 49 of [32] shows that the λ-dependent E^λ_{XC} can be obtained from E_{XC} via density scaling

$$E^\lambda_{XC} = \lambda^2 E_{XC}[\rho(\mathbf{r}/\lambda)/\lambda^3] \ . \tag{44}$$

The λ-dependent E^λ_{XC} itself is not so much of interest to us. What we are interested in is the λ-decomposition of E_{XC}, i.e. $E_{XC} = \int_0^1 d\lambda\, V_{XC,\lambda}$. Equations 18 and 44 show that $V_{XC,\lambda}$ can be obtained from E_{XC} [41,42,32]

$$V_{XC,\lambda} = \frac{d}{d\lambda}\left\{\lambda^2 E_{XC}[\rho(\mathbf{r}/\lambda)/\lambda^3]\right\} \ . \tag{45}$$

Any density functional approximation (DFA) E^{DFA}_{XC} to $E^{\lambda=1}_{XC}$ can be plugged into the right hand side of Eq. 45, and E^{DFA}_{XC} can then be written as $E^{DFA}_{XC} = \int_0^{\lambda=1} d\tilde{\lambda}\ V^{DFA}_{XC,\tilde{\lambda}}$. This formula facilitates the elimination of the density functional approximation to $V_{XC,\lambda}$ for small values of λ, where the local- and gradient-corrected approximations are often much less reliable than for large λ-values.

Table 2. Comparison of the exact ΔE_{X} ($= E_{\mathrm{X}}^{\mathrm{atoms}} - E_{\mathrm{X}}^{\mathrm{molecule}}$)-, $\Delta E_{\mathrm{XC},2}$-, and ΔE_{XC} contribution to the atomization energy, with GGA results. The GGA employed here is PW91. (Energies are in kcal/mol, 1eV=23.06 kcal/mol).

Molecule	$\Delta E_{\mathrm{X}}^{GGA}$	ΔE_{X}	$\Delta E_{\mathrm{XC},2}^{GGA}$	$\Delta E_{\mathrm{XC},2}$	$\Delta E_{\mathrm{XC}}^{GGA}$	ΔE_{XC}
CH_4	193	181	315	309	277	275
NH_3	125	97	232	239	201	197
H_2O	109	81	180	202	160	160
HF	71	49	105	126	96	94
CO	95	40	158	225	140	130
N_2	40	-34	123	199	99	85
NO	34	-43	104	177	83	65
F_2	6	-70	35	107	27	12
mae	47		43		9	

To illustrate the latter point, we study various contributions to the atomization energy. In Table 2 the differential exchange energy $\Delta E_{\mathrm{X}} = E_{\mathrm{X}}^{\mathrm{atoms}} - E_{\mathrm{X}}^{\mathrm{molecule}}$, the differential exchange-correlation energy (up to second order in λ) $\Delta E_{\mathrm{XC},2} = \Delta E_{\mathrm{X}} + \Delta E_{\mathrm{C},2}$, and the differential exchange-correlation energy ΔE_{XC}, are compared at the GGA and the exact levels. The "exact" ΔE_{XC} in Table 2 is obtained by subtracting $\Delta(E^{GGA} - E_{\mathrm{XC}}^{GGA})$ from the experimental atomization energies. We see that $\Delta E_{\mathrm{X}}^{GGA}$ is too large compared to the exact result. This is especially the case for the congested systems in the lower part of the table. Furthermore, $\Delta E_{\mathrm{XC},2}^{GGA}$ is usually smaller than the exact result, but the error made by GGA becomes smaller compared to the error for ΔE_{X}. The best results are obtained for ΔE_{XC}, the quantity needed to predict the experimental atomization energy.

To understand these findings [17–19,43–48], we note that the exchange-correlation energy of a system becomes more negative when the exchange-correlation hole becomes deeper and more short-ranged. The exact exchange hole $\rho_{\mathrm{XC},\lambda=0}(\mathbf{r},\mathbf{r}')$ in a covalent diatomic molecule extends over both nuclei and is very non-local. If the reference electron is at point $\mathbf{r}$, where $\mathbf{r}$ is close to the first nucleus, then the exchange hole around that electron has a significant value when $\mathbf{r}'$ is close to the second nucleus. However, the exchange hole in the GGA approximation is always localized around its electron [49], therefore E_{X}^{GGA} for the molecule is too negative. In atoms the exact exchange hole is always fairly localized around its electron and the GGA approximation to the hole is very accurate. Now if we turn on the electron-electron repulsion, the exact hole in the molecule becomes deeper around the reference electron at point $\mathbf{r}$ and localizes rapidly compared to $\rho_{\mathrm{XC},\lambda}^{GGA}(\mathbf{r},\mathbf{r}')$, so that $V_{\mathrm{XC},\lambda}^{\mathrm{molecule}}$ as a function of λ, drops more rapidly than its GGA counterpart. This expectation is confirmed by comparison of $\Delta E_{\mathrm{XC},2}^{GGA}$ with $\Delta E_{\mathrm{XC},2}$. For larger values of λ

the exact- and the GGA exchange-correlation hole are both localized and we expect the GGA to work accurately.

At this point it is useful to introduce the notions of static- and dynamic correlation. Static- and dynamic correlation are frequently- used terms in the electronic structure theory of atoms and molecules. Nevertheless they are not clearly defined. A system is often said to have static correlation if the correlation energy cannot be obtained from second-order perturbation theory with a HF zeroth-order reference system. Examples of molecules with static correlation are F_2 and O_3. The qualitative amount of static correlation depends on the reference system. In the present context the reference system is the Kohn-Sham non-interacting problem, and there might be cases where the density-functional perturbation expansion and the expansion around the HF reference system deviate in the prediction of static correlation.

In the present work we use the terms static and dynamic in connection with density-functional perturbation theory. Factors which give rise to a slowly-convergent perturbation expansion (or to static correlation) are small gaps between the highest-occupied and lowest-unoccupied orbitals, so that the energy denominators in the perturbation expressions are small. This situation is realized for stretched covalent bonds, where the binding- and antibinding orbitals are close in energy. In this case the second-order perturbation expansion gives a much too negative correlation energy. Another reason for a slowly-convergent perturbation expansion is that several electron pairs occupy the same region of space, such as in multiply-bonded systems (congested systems). Inter pair interactions become important, and they get accounted for only in higher-order perturbation theory. For multiply-bonded systems we also obtain a much too negative correlation energy in second order. A system is said to have primarily dynamic correlation if second-order perturbation theory accounts for the electron correlation. An example of such a system is the Ne atom [50].

In Table 2 we see that the congested systems indeed show a particularly large overestimation of ΔE_{XC} by the second-order approximation $\Delta E_{\mathrm{XC},2}$. This is due to the strong static correlation present in these molecules. We also see that the density functional underestimates the static correlation for these systems. On the other hand, $\Delta E_{\mathrm{X}}^{GGA}$ is strongly overestimated. In congested systems the overlapping orbitals give rise to complicated and highly delocalized exchange holes which are poorly described by LSD and GGA [50, 51]. Obviously we get a cancellation of errors between approximate exchange and correlation, so that $\Delta E_{\mathrm{XC}}^{GGA}$ is more accurate than its individual components. Note, however, that the systems with strong static correlation (lower part of Table 2) have a disturbingly large error in ΔE_{XC}.

The picture which emerges from Table 2 and from the discussion of $\rho_{\mathrm{XC},\lambda}(\mathbf{r},\mathbf{r}')$ is that $\Delta V_{\mathrm{XC},\lambda}^{GGA}$ is much too large at $\lambda = 0$ ($\Delta V_{\mathrm{XC},\lambda=0} = \Delta E_{\mathrm{X}}$) and that $\Delta V_{\mathrm{XC},\lambda}^{GGA}$ increases too slowly as λ increases. The overestimation of $\Delta V_{\mathrm{XC},\lambda=0}$ and the slow increase of $\Delta V_{\mathrm{XC},\lambda}$ as a function of λ lead to an effective error cancellation at large λ values. However, we do not want to leave the range of λ values which contribute to E_{XC}^{GGA}, since GGA's are designed to approximate $E_{\mathrm{XC}}^{\lambda=1}$. We therefore assume that $V_{\mathrm{XC},\lambda}^{GGA}$ is most accurate for $\lambda \approx 1$.

Before we continue, we recapitulate what we have learned about the adiabatic connection. We saw how to construct $V_{\mathrm{XC},\lambda}$ in term of a power series expansion. Furthermore, utilizing Eq. 44, we can λ-decompose any given approximate functional for E_{XC}. We also showed that $V_{\mathrm{XC},\lambda}^{GGA}$ is accurate for large λ-values ($\lambda \approx 1$), and that serious errors are made by $V_{\mathrm{XC},\lambda=0}^{GGA}$.

In the following section we take the most accurate and computationally inexpensive ingredients from both the perturbation- and the DFT approach and combine low-order perturbation theory with accurate DFT approximation to $V_{\mathrm{XC},\lambda=1}$.

5 The [1/1]-Padé model for the adiabatic connection

To construct an approximation to $V_{\mathrm{XC},\lambda}$, we first derive properties of the exact $V_{\mathrm{XC},\lambda}$. Second-order Rayleigh-Schrödinger perturbation theory always gives a negative contribution to the energy, therefore $E_{\mathrm{C},2}$ is negative and this implies that

$$\frac{dV_{\mathrm{XC},\lambda}}{d\lambda} < 0 \ . \tag{46}$$

Furthermore, since

$$\langle \psi_\lambda | \hat{V}_{ee} | \psi_\lambda \rangle = U + V_{\mathrm{XC},\lambda} \geq 0 \ , \tag{47}$$

we find [52]

$$V_{\mathrm{XC},\lambda} \geq -U \ . \tag{48}$$

The $V_{\mathrm{XC},\lambda}$ curve is a monotonically falling curve, which is bound from below and thus approaches a constant value as $\lambda \to \infty$. A simple ansatz for $V_{\mathrm{XC},\lambda}$ which satisfies these constraints is the [1/1]-Padé [53,33]

$$V_{\mathrm{XC},\lambda}^{[1/1]} = \tilde{a} \left[\frac{1 + \lambda \tilde{b}}{1 + \lambda \tilde{c}} \right] \ . \tag{49}$$

The parameters $\tilde{a}$, $\tilde{b}$, and $\tilde{c}$ are fixed by the three conditions

$$\mathrm{I}: \quad V_{\mathrm{XC},\lambda=0}^{[1/1]} = E_{\mathrm{X}}, \tag{50}$$

$$\mathrm{II}: \quad V_{\mathrm{XC},\lambda=1}^{[1/1]} = V_{\mathrm{XC},\lambda=1}^{GGA}, \tag{51}$$

$$\mathrm{III}: \quad \left.\frac{dV_{\mathrm{XC},\lambda}^{[1/1]}}{d\lambda}\right|_{\lambda=1} = \left.\frac{dV_{\mathrm{XC},\lambda}^{GGA}}{d\lambda}\right|_{\lambda=1} \ . \tag{52}$$

Condition I is chosen to eliminate the E_{X}^{GGA} contribution to the adiabatic connection, since local- and gradient-corrected approximations to E_{X} often do not work well. Conditions II and III are based on the observation that local- and gradient-corrected approximations work best for exchange and correlation together, i.e., for $V_{\mathrm{XC},\lambda\approx 1}$.

We also use the [1/1]-Padé ansatz to represent $\Delta V_{\mathrm{XC},\lambda}^{GGA}$. In this case the conditions I and III are replaced by

$$\mathrm{I}' : V_{\mathrm{XC},\lambda=0}^{[1/1]} = E_{\mathrm{X}}^{GGA} , \tag{53}$$

$$\mathrm{III}' : E_{\mathrm{XC}}^{GGA} = \int_0^1 d\lambda \, V_{\mathrm{XC},\lambda}^{GGA} . \tag{54}$$

From this $V_{\mathrm{XC},\lambda}^{GGA}$ curve we evaluate the slope at $\lambda = 1$ and use it as an input for the construction of the [1/1]-Padé model. The GGA curve (satisfying conditions I′, II, and III′) and the [1/1]-Padé curve (satisfying conditions I – III) obtained in this way can be characterized by the set of three numbers $E_{\mathrm{X}}^{GGA}, V_{\mathrm{XC},\lambda=1}^{GGA}, dV_{\mathrm{XC},\lambda=1}^{GGA}/d\lambda$, and $E_{\mathrm{X}}, V_{\mathrm{XC},\lambda=1}^{GGA}, dV_{\mathrm{XC},\lambda=1}^{GGA}/d\lambda$, respectively. The [1/1]-Padé representation of $V_{\mathrm{XC},\lambda}^{GGA}$ usually accurately reproduces $\Delta V_{\mathrm{XC},\lambda}^{GGA}$ calculated from Eq. 44, and is simpler to deal with in algebraic manipulations.

The analysis of the adiabatic connection in the previous section was carried out for the atomization process, where the contributions from the core electrons cancel out. The λ-dependence of $V_{\mathrm{XC},\lambda}$ for the core electrons is essentially linear, since the zeroth-order Hamiltonian $\hat{T} + \hat{v}_{\mathrm{KS}}$ dominates in the core region and correlation is only a weak perturbation, so that second-order density-functional perturbation theory gives an accurate correlation energy. In the core, exchange dominates correlation and we do not expect to find an effective error cancellation between exchange and correlation, as is the case for the valence electrons. In order to eliminate core contributions, we apply the [1/1]-Padé to the energy difference upon atomization and not to the total exchange-correlation energies.

As an illustration of the [1/1]-Padé model, we consider the atomization of the NO molecule. In Fig. 1 the adiabatic connection for the atomization of NO is shown, both in GGA and in the [1/1]-Padé approximation. The [1/1]-Padé approximation correctly describes the strong static-correlation contribution of this process. $-\Delta V_{\mathrm{XC},\lambda}^{[1/1]}$ rapidly drops from its $\lambda = 0$ value to the $\lambda = 1$ value. The slope of the [1/1]-Padé curve at $\lambda = 0$ is much bigger than the slope of the GGA curve. If we write

$$\left.\frac{d\Delta V_{\mathrm{XC},\lambda}^{GGA}}{d\lambda}\right|_{\lambda=1} = \alpha^{GGA} \left(\Delta V_{\mathrm{XC},\lambda=1}^{GGA} - \Delta E_{\mathrm{X}}^{GGA}\right) , \tag{55}$$

where α^{GGA} is a constant to be determined from the above equation, then $d\Delta V_{\mathrm{XC},\lambda}^{GGA}/d\lambda|_{\lambda=0}$ in the [1/1]-Padé representation is given by

$$\left.\frac{d\Delta V_{\mathrm{XC},\lambda}^{GGA}}{d\lambda}\right|_{\lambda=0} = \frac{1}{\alpha^{GGA}} \left(\Delta V_{\mathrm{XC},\lambda=1}^{GGA} - \Delta E_{\mathrm{X}}^{GGA}\right) . \tag{56}$$

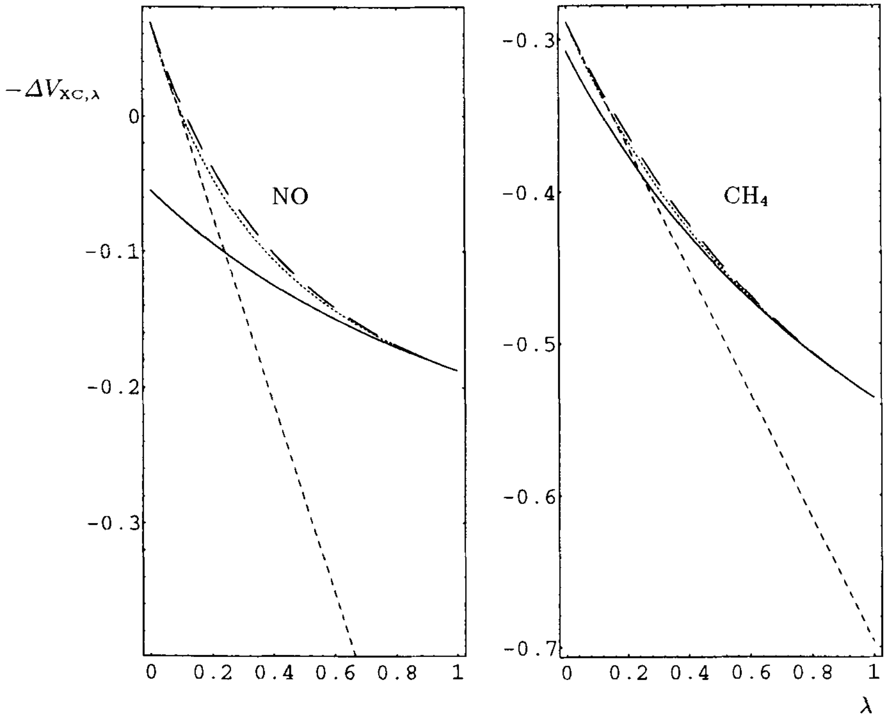


Fig. 1. Adiabatic connection for the atomization of NO and CH_4. Shown are $-\Delta V_{XC,\lambda}^{GGA}$ (solid line), the [1/1]-Padé (dotted line), the [2/2]-Padé (long dashes), and $-\Delta V_{XC,\lambda}$ obtained from second-order density functional perturbation theory (dashed line). (Energy in hartrees.)

The parameter α measures the bending of the adiabatic connection curve. α tends to 1 for a straight-line connection (no static correlation), and to 0 for maximum bending (strong static correlation). For $\alpha = 0$, the [1/1]-Padé curve immediately assumes its $\lambda = 1$ value for any $\lambda > 0$.

For the [1/1]-Padé model, employing exact exchange, we have the same relations

$$\left.\frac{d\Delta V_{XC,\lambda}^{[1/1]}}{d\lambda}\right|_{\lambda=1} = \left.\frac{d\Delta V_{XC,\lambda}^{GGA}}{d\lambda}\right|_{\lambda=1} \tag{57}$$

$$= \alpha^{[1/1]}\left(\Delta V_{XC,\lambda=1}^{GGA} - \Delta E_X\right) , \tag{58}$$

$$\left.\frac{d\Delta V_{XC,\lambda}^{[1/1]}}{d\lambda}\right|_{\lambda=0} = \frac{1}{\alpha^{[1/1]}}\left(\Delta V_{XC,\lambda=1}^{GGA} - \Delta E_X\right) , \tag{59}$$

where $\alpha^{[1/1]}$ is obtained from condition III

$$\alpha^{[1/1]} = \alpha^{GGA} \frac{\Delta V_{\mathrm{XC},\lambda=1}^{GGA} - \Delta E_{\mathrm{X}}^{GGA}}{\Delta V_{\mathrm{XC},\lambda=1}^{GGA} - \Delta E_{\mathrm{X}}} . \tag{60}$$

A big error in the exchange energy indicates strong static correlation, and the coefficient $\alpha^{[1/1]}$ becomes small, as can be seen from Eq. 60. Small $\alpha^{[1/1]}$ on the other hand means that the $\Delta V_{\mathrm{XC},\lambda}^{[1/1]}$ shows a big curvature compared to $\Delta V_{\mathrm{XC},\lambda}^{GGA}$.

In Fig. 1 we also show the adiabatic connection for the atomization of CH_4. In CH_4 (uncongested system) the bonds are well separated from each other and this means that the error in $\Delta E_{\mathrm{X}}^{GGA}$ is much smaller compared to multiply-bonded systems. Furthermore, since the binding electrons pairs are separated, we have little static correlation.

6 Applications of the [1/1]-Padé model

The most interesting application of the [1/1]-Padé model is to calculate $\Delta E_{\mathrm{XC}} = \int_0^1 d\lambda\ \Delta V_{\mathrm{XC},\lambda}^{[1/1]}$. Table 3 shows that the overbinding tendency of GGA gets reduced by the [1/1]-Padé model for the adiabatic connection. The most significant improvements are obtained for the congested systems in the lower part of Table 3. In the more uncongested systems such as NH_3 or CH_4, GGA's typically make small errors and the [1/1]-Padé does not lead to an improvement.

The slope at $\lambda = 0$ obtained from the [1/1]-Padé is given by

$$\left.\frac{d\Delta V_{\mathrm{XC},\lambda}^{[1/1]}}{d\lambda}\right|_{\lambda=0} = \frac{(\Delta V_{\mathrm{XC},\lambda=1}^{GGA} - \Delta E_{\mathrm{X}})^2}{d\Delta V_{\mathrm{XC},\lambda=1}^{GGA}/d\lambda} . \tag{61}$$

In Table 4 this prediction for $d\Delta V_{\mathrm{XC},\lambda}/d\lambda|_{\lambda=0}$ is compared to exact calculations of $d\Delta V_{\mathrm{XC},\lambda}/d\lambda|_{\lambda=0}$ with density-functional perturbation theory, and it turns out that $d\Delta V_{\mathrm{XC},\lambda}^{[1/1]}/d\lambda|_{\lambda=0}$ is usually in good agreement with the exact results. This provides a further verification of the [1/1]-Padé model, in addition to the atomization energies obtained from $\Delta V_{\mathrm{XC},\lambda}^{[1/1]}$. The straight lines in Fig. 1 have the exact slope of the adiabatic-connection curve at $\lambda = 0$, and illustrate how nicely the [1/1]-Padé reproduces this slope. Similar figures can be found in Ref. [33].

7 Using density functional perturbation theory to improve the [1/1]-Padé model

The [1/1]-Padé interpolation also has its limitations. First of all, the accuracy of the procedure is limited by the accuracy of the input data $\Delta V_{\mathrm{XC},\lambda=1}^{GGA}$ and $d\Delta V_{\mathrm{XC},\lambda}^{GGA}/d\lambda|_{\lambda=1}$. It has been shown that these input data are responsible for the remaining error ($\sim$ 4 kcal/mol) in the atomization energy of H_2,

Table 3. Comparison of atomization energies obtained from GGA and various hybrid approximations with experimental results. [1/1] and [2/2] denote the [1/1]-Padé and [2/2]-Padé, respectively. In the column $a = 0.25$, results obtained from Eq. 66 with $a = 0.25$ are listed. The GGA employed here is PW91. (Energies are in kcal/mol, 1eV=23.06 kcal/mol).

Molecule	*GGA*	[1/1]	[2/2]	exact	a=0.25
H_2	105	105	105	109	105
LiH	53	52	54	58	52
Li_2	20	20	20	24	19
LiF	137	131	133	139	130
Be_2	10	8	7	3	6
CH_4	422	419	418	419	419
NH_3	301	294	294	297	294
H_2O	233	226	226	232	226
OH	110	106	106	107	106
HF	143	138	138	141	137
Cl_2	64	59	57	58	57
P_2	120	111	111	117	109
mae	4	4	4	-	4
B_2	77	70	69	71	65
CN	197	181	181	179	177
CO	269	257	257	259	255
N_2	242	227	226	229	224
NO	171	155	155	153	152
O_2	143	128	126	121	123
O_3	185	150	146	146	137
F_2	54	41	38	39	35
mae	18	3	2	-	4

where $\Delta V^{GGA}_{XC,\lambda=1}$ is too small by about 14 mhartree [49]. For Li_2, LiH, and LiF, the GGA probably also predicts too little correlation energy, and this error is not compensated by an appropriate overestimation of ΔE_X. Furthermore, the interpolation procedure itself has its limitations. To improve on the latter point we can include further accurate input data in a more flexible interpolation scheme.

An obvious extension of the [1/1]-Padé scheme is to use the exact $d\Delta V_{XC,\lambda}/d\lambda|_{\lambda=0}$ from density-functional perturbation theory and to include this information in a [2/2]-Padé ansatz for $\Delta V_{XC,\lambda}$.

$$\Delta V^{[2/2]}_{xc,\lambda} = a\left[\frac{1+\lambda b+\lambda^2 c}{1+\lambda d+\lambda^2 e}\right] . \qquad (62)$$

The additional two parameters in this ansatz are fixed by the conditions

Table 4. Comparison of $\Delta V'_{\mathrm{XC},\lambda=0} = d\Delta V_{\mathrm{XC},\lambda}/d\lambda|_{\lambda=0}$ obtained from $\Delta V^{GGA}_{\mathrm{XC},\lambda}$ and from $\Delta V^{[1/1]}_{\mathrm{XC},\lambda}$ with exact results from density-functional perturbation theory. The GGA employed here is PW91. (Atomic units are used.)

Molecule	$\Delta V^{GGA}_{\mathrm{XC},\lambda=0}{}'$	$\Delta V^{[1/1]}_{\mathrm{XC},\lambda=0}{}'$	$\Delta V^{\mathrm{exact}}_{\mathrm{XC},\lambda=0}{}'$
CH_4	.389	.456	.407
NH_3	.339	.506	.453
H_2O	.226	.398	.386
HF	.109	.244	.245
CO	.200	.596	.592
N_2	.265	.790	.742
NO	.221	.820	.701
F_2	.091	.895	.566
mae	.281	.077	–

$$\mathrm{IV} : 0 \leq \left[\frac{-d^2 \Delta V^{[2/2]}_{\mathrm{XC},\lambda}}{d\lambda^2} \right]_{\lambda=0} = \min, \tag{63}$$

$$\mathrm{V} : \Delta V^{[2/2]}_{\mathrm{XC},\lambda} \text{ is analytic for } 0 \leq \lambda \leq 1 \ . \tag{64}$$

Conditions IV and V ensure that the [2/2]-Padé curve tries to follow the straight line with the exact slope, up to second order in λ.

For systems where ΔE^{GGA}_x shows a very large error, the [1/1]-Padé curve becomes somewhat unrealistic and the slope at $\lambda = 0$ obtained from the [1/1]-Padé curve does not accurately reproduce the exact slope obtained from density-functional perturbation theory. Examples for such systems are F_2 and O_3 (see Table 4 and also Ref. [33]). The atomization energies obtained with the [2/2]-Padé are also listed in Table 3, and we see that [2/2]-Padé scheme further improves the atomization energies for F_2 and O_3.

For NO and CH_4 the [2/2]-Padé curves are plotted in Fig. 1. They are virtually identical to the [1/1]-Padé curves for these systems.

8 Relation between the nonempirical- and empirical hybrid schemes

8.1 Empirical hybrid schemes

Hybrid schemes, which mix a fraction of exact exchange with density functional approximations to exchange and correlation, have been invented by Becke [17–19]. Motivated by the adiabatic connection formula, Becke proposed the 3-parameter mixing formula [18]

$$E_{\mathrm{XC}}^{BeckeI} = E_{\mathrm{XC}}^{LSD} + a_0(E_{\mathrm{X}} - E_{\mathrm{X}}^{LSD}) \\ + a_x(E_{\mathrm{X}}^{GGA} - E_{\mathrm{X}}^{LSD}) + a_c(E_{\mathrm{C}}^{GGA} - E_{\mathrm{X}}^{LSD}) \ . \tag{65}$$

The crucial parameter for the success of this approach is a_0. a_0 determines how much LSD exchange is replaced by exact exchange to cure the problems associated with density functional approximations to E_{X}. The empirical parameters a_0, a_x, and a_c are obtained from a fit to atomization energies, ionization potentials, proton affinities, and total energies [18], and their values are 0.2, 0.72, and 0.81, respectively. More recently [19] Becke introduced a formally much simpler 1-parameter mixing formula

$$E_{\mathrm{XC}}^{BeckeII} = a(E_{\mathrm{X}} - E_{\mathrm{X}}^{GGA}) + E_{\mathrm{XC}}^{GGA} \ . \tag{66}$$

The mixing coefficient a is fitted to atomization energies, ionization potentials, and proton affinities. As in the case of Eq. 65, a is dependent on the data set used and on the choice of the density functional approximation to $E_{\mathrm{XC}}[\rho]$. The value of a reported by Becke is 0.16 or 0.28, depending on the GGA used.

Here we show how the fixed mixing coefficient can be estimated, starting from the [1/1]-Padé hybrid scheme. To arrive at a system-independent mixing coefficient, we use the fact that the bending of the $\Delta V_{\mathrm{XC},\lambda}^{GGA}$ curve, as measured by α^{GGA} (Eq. 55), turns out to be fairly system independent for a large number of molecules. We will see that a fixed mixing coefficient therefore gives good results for this class of molecules. However, systems for which $\Delta V_{\mathrm{XC},\lambda}^{GGA}$ does not show the typical bending behavior are not accurately described by this fixed mixing coefficient. Furthermore, we show that the validity of a fixed mixing coefficient relies on the fact that the error in $\Delta E_{\mathrm{X}}^{GGA}$ is sufficiently small.

8.2 Estimation of the exact-exchange mixing coefficient

To keep the discussion simple, we focus on Eq. 66, which has only an exchange-mixing parameter, and estimate the value of this parameter. Applied to a chemical process, this formula can be rewritten in terms of energy differences

$$\Delta E_{\mathrm{XC}}^{BeckeII} = a(\Delta E_{\mathrm{X}} - \Delta E_{\mathrm{X}}^{GGA}) + \Delta E_{\mathrm{XC}}^{GGA} \ . \tag{67}$$

We define the quantity δ^{hyb} by

$$\delta^{hyb} = \frac{\Delta E_{\mathrm{XC}}^{hyb} - \Delta E_{\mathrm{XC}}^{GGA}}{\Delta V_{\mathrm{XC},\lambda=1}^{GGA}} \ , \tag{68}$$

where *hyb* stands either for BeckeII or [1/1]. $\delta^{[1/1]}$ depends on 4 variables: ΔE_{X}, $\Delta E_{\mathrm{X}}^{GGA}$, $\Delta V_{\mathrm{XC},\lambda=1}^{GGA}$ and $d\Delta V_{\mathrm{XC},\lambda}^{GGA}/d\lambda|_{\lambda=1}$, but, due to the normalization performed (i.e., the division by $\Delta V_{\mathrm{XC},\lambda=1}^{GGA}$), only 3 variables can be varied independently. To further simplify the discussion, we write $d\Delta V_{\mathrm{XC},\lambda}^{GGA}/d\lambda|_{\lambda=1}$ as

Table 5. The reduced exchange energy $\Delta\tilde{E}_{X}^{GGA} = \Delta E_{X}^{GGA}/\Delta V_{XC,\lambda=1}^{GGA}$ and the error in the reduced exchange energy $\Delta\Delta\tilde{E}_{X}^{GGA} = (\Delta E_{X} - \Delta E_{X}^{GGA})/\Delta V_{XC,\lambda=1}^{GGA}$ for the atomization of molecules. The bending-parameter α^{GGA} and the effective exact-exchange mixing coefficients for the [1/1]-Padé- and [2/2]-Padé schemes are also listed. The GGA employed here is PW91.

Molecule	$\Delta\tilde{E}_{X}^{GGA}$	$-\Delta\Delta\tilde{E}_{X}^{GGA}$	α^{GGA}	$a^{[1/1]}$	$a^{[2/2]}$
H_2	0.43	0.01	0.63	0.259	0.065
LiH	0.51	0.06	0.64	0.253	0.023
Li_2	0.22	0.16	0.42	0.193	0.194
LiF	0.75	0.18	0.69	0.231	0.216
Be_2	0.63	0.49	0.45	0.157	0.320
CH_4	0.57	0.04	0.59	0.244	0.300
NH_3	0.49	0.11	0.61	0.240	0.260
H_2O	0.55	0.14	0.62	0.237	0.238
OH	0.50	0.20	0.63	0.234	0.247
HF	0.62	0.19	0.64	0.232	0.232
Cl_2	0.46	0.36	0.51	0.194	0.253
P_2	0.02	0.53	0.52	0.203	0.206
B_2	0.66	0.52	0.46	0.154	0.159
CN	0.41	0.56	0.58	0.200	0.200
CO	0.55	0.32	0.61	0.214	0.215
N_2	0.28	0.52	0.61	0.215	0.217
NO	0.29	0.66	0.60	0.205	0.214
O_2	0.39	0.74	0.58	0.191	0.212
O_3	0.14	1.40	0.59	0.180	0.203
F_2	0.15	1.82	0.63	0.174	0.204
averages	0.43	0.45	0.58	0.211	0.209

as

$$\left.\frac{d\Delta V_{XC,\lambda}^{GGA}}{d\lambda}\right|_{\lambda=1} = \alpha^{GGA}(\Delta V_{XC,\lambda=1}^{GGA} - \Delta E_{x}^{GGA}) \ . \tag{69}$$

Most molecules in Table 5 have α^{GGA} values which scatter only little around the average value of 0.58. The molecules with exceptionally low or high α^{GGA} are Li_2, B_2, Be_2, and LiF. To a good approximation we set $\alpha^{GGA} = 0.58$. $\delta^{[1/1]}$ now depends only on two independent variables, which we choose as $\Delta\Delta\tilde{E}_{X}^{GGA} = (\Delta E_{x} - \Delta E_{x}^{GGA})/\Delta V_{XC,\lambda=1}^{GGA}$ and $\Delta\tilde{E}_{X}^{GGA} = \Delta E_{x}^{GGA}/\Delta V_{XC,\lambda=1}^{GGA}$. $\delta^{BeckeII}$ takes the simple form

$$\begin{aligned}\delta^{BeckeII} &= \frac{\Delta E_{XC}^{BeckeII} - \Delta E_{XC}^{GGA}}{\Delta V_{XC,\lambda=1}^{GGA}} \\ &= a\Delta\Delta\tilde{E}_{X}^{GGA} \ ,\end{aligned} \tag{70}$$

which is independent of $\Delta\tilde{E}_{X}^{GGA}$ and α^{GGA}. In Figs. 2 and 3 we plot $\delta^{[1/1]}$

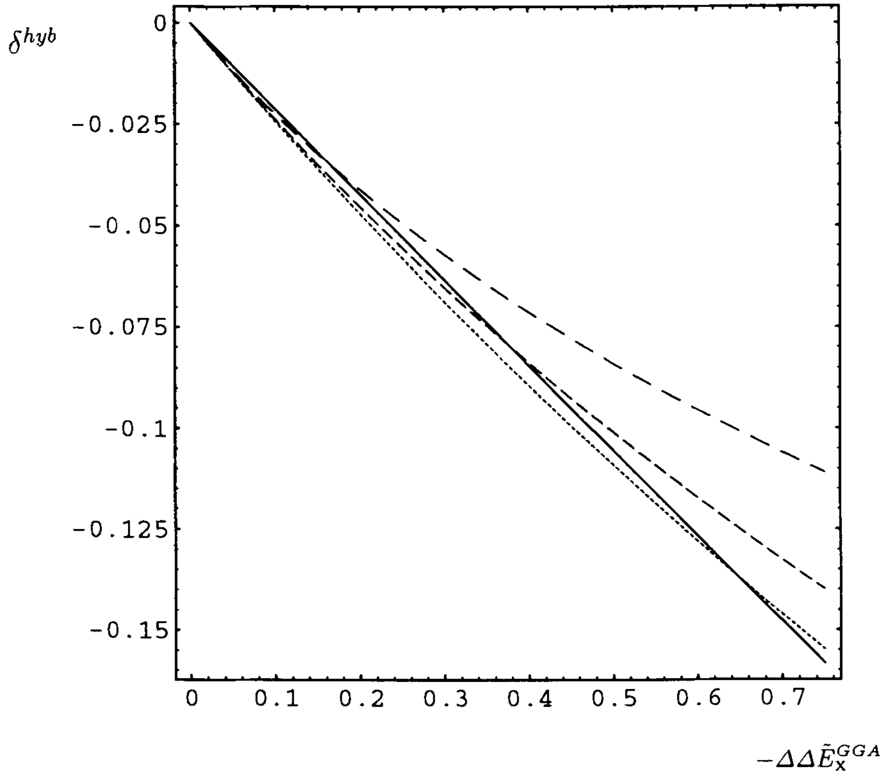


Fig. 2. Shown is $\delta^{[1/1]}$ for various values of $\Delta\tilde{E}_{\mathrm{x}}^{GGA} = 0.02$ (dotted line), $\Delta\tilde{E}_{\mathrm{x}}^{GGA} = 0.43$ (short dashes), $\Delta\tilde{E}_{\mathrm{x}}^{GGA} = 0.75$ (long dashes). The value of α^{GGA} =0.58 has been used for these curves. The straight solid line has a slope of -0.21.

for different values of $\Delta\tilde{E}_{\mathrm{x}}^{GGA}$. The various curves cover the range of typical $\Delta\tilde{E}_{\mathrm{x}}^{GGA}$ values: $\Delta\tilde{E}_{\mathrm{x}}^{GGA} \in [0.02, 0, 75]$ (see Table 5). Also in Table 5, we list the values of $-\Delta\Delta\tilde{E}_{\mathrm{x}}^{GGA}$ for a number of molecules. This quantity typically lies in the range of 0-0.7. Exceptions are F_2 and O_3, which show strong static correlation and hence large $-\Delta\Delta\tilde{E}_{\mathrm{x}}^{GGA}$. To discuss the curves, we define an effective mixing coefficient $a^{[1/1]}$ for the [1/1]-Padé

$$\Delta E_{\mathrm{xc}}^{[1/1]} = a^{[1/1]}(\Delta E_{\mathrm{x}} - \Delta E_{\mathrm{x}}^{GGA}) + \Delta E_{\mathrm{xc}}^{GGA} \ . \tag{71}$$

This effective mixing coefficient can be read off the curves in Figs. 2 and 3. $a^{[1/1]}$ is simply the negative slope of a line joining the points $(0, 0)$ and

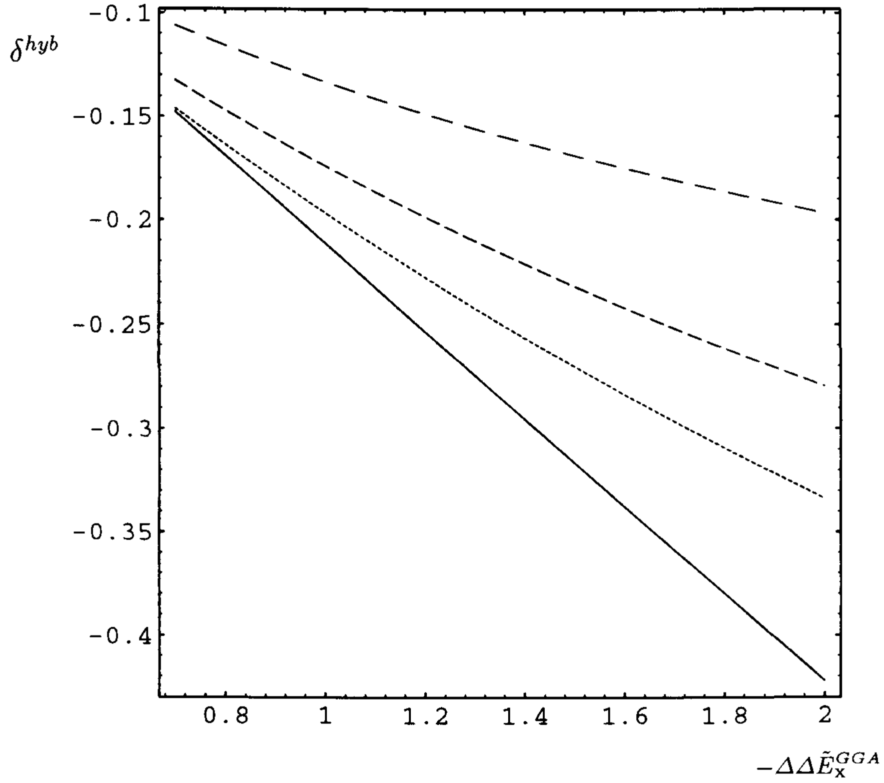


Fig. 3. The same curves are shown as in Fig. 2, but for $-\Delta\Delta\tilde{E}_x^{GGA}$ values between 0.7 and 2.

$(-\Delta\Delta\tilde{E}_x^{GGA}, \delta^{[1/1]})$. Now we want to obtain an approximation for the coefficient multiplying $\Delta\Delta\tilde{E}_x^{GGA}$ in Eq. 70. We expand $\delta^{[1/1]}$ in powers of $\Delta\Delta\tilde{E}_x^{GGA}$

$$\delta^{[1/1]} = \frac{\alpha^{GGA}(1 - {\alpha^{GGA}}^2 + 2\alpha^{GGA}\ln\alpha^{GGA})}{(1-\alpha^{GGA})^3}\Delta\Delta\tilde{E}_x^{GGA} + \ldots \quad (72)$$

Note that the first-order term is independent of $\Delta\tilde{E}_x^{GGA}$. For $\alpha^{GGA} = 0.58$, the coefficient of the term linear in $\Delta\Delta\tilde{E}_x^{GGA}$ in Eq. 72 becomes 0.25. This mixing coefficient agrees with the one obtained in Ref. [54], and falls in the range of the empirically determined mixing coefficients. A plot of

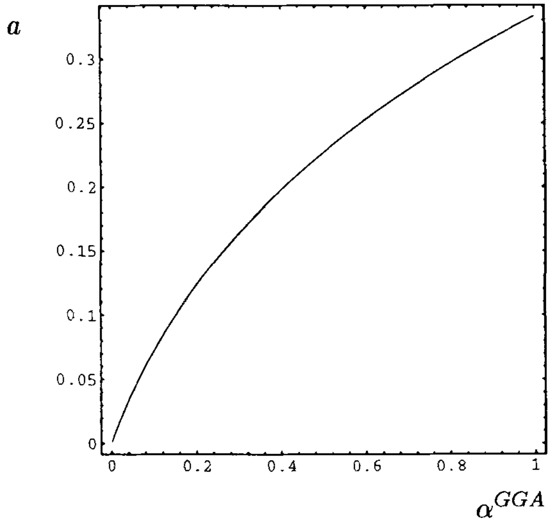


Fig. 4. The exact-exchange mixing coefficient a obtained from the [1/1]-Padé as a function of the bending parameter α^{GGA}.

$$a(\alpha^{GGA}) = \frac{\alpha^{GGA}(1 - \alpha^{GGA^2} + 2\alpha^{GGA} \ln \alpha^{GGA})}{(1 - \alpha^{GGA})^3} \tag{73}$$

is shown in Fig. 4. In the range of α-values of the systems in Table 5 ($\alpha^{GGA} \in [0.42, 0.69]$), the mixing coefficient depends only weakly on α^{GGA}.

It is clear from Figs. 2 and 3 that the mixing coefficient of 0.25 is an upper bound on $a^{[1/1]}$ for $\alpha^{GGA} = 0.58$. Therefore we expect that the quarter mixing in general leads to an overcorrection of the GGA atomization energies. This expectation is confirmed by the $a = 0.25$ atomization energies listed in Table 3. An empirical fit of the mixing parameter to the atomization energies of the systems listed in Table 5 would result in a lower value for a.

For comparison, a straight line is also drawn in Figs. 2 and 3. This line has a slope of 0.21, which is the average of the effective mixing coefficients obtained from the [1/1]-Padé model. This value of a is very close to the empirical value of 0.2 used in Eq. 65, and falls in the range of empirical values used in Eq. 66. Furthermore, it is interesting to note that a mixing coefficient $a^{[1/1]} = 0.22$ is obtained by inserting the average values of α^{GGA}, $\Delta\Delta\tilde{E}_{\mathrm{x}}^{GGA}$, and $\Delta\tilde{E}_{\mathrm{x}}^{GGA}$ into the [1/1]-Padé model.

As can be seen from Table 5 and from Figs. 2 and 3, a straight line gives a very good approximation to the predictions of the [1/1]-Padé approximation for a, in the range of physical combinations of $\Delta\Delta\tilde{E}_{\mathrm{x}}^{GGA}$ and $\Delta\tilde{E}_{\mathrm{x}}^{GGA}$.

Molecules with large values of $\Delta\tilde{E}_{\mathrm{X}}^{GGA}$, which causes a non-linear behavior of the $\delta^{[1/1]}$ curve as a function of $\Delta\Delta\tilde{E}_{\mathrm{X}}^{GGA}$, usually have quite small $-\Delta\Delta\tilde{E}_{\mathrm{X}}^{GGA}$, so that the straight line is still a good approximation to the [1/1]-Padé model. A big value of $\Delta\tilde{E}_{\mathrm{X}}^{GGA}$ means that we have little correlation contribution to the atomization process. This is typical for uncongested systems, where the binding electron pairs are separated from each other. These systems tend to have only a small $-\Delta\Delta\tilde{E}_{\mathrm{X}}^{GGA}$, compared to the congested molecules, since the binding electron pairs have only little overlap with each other and the orbital-nodality problem does not cause problems for the GGA approximation to ΔE_{X} [43,51]. On the other hand, molecules with large values of $-\Delta\Delta\tilde{E}_{\mathrm{X}}^{GGA}$ tend to have small $\Delta\tilde{E}_{\mathrm{X}}^{GGA}$, so that the mixing coefficient obtained from the linear approximation to $\delta^{[1/1]}$ is again very close to the one obtained from the [1/1]-Padé model. A large $-\Delta\Delta\tilde{E}_{\mathrm{X}}^{GGA}$ can usually be found in the congested systems, where the bond orbitals are in the same region of space and the orbital nodality causes difficulties for local- and gradient-corrected functionals [51]. The interacting electron pairs in congested systems give rise to a large $\Delta V_{\mathrm{C},\lambda=1}$ contribution to the atomization energy, and hence to a small $\Delta\tilde{E}_{\mathrm{X}}^{GGA}$. These considerations show that for a large class of systems the fixed mixing coefficient is a good approximation to $a^{[1/1]}$.

In addition to the dependence on $\Delta\Delta\tilde{E}_{\mathrm{X}}^{GGA}$, the [1/1]-Padé approximation to ΔE_{XC} also depends on $\Delta\tilde{E}_{\mathrm{X}}^{GGA}$ and on the coefficient α^{GGA}. To discuss these dependencies and to discuss the difference between the empirical- and nonempirical hybrid schemes, we consider the limits $\Delta\tilde{E}_{\mathrm{X}}^{GGA} \to -\infty$, $\alpha^{GGA} \to 0$, or $\Delta\tilde{E}_{\mathrm{X}}^{GGA} \to 1$, and $\Delta\Delta\tilde{E}_{\mathrm{X}}^{GGA} \to -\infty$.

For $\Delta\tilde{E}_{\mathrm{X}}^{GGA} \to -\infty$, the $\delta^{[1/1]}$ curve approaches a straight line

$$\lim_{\Delta\tilde{E}_{\mathrm{X}}^{GGA}\to-\infty} \delta^{[1/1]} = \frac{\alpha^{GGA}(1-{\alpha^{GGA}}^2+2\alpha^{GGA}\ln\alpha^{GGA})}{(1-\alpha^{GGA})^3}\Delta\Delta\tilde{E}_{\mathrm{X}}^{GGA} \ . \tag{74}$$

In this limit the slope of the straight line is only a function of the parameter α^{GGA}. The systems which are closest to the $\Delta\tilde{E}_{\mathrm{X}}^{GGA} \to -\infty$ limit are the congested molecules, thus we expect that a mixing coefficient $a = 0.25$, which is obtained by using a value of 0.58 for α^{GGA} on the right-hand-side of Eq. 74, will work well for most congested systems. These systems are the ones which show the biggest errors in the GGA atomization energies.

In the limit $\Delta\tilde{E}_{\mathrm{X}}^{GGA} \to 1$ or $\alpha^{GGA} \to 0$, we have negligible $\Delta V_{\mathrm{C},\lambda=1}^{GGA}$ contribution and therefore no dynamic correlation. $-\Delta V_{\mathrm{XC},\lambda}$ should drop immediately from $-\Delta E_{\mathrm{X}}$ to $-\Delta E_{\mathrm{X}}^{GGA}$, as it does in the [1/1]-Padé model. $a^{[1/1]}$ is zero in this case.

The [1/1]-Padé hybrid shows a more complicated non-linear dependence on $\Delta\Delta\tilde{E}_{\mathrm{X}}^{GGA}$ than $\delta^{BeckeII}$, and in the limit $\Delta\Delta\tilde{E}_{\mathrm{X}}^{GGA} \to -\infty$ we find

$$\lim_{\Delta\Delta\tilde{E}_{\mathrm{X}}^{GGA}\to-\infty} \left(\frac{\delta^{[1/1]}}{\Delta\Delta\tilde{E}_{\mathrm{X}}^{GGA}}\right) = 0 \ . \tag{75}$$

The result of Eq. 75 is the one which we expect in the limit where the static correlation contribution to the atomization process goes to infinity. Any mixing of exact exchange must be avoided, and this is assured by the [1/1]-Padé model but not by the Becke hybrid. A physical system which is close to this limit is O_3, which is not accurately described by empirical hybrid schemes.

9 Application of hybrid schemes to ionization potentials of atoms

The motivation for the hybrid schemes described here is that the exchange-correlation hole in molecules, at small values of the coupling-constant λ, is not well approximated by local- and gradient-corrected approximations. The question which arises is how the hybrid schemes perform for properties of atoms, for instance, where the exchange hole is less nonlocal and better approximated by GGA than in molecules. To address this question, we study the ionization potentials of atoms. The ionization potential I is defined by

$$I = E^{\text{ion}} - E^{\text{atom}} \ . \tag{76}$$

From Table 6 we see that GGA improves upon LSD, and furthermore that the inclusion of exact exchange with the [1/1]-Padé and the $a = 0.25$ hybrid does not lead to a further significant improvement. This shows that the accuracy of the GGA is not limited by the accuracy of the small-λ behavior of $V^{GGA}_{\text{xc},\lambda}$. Interestingly, the [2/2]-Padé reduces the mean absolute error to

Table 6. Ionization potentials for atoms. LSD, GGA, and various hybrid calculations are compared to experimental values. [1/1] and [2/2] denote the [1/1]-Padé and [2/2]-Padé, respectively. In the column $a = 0.25$, results obtained from Eq. 66 with a =0.25 are listed. The GGA employed here is PBE. (All energies are in kcal/mol, 1eV=23.06 kcal/mol.)

Atom	LSD	GGA	[1/1]	[2/2]	exact	$a = 0.25$
He	560	564	564	565	567	564
Li	126	129	128	125	124	128
Be	208	208	207	208	215	207
B	199	200	199	193	191	199
C	269	266	265	262	260	265
N	345	339	339	337	335	339
O	321	323	318	316	314	318
F	414	407	402	402	402	402
Ne	510	497	493	493	497	493
mae	8	5	4	2	–	4

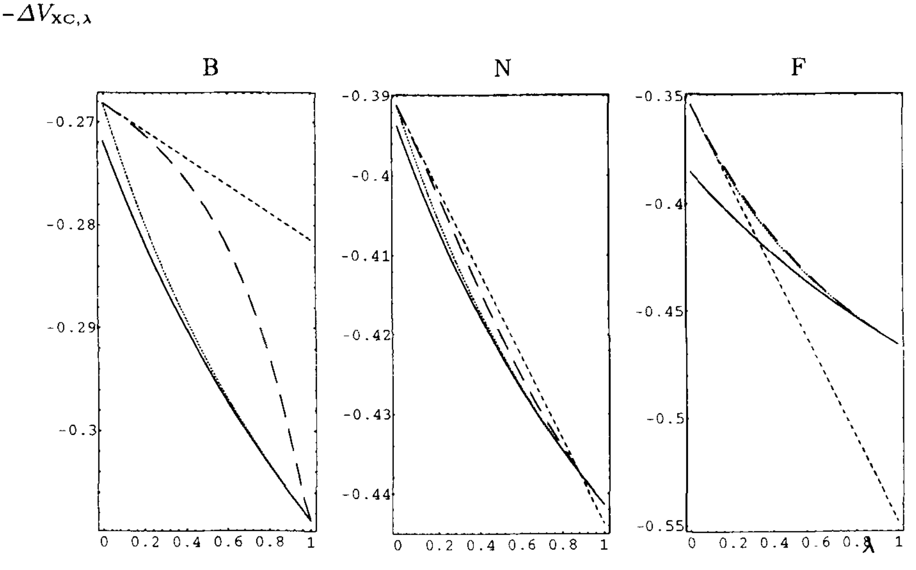


Fig. 5. Adiabatic connection for the ionization of B, N, and F. $-\Delta V_{XC,\lambda} = -(V^{ion}_{XC,\lambda} - V^{atom}_{XC,\lambda})$ is shown for GGA (solid line), the [1/1]-Padé model (dotted line), and the [2/2]-Padé model (long dashes). The dashed line is $\Delta V_{XC,\lambda}$ in second-order density functional perturbation theory. (Energy in hartrees.)

2 kcal/mol. To understand these findings, we plot the adiabatic connection for the ionization process of B, N, and F in Fig. 5. The GGA shows a small error in ΔE_X for the B atom, and correspondingly only a small improvement is obtained from the [1/1]-Padé and the quarter mixing. The $\Delta V_{XC,\lambda}$ from second-order density functional perturbation theory indicates that the correlation contribution to the ionization potential is overestimated, i.e., the $-\Delta V^{GGA}_{XC,\lambda=1}$ value in Fig. 5 is probably much too negative. The [2/2]-Padé removes a large fraction of the error in the correlation contribution. The error in ΔE_C is coming from the self-correlation error [55] made by the GGA for the 2p electron. This self-correlation error of the GGA can of course not be cured by using exact exchange. For the N atom we find again a small error in ΔE_X, so that the hybrids employing exact exchange do not improve the ionization potential. The [2/2]-Padé hybrid removes part of the self-correlation error. For the ionization of the F atom, we obtain adiabatic-connection curves similar to the ones obtained for atomization processes. We have a significant error in ΔE_X which gets compensated by an underestimation of the correlation energy, so that $\Delta V_{XC,\lambda\approx 1}$ is probably very accurate. The [1/1]-Padé improves upon GGA, and the slope of the [1/1]-Padé curve at $\lambda = 0$ agrees

with the slope obtained from second-order perturbation theory, therefore the [2/2]-Padé curve lies on-top of the [1/1]-Padé curve.

The hybrid schemes make the largest error for the ionization potential of Be. This atom has a large static-correlation energy, due to the near degeneracy of the 2s- and 2p levels. E_{X}^{GGA} does not account for the static correlation in this system. The exact exchange-correlation hole in Be does not have the highly nonlocal character of the exchange-correlation hole in typical covalent molecules, and the error made by $\Delta E_{\mathrm{X}}^{GGA}$ is fairly small.

10 Self-consistency in hybrid calculations

The hybrid calculations reported here are done in a post-GGA manner. The densities, the orbitals, and the orbital energies from a self-consistent Kohn-Sham calculation are used in the subsequent hybrid calculations. Until now, no self-consistent implementation of hybrid schemes has been reported. Usually the empirical-hybrid calculations are made self-consistent with respect to the occupied orbitals [20], but not with respect to the density.

A self-consistent hybrid scheme, employing exact exchange, requires the calculation of the derivative of the orbital-dependent exchange functional

$$\begin{aligned} E_x = & -\frac{1}{2}\sum_{\sigma}\int d^3r d^3r' \\ & \times |\sum_{i}^{occ} \phi_i^{KS*}(\mathbf{r}',\sigma)\phi_i^{KS}(\mathbf{r},\sigma)|^2/|\mathbf{r}'-\mathbf{r}| \end{aligned} \tag{77}$$

with respect to the density [56–61]. σ in the above equation is the spin variable.

11 Summary

Nonempirical local- and gradient-corrected approximations to E_{XC} are derived from the homogeneous- or slowly-varying electron gas, where the coupling-constant averaged exchange-correlation hole $\rho_{\mathrm{XC}}(\mathbf{r},\mathbf{r}') = \int_0^1 d\lambda\ \rho_{\mathrm{XC},\lambda}(\mathbf{r},\mathbf{r}')$ is localized around the reference electron. In molecules $\rho_{\mathrm{XC}}(\mathbf{r},\mathbf{r}')$ is often highly non-local and therefore not correctly described by LSD and GGA. As a consequence $E_{\mathrm{XC}}^{\mathrm{molecule}}$ is too negative in LSD and GGA. The coupling-constant decomposition of $\rho_{\mathrm{XC}}(\mathbf{r},\mathbf{r}')$ facilitates the elimination of the most long-ranged components of $\rho_{\mathrm{XC}}(\mathbf{r},\mathbf{r}')$, which are coming from the lower end of the coupling-constant integration. For small values of the coupling constant, $V_{\mathrm{XC},\lambda} = 1/2\int d^3rd^3r'\ \ \rho(\mathbf{r})\rho_{\mathrm{XC},\lambda}(\mathbf{r},\mathbf{r}')\ /|\mathbf{r}-\mathbf{r}'|$ can be calculated from low-order density-functional perturbation theory [32,33]. By interpolating between $\Delta V_{\mathrm{XC},\lambda}$ for small λ and $\Delta V_{\mathrm{XC},\lambda}^{GGA}$ for large λ, we obtained a nonempirical hybrid scheme, which improves upon GGA. The nonempirical- [33,54,62] and empirical hybrid schemes [17–19] rely on the accuracy of $\Delta V^{GGA}_{\mathrm{XC},\lambda\approx 1}$, and cases where $\Delta V^{GGA}_{\mathrm{XC},\lambda\approx 1}$ is seriously in error, such as for the atomization of LiF or the ionization of Be, remain unsolved problems.

The remarkable success of the empirical hybrid schemes can be explained by starting from the [1/1]-Padé model, using the observation that the bending of the $\Delta V^{GGA}_{\mathrm{XC},\lambda}$ curve, as measured by the parameter α^{GGA}, is fairly system-independent. For molecules with very strong static correlation the [1/1]-Padé interpolation for $\Delta V_{\mathrm{XC},\lambda}$ becomes uncertain and hybrids employing only exact exchange begin to fail. The inclusion of second-order density-functional perturbation theory becomes necessary to describe the small-λ behavior of $\Delta V_{\mathrm{XC},\lambda}$ correctly.

The nonempirical hybrid schemes described here (and also others [54,62]) yield energy differences by construction. The application of these schemes to total energies is in general not promising, since the prerequisite of an effective error cancellation between exchange and correlation is often not met if the energy has large core contributions. In the core, exchange dominates correlation. Empirical hybrid schemes (Eqs. 65 and 66) have a well-defined total energy; this however does not mean that the fixed mixing coefficient of about 0.2 is appropriate for the high-density core electrons. Core electrons usually do not matter in the empirical hybrid schemes, since their contribution to ΔE_{XC} cancels out for valence-electron processes. This cancellation is possible because the empirical hybrid schemes are linear in the inputs E_{X}, E^{GGA}_{XC}, etc., which is not the case for the nonempirical schemes, where we have to use energy differences as inputs, to ensure cancellation of core contributions. An optimal hybrid method would combine a well-defined energy with a variable exact-exchange mixing.

12 Appendix

The calculations reported in this paper are performed with a modified version of the CADPAC program [63]. The electron densities are obtained from unrestricted Kohn-Sham calculations in the GGA approximation. Nonspherical densities and Kohn-Sham potentials have been used for open-shell atoms [64]. The experimental geometries employed in our work are taken from [65], and the experimental atomization energies and the zero point energies are from Refs. [66]. The gaussian basis sets used are of triple-zeta quality with up to $l+2$-type polarization functions for the first- and second-row elements and $l+1$-type polarization functions for the third-row elements. l is the angular momentum number of the highest occupied orbital in the atom.

13 Acknowledgements

Helpful discussions with Prof. J. P. Perdew are acknowledged. Furthermore, I would like to thank M. E. Kieh for inspiring discussions. This work was supported by the Deutsche Forschungsgemeinschaft and the National Science Foundation under grant DMR95-21353.

References

1. M. Born and R. Oppenheimer, Ann. Phys. **84**, 457 (1927).
2. M. Born and K. Huang, *Dynamical Theory of Crystal Lattices* (Oxford Univ. Press, New York, 1956).
3. A. Szabo and N. S. Ostlund, *Modern Quantum Chemistry* (Macmillan Publishing, New York, 1982).
4. W. Kohn and L. J. Sham, Phys. Rev. **140**, A1133 (1965).
5. J. P. Perdew, K. Burke, and M. Ernzerhof, Phys. Rev. Lett. **77**, 3865 (1996); ibid, **78**, 1396 (1997)(E).
6. K. Burke, J. P. Perdew, and Y. Wang, in *Electronic Density Functional Theory: Recent Progress and New Directions*, edited by J. F. Dobson, G. Vignale, and M. P. Das (Plenum, NY, 1997).
7. A. D. Becke, Phys. Rev. A **38**, 3098 (1988).
8. C. Lee, W. Yang, and R. G. Parr, Phys. Rev. B **37**, 785 (1988).
9. J. P. Perdew and Y. Wang, Phys. Rev. B **33**, 8800 (1986).
10. J. P. Perdew, in *Electronic Structure of Solids' 91*, edited by P. Ziesche and H. Eschrig (Akademie Verlag, Berlin, 1991).
11. J. P. Perdew, K. Burke, and Y. Wang, Phys. Rev. B **54**, 16533 (1996).
12. J. P. Perdew, M. Ernzerhof, A. Zupan, and K. Burke, Nonlocality of exchange and correlation: Physical origins and chemical consequences, submitted to J. Chem. Phys.
13. *Chemical Applications of Density-Functional Theory*, edited by B. B. Laird, R. B. Ross, and T. Ziegler (ACS Books, Washington, 1996).
14. *Density Functional Theory III*, Vol. 182 of *Topics in Current Chemistry*, edited by R. F. Nalewajski (Springer-Verlag, Berlin, 1996).
15. *Density Functional Theory IV*, Vol. 183 of *Topics in Current Chemistry*, edited by R. F. Nalewajski (Springer-Verlag, Berlin, 1996).
16. A. D. Becke and K. E. Edgecombe, J. Chem. Phys. **92**, 5397 (1990).
17. A. D. Becke, J. Chem. Phys. **98**, 1372 (1993).
18. A. D. Becke, J. Chem. Phys. **98**, 5648 (1993).
19. A. D. Becke, J. Chem. Phys. **104**, 1040 (1996).
20. P. J. Stephens, F. J. Devlin, C. F. Chabalowski, and M. J. Frisch, J. Phys. Chem. **98**, 11623 (1994).
21. L. A. Curtiss, K. Raghavachari, P. C. Redfern, and J. A. Pople, J. Chem. Phys. **106**, 1063 (1997).
22. H. Hellmann, *Einführung in die Quantenchemie* (Deuticke, Leipzig, 1937).
23. R. P. Feynman, Phys. Rev. **41**, 721 (1939).
24. D. C. Langreth and J. P. Perdew, Solid State Comm. **17**, 1425 (1975).
25. D. C. Langreth and J. P. Perdew, Phys. Rev. B **15**, 2884 (1977).
26. O. Gunnarsson and B. I. Lundqvist, Phys. Rev. B **13**, 4274 (1976).
27. P. Hohenberg and W. Kohn, Phys. Rev. **136**, B864 (1964).
28. M. Levy, Proc. Natl. Acad. Sci. USA **76**, 6062 (1979).
29. R. G. Parr and W. Yang, *Density-Functional Theory of Atoms and Molecules* (Oxford University Press, Oxford, 1989).

30. R. M. Dreizler and E. K. U. Gross, *Density Functional Theory* (Springer Verlag, Berlin, 1990).
31. J. P. Perdew and S. Kurth, Density Functionals for Non-Relativistic Coulomb Systems, this book.
32. A. Görling and M. Levy, Phys. Rev. B **47**, 13105 (1993).
33. M. Ernzerhof, Chem. Phys. Lett. **263**, 499 (1996).
34. S. H. Werden and E. R. Davidson, *Local Density Approximation in Quantum Chemistry and Solid State Physics* (Plenum Press, New York, 1982).
35. A. Görling, Phys. Rev. A **46**, 3753 (1992).
36. A. Görling and M. Ernzerhof, Phys. Rev. A **51**, 4501 (1995).
37. Q. Zhao and R. G. Parr, Phys. Rev. A **46**, 2337 (1992).
38. R. van Leeuwen and E. J. Baerends, Phys. Rev. A **49**, 2421 (1994).
39. M. Levy, M. Ernzerhof, and A. Görling, Phys. Rev. A **53**, 3963 (1996).
40. C. Filippi, C. J. Umrigar, and X. Gonze, Phys. Rev. A **54**, 4810 (1996).
41. M. Levy and J. P. Perdew, Phys. Rev. A **32**, 2010 (1985).
42. M. Levy, N. H. March, and N. C. Handy, J. Chem. Phys. **104**, 1989 (1996).
43. M. Ernzerhof, J. P. Perdew, and K. Burke, Int. J. Quantum Chem. **64**, 285 (1997).
44. J. P. Perdew, A. Savin, and K. Burke, Phys. Rev. A **51**, 4531 (1995).
45. J. P. Perdew, M. Ernzerhof, K. Burke, and A. Savin, Int. J. Quantum Chem. **61**, 197 (1997).
46. J. C. Slater and K. H. Johnson, Phys. Rev. B **5**, 844 (1971).
47. M. Cook and M. Karplus, J. Phys. Chem. **91**, 31 (1987).
48. V. Tschinke and T. Ziegler, J. Chem. Phys. **93**, 8051 (1990).
49. M. Ernzerhof, K. Burke, and J. P. Perdew, in *Recent Developments in Density Functional Theory, Theoretical and Computational Chemistry,* edited by J. M. Seminario (Elsevier, Amsterdam, 1997).
50. M. Ernzerhof, J. P. Perdew, and K. Burke, in *Density Functional Theory I,* Vol. 180 of *Topics in Current Chemistry,* edited by R. F. Nalewajski (Springer-Verlag, Berlin, 1996).
51. O. Gunnarsson and R. O. Jones, Phys. Rev. B **31**, 7588 (1985).
52. M. Levy, Phys. Rev. A **43**, 4637 (1991).
53. G. A. Baker, *Essentials of Padé approximants* (Academic Press, New York, 1975).
54. J. P. Perdew, M. Ernzerhof, and K. Burke, J. Chem. Phys. **105**, 9982 (1996).
55. J. P. Perdew and M. Ernzerhof, in *Electronic Density Functional Theory: Recent Progress and New Directions*, edited by J. F. Dobson, G. Vignale, and M. P. Das (Plenum, NY, 1997).
56. R. Sharp and G. Horton, Phys. Rev. **90**, 317 (1953).
57. J. Talman and W. Shadwick, Phys. Rev. A **14**, 36 (1976).
58. J. Krieger, Y. Li, and G. Iafrate, Phys. Rev. A **45**, 101 (1992).
59. J. Krieger, Y. Li, and G. Iafrate, Phys. Rev. A **46**, 5453 (1992).
60. J. Krieger, Y. Li, and G. Iafrate, in *Density Functional Theory,* edited by R. Dreizler and E. Gross (Plenum Press, New York, 1995).
61. T. Grabo and E.K.U. Gross, Chem. Phys. Lett. **240**, 141 (1995).
62. K. Burke, M. Ernzerhof, and J. P. Perdew, Chem. Phys. Lett. **265**, 115 (1997).
63. CADPAC6: The Cambridge Analytical Derivatives Package Issue 6.0 Cambridge (1995). A suite for quantum chemistry programs developed by

R.D. Amos, with contributions from I.L. Alberts, J.S. Andrews, S.M. Colwell, N.C. Handy, D. Jayatilaka, P.J. Knowles, R. Kobayashi, G.J. Laming, A.M. Lee, P.E. Maslen, C.W. Murray, P. Palmieri, J.E. Rice, J. Sanz, E.D. Simandiras, A.J. Stone, M.-D. Su, and D.J. Tozer.

64. F.W. Kutzler and G.S. Painter, Phys. Rev. Lett. **59**, 1285 (1987).
65. B.G. Johnson, P.M.W. Gill and J.A. Pople, J. Chem. Phys. **98**, 5612 (1993). Geometry of Be: V.E. Bondybey and J.H. English, J. Chem. Phys. **80**, 568 (1984); O_3: M.V. Rama Krishna and K.D. Jordan, Chem. Phys. Lett. **115**, 423 (1987); B_2: K.P. Huber and G. Herzberg, *Molecular Spectra and Molecular Structure IV: Constants of Diatomic Molecules*, (Van Nostrand Reinhold, New York, 1979).
66. J.A. Pople, M. Head-Gordon, D.J. Fox, K. Raghavachari, L.A. Curtiss, J. Chem. Phys. **90**, 5622 (1989); For Be, O_3, and B_2 see Ref. [65].

Density Polarization Functional Theory

Richard M. Martin

[a]Department of Physics and [b]Materials Research Laboratory, University of Illinois, Urbana, Illinois 61801

Abstract. At the center of density functional theory (DFT) are the proofs by Hohenberg and Kohn, which show that all properties of quantum many-body systems are functionals of the ground state density, and the Kohn-Sham construction, in which the exchange-correlation energy is a functional only of the density. DFT has been widely assumed to apply directly to the static dielectric properties of insulators. However, in 1995, Godby, Ghosez, and Godby pointed out that the assumptions of HK do not strictly apply to the case of a crystal in a finite electric field, since there is no ground state, and they argued that the description of intrinsic bulk dielectric phenomena in a crystal requires a functional of both the bulk density and the polarization. Here we summarize the status of recent work, especially a detailed exposition given elsewhere by the present author and G. Ortiz. The primary goal is to construct a density-polarization functional theory that will provide a fundamental basis for the theory of dielectrics, which is formulated in terms of polarization and electric fields. The consequences of the ideas presented here are: 1) it is essential to use polarization in order to describe the long wavelength limit; 2) physically meaningful changes in polarization can be derived directly from the wavefunction; and 3) DFT must be generalized to a density-polarization functional theory in order to fully describe the dielectric behavior of materials.

1 Introduction

The purpose of these lecture notes is to survey recent work that brings together two of the basic aspects of the theory of condensed matter: the description of dielectric phenomena [1–4] and density functional theory [5–8]. The consequence of this analysis is that in order to describe intrinsic bulk dielectric properties of an insulator in terms of only properties defined in the bulk, even in the static low frequency limit, the bulk density alone is not sufficient; *in general, the state of the many-electron quantum system is a functional not only of the bulk density, but also of the macroscopic polarization.* Since the exchange and correlation are in general functionals of the polarization as well as the density, this also leads to changes in the form of the Kohn-Sham (KS) [6] equations needed to describe polarized dielectrics. Since this is a controversial idea for a conference on density functional theory, it is important to review the ideas and state clearly the particular circumstances in which it is argued that the density is not sufficient. In the opinion of the author, this exercise is especially valuable because it requires a thorough analysis of the assumptions behind density functional theory, as well as the problems associated with the theory of dielectrics, which deepens our understanding of both areas. This summary is based primarily upon work with G. Ortiz which is described in more detail in a set of papers, [9–12] especially a paper by Martin and Ortiz to be published in 1997 which will be referred to here as MO [12].

The recent episode in this drama was initiated in 1995 by Gonze, Ghosez, and Godby [13] (GGG), who proposed that the case of an infinite periodic insulating solid in an electric field is an interesting example where the "original proof of Hohenberg and Kohn does not apply". They concluded that the state of the crystal in the presence of an "applied perturbing potential is not a unique functional of the periodic density change; it depends also on the change in the macroscopic polarization." The reason that the original proof of HK is argued to be invalid is because there is no lowest energy ground state for a system in the presence of a constant electric field, which decreases indefinitely. This is a well-known catastrophe that occurs for any Coulomb system whose particles are allowed to move in infinite space [14–17]. Since the HK arguments were cast in terms of ground state densities, at the very least this leads to the need to revisit the proofs of HK and KS, and to devise an approach that is clearly applicable to dielectric phenomena in crystals.

Furthermore, this argument connects directly to the basic theory of dielectrics. In general the state of an extended system in the presence of an electric field is a metastable state, and theoretical formulations which claim to be rigorous must take this into account. In fact, there is a long history of difficulty in defining the polarization in an infinite crystal, [18–24] and only recently has the formalism been developed to determine the macroscopic polarization directly from the wavefunction for the electrons [25,9,24,26,10]. Thus it is essential to review the theory of polarization in dielectrics to establish the formulation.

The proposal that the HK functional is intrinsically a function of the polarization has led to considerable controversy, [27–29,10,30,31] with claims [29] and counterclaims [10,30,32] about the nature of ground state correlations in insulators , and a discussion of the problems encountered in taking the infinite system limit within the framework of the Kohn-Sham (KS) theory [31]. This leads us to a reinvestigation of the nature of the fundamental quantity that determines the energy of the many-body electron system – the exchange-correlation hole.

A final comment is needed to set the stage for the present discussion. The present proposals are NOT a refutation of the HK analysis in cases where it should apply. In fact, one of the important results of the recent work is a clarification of the situations where it should apply. The primary goal of the new formulation is to describe the intrinsic bulk properties of extended insulating systems in terms of only the bulk electron density and the bulk polarization. The distinction from the original work of HK is that they consider the density of the *entire system including surfaces*, which leads to great difficulties in constructing a theory of the intrinsic bulk properties of macroscopic bodies that is manifestly independent of the size, shape, and state of the surface. The present analysis in terms of *bulk density and polarization* is a natural way to accomplish this goal [12].

The approach taken in this paper is to describe first the nature of polarization in insulators and to present the reasons why changes in polarization are measurable, physical quantities associated with the bulk of the body, which are independent of the density in the bulk. We then describe the modern theory of polarization which relates the changes in polarization directly to the bulk wavefunctions for the electrons. This is followed by a survey of density functional theory and the modifications that must be made in order to describe the intrinsic properties of dielectric phenomena in extended materials. Finally, we discuss the modified form of the KS equations, which involve a non-local operator in addition to the usual local potential.

2 Dielectric Theory and Polarization

2.1 Why Is Polarization Important, but Problematic ?

In the classical theory of dielectrics it is well known that it is preferable to write the electrostatic energies in terms of the polarization and electric fields, rather than charge density and potentials [1,2]. The energy is written as the energy of a reference state of the material plus changes due to external fields. The purpose of this section is to describe the formulation in terms of polarization, along with the corresponding expressions in terms of the density. The primary points are: 1) the difficulty in defining the desired quantity – the polarization – directly from the density, and 2) the need to have a formulation which makes it possible to determine changes in polarization directly from the quantum mechanical wavefunctions. This sets the stage for the modern quantum theory of polarization and the special role of polarization in the functionals that will be discussed later.

Although electric polarization plays a central role in the dielectric phenomena of matter, formulation of the polarization in terms of the nuclei and electrons that make up the material is often considered only in simplified models. If the material is modeled as a sum of neutral polarizable units, there is no fundamental problem: the polarization is the sum of the dipole moments per unit volume and each dipole moment is well defined. This is the model presented in Ashcroft and Mermin [3] (see Chapter 27) or Kittel [4] (see Chapter 13) and is the basis for the famous Clausius-Mossotti relations [2–4]. This is illustrated in Fig. 1, where is shown a collection of isolated units (molecules) each of which is polarized by external fields and microscopic fields due to the neighbors. The key feature of the model is that one can relate the dipole moment of a unit to its intrinsic properties *including* its response to local electric fields acting on the unit from all sources outside the unit.

However, a quantum mechanical view of extended matter is quite different, at least at first sight. All the electrons are equivalent and an electron

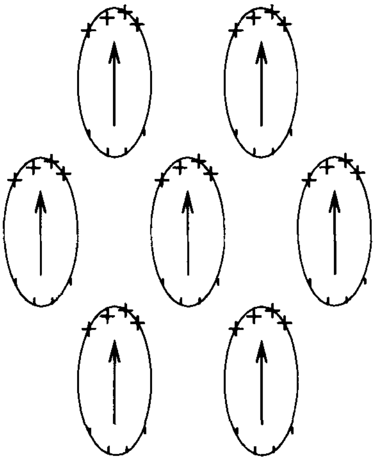


Fig. 1. Model for polarizable units or molecules in a dielectric assumed to be isolated with no flow of charge between the units, as in the Clausius-Mossotti model.

cannot be assigned to a given molecule or unit cell of a crystal. The charge density of the electrons is a continuous function of position and there is no way to uniquely "cut" the density and derive the dipole moment of a cell [20]i. This is illustrated in Fig. 2, which shows one possible division of the density in to "bulk" and "surfaces". However, any displacement of the dividing lines within the bulk region is an equally good definition. In a solid with no center of inversion, there is no unique choice [20].

The polarization $\mathbf{P}$ is the average value of the polarization density field $\mathbf{P}(\mathbf{r})$, which satisfies the equation

$$\nabla \cdot \mathbf{P}(\mathbf{r}) = -n(\mathbf{r}) \ , \tag{1}$$

where $n(\mathbf{r})$ is the charge density of the system. In a finite system, if $\mathbf{P}(\mathbf{r})$ vanishes outside the sample, it is straightforward to show that $\mathbf{P}$ is the dipole moment of $n(\mathbf{r})$,

$$\mathbf{P} = \frac{1}{\Omega}\int_{\Omega} d^3r\, \mathbf{P}(\mathbf{r}) = \frac{1}{\Omega}\int_{\Omega} d^3r\, \mathbf{r}\, n(\mathbf{r}) \ . \tag{2}$$

Thus, $\mathbf{P}$ is uniquely determined by $n(\mathbf{r})$ *and* the condition $\mathbf{P}(\mathbf{r}) = 0$ outside the volume of the sample Ω.

In an extended system, however, the interpretation of this integral suffers from difficulties. In particular, there has been much debate and discussion [18–24] about whether or not it is possible to uniquely define a "macroscopic bulk polarization" which is a property only of the interior of the material independent of surface termination. The difficulty originates in the factor of

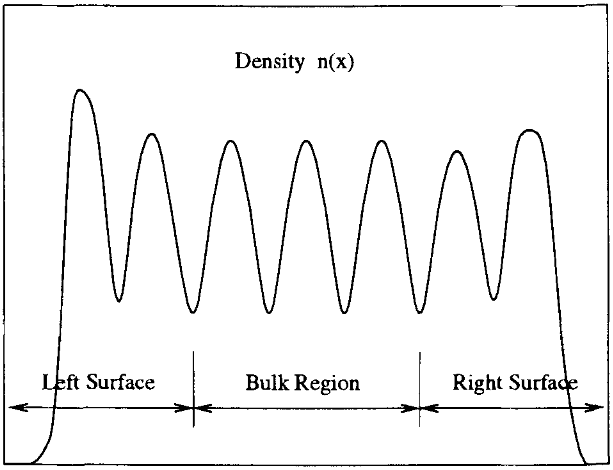


Fig. 2. Sketch of density in a finite crystal. The central section is representative of the bulk and is periodic. The vertical lines indicate one division of "bulk" and "surface"; however, this in completely arbitrary and it is impossible to draw unique boundaries which divide "bulk" and "surface" in a crystal which does not have inversion symmetry.

the position vector **r** in the definition above: a charge distribution on the boundary of the finite system can lead to a contribution to the polarization per unit volume **P** which does not vanish in the thermodynamic limit in which the finite system is taken to infinite size. This corresponds to the fact that the total polarization of a finite sample depends upon the charge state of the surfaces; therefore it cannot be a bulk property. It is impossible, strictly from knowledge of only the density in the bulk, to uniquely define an intrinsic bulk polarization or changes in the bulk polarization between two states of the material.

There is an additional quantity needed to determine the change in bulk polarization – the integrated polarization current that flows through the bulk. As shown years ago, [20] the change in polarization between any two physical states $\Delta\mathbf{P}$ is uniquely given by the change in dipole moment of a unit cell plus a surface integral which physically is the contribution of the polarization current which flows through the cell. This is illustrated in Fig. 3, which shows the change in density between two states including charge flow through each unit cell in the bulk. For any particular geometry of the sample there is a relation of the integrated current to the charge accumulated at the surfaces; however, it is much more convenient to consider the polarization current itself rather than the surface charges which depend upon the detailed geometry. The quantity $\Delta\mathbf{P}$ was shown to be independent of the choice of cell; however,

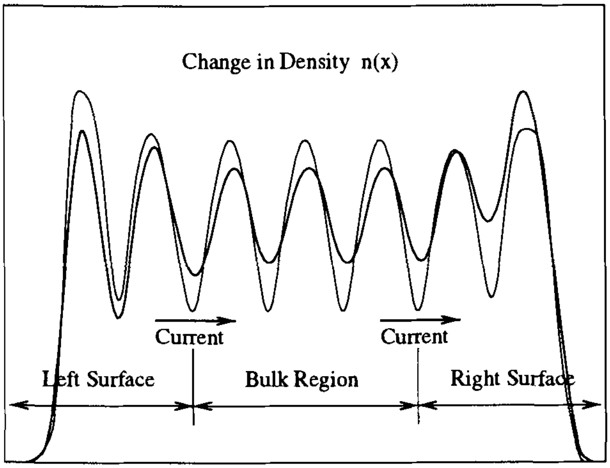


Fig. 3. Sketch of the change in density of Fig. 2 due to a change in the Hamiltonian, e.g., if atoms are displaced. In addition to changes in the periodic density in the bulk, there is a net current which flows through the bulk. Only if the current flow is taken into account is the total dipole moment of the finite crystal, and the average polarization given correctly.

no prescription for calculating the surface integral was given [20]. The key point for our purposes is that the integrated current is a physically measurable quantity. It is independent of the density within the bulk and it must be considered as a physical variable needed to fully specify the state of a dielectric. In fact, the formulation of the change in polarization as an integrated current is exactly the approach used in describing the response of a dielectric to transverse fields, [33] and is the basis for the new formulation [25,24,9,10] in terms of a Berry's phase, which is described in section IIc below. Let us now summarize some of the reasons why it is preferable to describe dielectrics in terms of polarization and electric field rather than density and potential [1,2]. The basic reason is that the electric field is the measurable, physical quantity which causes the displacement of charge, i.e., polarization in a region. The physical meaning of the polarization density and the reason why it is the appropriate variable in insulators has been stated clearly by Landau and Lifshitz [1]. The polarization density field defined by Eq. (1) embodies the property that in an insulator charge cannot flow over large distances. A change of $\mathbf{P}(\mathbf{r})$ in a localized region (i.e., $\delta\mathbf{P}(\mathbf{r}) = \mathbf{0}$ outside the region) corresponds to a change in charge density only in that region and it is simple to show using Eq. (1) that the integrated charge in the region is conserved, $\int d\mathbf{r}\ \delta n(\mathbf{r}) = 0$, i.e., local changes in $\mathbf{P}(\mathbf{r})$ describe local *polarization* of the charge density. Furthermore, the change in the electric dipole moment of the finite region is well-defined and is given by $\int d\mathbf{r}\ \mathbf{r}\ \delta n(\mathbf{r})$. In contrast, in metals the charge in a region can vary because charge can flow over macroscopic distances to fully screen any potential, and the density is the appropriate variable.

2.2 Total Energies in Coulomb Systems

In general the energies in Coulomb systems are very non-local functionals of the density, since the potentials and electric fields depend upon the charge density everywhere. In the KS theory [6] this is recognized in the Hartree term, which is explicitly left as the full non-local expression. We revisit this issue, however, because in the later sections it is argued that not only the Hartree term but also the exchange and correlation are in general expected to have very non-local dependence upon the density. If one considers a crystal as a large finite system and derives its properties by taking the thermodynamic limit, then the electric field inside an insulating crystal will depend in detail on the charges on the surface of this finite system, even in the thermodynamic limit. This can be seen by considering the total Coulomb energy of the charge density $n_T(\mathbf{r}) = n(\mathbf{r}) + n_+(\mathbf{r})$ (where we consider together the Coulomb energies of the electrons and the positive fixed charges, the nuclei, in order to have well-defined energies per particle in the thermodynamic limit). The potential terms involving Coulomb interactions can be written in either of two forms involving either potentials and densities or polarizations and electric fields:

$$E_{Coul} = \frac{1}{2}\int d\mathbf{r}\ V_{int}(\mathbf{r}) n_T(\mathbf{r}) = -\frac{1}{2}\int d\mathbf{r}\ \mathbf{E}(\mathbf{r}) \cdot \mathbf{P}(\mathbf{r}) \ , \tag{3}$$

where $\mathbf{E}(\mathbf{r})$ is the conventional symbol in dielectric theory for the total internal electric field, related to the potential by $\nabla V_{int}(\mathbf{r}) = -\mathbf{E}(\mathbf{r})$. Also the total internal field can be written as $\mathbf{E} = \mathbf{D} + \mathbf{E}_H$, where $\mathbf{D}(\mathbf{r})$ is the electric field due to external charges $\nabla v_{ext}(\mathbf{r}) = -\mathbf{D}(\mathbf{r})$, and $\mathbf{E}_H(\mathbf{r})$ is the electric field due to the electronic charge distribution, $\nabla V_H(\mathbf{r}) = -\mathbf{E}_H(\mathbf{r})$. This is consistent with the definitions in density functional theory where one considers the potential external to the electrons which is due to the nuclei; nevertheless the long range part of the "external" potential includes both nuclear and electronic contributions since the potential is convergent only if the system is neutral [34].

The key point for our purposes is that the total energy of an insulator in general involves long wavelength polarizations and electric fields. Of course, such terms are absent in a metal since there can be no average electric fields in equilibrium. This is discussed in great detail in MO [12] where we consider the Fourier components of the density $n(\mathbf{q})$ and the polarization $\mathbf{P}(\mathbf{q})$. In the $\mathbf{q} \to 0$ limit there is a term

$$E_{tot} = -\Omega\ \mathbf{E}_{mac} \cdot \mathbf{P}_{mac} + \dots \tag{4}$$

where Ω is the total volume of the system, $\mathbf{E}_{mac}$ is the macroscopic electric field and $\mathbf{P}_{mac}$ is the macroscopic polarization density. The other terms denoted by ... include all the short wavelength and non-Coulomb terms. This equation is useful because the polarization density $\mathbf{P}(\mathbf{q})$ is regular at the origin and $\mathbf{P}_{mac}$ is the limit, which is well defined. On the other hand, the density $n(\mathbf{q}) = \mathbf{q} \cdot \mathbf{P}(\mathbf{q})/q^2$ is non-analytic as $\mathbf{q} \to 0$. As long as one works at finite wavelength one can use either $n(\mathbf{q})$ or $\mathbf{P}(\mathbf{q})$; however, in the limit $\mathbf{q} \equiv 0$ the polarization is the macroscopic value whereas the density $n(\mathbf{q}) = 0$ is fixed by charge neutrality.

The conclusion of this section is that polarization is the key quantity in dielectric theory which makes possible a local, causal description of dielectric phenomena instead of the extremely non-local relations that occur if one attempts to describe the effects in terms of the density alone. Although the macroscopic polarization in a region is not determined solely by the density in that region, changes in the polarization are determined by the charge density in the region *plus integrated current flow through the region*. Finally, the total energy of the system depends upon the average polarization if there is an average electric field in the region.

2.3 Polarization, Berry's Phases, and Wannier-Like Functions

Recently [25,24,9], there has been a breakthrough providing a new approach for calculation of polarization in crystalline dielectrics. The new approach has been elegantly formulated using concepts of topology and differential geometry [35] leading to a Berry's phase [36] formulation. This was realized by King-Smith and Vanderbilt [25] who built upon earlier work of Thouless and coworkers [37,38]. A review has been given recently by Resta [24], a short summary by Martin and Ortiz [10] and the extension to interacting many-body systems has been given by Ortiz and Martin [9].

The authors of Refs. [25,24] consider insulators in the single-body, mean-field approximation, where the integrals over filled bands are over the complete Brillouin zone and they take advantage of periodicity in $\mathbf{k}$ space. (It has also been shown [9] that the polarization in a full many-body problem can be expressed by formulas which have similar structure; however, the independent-electron formulation is sufficient to illustrate the ideas). Consider the change in polarization when a parameter of the Hamiltonian, λ, is changed adiabatically (*e.g.*, when atoms are displaced), always requiring that the macroscopic electric field vanishes. The change in polarization is found from

$$\Delta \mathbf{P} = \int_0^1 d\lambda \, \frac{\partial \mathbf{P}}{\partial \lambda} \ . \tag{5}$$

The polarization involves also an integral over the Brillouin zone (*i.e.*, over all possible Bloch boundary conditions), and the wavevector $\mathbf{k}$ can be considered as a parameter in the gauge-transformed Hamiltonian, $H_{\mathbf{k}}(\lambda)$, whose eigenfunctions are the strictly periodic part of the Bloch functions $u^{\lambda}_{\mathbf{k}n}$, where n is the band index. Thus the final expression involves summation over occupied bands and integration over **k** and λ. King-Smith and Vanderbilt [25] showed that the adiabatic perturbation expressions for the change in polarization can be cast in terms of a Berry's phase [36] involving the phases of the eigenfunctions, with **k** and λ playing the role of the slowly changing parameters in the approach of Berry. The resulting expression for the electronic contribution to $\Delta\mathbf{P}$ of a spin-unpolarized system has the form [25]

$$\Delta\mathbf{P}_j = -|e|\frac{4}{(2\pi)^3}\,\mathrm{Im}\int_{BZ} d\mathbf{k}\int_0^1 d\lambda\,\sum_n\left\langle\frac{\partial u_{\mathbf{k}n}^{\lambda}}{\partial k_j}\middle|\frac{\partial u_{\mathbf{k}n}^{\lambda}}{\partial\lambda}\right\rangle\,, \tag{6}$$

where the sum is over all filled bands with j representing a particular crystal direction. By defining a reduced dimensionless vector $\xi = \frac{a}{2\pi}\mathbf{k}$ (where a has dimensions of length), the right hand side is easily shown to be a factor $2ea/volume$ (where the factor of 2 is included if it is assumed [25] that the system is non-magnetic and both up and down spin electrons contribute equally) multiplied by a dimensionless quantity which is gauge-independent and is precisely the Berry's phase. Using Stokes's theorem this integral can be converted into an integral along the closed path defined as the boundary of the region of (ξ, λ) space, as shown in Fig. 4.

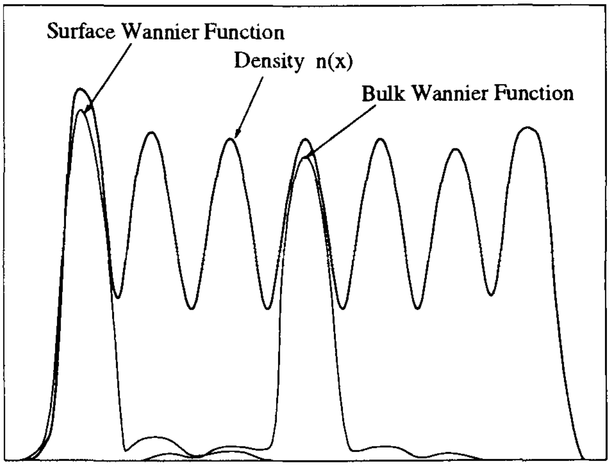


Fig. 4. The change in polarization is given by the surface integral over the area described by $\int d\lambda dk$, which can be converted into a line integral around the contour shown. By particular choice of gauge the contributions of the integrals over λ at the zone boundaries cancel, and one is left with only the integral over k at $\lambda = 1$ minus the integral at $\lambda = 0$.

Although the complete expression depends upon the path in (ξ, λ) space, [9] a choice of phases of the wavefunctions as a function of $\mathbf{k}$ so that $u^{\lambda}_{\mathbf{k}n}(\mathbf{r}) = \exp[i\mathbf{G} \cdot \mathbf{r}]\ u^{\lambda}_{\mathbf{k}+\mathbf{G},n}(\mathbf{r})$ (where $\mathbf{G}$ is a reciprocal lattice vector) leads to the simplest form dependent only upon the end-points $\lambda = 0$ and $\lambda = 1$, with, however, the disadvantage that the Berry's phase is uncertain by any multiple of 2π. The result including a factor of 2 for spin for $\Delta \mathbf{P}_j$ is [25,24,10]

$$\Delta \mathbf{P}_j = i\frac{-2|e|}{(2\pi)^3} \int_{BZ} d\mathbf{k} \sum_n \left[\langle u^{\lambda=1}_{\mathbf{k}n} \,| \partial_{k_j} u^{\lambda=1}_{\mathbf{k}n} \rangle - \langle u^{\lambda=0}_{\mathbf{k}n} \,| \partial_{k_j} u^{\lambda=0}_{\mathbf{k}n} \rangle \right]$$
$$+(\text{integer}) \times \frac{-2|e|}{area} \;, \qquad (7)$$

where the term *area* is the cell volume divided by the length of the unit cell in the direction j, which is the area of a cell surface perpendicular to the direction j.

King-Smith and Vanderbilt [25] showed that one can derive practical expressions for the Berry's phase in terms of differences of phases at discrete values of $\mathbf{k}$, which provide *unique* results even though they are derived from the phases of wavefunctions found by a standard diagonalization routine in which the phases are arbitrary. This paradoxical result stems from the fact that the only requirement on the phases of the periodic parts of the Bloch functions $u^{\lambda}_{\mathbf{k}n}(\mathbf{r})$ is that they obey the relation $u^{\lambda}_{\mathbf{k}n}(\mathbf{r}) = \exp[i\mathbf{G} \cdot \mathbf{r}]\ u^{\lambda}_{\mathbf{k}+\mathbf{G},n}(\mathbf{r})$. This can be easily accomplished merely by *defining* the function $u^{\lambda}_{\mathbf{k}n}(\mathbf{r})$ for $\mathbf{k}$ on one boundary of the Brillouin Zone to be the function actually found at the point $\mathbf{k} + \mathbf{G}$ on the opposite boundary multiplied by the phase factor $\exp[i\mathbf{G} \cdot \mathbf{r}]$. This provides an expression for the change in polarization between two states which is unique except for the addition of any integral multiple of 2π in the Berry's phase, as discussed before. For a continuous change, such as the induced polarization when atoms are displaced, uncertainties by possible integral multiples of 2π can be avoided by defining small enough changes (changes in the value of λ) where one always knows that the change in polarization is given by the smallest value, *i.e.*, with the integer equal zero. This effectively defines the *path* for the λ parameter.

Note that the geometric Berry's phase is nonzero only if the periodic functions $u^{\lambda}_{\mathbf{k}n}$ are complex; this occurs if there is no center of inversion, which is of course exactly the condition that there may be a non-zero polarization! Hence, *the change in macroscopic polarization between two different insulating states can be regarded as a measure of the phase difference between their initial and final many-body wavefunctions* (which, in all mean-field approaches, are Slater determinants of single-body functions $u^{\lambda}_{\mathbf{k}n}$). The new formulation is in fact related to the integrated current that flows through the insulator illustrated in Fig. 2, which is apparent since it involves the phases of the wavefunctions.

The integer multiples of 2π in the Berry's phase have an interesting interpretation. They correspond to the transport of an integer number of electrons

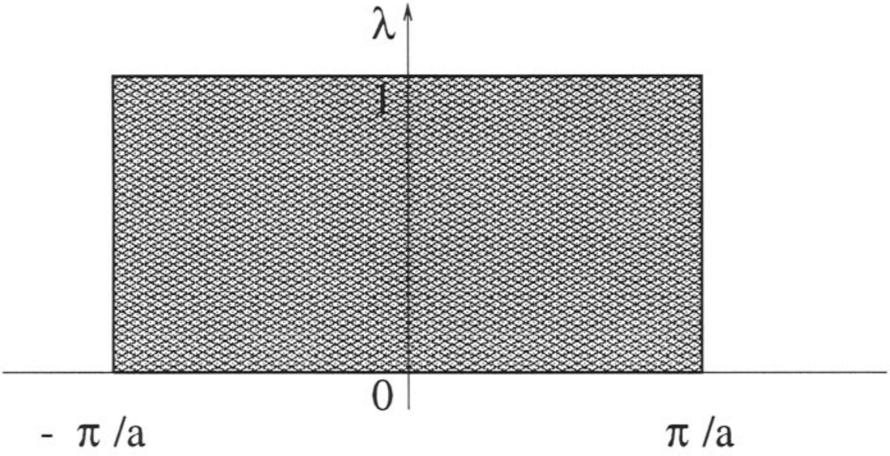


Fig. 5. Sketch of typical densities of localized Wannier-like functions which describe the occupied electronic states of an insulator. There is one function per cell per band; in the bulk the Wannier-like states are translations of the one shown $W(\mathbf{r}+\mathbf{R})$ and in the surface regions there are different functions. So long as the bulk is insulating, the surface functions are localized to the surface and decay exponentially into the bulk, independent of the nature and perfection of the surface. The change in polarization is given by the average change in dipole moment of each function per unit volume. In the thermodynamic limit the contribution of the surface is always negligible and the change in macroscopic polarization in given by the change in dipole moment of the bulk Wannier-like functions.

across the entire crystal, which of course leaves the bulk invariant. This is the aspect of the transport that was emphasized by Niu and Thouless [38] which leads to quantized charge transport in an insulator. In contrast, the work described here emphasizes the changes in the Berry's phase by fractions of 2π, which is the physical polarization of the crystal.

There is another (simpler to describe) interpretation of the expressions for the polarization. King-Smith and Vanderbilt [25] also showed that a physical understanding of the Berry's phase expressions can be gained by transforming to a basis of localized Wannier-like functions, which is possible in an insulator. The total density can be expressed in terms of overlapping densities of the Wannier-like functions,

$$n(\mathbf{r}) = \sum_{\mathbf{R}} \sum_{i} |W_i(\mathbf{r} - \mathbf{R})|^2 \ , \tag{8}$$

which is illustrated in Fig. 5. The important result for our purposes is that the change of polarization is simply the change in dipole moments of the Wannier-like functions,

$$\Delta \mathbf{P} = \frac{1}{\Omega_{cell}} \sum_i \int d^3 r \; \mathbf{r} \Delta |W_i(\mathbf{r})|^2 \; . \tag{9}$$

Even though electrons are indistinguishable, the fact that each one-body Wannier-like function can accept only one electron (of each spin) means that the charge of each is fixed and it acts as a polarizable object. The essential feature added by the expressions in terms of the geometric quantum phase, is the proof that the change in dipole moment of each neutral set of Wannier-like functions and nuclei in one unit cell is gauge-invariant and unique, despite the well-known non-uniqueness of Wannier functions themselves [39]. [Note: it has been pointed out to the author by D. Vanderbilt that a proof of uniqueness of the total dipole moments of the Wannier functions is contained in the 1962 article by Blount [39].] Therefore, this work provides a rigorous derivation of the Clausius-Mossotti-type polarizability models, like those discussed in textbooks (although the building blocks have a different meaning), [2,1,3,4] and provides explicit formulas for actual calculations. Furthermore, Nunes and Vanderbilt [40] have shown that the formalism of constrained "cut-off" Wannier-like functions developed for the "linear scaling methods" [41,42] can be used to carry out calculations of polarized crystals in the presence of macroscopic external electric fields.

In summary, changes in bulk polarization between two states of a crystal can be calculated directly from the wavefunctions in the bulk, using either the Berry's phase formulation in terms of Bloch eigenstates, or in terms of Wannier-like functions. The changes are physically measurable since they are related to the integrated charge flow due to bulk currents.

2.4 Further Issues in Dielectrics

For a homogeneous electric field which is finite (but weak enough that it does not cause breakdown) the state of an infinite crystal described by the dielectric theory is metastable [14–17]. The spectrum of energies is continuous and non-bounded below [16]. Nevertheless, the polarizability can be defined in a perturbation theory expansion about the zero-field ground state. Although the formulas are asymptotically convergent, [17] they are well-behaved at any finite order. Thus the ground state is stable about the minimum, but at any finite field the state described by dielectric theory is metastable, with a lifetime which is very long for typical field strengths well-below dielectric breakdown [17].

For independent electrons in an electric field, stable solutions for the long-lived states can be derived [17] by restricting the Hilbert space of wavefunctions to a subspace which is "slightly deformed" in a continuous manner from a zero-field band subspace [17]. Martin and Ortiz [12] have proposed that similar constraints make the full interacting many-body system stable, and have used this in the formulation of functionals of polarization at finite fields. However, to the knowledge of the author, such a formulation has never been proven rigorously for the full many-body problem. The logic for this assumption is very similar to the description of the insulating state by Kohn [43] in which the state is described for many non-interacting electrons, and the basic properties are proposed to carry over to the state of many interacting electrons assuming adiabatic continuation.

3 Aspects of Density Functional Theory

3.1 The Hohenberg-Kohn (HK) Theorems

Density functional theory [5–8] (DFT) is the foundation of much of current research on the electronic properties of condensed matter [44] and is widely used in studies of molecules and other finite systems [7,45]. The most celebrated tenet of DFT is that the ground state density $n_0(\mathbf{r})$ is in principle sufficient to determine all the properties of the many-body system of interacting electrons when there is time-reversal symmetry. In this section we review salient aspects of DFT as derived by Hohenberg and Kohn (HK) and others in recent decades [7,8]. We focus particularly upon aspects that relate to the role of polarization discussed in the present paper. The starting point of DFT is that the total ground state energy of a finite system of N non-relativistic spin-unpolarized electrons can be written in terms of an intrinsic Hamiltonian for the electrons plus the effects of an external potential, v_{ext},

$$E_{tot} \equiv \int d\mathbf{r}\ v_{ext}(\mathbf{r}) n_0(\mathbf{r}) + \langle \Psi_0 | -\frac{1}{2}\sum_{i=1}^{N} \nabla_i^2 | \Psi_0 \rangle + \langle \Psi_0 | V_{e-e} | \Psi_0 \rangle \ , \qquad (10)$$

where Ψ_0 is the ground state wavefunction of the many-electron system, the second term is the kinetic energy and the final term is the expectation value of the electron-electron interaction.

The reason why the density has a special role in the formulation of HK is that the total energy of the system depends upon effects external to the electron system only through the integral of the product of the density $n_0(\mathbf{r})$ and the external potential $v_{ext}(\mathbf{r})$ which is the first term in Eq. (10). Using only this information, HK showed that, in a finite system of a fixed number of electrons, two different external potentials differing by more than a constant cannot lead to the same ground state density. Thus, there is a one-to-one correspondence between the ground state density $n_0(\mathbf{r})$ and the external potential $v_{ext}(\mathbf{r})$, and the complete Hamiltonian is uniquely specified by the density, since all other terms in the Hamiltonian are known for a given type of particle. Furthermore, they showed that the construction of a DFT is possible in principle based upon the proof that there exists a universal functional of the density $F[n(\mathbf{r})]$, independent of $v_{ext}(\mathbf{r})$, such that the expression

$$E[n(\mathbf{r})] \equiv \int d\mathbf{r}\; v_{ext}(\mathbf{r})n(\mathbf{r}) + F[n(\mathbf{r})] \;, \tag{11}$$

has as its minimum value the correct ground state

$$E[n_0(\mathbf{r})] = E_{tot} = \min_n E[n(\mathbf{r})] \;. \tag{12}$$

The functional $F[n(\mathbf{r})] = T[n(\mathbf{r})] + \langle V_{e-e}\rangle$ is the HK functional for the internal energy, which is determined by the density alone.

3.2 The Kohn-Sham Ansatz

Essentially all actual DFT calculations are based upon the ansatz of Kohn and Sham (KS), [6] who separated the HK functional into three parts,

$$F[n(\mathbf{r})] \equiv T_{ind}[n(\mathbf{r})] + E_H[n(\mathbf{r})] + E_{xc}[n(\mathbf{r})] \;. \tag{13}$$

The first term is the kinetic energy of independent fermions with mass m_e (m_e=1 in Hartree atomic units) and density $n(\mathbf{r})$; the second is the average Coulomb Hartree energy, which is easily expressed in terms of the density

$$E_H[n(\mathbf{r})] = \frac{1}{2}\int\int d\mathbf{r}d\mathbf{r}'\; n(\mathbf{r})\frac{1}{|\mathbf{r}-\mathbf{r}'|}n(\mathbf{r}') \;; \tag{14}$$

and the third is defined to be the sum of all remaining terms which involve exchange and correlation

$$E_{xc}[n(\mathbf{r})] = T[n(\mathbf{r})] - T_{ind}[n(\mathbf{r})] + \langle V_{e-e}\rangle - E_H[n(\mathbf{r})] \;, \tag{15}$$

The ansatz made by KS is that the exact density can be represented by the density of non-interacting electrons (that is, interacting v-representable densities are assumed to be also non-interacting v-representable. Although this has never been proven or disproved in general, we will continue with the assumption that the density can be represented in this way.) If the ground state of the non-interacting system is non-degenerate, then the density is given by a sum over one-body occupied orbitals, $n(\mathbf{r}) = \sum_{i \in occupied} |\psi_i(\mathbf{r})|^2$, and the kinetic energy is found from the orbitals without explicitly constructing a functional of the density.

The genius of the KS ansatz is that the first two terms in Eq. (13) can be calculated exactly by the usual techniques of independent electron theory without ever explicitly constructing the functional, and the final exchange-correlation (x-c) term $E_{xc}[n(\mathbf{r})]$ is more amenable to physical interpretation and approximations than the original HK functional. The key points of the KS approach for our purposes are that 1) the exact x-c energy is expressed as a functional of the density alone, and 2) it is advantageous to formulate the theory in such a way that the functional relationships are as *local* as possible.

So long as the x-c energy E_{xc} is a functional only of the density $n(\mathbf{r})$, then KS showed that the (stationary) minimum energy solution can be found by the variational principle. Since the density is determined by the orbitals, the practical form for the variational equations is,

$$\frac{\delta E[n(\mathbf{r})]}{\delta\psi_i^*(\mathbf{r})} = \frac{\delta T_{ind}[\psi_i]}{\delta\psi_i^*(\mathbf{r})} + \int d\mathbf{r}' \left(\frac{\delta(E_H + E_{xc})}{\delta n(\mathbf{r}')} + v_{ext}(\mathbf{r}') \right) \frac{\delta n(\mathbf{r}')}{\delta\psi_i^*(\mathbf{r})} = 0 \ , \tag{16}$$

subject to the orthonormalization constraint $\langle\psi_i|\psi_j\rangle = \delta_{ij}$. This leads to the self-consistent independent-particle Schrödinger-like equations,

$$\hat{H}_{KS}\psi_i(\mathbf{r}) = \left(-\frac{1}{2}\nabla^2 + V_{KS}(\mathbf{r})\right)\psi_i(\mathbf{r}) = \epsilon_i\psi_i(\mathbf{r}) \ , \tag{17}$$

in which each non-interacting KS "electron" moves in an effective local potential $V_{KS}(\mathbf{r}) = v_{ext}(\mathbf{r}) + V_H(\mathbf{r}) + V_{xc}(\mathbf{r})$, where $V_H(\mathbf{r})$ and $V_{xc}(\mathbf{r})$ are the Hartree and x-c potentials. Thus the effective potential $V_{KS}(\mathbf{r})$ is a simple local function of position. Knowledge of the x-c functional $E_{xc}[n]$ would lead to the exact ground state energy and density of the many-body system in term of the simple non-interacting electron KS equations.

4 Exchange-Correlation Hole

The exchange-correlation (x-c) energy is the key functional in the KS approach that takes into account the fermion character and the correlations among the electrons: calculations using this approach are accurate only to the extent that the functional accurately captures the essential features of exchange and correlation in the actual system being considered. In this section we summarize ideas on the nature of exchange and correlation, the definition of the x-c energy within the KS approach, and expected dependence upon the density. The main purpose of this section is to show that in an insulator one *expects* a very-non local dependence of the x-c energy upon the density, much more non-local than in a metal. These arguments are given in more detail in MO and are also discussed in a simpler physical context in a comment [11] and reply [32] published in Physical Review Letters this year.

The $E_{xc}[n(\mathbf{r})]$ functional is directly related to the ground state wavefunction, at least in principle, by the Levy constrained search method [46,47] in which one searches over all antisymmetric, normalized N-particle functions Ψ which yield the prescribed density $n(\mathbf{r})$ ($n(\mathbf{r}) = \mathrm{N} \sum_\sigma \int d\mathbf{x}^{\mathrm{N}-1}|\Psi(\mathbf{r},\sigma,\mathbf{x}^{\mathrm{N}-1})|^2$ is assumed to be only N-representable; $\mathbf{x} = (\mathbf{r},\sigma)$, and σ represents a spin variable), and minimizes $\langle -\frac{1}{2}\sum_{i=1}^{\mathrm{N}} \nabla_i^2 + V_{e-e}\rangle$, i.e.,

$$E_{xc}[n(\mathbf{r})] = \min_{\Psi\to n(\mathbf{r})} \langle\Psi| -\frac{1}{2}\sum_{i=1}^{\mathrm{N}} \nabla_i^2 + V_{e-e}|\Psi\rangle - T_{ind}[n(\mathbf{r})] - E_H[n(\mathbf{r})] \ , \tag{18}$$

where $T_{ind}[n(\mathbf{r})]$ is the independent particle kinetic energy which is a functional of the density [6]. The x-c energy can also be formulated [48–50,7] in terms of potential energies only by a coupling constant integration. In this approach E_{xc} is given directly by integrating the derivative of the total energy with respect to the interaction strength λe^2 using the Hellmann-Feynman theorem. This is done by considering the set of Hamiltonians H_λ with the electron-electron interaction scaled by a factor of λ which increases from 0 to 1, zero coupling strength representing the non-interacting KS electrons. The ground state density is required to remain constant as λ is varied, which is accomplished by varying $v_{ext}(\lambda)$. The resulting expression is [7]

$$E_{xc}[n(\mathbf{r})] = \frac{1}{2}\int d\mathbf{r}\; n(\mathbf{r}) \int d\mathbf{u}\; \frac{1}{|\mathbf{u}|} n_{xc}(\mathbf{r}, \mathbf{r}+\mathbf{u}) \ , \tag{19}$$

where

$$n_{xc}(\mathbf{r}, \mathbf{r}+\mathbf{u}) = \int_0^1 d\lambda \; n_{xc,\lambda}(\mathbf{r}, \mathbf{r}+\mathbf{u}) =$$
$$\frac{1}{n(\mathbf{r})} \int_0^1 d\lambda \left(\langle \Psi_0^\lambda | \delta\hat{n}(\mathbf{r}) \delta\hat{n}(\mathbf{r}+\mathbf{u}) | \Psi_0^\lambda \rangle - n(\mathbf{r}) \delta(\mathbf{u}) \right) , \tag{20}$$

and $\delta\hat{n}(\mathbf{r}) = \hat{n}(\mathbf{r}) - n(\mathbf{r})$ is the density fluctuation operator. Note that $n_{xc}(\mathbf{r}, \mathbf{r}+\mathbf{u})$ is a correlation function giving the probability of finding an electron at $\mathbf{r}+\mathbf{u}$ if there is one at $\mathbf{r}$; however, it is not the x-c hole at the actual value of e^2 but is the average of the correlation function defined by the right-hand side of Eq. (20) for e^2 ranging from 0 to its actual value with the density $n(\mathbf{r})$ kept constant. The wavefunction Ψ_0^λ is the ground state of the Hamiltonian H_λ with density $n(\mathbf{r})$.

The dependence of the x-c hole $n_{xc}(\mathbf{r}, \mathbf{r}+\mathbf{u})$ and the energy $E_{xc}[n]$ upon the density $n(\mathbf{r})$ can arise from two sources. One is the implicit dependence upon the density through the correlations and the Pauli exclusion principle among the electrons in the many-body wavefunction in Eq. (20) as a function of λ. This must be calculated from the many-body function over the range of the x-c hole, i.e., the range of $\mathbf{u}$ over which $n_{xc}(\mathbf{r}, \mathbf{r}+\mathbf{u})$ must be explicitly considered.

In addition, however, $E_{xc}[n]$ may depend upon the density through effects of potentials (or electric fields) generated by the density. The key point for our purposes [12] is that in an insulator the properties at any point can depend upon the density at distant points through long-distance electric fields or potentials. The longest range effects are due to slowly varying long-range electric fields. Of course, the wavefunction of the many-body electron system is affected by an electric field, and so we expect the shape of the average x-c hole $n_{xc}(\mathbf{r}, \mathbf{r}+\mathbf{u})$ to be modified (*polarized*), with consequent changes in the x-c energy due to changes in the angle-averaged dependence of the hole upon the distance $|\mathbf{u}|$. Thus we expect a very *non-local* dependence of $E_{xc}[n]$ upon the density. Note that this does *not* require the x-c hole itself to have any anomalous long-range behavior as was proposed by Resta [29] (see the comment [11]). There are no such effects in a metal since static electric fields must vanish in its interior.

In summary, we have established that in an insulator the exact x-c energy *is expected to be an extremely non-local functional of the density.* This undesirable situation is the basis for introducing a different way to approach the problem in terms of *both the density and the polarization* which is discussed below. The reasoning is that the x-c hole is expected to be a function of the polarization of the wavefunction in the local region, giving a more local description.

5 Density-Polarization Functional Theory

The important prerequisites for formulating the *Density-Polarization Functional Theory* (DPFT) are now in place. We will give here only a summary with the details left to the more complete discussion in MO [12]. DPFT was first proposed by GGG [13] and the present description is consistent with their initial description. Note that one result is a derivation of important cases where the original analysis of HK applies directly to the bulk, i.e., the bulk density is sufficient, even though the description of an insulator in the general case requires a functional not only of the density but also the macroscopic polarization.

We consider the case of a crystal and we wish to formulate the theory in a way that is valid in the thermodynamic limit. One procedure is to consider finite wavevectors, i.e., spatially varying density $n(\mathbf{r})$ and polarization density field $\mathbf{P}(\mathbf{r})$. As long as there is no average polarization, then it is straightforward to work with either $n(\mathbf{r})$ or $\mathbf{P}(\mathbf{r})$, since they are related by Eq. (1). However, if we take the limit of long-wavelength variations, $\mathbf{P}(\mathbf{r})$ approaches the macroscopic polarization $\mathbf{P}_{mac}$ whereas the density is non-analytic since the average value is fixed to be zero by charge neutrality. It is in the limit that the formulation in terms of both $n(\mathbf{r})$ and $\mathbf{P}_{mac}$ is essential.

The most important point is that in general in dielectrics there is a term in the energy which involves the macroscopic polarization $\mathbf{P}_{mac}$ (see Eq. (4)), in addition to other terms which involve the short wavelength variations of the density. Using exactly the same reasoning as in the original HK arguments, it follows that the internal energy of the electrons must be a functional of the polarization. As discussed in MO, section V, one must be careful to treat properly all the Coulomb terms in the limit and one must restrict the wavefunction to a space in which the solution is stable. The approach of MO leads to the formulation in which the wavefunctions Ψ must be the lowest energy states of the *auxiliary* system Hamiltonian

$$\widetilde{H}_{\mathrm{N}} = -\frac{1}{2}\sum_{i=1}^{\mathrm{N}} \nabla_i^2 + \widetilde{V}_{ext} + V_{e-e} - \Omega\, \mathbf{E}_{mac} \cdot \widehat{\mathbf{P}}_{mac} \ , \tag{21}$$

for a given *generalized Bloch condition* $\mathbf{k}$ on the boundary of the domain Ω, and where $\widetilde{V}_{ext} = \sum_{\mathbf{G}\neq\mathbf{0}} V_{ext}(\mathbf{G}) \exp(i\mathbf{G}\cdot\mathbf{r})$. In Eq. (21) $\widehat{\mathbf{P}}_{mac}$ represents a "polarization operator" that is bounded and whose precise definition depends upon the order $(\mathcal{O}(\mathbf{E}_{mac}^{\mathcal{N}}))$ of the "deformed" subspaces generated by the recursive algorithm of Nenciu [17]. This procedure defines the map $(\widetilde{V}_{ext}; \mathbf{E}_{mac}) \rightarrow \Psi$, and where Ψ represents the *metastable* solution.

In order to fully generalize the HK theorems to the present problem, we must show that there is a map $(n(\mathbf{r}); \mathbf{P}_{mac}) \rightarrow (\widetilde{V}_{ext}; \mathbf{E}_{mac})$, which tells us that knowledge of $n(\mathbf{r})$ and $\mathbf{P}_{mac}$ completely determines the "Hamiltonian" of the system. We must also show that the "ground state" expectation value of *any* observable is a unique functional of the periodic density and a function of the macroscopic polarization. This is accomplished [12] by showing that, for each Bloch boundary condition $\mathbf{k}$, two different external potentials $\widetilde{V}_{ext}$ differing by more than a constant lead to different "ground states" $\Psi_{\mathbf{k}}$, and that two different macroscopic fields $\mathbf{E}_{mac}$ lead also to two different $\Psi_{\mathbf{k}}$'s using *reductio ad absurdum*. Consider the ground states $\Psi_{\mathbf{k}}$ and $\Psi'_{\mathbf{k}}$ of two different Hamiltonians $\widetilde{H}_{\mathrm{N}}$ and $\widetilde{H}'_{\mathrm{N}}$, which involve respectively external potentials $\widetilde{V}_{ext}$ and $\widetilde{V}'_{ext}$ and macroscopic fields $\mathbf{E}_{mac}$ and $\mathbf{E}'_{mac}$. We assume that $\Psi_{\mathbf{k}} \neq \Psi'_{\mathbf{k}}$, but that $n_{\mathbf{k}}(\mathbf{r}) = n'_{\mathbf{k}}(\mathbf{r})$ and $\mathbf{P}_{\mathbf{k},mac} = \mathbf{P}'_{\mathbf{k},mac}$. Then,

$$\widetilde{E}_0(\mathbf{k}) = \langle \Psi_{\mathbf{k}} | \widetilde{H}_{\mathrm{N}} | \Psi_{\mathbf{k}} \rangle < \langle \Psi'_{\mathbf{k}} | \widetilde{H}_{\mathrm{N}} | \Psi'_{\mathbf{k}} \rangle \tag{22}$$

implies

$$\begin{aligned} \widetilde{E}_0(\mathbf{k}) < \widetilde{E}'_0(\mathbf{k}) &+ \int d\mathbf{r}\; n_{\mathbf{k}}(\mathbf{r})\; (\tilde{v}_{ext}(\mathbf{r}) - \tilde{v}'_{ext}(\mathbf{r})) \\ &+ \Omega\; (\mathbf{E}'_{mac} - \mathbf{E}_{mac}) \cdot \mathbf{P}_{\mathbf{k},mac}. \end{aligned} \tag{23}$$

But using the corresponding argument for $\widetilde{E}'_0(\mathbf{k})$

$$\begin{aligned} \widetilde{E}'_0(\mathbf{k}) < \widetilde{E}_0(\mathbf{k}) &- \int d\mathbf{r}\; n_{\mathbf{k}}(\mathbf{r})\; (\tilde{v}_{ext}(\mathbf{r}) - \tilde{v}'_{ext}(\mathbf{r})) \\ &- \Omega\; (\mathbf{E}'_{mac} - \mathbf{E}_{mac}) \cdot \mathbf{P}_{\mathbf{k},mac}\ , \end{aligned} \tag{24}$$

which leads to the obvious contradiction

$$\widetilde{E}'_0(\mathbf{k}) + \widetilde{E}_0(\mathbf{k}) < \widetilde{E}'_0(\mathbf{k}) + \widetilde{E}_0(\mathbf{k})\ , \tag{25}$$

meaning that $(n_{\mathbf{k}}(\mathbf{r}); \mathbf{P}_{\mathbf{k},mac}) \rightarrow \Psi_{\mathbf{k}}$ for each boundary condition specified by $\mathbf{k}$. Finally, we can derive the generalization of the HK theorem by noting that the density $n(\mathbf{r})$ and the polarization $\mathbf{P}_{mac}$ are averages over the boundary conditions $\mathbf{k} \in$ 1BZ of Ω, which are well defined [9] in the limit of large Ω. Similarly, other expectation values are averages over the many-body wavefunction as a function of the boundary conditions.

The key point of the reasoning leading to the density-polarization theory is that external effects couple to both density and polarization; so the proof follows that no two different combinations of periodic density *and* macroscopic polarization can lead to the same external potential *and* electric field. Therefore, the periodic density *and* macroscopic polarization completely determine all properties of the system.

There is, however, a special case of great importance which should be considered. If the macroscopic electric field is zero, then the energy is independent of the macroscopic polarization, and it follows by simple reasoning that the original proof of HK applies: the density determines the short wavelength components of the potential and therefore all aspects of the system. This is understandable since the requirement that the macroscopic electric field vanishes really is the requirement that there are no effects in the bulk from the surface, since the only long range effects of any aspects extrinsic to the bulk are due to the macroscopic electric field.

It follows therefore that the case in which the polarization is essential in the functional is for *changes in the polarization from the value it has with zero macroscopic electric field*, $\delta \mathbf{P}_{mac} \equiv \mathbf{P}_{mac} - \mathbf{P}^0_{mac}$. Finally, the generalization of the HK functional F (Eq. (11)) to define a universal functional applicable to systems with polarization $\delta \mathbf{P}_{mac} \neq \mathbf{0}$ must have the form $\widetilde{F}[n; \delta \mathbf{P}_{mac}]$. The total energy functional can be written

$$E[n; \mathbf{P}_{mac}] = -\Omega\ \mathbf{E}_{mac} \cdot \mathbf{P}_{mac} + E'_{ext}[n] + \widetilde{F}[n; \delta \mathbf{P}_{mac}] \ , \tag{26}$$

where the first term involves the entire polarization and not only the change.

5.1 Example of a Case Where Two Potentials Lead to the Same Density

In this section we give a very brief discussion of an explicit example in which two potentials lead to the same periodic density in a crystal. The specific example is given in MO, section IV, and it is related to the example of GGG. The case is a crystal in a macroscopic electric field, treated up to second order in perturbation theory, i.e., linear response. This is an explicit counterexample to the naive interpretation of the HK theorem, and it is also sufficient to show that the modification to the HK theorem is *not* simply due to long-range Hartree-type terms; there are fundamental changes in the quantum mechanical behavior of the many-electron system as described in MO.

Consider the linear response of a crystal to externally applied fields $\mathbf{D}(\mathbf{q}+\mathbf{G})$, where $\mathbf{G}$ is any of the reciprocal lattice vectors. To linear order the response is conventionally defined in terms of the changes in the polarization density $\delta \mathbf{P}(\mathbf{q}+\mathbf{G})$. We do not need to consider any other Fourier components such as $\delta \mathbf{P}(2\mathbf{q}+\mathbf{G})$, which are higher order, in the externally applied field. The linear response is given by [33,51]

$$\delta \mathbf{D} = \delta \mathbf{E} + 4\pi\ \delta \mathbf{P} = \epsilon\ \delta \mathbf{E} \ , \tag{27}$$

or

$$4\pi\,\delta\mathbf{P} = \delta\mathbf{D} - \delta\mathbf{E} = (\epsilon - 1\!\!1)\delta\mathbf{E} = (1\!\!1 - \epsilon^{-1})\delta\mathbf{D}\ , \tag{28}$$

where ϵ is a matrix in $\mathbf{q}+\mathbf{G}, \mathbf{q}+\mathbf{G}'$. Our goal is to find an external field which can be applied with no change in the periodic part of the charge density, density $\delta n(\mathbf{G}) = 0$. Consider the long-wavelength limit where each vector has a macroscopic part $\delta\mathbf{P}(\mathbf{0}) = \delta\mathbf{P}_{mac}$ and periodic parts $\delta\mathbf{P}(\mathbf{G})$, $\mathbf{G} \neq \mathbf{0}$, and similarly for $\delta\mathbf{D}$ and $\delta\mathbf{E}$. If we require $\delta\mathbf{P}(\mathbf{G}) = \mathbf{0}$ for $\mathbf{G} \neq 0$, then $\delta n(\mathbf{G}) = \mathbf{G}\cdot\delta\mathbf{P}(\mathbf{G}) = 0$. Since the average value $n(\mathbf{0})$ is fixed by charge neutrality, there is no change in any Fourier component of $n(\mathbf{G})$ if

$$4\pi\,\delta\mathbf{P}(\mathbf{G}) = \sum_{\mathbf{G}'}(1\!\!1 - \epsilon^{-1})_{\mathbf{G},\mathbf{G}'}\ \delta\mathbf{D}(\mathbf{G}') = \mathbf{0}\ \ , \mathbf{G} \neq \mathbf{0}\ , \tag{29}$$

or

$$\delta\mathbf{D}(\mathbf{G}) = (1\!\!1 - \epsilon^{-1})^{-1}_{\mathbf{G},0}\ 4\pi\,\delta\mathbf{P}(0)\ . \tag{30}$$

The last equation explicitly defines the change in external field $\delta\mathbf{D}(\mathbf{G})$ that must be applied to keep all periodic parts of the charge density constant. Note that it involves both an average applied macroscopic field $\delta\mathbf{D}(\mathbf{0})$ and periodic parts $\delta\mathbf{D}(\mathbf{G})$, $\mathbf{G} \neq \mathbf{0}$. Thus, we have explicitly constructed a case where different external potentials lead to the same density, but different macroscopic polarizations $\mathbf{P}_{mac} = \mathbf{P}(\mathbf{0}) = \mathbf{P}(\mathbf{G} \equiv \mathbf{0})$.

It is also shown in MO that analysis of the energy of the dielectric demonstrates that the effect is not simply a change in the average Hartree Coulomb energies, but also the kinetic energy and the exchange-correlation energy of the electrons also are different for the two cases, even though the density is the same. Therefore, we have arrived at the desired result that in the exact KS theory for an insulating crystal *both* $T_{ind}[n(\mathbf{r}); \delta\mathbf{P}_{mac}]$ and $E_{xc}[n(\mathbf{r}); \delta\mathbf{P}_{mac}]$ must in general be functions of the polarization.

5.2 Generalized Kohn-Sham (KS) Equations

The next step in the theory is to generalize the Kohn-Sham (KS) approach to include the effects of polarization. Again, we will only summarize results. In the original theory the KS equations are derived by minimizing the energy with respect to variations of the density, subject to the constraint of conservation of the total charge. It therefore follows that in the generalized theory in which the polarization is an independent variable, the complete variational equations must have the form:

$$\frac{\delta E[n(\mathbf{r}); \delta\mathbf{P}_{mac}]}{\delta n(\mathbf{r})} = 0;\ \frac{\partial E[n(\mathbf{r}); \delta\mathbf{P}_{mac}]}{\partial\mathbf{P}_{mac}} = \mathbf{0}\ . \tag{31}$$

In the case where the macroscopic electric field is zero, the solution is a true ground state; however, just as we have seen before, if there is a nonzero homogeneous electric field, then the variational minimum must be found within a constrained space.

In the KS approach the density is determined by the independent electron KS orbitals $\psi_{i,\mathbf{k}}(\mathbf{r})$. Therefore the generalization requires that the polarization also be expressed in terms of the KS orbitals, and one approach is to use the dependence upon the Bloch boundary conditions as was described in the section on polarization. This is possible at least for the lowest order effects in terms of the polarization, which is all that is need for the variational analysis [12]. For a crystal one can always write $\psi_{i,\mathbf{k}}(\mathbf{r}) = \exp[i\mathbf{k}\cdot\mathbf{r}]\ \phi_{i,\mathbf{k}}(\mathbf{r})$, with $\mathbf{k}$ considered a continuous parameter and $\phi_{i,\mathbf{k}}(\mathbf{r}+\mathbf{R}) = \phi_{i,\mathbf{k}}(\mathbf{r})$, where $\mathbf{R}$ is a periodic lattice vector. In terms of the orbitals the Euler-Lagrange equations can be written

$$\begin{aligned}\frac{\delta E[n(\mathbf{r});\delta\mathbf{P}_{mac}]}{\delta\phi^*_{i,\mathbf{k}}(\mathbf{r})} &= \frac{\delta T_{ind}[\phi_{i,\mathbf{k}}]}{\delta\phi^*_{i,\mathbf{k}}(\mathbf{r})} \\ &+ \int d\mathbf{r}'\left(\frac{\delta(E_H+E_{xc})}{\delta n(\mathbf{r}')} + \tilde{v}_{ext}(\mathbf{r}')\right)\frac{\delta n(\mathbf{r}')}{\delta\phi^*_{i,\mathbf{k}}(\mathbf{r})} \\ &+ \left(-\Omega\ \mathbf{E}_{mac} + \frac{\partial E_{xc}}{\partial\mathbf{P}_{mac}}\right)\cdot\frac{\delta\mathbf{P}_{mac}}{\delta\phi^*_{i,\mathbf{k}}(\mathbf{r})} = 0\ ,\end{aligned} \tag{32}$$

subject to the constraint $\langle\phi_{i,\mathbf{k}}|\phi_{j,\mathbf{k}}\rangle = \delta_{ij}$. In this expression the electric field $\mathbf{E}_{mac}$ is the derivative of the total average Coulomb energy with respect to the macroscopic polarization, and the x-c "electric field" [13] $\frac{\partial E_{xc}}{\partial\mathbf{P}_{mac}}$ appears explicitly in the modified KS equations; however, the dependence of the kinetic energy upon the polarization is considered as an explicit functional of the orbitals, in analogy with the usual KS equations in which the independent particle kinetic energy is considered a functional of the orbitals rather than the density.

In these modified KS equations there is a new term which *cannot* be expressed as a local potential and its effect is not represented by any possible gradient approximation [7,8,52]. Instead it is proportional to the functional derivative of $\mathbf{P}_{mac}$ with respect to each orbital $\phi^*_{i,\mathbf{k}}(\mathbf{r})$. To the knowledge of the author no completely general expression for $\mathbf{P}_{mac}$ as a function of the orbitals has yet been found; however, for the case of zero macroscopic field and to linear order in an adiabatic perturbation expansion, [25,24,9,10] $\delta\mathbf{P}_{mac}$ can be expressed as a Berry's phase of the non-interacting (non-degenerate) many-body wavefunction (which is a product of Slater determinants of $\phi_{i,\mathbf{k}}$'s, one for each spin component) [25,24,9]. It follows that the exact KS equations contain a term of the form [13,28]

$$\frac{\delta\mathbf{P}_{mac}}{\partial\phi^*_{i,\mathbf{k}}(\mathbf{r})} = \frac{i}{(2\pi)^3}\ \nabla_{\mathbf{k}}\ \phi_{i,\mathbf{k}}(\mathbf{r})\ . \tag{33}$$

Thus the modified form of the KS equations for the orbitals is

$$\hat{H}_{KS}(\mathbf{k})\phi_{i,\mathbf{k}}(\mathbf{r}) = \left(-\frac{1}{2}(\nabla+i\mathbf{k})^2 + V_{KS}(\mathbf{r})\right.$$

$$- \Omega \left(\mathbf{E}_{mac} + \mathbf{E}_{xc,mac}\right) \frac{i}{(2\pi)^3} \nabla_{\mathbf{k}} \Big) \phi_{i,\mathbf{k}}(\mathbf{r}) = \epsilon_{i,\mathbf{k}} \, \phi_{i,\mathbf{k}}(\mathbf{r}) \, . \tag{34}$$

This is the same as the usual form of the equations, Eq. (17), explicitly written in terms of $\phi_{i,\mathbf{k}}(\mathbf{r})$, except that there is an additional term involving the derivative of the periodic orbitals with respect to the wavevector **k**.

This generalized form of the KS equations has been proposed by GGG (in linear response theory) and by MO. Furthermore, MO have proposed that the polarization dependence of the x-c energy should be derivable from the properties of the x-c hole in a polarized material, which is closely related to the changes in the x-c hole in the presence of electric fields. Aulbur, et. al., [28] have carried out calculations of dielectric response with this form of the equations using an a approximate expression in which a term involving polarization is added to the usual local density approximation.

6 Future Challenges and Conclusions

In this paper, we have argued that the usual DFT theory of Hohenberg and Kohn (HK), as well as the independent particle ansatz of Kohn and Sham (KS), must be generalized in order to describe polarized states of condensed matter. The goal of the present work is to describe the intrinsic properties of the bulk in terms only of information in the bulk region, with no additional information needed from extrinsic regions such as the surface. In this case, it is clear from previous work on dielectrics that the density is not sufficient, and that one way to describe the desired effects is by using the polarization. We have argued that: 1) changes in the bulk polarization can be found as intrinsic properties of the bulk; 2) they are determined by measurable physical integrated currents; 3) they are independent of the density in the bulk region; and 4) the total energy of the system involves the average polarization if there are macroscopic electric fields. It follows from reasoning analogous to the original proofs of HK that the internal energy must be a functional of both the bulk density and the bulk macroscopic polarization. We note that the effects cannot be incorporated in any theory based upon the metallic state or in any gradient approximations [52].

We have also argued that if one insists upon maintaining a description in terms of the density then one must include extremely non-local effects into the functional for exchange and correlation in order to have an exact Kohn-Sham theory, i.e., to capture all the aspects of exchange and correlation. This is illustrated by the analysis of a recent paper by Gonze, Ghosez, and Godby [31] which argues that in the usual KS approach in terms of densities, one must arrive at "exchange-correlation electric fields" which extend through the material and depend upon the size and shape of the macroscopic body.

Nevertheless, the proof of all aspects of the density-polarization theory is yet to be derived. The key problems that are unresolved are the ways to treat the constrained space required to stabilize the metastable solutions which exist in the presence of electric fields. Furthermore expressions for the polarization valid to all orders are at present unknown (private communication, G. Ortiz). To date there has been no derivation of the polarization dependence of the HK or the KS functionals from the fundamental dependence of the electronic exchange and correlation, that has been proposed by MO to be the essential problem in making a useful generalized Kohn-Sham approach. Future investigations can add to our understanding of the nature of correlations in many-body problems.

There are other challenges and opportunities as well. The theory described here must have a relation to another well-known generalization of DFT - current-density functional theory [53,8,54]. This is because the change in polarization is an integrated current in the special case of the *polarization current* in an insulator. The Berry's phase approach explicitly expresses these polarization currents in terms of the adiabatic evolution of the ground state of the insulator. Making the connection to current-density functional theory could open new avenues for interpretation and understanding of the dielectric phenomena of condensed matter. These are just examples of possible future challenges and opportunities for future understanding and practical calculation methods for the many-electron problem in condensed matter.

7 Acknowledgements

This work was supported by the U.S. NSF grant no. DMR-94-22496 and the Department of Energy under grant DEFG02-96-ER45439. The initial phases of this work were done at the Max Planck Institut für Festkörperforschung in Stuttgart supported by the Alexander von Humboldt Stiftung. All aspects of this work were done in collaboration with Gerardo Ortiz. I gratefully acknowledge conversations with O. K. Andersen, S. Baroni, D. Ceperley, C. Filippi, H. Kim, E. Koch, W. Kohn, P. Ghosez, R. Godby, X. Gonze, O. E. Gunnarsson, R. Resta, J. Shumway, I. Souza, and D. Vanderbilt.

References

1. L. D. Landau and E. M. Lifshitz, *Electrodynamics of Continuous Media* (Pergamon Press, Oxford, England, 1960).
2. J. D. Jackson, *Classical Electrodynamics* (Wiley, New York, 1962).
3. N. Ashcroft and N. Mermin, *Solid State Physics* (W.B. Saunders Company, New York, 1976).
4. C. Kittel, *Introduction to Solid State Physics* (John Wiley and Sons, New York, 1996).
5. P. Hohenberg and W. Kohn, Phys. Rev. **136**, B864 (1964).
6. W. Kohn and L. J. Sham, Phys. Rev. **140**, A1133 (1965).
7. R. G. Parr and W. Yang, *Density-Functional Theory of Atoms and Molecules* (Oxford University Press, New York, 1989).

8. R. M. Dreizler and E. K. U. Gross, *Density Functional Theory: An Approach to the Quantum Many-body Problem* (Springer, Berlin, 1990).
9. G. Ortiz and R. M. Martin, Phys. Rev. B **49**, 14202 (1994).
10. R. M. Martin and G. Ortiz, Solid State Commun. **102**, 121 (1997).
11. R. M. Martin and G. Ortiz, Phys. Rev. Lett. **78**, 2028 (1997).
12. R. M. Martin and G. Ortiz, accepted in Phys. Rev. B (1997).
13. X. Gonze, P. Ghosez, and R. W. Godby, Phys. Rev. Lett. **74**, 4035 (1995).
14. C. Zener, Proc. Roy. Soc. (London) **145**, 523 (1934).
15. G. Wannier, Rev. Mod. Phys. **34**, 645 (1962).
16. J. Avron and J. Zak, Phys. Rev. B **9**, 658 (1974).
17. G. Nenciu, Rev. Mod. Phys. **63**, 91 (1991).
18. R. Landauer, J. Chem. Phys. **32**, 1784 (1960).
19. J. F. Woo and R. Landauer, Phys. Rev. B **6**, 4976 (1972).
20. R. M. Martin, Phys. Rev. B **9**, 1998 (1974).
21. C. Kallin and B. I. Halperin, Phys. Rev. B **29**, 2175 (1984).
22. A. K. Tagantsev, Phase Transitions **35**, 119 (1991).
23. A. K. Tagantsev, Phys. Rev. Lett. **69**, 389 (1992).
24. R. Resta, Rev. Mod. Phys. **66**, 899 (1994).
25. R. D. King-Smith and D. Vanderbilt, Phys. Rev. B **47**, 1651 (1993).
26. G. Ortiz, P. Ordejón, R. M. Martin, and G. Chiappe, Phys. Rev. B - in press (1996).
27. I. I. Mazin and R. E. Cohen, Ferroelectrics **194**, 263 (1997).
28. W. G. Aulbur, L. Jonsson, and J. W. Wilkins, Phys. Rev. B **54**, 8540 (1996).
29. R. Resta, Phys. Rev. Lett. **77**, 2265 (1996).
30. X. Gonze, P. Ghosez, and R. W. Godby, Phys. Rev. Lett. **78**, 2029 (1997).
31. X. Gonze, P. Ghosez, and R. W. Godby, Phys. Rev. Lett. **78**, 294 (1997).
32. R. Resta, Phys. Rev. Lett. **78**, 2030 (1997).
33. V. Ambegaokar and W. Kohn, Phys. Rev. **117**, 423 (1960).
34. J. Ihm, A. Zunger, and M. L. Cohen, J. Chem. Phys. **12**, 4409 (1979).
35. C. Nash and S. Sen, *Topology and Geometry for Physicists* (Academic Press, London, 1983).
36. M. V. Berry, Proc. Roy. Soc. London **A 392**, 45 (1984).
37. D. J. Thouless, Phys. Rev. B **27**, 6083 (1983).
38. Q. Niu and D. J. Thouless, J. Phys. A **17**, 2453 (1984).
39. G. Blount, in *Solid State Physics*, edited by H. Ehrenreich, F. Seitz, and D. Turnbull (Academic, New York, 1962), p. 305.
40. R. W. Nunes and D. Vanderbilt, Phys. Rev. Lett. **73**, 712 (1994).
41. F. Mauri, G. Galli, and R. Car, Phys. Rev. B **47**, 9973 (1993).
42. P. Ordejón, D. A. Drabold, M. P. Grumbach, and R. M. Martin, Phys. Rev. B **48**, 14646 (1993).
43. W. Kohn, Phys. Rev. **133**, A171 (1964).
44. *Theory of the Inhomogeneous Electron Gas*, edited by S. Lundqvist and N. H. March (Plenum, New York, 1983).
45. R. O. Jones and O. Gunnarsson, Rev. Mod. Phys. **61**, 689 (1989).
46. M. Levy, Proc. Nat. Acad. Sci. USA **76**, 6062 (1979).
47. M. Levy and J. P. Perdew, in *Density Functional Methods in Physics*, edited by R. M. Dreizler and J. da Providencia (Plenum, New York, 1985), p. 11.
48. J. Harris and R. O. Jones, J. Phys. F **4**, 1170 (1974).

49. D. C. Langreth and J. P. Perdew, Solid State Commun. **17**, 1425 (1975).
50. O. Gunnarsson and B. I. Lundqvist, Phys Rev. B **13**, 4274 (1976).
51. R. Pick, M. H. Cohen, and R. M. Martin, Phys. Rev. B **1**, 910 (1970).
52. J. P. Perdew and S. Kurth, this volume.
53. G. Vignale and M. Rasolt, Phys. Rev. B **37**, 10685 (1988).
54. K. Burke and E. K. U. Gross, this volume.

A Guided Tour of Time-Dependent Density Functional Theory

Kieron Burke[1] and E.K.U. Gross[2]

[1] Department of Chemistry, Rutgers University, 315 Penn Street, Camden, NJ 08102

[2] Institut für Theoretische Physik, Universität Würzburg, Am Hubland, D-97074 Würzburg, Germany

Abstract. This chapter briefly reviews progress to date in the density functional theory of time-dependent phenomena. We discuss the fundamental theorems and their relation to other approaches. Several exact conditions are treated. We review the special case of the linear response to a weak external potential and look specifically at the linear response of the uniform electron gas, which is important for local density approximations. We discuss recent suggestions for functional approximations, including both the local current approximation and the local-with-memory density approximation . We review applications of the theory in three situations: Beyond linear response, linear response, and excitation energies, and conclude with a brief outlook.

1 Introduction and User's Guide

Density functional theory is the study of the one-to-one correspondence [1] between an interacting many-body system, and a fictitious non-interacting analog, the Kohn-Sham system [2], whose equations are much easier to solve numerically. This mapping is exact in principle, but must be approximated in practice. Over the last several years, there has been an explosion of interest in density functional theory, driven largely by its applications in quantum chemistry [3]. This is due to recent progress in the accuracy of available approximations [4], and because of the wealth of chemical problems that can be tackled with such a computationally inexpensive tool [5].

Much of this interest has focussed on the problem of N electrons in the ground state of an external time-independent potential. Accurate solution to this problem yields predictions of atomic energies, reaction energies in chemistry, cohesive energies in solids, vibrational energies, phonon spectra, activation barriers, rotational energies, etc. [6]. Much of the other chapters in this book are devoted to this subject.

But the basic idea of mapping an interacting problem onto a non-interacting one via the density is extremely general, and has been applied to many circumstances beyond the original ground-state problem mentioned above. In the present article, we briefly survey the progress that has been made on the time-dependent problem. For those less familiar with the time-dependent case, we emphasize the comparison between it and the static ground-state case. For those more familiar with the time-dependent case, this article should serve as an update of an earlier, more comprehensive review [7]. Throughout this article, we use atomic units, in which $e^2 = \hbar = m = 1$. We also use the shorthand $x = (\mathbf{r}, t)$ for the pair $\mathbf{r}$ and t, which makes the equations shorter and more legible. In the remainder of this introduction, we first demonstrate how far time-dependent density functional theory has come in the last few years (section 1.1), and then give an overview of each of the remaining sections in this chapter.

1.1 Why time-dependent density functional theory is exciting

To give some feeling for the power of time-dependent density functional theory, consider the He atom as a prototypical system. From a ground-state point of view, this system might be regarded as rather dull. Not much chemistry occurs with He.[1] Furthermore, most physicists are more interested in extended systems, whether or not they are metallic, superconducting, etc. What can you do with a ground-state density functional theory of the He atom? You can examine how well your favorite exchange-correlation energy approximations perform. Among them might be the local density approximation (LDA) [2], using the latest uniform gas input [12], and the Colle-Salvetti approximation [13], an orbital-dependent correlation approximation. You can construct the exact Kohn-Sham potential [14], examine how good the functional derivative of your favorite energy approximation is, and argue about the implications [15]. You might turn on a electric field, and calculate the static polarizability, or hyperpolarizability [16].

Now consider the He atom when you have a fully time-dependent density functional theory [17–20]. In the most dramatic case, you can apply a strong laser field, whose strength is comparable to the electron-nuclear electrostatic field. You can then watch the system evolve, see if any electrons are ionized, watch them oscillate back and forth, calculate the induced dipole moment, etc. You compare results with experiment, or other more computationally demanding techniques, and find the most reliable functional approximations. You can then search the vast parameter space of the problem on the computer, guiding the experimentalists in their search for greater gain, etc. As an example of interest to the density functional community, Fig. 1 displays the harmonic generation spectrum of He for a 616 nm, 7×10^{14} W/cm^2 laser pulse, with and without correlation [19]. While the overall pattern is the same in both cases, the effect of correlation reduces the peak heights by a factor of 2 or 3. There are few examples in the ground-state lexicon where correlation effects are so large.

If the complexity of this application is too much for you, you might wish to consider the more simple linear response regime, where the external time-

---

[1] Although see the recent interest in van der Waals dimers[8–11].

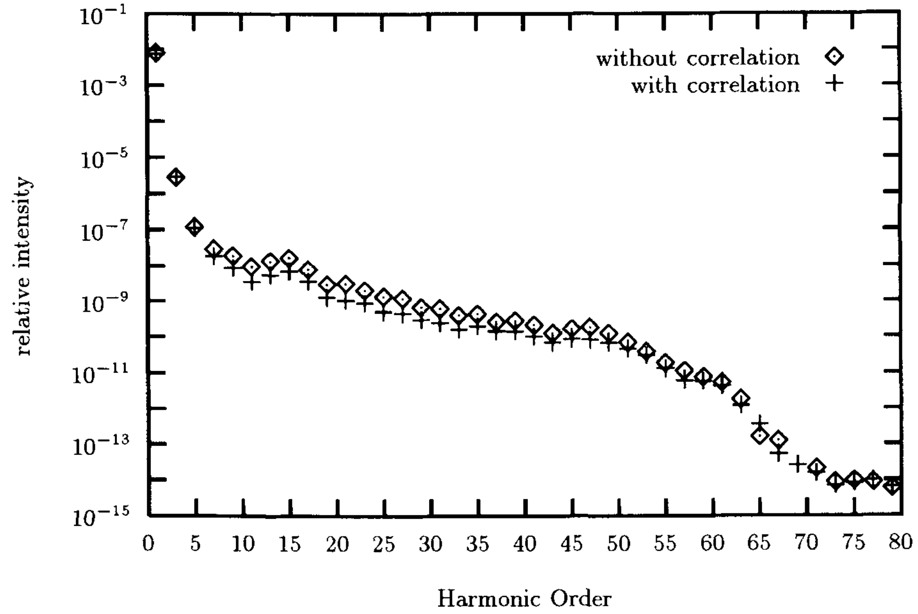


Fig. 1. *Harmonic spectrum of He with and without correlation (see text). The laser parameters are* $\lambda = 616$ *nm and* $I = 7.0 \times 10^{14}$ *W/cm*^{2}*[19].*

dependent potential is weak relative to the ground-state Kohn-Sham potential. Then the formalism gives all the response functions as a function of frequency ω, including the static limit ($\omega \to 0$) as a special case. Such quantities are related to photoabsorption for our simple He atom [21], bulk and surface plasmon dispersions for metals [22–25], etc.

To appreciate how sophisticated the frequency dependence of these response functions is, note that a single response function has poles at *all* the excitation energies of the system [26]. Thus a simple application is to extract these energies from the response function, producing a very natural excited-state density functional theory. Many of the excitation energies of He [27] (and other atoms [28]) have been calculated this way, testing various approximations. Fig. 2 shows the *exact* Kohn-Sham eigenvalues of the He atom[29] (left-hand side), which typically lie between the experimental spin-split levels (right-hand side). In the middle, we show the spin-split levels calculated within linear response, using the exact Kohn-Sham potential, and an approximation to to the response kernel (defined in section 6.3). Clearly, progress is being made toward an accurate, reliable calculational scheme for the excitation energies of a system.

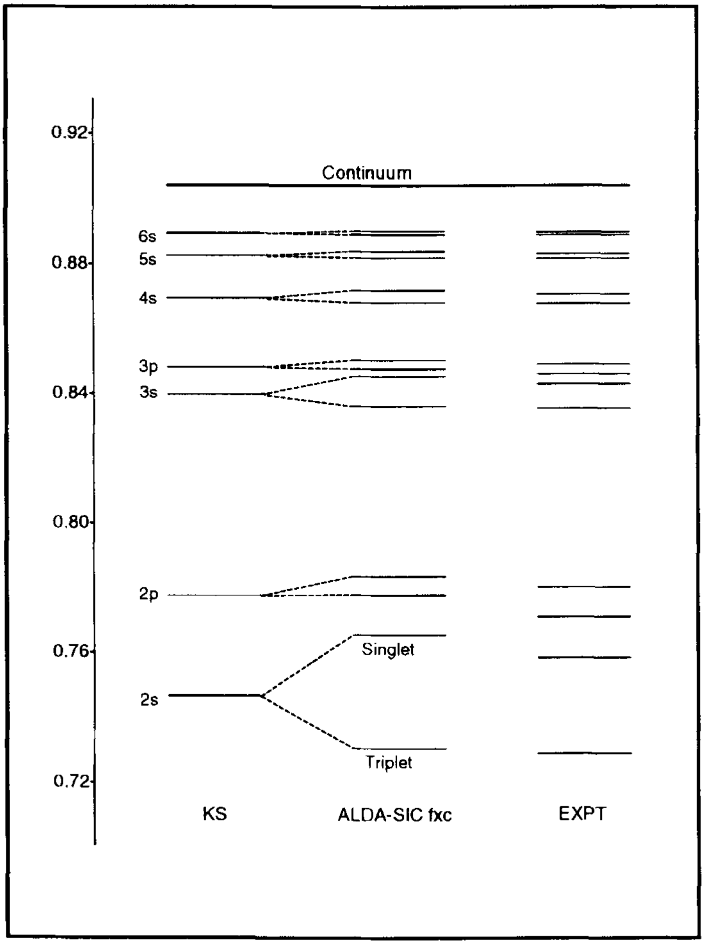


Fig. 2. *Typical excitation energies from the ground-state of He, including the orbital energies of the exact Kohn-Sham potential (KS), the time-dependent OEP spin-split correction calculated within the self-interaction corrected adiabatic local density approximation[30], and experiment [27]. (All energies in hartree.)*

1.2 Fundamentals

In section 2 of this chapter, we review fundamental ideas. To find the ground state energy of He or any other electronic system in density functional theory, one first establishes the Hohenberg-Kohn theorem [1]. This was done thirteen years ago for the time-dependent problem by Runge and Gross [31]. Application of this theorem to a non-interacting system immediately yields the Kohn-Sham equations [2], which are simply generalized to the time-dependent case.

In the ground state, the energy satisfies a minimization principle, and the value of the energy is of great interest in most applications. The time-dependent analog of the energy is the action, which satisfies a stationarity condition. However, the action plays a much less significant role in the time-dependent problem, where its actual value is never of interest, but only its functional derivatives are, in striking contrast to the ground-state energy.

A fruitful area of progress in the ground-state case has been the relation of density functional to standard perturbation techniques. The Sham-Schlüter equation [32] relates the exchange-correlation potential of density functional theory to the exchange-correlation contribution to the self-energy of many-body theory, which was useful in understanding the difference between the Kohn-Sham gap and physical gap in bulk semiconductors and insulators. We discuss the time-dependent generalization of this equation [33].

The optimized effective potential (OEP) formalism has been developed as a method for calculating the Kohn-Sham potential of explicitly orbital-dependent functionals [34,35]. The (highly accurate) Krieger-Li-Iafrate (KLI) approximation [36] makes this a practical scheme for calculations. We show the time-dependent equivalent of this approach. This scheme includes exact exchange by construction, since the Fock integral is an explicit functional of the orbitals. In time-dependent and excited-state problems, this can be much more important than in the ground state, given the emphasis on unoccupied orbital energies, in contrast to the total ground-state energy.

Lastly in this section, we generalize all results to include spin-dependence. This is very useful for the accurate treatment of open-shell atoms, e.g., Li.

1.3 Exact conditions

In section 3 of the chapter, we discuss exact conditions satisfied by the time-dependent exchange-correlation potential. An extremely useful guide to building accurate approximations to the ground-state energy has been the study of conditions satisfied by *exact* functionals. Our next section is rather short, where we discuss exact conditions which must be satisfied by the time-dependent exchange-correlation potential. These may be divided into two kinds: those that have a ground-state analog (e.g., zero net exchange-correlation force and torque, [37,38]) and those that have not (e.g. translational invariance [39]). The brevity of this section suggests that much work is yet to be done in this area.

1.4 Linear response

In section 4, we develop the special case of linear response, i.e., what happens when the time-dependent external potential can be treated as a weak perturbation of the ground state. This is formally analogous to the linear response theory of the ground-state, only now there is a non-zero frequency ω in the perturbation. This leads to a much richer variety of behavior in the response functions. We introduce the exchange-correlation kernel, which characterizes the deviations from the bare Kohn-Sham response, and discuss its exact restraints. We also show the OEP equation for this kernel. We generalize these results to the response to a time-dependent *vector* potential [40], which we later need for the development of a frequency-dependent local density (type) approximation [40]. We end this section with some results for the exchange-correlation kernel of a uniform gas [41,42].

1.5 Approximate functionals

In section 5, we present a variety of approximate functionals. These fall into two broad categories: the majority are LDA-type, using uniform (or slowly-varying) gas input, and the rest are perturbative in the Coulomb repulsion, e.g., exact exchange-only. This mirrors the situation in ground-state density functional theory. The LDA-type appproximations have gone through a very interesting development, in which appreciation of the essential differences between the time-development in an inhomogeneous system and that in a uniform system has been crucial to the construction of approximate functionals. Dobson [39] observed that the two distinct components of fluid flow, a compressive piece and a rigid displacement, *must* have different frequency-dependences, if certain exact constraints are to be satisfied. This has led to two of the latest approximations: the local current-density approximation (LCDA) of Vignale and Kohn for the high-frequency linear response [40], and the local-with-memory density approximation (LMDA) of Dobson, Bünner, and Gross [43], which goes beyond linear response. Lastly, we discuss a simple exchange-only approximation to the KLI exchange-correlation kernel, which is used in practical calculations [44].

1.6 Applications, including excitation energies

In section 6.3, we discuss (in more detail than in section 1.1 above) recent applications of density functional theory. The most exciting is perhaps the phenomena seen when atoms are subjected to superintense laser pulses, (section 6.1) as illustrated in Fig. 1. For these problems, it appears that density functional theory is the *only* practical way to perform calculations which include electronic correlations.

A more standard application of time-dependent density functional theory has been in the photoresponse of a large variety of electronic systems [45], mostly in the linear regime. More recent has been the application of linear response theory to the problem of finding excitation energies [26]. This can be done exactly in principle, since the poles of the interacting susceptibility are at the true excitation energies of the system. We discuss the accuracy of the various approximations needed to make this a practical scheme.

2 Fundamentals

2.1 Hohenberg-Kohn Theorem

The analog of the Hohenberg-Kohn theorem for time-dependent problems is the Runge-Gross theorem [31]. Consider N non-relativistic electrons, mutually interacting via the Coulomb repulsion, in a time-dependent external potential. The Runge-Gross theorem states that the densities $n(x)$ and $n'(x)$ evolving from a common initial state $\Psi_0 = \Psi(t_0)$ under the influence of two potentials $v(x)$ and $v'(x)$ (both Taylor expandable about the initial time t_0) are always different provided that the potentials differ by more than a purely time-dependent ($\mathbf{r}$-independent) function:

$$v(x) \neq v'(x) + c(t) \quad . \tag{1}$$

Thus there is a one-to-one mapping between densities and potentials.

We can prove this theorem by first showing that the corresponding current densities:

$$\mathbf{j}(x) = \langle \Psi(t) | \, \hat{\mathbf{j}}_{\mathrm{p}}(\mathbf{r}) \, | \Psi(t) \rangle \quad , \tag{2}$$

where

$$\hat{\mathbf{j}}_{\mathrm{p}}(\mathbf{r}) = \frac{1}{2i} \sum_{j=1}^{N} \left(\nabla_j \delta(\mathbf{r} - \mathbf{r}_j) + \delta(\mathbf{r} - \mathbf{r}_j) \nabla_j \right) \tag{3}$$

is the paramagnetic current density operator, must differ. The equation of motion for the difference of the two current densities is:

$$\begin{aligned} \frac{\partial}{\partial t} (\mathbf{j}(x) - \mathbf{j}'(x)) \bigg|_{t=t_0} &= -i \langle \Psi_0 | \left[\hat{\mathbf{j}}_{\mathrm{p}}(\mathbf{r}), \hat{H}(t_0) - \hat{H}'(t_0) \right] | \Psi_0 \rangle \\ &= -n_0(\mathbf{r}) \nabla \left(v(\mathbf{r}, t_0) - v'(\mathbf{r}, t_0) \right) \quad , \end{aligned} \tag{4}$$

with the initial density

$$n_0(\mathbf{r}) = n(\mathbf{r}, t_0) \quad , \tag{5}$$

since the Hamiltonians differ only in their external potentials. Repeated use of the equation of motion yields, after some algebra [7],

$$\left(\frac{\partial}{\partial t} \right)^{k+1} (\mathbf{j}(x) - \mathbf{j}'(x)) \bigg|_{t=t_0} = -n_0(\mathbf{r}) \nabla w_k(\mathbf{r}) \tag{6}$$

with

$$w_k(\mathbf{r}) = \left(\frac{\partial}{\partial t} \right)^{k} (v(x) - v'(x)) \bigg|_{t=t_0} \quad . \tag{7}$$

If (1) holds, and the potentials are Taylor expandable about t_o, then there must be some finite k for which the right hand side of (4) does *not* vanish, so that

$$\mathbf{j}(x) \neq \mathbf{j}'(x) \quad . \tag{8}$$

To extend this result to the densities, we use continuity,

$$\frac{\partial}{\partial t}\left(n(x) - n'(x)\right) = -\nabla \cdot \left(\mathbf{j}(x) - \mathbf{j}'(x)\right) \tag{9}$$

and calculate the $(k+1)$th time derivative of (9) at $t = t_0$:

$$\left(\frac{\partial}{\partial t}\right)^{k+2} \left(n(x) - n'(x)\right)\Big|_{t=t_0} = \nabla \cdot \left(n_0(\mathbf{r})\nabla w_k(\mathbf{r})\right) \tag{10}$$

To see that the right-hand-side cannot vanish for all k, note the following identity:

$$\int d^3r\ \nabla \cdot \left(w_k(\mathbf{r}) n_0(\mathbf{r}) \nabla w_k(\mathbf{r})\right) =$$

$$\int d^3r\ \left\{w_k(\mathbf{r})\nabla \cdot \left(n_0(\mathbf{r})\nabla w_k(\mathbf{r})\right) + n_0(\mathbf{r})[\nabla w_k(\mathbf{r})]^2\right\} \quad . \tag{11}$$

The left-hand side may be transformed into a surface integral via Green's theorem, and vanishes for physically realistic potentials (i. e., potentials arising from normalizable external charge densities), because for such potentials the quantities $w_k(\mathbf{r})$ fall off at least as $1/r$. Since there must exist some k for which $\nabla w_k(\mathbf{r})$ does not vanish everywhere, the second integral on the right must be non-zero. Thus the right-hand side of (10) cannot vanish everywhere for all values of k, and the densities $n(\mathbf{r}, t)$ and $n'(\mathbf{r}, t)$ will become different infinitesimally later than t_0.

Note that the difference between $n(x)$ and $n'(x)$ is non-vanishing already in first order of $v(x) - v'(x)$, ensuring the invertibility of the linear response operators of section 4.

Since the density determines the potential up to a time-dependent constant, the wavefunction is in turn determined up to a time-dependent phase, which cancels out of the expectation value of any operator.

2.2 Kohn-Sham Equations

The Runge-Gross theorem can also be applied to a fictitious system of non-interacting electrons having the same density as the physical system, thereby establishing the *uniqueness* of the Kohn-Sham potential (but not its existence) for an arbitrary $n(x)$. Assuming $v_s[n]$ exists, the density of the interacting system is

$$n(x) = \sum_{j=1}^{N} |\varphi_j(x)|^2, \tag{12}$$

with orbitals $\varphi_j(x)$ satisfying the time-dependent KS equation

$$i\frac{\partial \varphi_j(x)}{\partial t} = \left(-\frac{\nabla^2}{2} + v_s[n](x)\right)\varphi_j(x) \quad . \tag{13}$$

We may then define the exchange-correlation potential as

$$v_{\rm XC}[n](x) = v_{\rm S}[n](x) - v(x) - v_{\rm H}[n](x) \ , \tag{14}$$

where $v(x)$ is the external time-dependent field, and $v_{\rm H}[n](x)$ is the time-dependent Hartree potential generated by $n(x)$:

$$v_{\rm H}(x) = \int d^3r' \, \frac{n(\mathbf{r}', t)}{|\mathbf{r} - \mathbf{r}'|} \ . \tag{15}$$

As in the static case, the great advantage of the time-dependent KS scheme lies in its computational simplicity compared to other methods such as time-dependent Hartree-Fock or time-dependent configuration interaction [46–53]. In contrast to time-dependent Hartree-Fock, the effective single-particle potential $v_{\rm S}$ is a *local* potential, i.e., a multiplicative operator in configuration space, which, in principle, contains all correlation effects.

An important difference between ground-state density functional theory and the time-dependent formalism developed above is that in the time-dependent case the 1–1 correspondence between potentials and densities can be established only for a *fixed* initial many-body state Ψ_0, so that functionals depend implicitly on Ψ_0. However, if both Ψ_0 and the initial KS determinant Φ_0 are non-degenerate ground states, then they are uniquely determined by the density, and all quantities are functionals of the density alone.

2.3 Stationary-action principle

At this point, it is customary to define an action functional, which has a stationary point at the solution of the time-dependent Schrödinger equation, with initial condition $\Psi(t_0) = \Psi_0$. A standard choice has been [7]:

$$\mathcal{A} = \int_{t_0}^{t_1} dt \, \langle \Psi(t) | i \frac{\partial}{\partial t} - \hat{H}(t) | \Psi(t) \rangle \ . \tag{16}$$

which may be considered as a functional of $n(x)$. Variation of this functional with respect to $n(x)$ yields an Euler equation whose solution should yield the physical density. In the Kohn-Sham scheme,

$$\mathcal{A}_{\rm S} = \int_{t_0}^{t_1} dt \, \langle \Psi(t) | i \frac{\partial}{\partial t} - \hat{H}_{\rm S}(t) | \Psi(t) \rangle \ , \tag{17}$$

where $H_{\rm S}$ is the non-interacting Kohn-Sham Hamiltonian. An exchange-correlation action functional can then be readily identified as

$$\mathcal{A}_{\rm XC}[n] = \mathcal{B}_{\rm S}[n] - \mathcal{B}[n] - \frac{1}{2} \int_{t_0}^{t_1} dt \int d^3r \int d^3r' \, \frac{n(\mathbf{r}, t)\, n(\mathbf{r}', t)}{|\mathbf{r} - \mathbf{r}'|} \ , \tag{18}$$

where

$$\mathcal{B}[n] = \mathcal{A}[n] + \int_{t_0}^{t_1} dt \int d^3r\, n(x) v(x) \tag{19}$$

is the universal (i.e., potential-independent) contribution to $\mathcal{A}$, and likewise for $\mathcal{B}_\mathrm{S}$. Employing these definitions and the Euler equation, we find

$$v_\mathrm{XC}(x) = \delta\mathcal{A}_\mathrm{XC}/\delta n(x) \ . \tag{20}$$

However, some difficulties with the action defined as in (16) have been noticed (see section 4.1). Here we give a simple demonstration of the inadequacy of (16). Since the stationary point satisfies the Schrödinger equation, $\mathcal{A} = 0$ at this point. This is likewise true in the Kohn-Sham system, so that $\mathcal{A}_\mathrm{S} = 0$ also. Inserting the definitions of $\mathcal{B}$ and $\mathcal{B}_\mathrm{S}$ into (18), and using the definition of the exchange-correlation potential, and the fact that $\mathcal{A} = \mathcal{A}_\mathrm{S} = 0$ at the stationary point, we find:

$$\mathcal{A}_\mathrm{XC}[n] = \int_{t_0}^{t_1} dt \int d^3r \ n(x) \ \{v_\mathrm{H}[n](x)/2 + v_\mathrm{XC}[n](x)\} \ . \tag{21}$$

Functional differentiation of this suspiciously simple result yields a result which is only true for the one-electron case:

$$v_\mathrm{H}(x) = -\int_{t_0}^{t_1} dt' \int d^3r' \ n(x') \frac{\delta v_\mathrm{XC}(x)}{\delta n(x')} \ . \tag{22}$$

From this simple exercise, we conclude that the definition of the action of (16) is inadequate.

Note that the same reasoning may be applied to the ground state problem, but with different results. The energy is analogous to the action, but in this case, the value of the energy at the minimum differs in the physical and KS systems, thereby avoiding the conundrum.

2.4 Relation to many-body theory

Van Leeuwen [33] has recently shown how to connect time-dependent density functional theory to the better-known language of many-body perturbation theory. However, because of the explicit time-dependence of the external potential, we are always dealing in non-stationary states, so that one cannot use the usual technique of time-ordered products and Wick's theorem [54]. This difficulty shows up whenever there is explicit time-dependence in the many-body problem, and Keldysh [55] devised a method used to get round it, by parametrizing the physical time $t(\tau)$ with a pseudotime τ.

Employing this formalism, one can derive the time-dependent Sham-Schlüter equation [32]:

$$\begin{aligned} &\int dy' \ G(y,y') \ G_\mathrm{S}(y',y) \ v_\mathrm{XC}(y') = \\ &\int d^4y' \int d^4y'' \ G_\mathrm{S}(y,y') \ \Sigma_\mathrm{XC}(y',y'') \ G(y'',y) \end{aligned} \tag{23}$$

where $y = (\mathbf{r}, \tau)$, $\int d^4y = \int_{-\infty}^{\infty} d\tau dt/d\tau \int d^3r$, and $G(y, y')$ is the one-particle Green's function, with G_S its Kohn-Sham analog, while Σ_{XC} is the exchange-correlation contribution to the self-energy [33].

Van Leeuwuen has also constructed an action functional which does not suffer the causality difficulty. In the Keldysh formalism, all response functions are symmetric in pseudotime, but become causal when transformed back to real time[56].

2.5 Explicit orbital dependence

Since the Kohn-Sham orbitals are implicit functionals of the density, we can consider the Kohn-Sham potential as an explicit functional of the orbitals, rather than of the density alone. The theory remains formally exact, but is written in a different way. Such a procedure can be useful for approximate functional development. In particular, with explicit orbital functionals, the exchange contribution becomes exact.

When the action is written as a functional of the orbitals, the condition used to determine the potential is $\delta A/\delta v_S(x) = 0$. Hence the name optimized effective potential (OEP), but we emphasize that this potential is simply the usual Kohn-Sham potential. Here we do not give a derivation of the time-dependent OEP equations, but instead we just state the results:

$$\int_{-\infty}^{t_1} dt \int d^3r \left(G^{\text{ret}}(x, x') \sum_{j=1}^{N} \phi_j(x)\phi_j^*(x') \left[v_{XC}(x') - u_{XC,j}(x') \right] + c.c. \right) = 0 \ , \tag{24}$$

where

$$u_{XC,j}(x) = \frac{1}{\phi_j^*(x)} \frac{\delta A_{XC}[\Phi]}{\delta \phi_j(x)} \ . \tag{25}$$

Here Φ denotes the Kohn-Sham wavefunction, and the kernel $G^{\text{ret}}(x, x')$ is the retarded one-particle Green's function of the system:

$$\left\{ i\partial/\partial t - \left(-\nabla'^2/2 + v_S(x') \right) \right\} G^{\text{ret}}(x, x') = \delta(x - x') \tag{26}$$

with initial condition $G^{\text{ret}}(x, x') = 0$ for $t' > t$. Explicitly,

$$iG^{\text{ret}}(x, x') = \sum_{j=1}^{\infty} \phi_j^*(x)\ \phi_j(x')\ \Theta(t - t') \ . \tag{27}$$

Note that the functions $u_{XC,j}$ are given as functional derivatives of the exchange-correlation action. While the definition (16) has serious difficulties for other purposes, it appears to yield the correct results in this context.

The solution of these time-dependent OEP equations (with some approximation for $A_{XC}[\Phi]$) is extremely computationally demanding. The Krieger-Li-Iafrate (KLI) approximation is very useful in this regard, because the error made in this approximation is expected to be so small relative to the errors made in functional approximations [36], that we can treat its results as 'exact.' By ignoring an orbital-dependent contribution which averages to zero over each orbital, one can rewrite the integral equation (24) so that it has an *analytic* solution for the exchange-correlation potential, Equation (22) of Ref. [44], which is used for practical calculations.

2.6 Spin-density functional theory

The above arguments can all be easily generalized to the common case of spin-density functional theory [7], in which the external field can differ for each of the two spin components along some fixed direction. This corresponds physically to the coupling to the spin of an electron of an external magnetic field which has only one non-zero component. Then we can establish a one-to-one mapping between spin densities $n_\uparrow(x)$ and $n_\downarrow(x)$ and potentials $v_\uparrow(x)$ and $v_\downarrow(x)$, and all quantities can be considered functionals of the two spin densities, e.g., the Kohn-Sham potential is $v_{\mathrm{S},\sigma}(x)[n_\uparrow, n_\downarrow]$, where $\sigma = \uparrow, \downarrow$.

In analogy to the ground-state problem, when $B = 0$, $v_\uparrow = v_\downarrow = v$, but the spin-dependent Kohn-Sham equations do not reduce to the regular Kohn-Sham equations, unless the system is spin-unpolarized. The great practical benefit of spin-density functional theory when $B = 0$ is the improvement in accuracy of approximations for spin-polarized systems (e.g., local spin density approximation over local density approximation).

3 Exact Conditions

3.1 Newton's third law

Consideration of the equation of motion for the $\hat{\mathbf{r}}$ operator leads to [38,57,7]

$$\langle\ddot{\mathbf{r}}\rangle = \frac{d^2}{dt^2}\int d^3r\, \mathbf{r}\, n(x) = -\int d^3r\; n(x)\; \nabla v_{\mathrm{ext}}(x) \quad , \tag{28}$$

i.e., the net external exchange-correlation force must be zero. The analogous statement is true for the Kohn-Sham system, which must reproduce the same equation of motion. Then, from the relation between the Kohn-Sham and external potentials (14), and the fact that

$$\int d^3r\; n(x)\; \nabla v_{\mathrm{H}}(x) = 0 \quad , \tag{29}$$

we find

$$\int d^3r\; n(x)\; \nabla v_{\mathrm{XC}}(x) = 0 \quad . \tag{30}$$

Similarly, one obtains

$$\int d^3r\; n(x)\; \mathbf{r} \times \nabla v_{\mathrm{XC}}(x) = 0 \quad , \tag{31}$$

i.e., the net exchange-correlation torque must vanish.

3.2 Translational invariance

Consider a rigid boost $\mathbf{X}(t)$ of a static density, $n(\mathbf{r})$. Then the exchange-correlation potential of the boosted static density will be that of the unboosted density, evaluated at the boosted point, i.e.,

$$v_{\mathrm{XC}}[n'](\mathbf{r},t) = v_{\mathrm{XC}}[n](\mathbf{r}-\mathbf{X}(t)), \qquad n'(\mathbf{r},t) = n(\mathbf{r}-\mathbf{X}(t)), \qquad \mathbf{X}(t_0) = 0 \tag{32}$$

This condition is universally valid [37] . It was first discovered [39] in the special case of a harmonic external potential driven by driving force $\mathbf{F}(t)$, where $\mathbf{X}(t)$ is the classical motion of an oscillator under this driving force. This is the harmonic potential theorem [58], which is an extension of the generalized Kohn theorem [59].

4 Linear Response

Consider an external potential of the form

$$v_{\mathrm{ext}}(x) = v_0(\mathbf{r}) + v_1(\mathbf{r},t)\ \Theta(t-t_0)\ , \tag{33}$$

where $\Theta(t) = 1$ for $t > 0$, but is zero otherwise, and assume that at times $t \leq t_0$ the system is in the *ground state* corresponding to $v_0(\mathbf{r})$. In this case, the initial density $n_0(\mathbf{r})$ can be obtained from the self-consistent solution of the static ground-state Kohn-Sham equations and, via the Hohenberg-Kohn theorem, the time-dependent density is a functional of the external potential alone, $n[v_{\mathrm{ext}}](x)$. We expand this functional in a Taylor series in $v_1(\mathbf{r},t)$:

$$n(x) = n_0(\mathbf{r}) + n_1(x) + \dots\ , \tag{34}$$

where the lower indices denote the orders in v_1. The first order density response n_1 is given by

$$n_1(x) = \int d^4x\ \chi(x,x')\, v_1(x') \tag{35}$$

with the density-density response function

$$\chi(x,x')[n_0] = \left.\frac{\delta n[v_{\mathrm{ext}}](x)}{\delta v_{\mathrm{ext}}(x')}\right|_{v_0} . \tag{36}$$

The same reasoning applies to the Kohn-Sham system, yielding

$$\chi_{\mathrm{S}}(x,x')[n_0] = \left.\frac{\delta n[v_{\mathrm{s}}](x)}{\delta v_{\mathrm{s}}(x')}\right|_{v_{\mathrm{s}}[n_0]}\ , \tag{37}$$

which can be found by inverting the dependence $v_{\mathrm{S}}[n]$, and can be expressed in terms of the static unperturbed Kohn-Sham orbitals ϕ_k:

$$\chi_S(\mathbf{r},\mathbf{r}';\omega)=\sum_{j,k}(f_k-f_j)\frac{\phi_j(\mathbf{r})\phi_k^*(\mathbf{r})\phi_j^*(\mathbf{r}')\phi_k(\mathbf{r}')}{\omega-(\epsilon_j-\epsilon_k)+i\eta}\ . \tag{38}$$

Here, (f_k, f_j) are the occupation numbers (0 or 1) of the KS orbitals. The summation in (38) ranges over both occupied and unoccupied orbitals, including the continuum states.

A Dyson-like equation may be derived between χ and χ_S. Using the chain rule, we write

$$\chi(x,x')=\int d^4x''\,\frac{\delta n(x)}{\delta v_S(x'')}\frac{\delta v_S(x'')}{\delta v_{ext}(x')}\bigg|_{n_0}\ . \tag{39}$$

We can write an expression of the second functional derivative above in terms of v_{XC} using its definition, (14):

$$\frac{\delta v_S(x)}{\delta v_{ext}(x')}\bigg|_{n_0}=\delta(x-x')+\int d^4x''\,\frac{\delta v_{HXC}(x)}{\delta n(x'')}\frac{\delta n(x'')}{\delta v_{ext}(x')}\ . \tag{40}$$

where $v_{HXC}=v_H+v_{XC}$ and $\delta v_H(x)/\delta n(x')=\delta(t-t')/|\mathbf{r}-\mathbf{r}'|$. Insertion of this result into (39), and using the definitions of the physical and Kohn-Sham response functions, we find

$$\chi(x,x')=\chi_S(x,x')+\int d^4x_3\int d^4x_4\ \chi_S(x,x_3)\frac{\delta v_{HXC}(x_3)}{\delta n(x_4)}\ \chi(x_4,x')\ . \tag{41}$$

This equation has the form of the time-dependent random phase approximation (RPA), and reduces to that approximation if $\delta v_{XC}/\delta n$ is ignored.

¿From (41), we can extract the *exact* linear response of the physical system, by multiplying χ by the perturbing potential, and integrating over x', to find the self-consistent linear response equations:

$$n_1(x)=\int d^4x'\chi_S(x,x')\ v_{S,1}(x')\ , \tag{42}$$

where

$$v_{S,1}(x)=v_1(x)+v_H(x)+\int d^4x'\frac{\delta v_{XC}(x)}{\delta n(x')}[n_0]n_1(x') \tag{43}$$

consists of the external perturbation v_1 and the Hartree- and exchange-correlation contributions to first order in the perturbing potential v_1.

The results here have been generalized to finite temperature in thermal equilibrium [60,61].

4.1 Exchange-correlation kernel

In these derivations, the quantity

$$f_{\text{XC}}(x,x') = \frac{\delta v_{\text{XC}}(x)}{\delta n(x')} \tag{44}$$

plays a key role, and is called the exchange-correlation kernel. In the theory of classical liquids [62], $f_{\text{HXC}} = \delta v_{\text{HXC}}/\delta n$ is the Ornstein-Zernike function. The Fourier transform with respect to time of f_{XC} is

$$f_{\text{XC}}(\mathbf{r},\mathbf{r}';\omega) = \int_{-\infty}^{\infty} dt\; e^{i\omega t}\; f_{\text{XC}}(\mathbf{r},t;\mathbf{r}',0) \tag{45}$$

and the relation to the static case is [7]

$$\lim_{\omega\to 0} f_{\text{XC}}(\mathbf{r},\mathbf{r}';\omega) = \frac{\delta^2 E_{\text{XC}}}{\delta n(\mathbf{r})\delta n(\mathbf{r}')}. \tag{46}$$

An exact formal representation of f_{XC} is readily obtained by solving (35) for v_1 and inserting the result in (43):

$$f_{\text{HXC}}(x,x') = \chi_{\text{S}}^{-1}(x,x') - \chi^{-1}(x,x') \tag{47}$$

where χ_{S}^{-1} and χ^{-1} stand for the kernels of the corresponding inverse integral operators whose existence on the set of densities specified by (33) and (35) follows from (10), as mentioned in section 2.1.

As a consequence of causality,

$$f_{\text{XC}}(x,x') = 0 \quad \text{for} \quad t' > t\ . \tag{48}$$

making f_{XC} asymmetric under interchange of x and x'. This is the difficulty in treating f_{XC} as a second functional derivative of the action defined by (16) in section 2.3.

4.2 Exact conditions

The kernel $f_{\text{XC}}(\mathbf{r},\mathbf{r}',\omega)$ is an analytic function of ω in the upper half of the complex ω-plane and approaches a real function $f_{\text{XC}}(\mathbf{r},\mathbf{r}';\infty)$ for $\omega\to\infty$ [63]. Therefore, the function $(f_{\text{XC}}(\mathbf{r},\mathbf{r}',\omega) - f_{\text{XC}}(\mathbf{r},\mathbf{r}';\infty)$ satisfies Kramers-Kronig relations :

$$\Re f_{\text{XC}}(\mathbf{r},\mathbf{r}',\omega) - f(\mathbf{r},\mathbf{r}',\infty) = \mathbb{P}\int \frac{d\omega'}{\pi}\frac{\Im f_{\text{XC}}(\mathbf{r},\mathbf{r}',\omega')}{\omega'-\omega} \tag{49}$$

$$\Im f_{\text{XC}}(\mathbf{r},\mathbf{r}',\omega) = -\mathbb{P}\int \frac{d\omega'}{\pi}\frac{\Re f_{\text{XC}}(\mathbf{r},\mathbf{r}',\omega') - f_{\text{XC}}(\mathbf{r},\mathbf{r}';\infty)}{\omega'-\omega}\ . \tag{50}$$

Also, since $f_{\text{XC}}(x,x')$ is real-valued,

$$f_{\rm XC}(\mathbf{r},\mathbf{r}';\omega) = f^*_{\rm XC}(\mathbf{r},\mathbf{r}';-\omega) \quad . \tag{51}$$

Besides that, the response functions χ_s and χ satisfy the symmetry relations [64]

$$\chi(\mathbf{r},\mathbf{r}';\omega) = \chi(\mathbf{r}',\mathbf{r};\omega) \tag{52}$$

provided that the unperturbed system has time-reversal symmetry. Equation (47) then implies that

$$f_{\rm XC}(\mathbf{r},\mathbf{r}';\omega) = f_{\rm XC}(\mathbf{r}',\mathbf{r};\omega) \quad . \tag{53}$$

The exact conditions on the potential of section (3) also yield conditions on $f_{\rm XC}$, when applied to an infinitesimal perturbation. Taking functional derivatives of (30) and (31) yields

$$\int d^3r\ n(\mathbf{r})\ \nabla f_{\rm XC}(\mathbf{r},\mathbf{r}';\omega) = -\nabla' v_{\rm XC}(\mathbf{r}') \tag{54}$$

and

$$\int d^3r\ n(\mathbf{r})\ \mathbf{r}\times\nabla f_{\rm XC}(\mathbf{r},\mathbf{r}';\omega) = -\mathbf{r}'\times\nabla' v_{\rm XC}(\mathbf{r}') \quad . \tag{55}$$

Taking $\int d^3r'\ n(\mathbf{r}')$ of each side, and using (30) and (31) once again, yields

$$\int d^3r \int d^3r'\ n(\mathbf{r})\ n(\mathbf{r}')\ \nabla f_{\rm XC}(\mathbf{r},\mathbf{r}';\omega) = 0 \tag{56}$$

and

$$\int d^3r\ n(\mathbf{r}) \int d^3r'\ n(\mathbf{r})\ n(\mathbf{r}')\ \mathbf{r}\times\nabla f_{\rm XC}(\mathbf{r},\mathbf{r}';\omega) = 0 \quad . \tag{57}$$

These equations place strong restrictions on the frequency dependence of $f_{\rm XC}$, since the right hand sides are all frequency independent [38].

4.3 Explicit orbital-dependence

One can also consider the linear response regime of the OEP equations, and ask what integral equation $f_{\rm XC}$ satisfies [27]. This turns out to be identical to the OEP equation itself, (24), but with $v_{\rm XC}(x)$ replaced by $f_{\rm XC}(x,x')$ and $u_{{\rm XC},j}(x)$ replaced by $g_{{\rm XC},j}(x,x')$, where

$$g_{{\rm XC},j}(x,x') = \left[\frac{1}{\phi^*_j(x')}\frac{\delta v_{\rm XC}(x)}{\delta\phi_j(x')}\right]_{\{\phi_k(\mathbf{r})\ \exp(-i\epsilon_k(t-t_0))\}} \tag{58}$$

and the $\{\phi_k(\mathbf{r})\}$ are the ground-state Kohn-Sham orbitals. Thus $g_{{\rm XC},j}$ involves a straightforward functional derivative of $v_{\rm XC}$, but evaluated on the time-evolved ground-state Kohn-Sham orbitals.

4.4 Tensor exchange-correlation kernel

This scalar linear response theory may be easily generalized to the time-dependent current response $\mathbf{j}_1(x)$ to a vector potential $\mathbf{a}_1(x)$ [65,40]. Following the reasoning above, we find

$$\mathbf{j}_1(\mathbf{r},\omega) = \int d^3r' \; \boldsymbol{\chi}_{\rm S}\,(\mathbf{r},\mathbf{r}';\omega)\cdot \mathbf{a}_{\rm S,1}(\mathbf{r}',\omega) \;, \tag{59}$$

where

$$\mathbf{a}_{\rm S,1} = \mathbf{a}_1 + \mathbf{a}_{\rm H} + \mathbf{a}_{\rm XC} \tag{60}$$

is the Kohn-Sham vector potential,

$$\mathbf{a}_{H} = \frac{1}{(i\omega)^2}\int d^3r' \; \nabla \frac{1}{|\mathbf{r}-\mathbf{r}'|}\nabla' \; \mathbf{j}_1(\mathbf{r}',\omega) \tag{61}$$

is the longitudinal vector Hartree potential, and

$$\mathbf{a}_{\rm XC}(\mathbf{r},\omega) = \int d^3r' \; \mathbf{f}_{\rm XC}(\mathbf{r},\mathbf{r}';\omega) \;\cdot \mathbf{j}_1(\mathbf{r}';\omega) \tag{62}$$

is a linear functional of $\mathbf{j}_1$. The boldface $\boldsymbol{\chi}$ and $\mathbf{f}_{\rm XC}$ indicates that these are now 3 × 3 tensors. The tensor exchange-correlation kernel is

$$\mathbf{f}_{\rm XC}(\mathbf{r},\mathbf{r}';\omega) = \boldsymbol{\chi}_{\rm S}^{-1}\,(\mathbf{r},\mathbf{r}';\omega) - \boldsymbol{\chi}^{-1}\,(\mathbf{r},\mathbf{r}';\omega) - \nabla \frac{\otimes}{|\mathbf{r}-\mathbf{r}'|}\nabla' \tag{63}$$

where $\otimes$ denotes the outer product of two vectors, and the Kohn-Sham susceptibility tensor is

$$\boldsymbol{\chi}_{\rm S}(\mathbf{r},\mathbf{r}';\omega) = n(\mathbf{r})\;\delta(\mathbf{r}-\mathbf{r}')\;\mathbb{1} + \sum_{j,k}(f_k - f_j)\frac{\phi_j(\mathbf{r})\phi_j^*(\mathbf{r}')\nabla\phi_k^*(\mathbf{r})\otimes\nabla\phi_k(\mathbf{r}')}{\omega-(\epsilon_j-\epsilon_k)+i\eta} \;. \tag{64}$$

This generalization is necessary for the construction of a local density type approximation for time-dependent problems, as in section 5.4.

4.5 Homogeneous gas

In order to investigate $f_{\rm XC}$ for the uniform gas, we consider (47) in the uniform case. Fourier transformation with respect to $(\mathbf{r}-\mathbf{r}')$ and $(t-t')$ leads to

$$f_{\rm XC}^{\rm unif}(n,q;\omega) = \frac{1}{\chi_{\rm S}^{\rm unif}(n,q;\omega)} - \frac{1}{\chi^{\rm unif}(n,q;\omega)} - \frac{4\pi}{q^2} \;, \tag{65}$$

where $\chi_{\rm S}^{\rm unif}$ is the Lindhard function [54]. In the theory of the uniform electron gas [63,66–68] $f_{\rm XC}^{\rm unif}(q,\omega)$ is proportional to the local field correction

$$\mathcal{G}(q,\omega) = -\frac{q^2}{4\pi}f_{\rm XC}^{\rm unif}(q,\omega) \;. \tag{66}$$

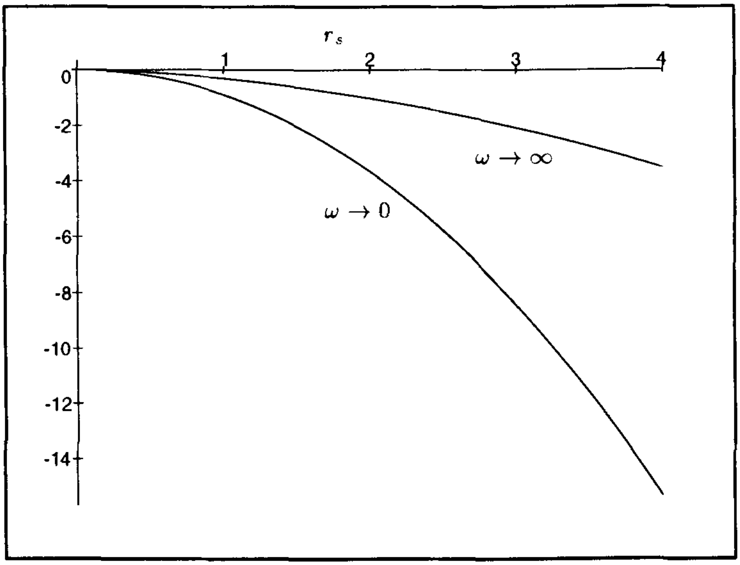


Fig. 3. *$f_{\rm XC}^{\rm unif}(q=0,\omega\to 0)$ and $f_{\rm XC}^{\rm unif}(q=0,\omega\to\infty)$ as functions of r_s.*

We focus here on the limit $q \to 0$, as this is used in the Gross-Kohn approximation (see section 5.2). Writing

$$f_\omega(n) = \lim_{q\to 0} f_{\rm XC}^{\rm unif}(q,\omega) \ , \tag{67}$$

then $f_0(n)$ is known from the compressibility sum-rule, while $f_\infty(n)$ is known from the third frequency-moment sum rule [42], each in terms of the accurately-known uniform gas exchange-correlation energy [69]. We plot these functions in Fig. 3. Furthermore, the imaginary part of $f_{\rm XC}^{\rm unif}$ exhibits the high-frequency behavior

$$\lim_{\omega\to\infty} \Im f_{\rm XC}^{\rm unif}(q,\omega) = -\frac{c}{\omega^{3/2}} \tag{68}$$

for any $q < \infty$ [70]. A second-order perturbation expansion [70,71] of the irreducible polarization propagator leads in the high-density limit to $c = 23\pi/15$, while others [72,73] find double this value; see also Ref. [74]. Some exact features of $\chi^{\rm unif}$ are known, which lead to some exact properties of $f_{\rm XC}^{\rm unif}$. Many of the corresponding properties for $q \to \infty$ are also known [7].

Taking into account the exact high- and low-frequency limits, Gross and Kohn [41] proposed the following smooth interpolation for $f_\omega(n)$:

$$\Im f_\omega(n) = -c\frac{\alpha^{5/3}(n)\ \omega}{(1+\alpha^{4/3}(n)\ \omega^2)^{5/4}} \tag{69}$$

where

$$\alpha(n) = (\gamma/c)(f_\infty(n) - f_0(n)) \ , \tag{70}$$

and $\gamma = (\Gamma(1/4))^2/(4\sqrt{2\pi})$. The real part can then be found in terms of elliptic integrals using the Kramers-Kronig relation (49). An extension of the parametrization (69) to non-vanishing q was given by Dabrowski [75]. The spin-dependent case was treated by Liu [76]. A similar interpolation for the exchange-correlation kernel of the 2-dimensional electron gas has been derived by Holas and Singwi [70].

More recently, Böhm, Conti, and Tosi [77,78] have calculated $f_{\text{XC}}(\omega)$ within several different approximations for the uniform gas. They find that two-plasmon processes induce a strong peak in the imaginary part just above twice the plasmon frequency, leading to highly nonmonotonic behavior in the real part. The resulting frequency dependence is very different from the smooth interpolation of (69).

Lastly, we discuss the response of the uniform gas to the vector potential of section 4.4. In the limit of small wavevector q, i.e., $q << k_F$ and $q << \omega/v_F$,

$$\mathbf{f}_{\text{XC}}^{\text{unif}}(\mathbf{q},\omega) = \omega^{-2}\left[f_\omega(n)\ \mathbf{q}\otimes\mathbf{q} + f_\omega^{\text{T}}(n)(q^2\mathbb{1} - \mathbf{q}\otimes\mathbf{q})\right] , \tag{71}$$

where $f_\omega^{\text{T}}(n)$ is the transverse exchange-correlation long wavelength response function of the uniform gas. Thus the exchange-correlation vector potential of a uniform gas in response to a slowly varying perturbation is

$$\mathbf{a}_{\text{XC}}^{\text{unif}}(\mathbf{r},\omega) = \omega^{-2}\left\{\nabla\left[f_\omega(n)\nabla\mathbf{j}_1(\mathbf{r},\omega)\right] - \nabla\times\left[f_\omega^{\text{T}}(n)\nabla\mathbf{j}_1(\mathbf{r},\omega)\right]\right\} \ . \tag{72}$$

The behavior of the transverse kernel $f_\omega^{\text{T}}(n)$ is also known in the limits $\omega \to 0$ [79] and $\omega \to \infty$ [40]. Conti, Nifosi, and Tosi have calculated this transverse component, again finding strong structure around twice the plasmon frequency [79].

5 Approximate Functionals

In this section, we sketch the historical evolution of functional approximations for v_{XC} and f_{XC}. This illustrates how exact conditions are used to refine and sophisticate approximate functionals. What is perhaps lacking is more calculational comparison between these various functionals. After all, the simplest approximation (ALDA) satisfies all exact conditions except, oddly enough, recovery of the frequency-dependent response of the uniform gas.

5.1 Adiabatic local density approximation

The simplest possible approximation of the time-dependent exchange-correlation potential is the adiabatic local density approximation (ALDA). It employs the functional form of the static LDA with a time-dependent density·

$$v_{\mathrm{XC}}^{\mathrm{ALDA}}[n](x) = v_{\mathrm{XC}}^{\mathrm{unif}}(n(x)) = \frac{d}{d\rho}\left(\rho\epsilon_{\mathrm{XC}}^{\mathrm{unif}}(\rho)\right)\Big|_{\rho=n(x)} \quad . \tag{73}$$

Here $\epsilon_{\mathrm{XC}}^{\mathrm{unif}}$ is the accurately known exchange-correlation energy per particle of the uniform electron gas. Naively, one might expect ALDA to be a good approximation only for nearly uniform densities, i.e., for functions $n(x)$ that are slowly varying both spatially and temporally. However, for the ground-state energy, LDA is moderately accurate for almost all systems [80], so there is room for optimism. For the time-dependent exchange-correlation kernel of (44), (73) leads to

$$f_{\mathrm{XC}}^{\mathrm{ALDA}}[n_0](x, x') = \delta(x - x')\frac{d^2}{d\rho^2}\left(\rho\epsilon_{\mathrm{XC}}^{\mathrm{unif}}(\rho)\right)\Big|_{\rho=n_0(\mathbf{r})} \quad . \tag{74}$$

The time Fourier-transform of the kernel has no frequency-dependence at all. This is not such a bad thing, as it means that translational invariance, (32), is satisfied, although in a rather crude fashion.

5.2 Frequency-dependent LDA

In order to incorporate the frequency-dependence of f_{XC} in some approximate fashion, Gross and Kohn [41] suggested using the frequency-dependent exchange-correlation kernel $f_{\mathrm{XC}}^{\mathrm{unif}}$ of the uniform electron gas in a local approximation:

$$f_{\mathrm{XC}}^{\mathrm{LDA}}[n_0](\mathbf{r}, \mathbf{r}'; \omega) = f_{\mathrm{XC}}^{\mathrm{unif}}(n_0(\mathbf{r}), |\mathbf{r} - \mathbf{r}'|; \omega) \quad . \tag{75}$$

The LDA of non-local quantities, such as response functions, always involves some ambiguity [2] as to where the inhomogeneous $n_0(\mathbf{r})$ is to be evaluated. To avoid this, Gross and Kohn [41] treated the case where $n_1(\mathbf{r}, \omega)$ is slowly varying on the length scale given by the range of $f_{\mathrm{XC}}^{\mathrm{unif}}(n_0(\mathbf{r}), |\mathbf{r} - \mathbf{r}'|; \omega)$. The change in the exchange-correlation potential is then

$$v_{\mathrm{XC}}(\mathbf{r}, \omega) = n_1(\mathbf{r}, \omega)\int d^3r'\, f_{\mathrm{XC}}^{\mathrm{unif}}(n_0(\mathbf{r}), |\mathbf{r} - \mathbf{r}'|; \omega) \quad . \tag{76}$$

Thus the Gross-Kohn approximation may be written as

$$f_{\mathrm{XC}}^{\mathrm{GK}}[n_0](\mathbf{r}, \mathbf{r}'; \omega) = \delta(\mathbf{r} - \mathbf{r}') f_{\mathrm{XC}}^{\mathrm{unif}}(n_0(\mathbf{r}), q = 0; \omega) \quad , \tag{77}$$

whereas ALDA uses the zero-frequency limit of this kernel for all ω. As we see in the next two sections, the GK approximation, by incorporating the full frequency dependence of the uniform gas exchange-correlation kernel, violates several of the exact conditions developed in section 3.

5.3 Following a fluid element

The difficulties with the frequency-dependent LDA of Gross and Kohn were first noted by Dobson [39]. Harmonic potential motion causes a static ground-state density to be rigidly boosted, as discussed in section 3.2, leading to a response kernel with no frequency dependence. This suggested the following general property of the time-evolution of an inhomogeneous system (in the linear response regime). Consider a fluid element at $x = (\mathbf{r}, t)$. It has evolved from an element at x', where $t' < t$, and in general, $\mathbf{r}' \neq \mathbf{r}$. The position $\mathbf{r}'$ of the fluid element at earlier t' is determined by the differential equation:

$$\dot{\mathbf{r}} = \mathbf{u}(x') = \mathbf{j}(x')/n(x') \tag{78}$$

where the dot denotes a time derivative, and $\mathbf{u}(x)$ is the fluid velocity, with initial condition $\mathbf{r}'_{t'=t} = \mathbf{r}$ (not to be confused with $u_{\mathrm{XC},j}$). This uniquely determines the fluid displacement history. The linearized continuity equation (9) then yields

$$\dot{n}_1(x) = \dot{n}_1^{\mathrm{comp}}(x) + \dot{n}_1^{\mathrm{rigid}}(x) = -n_0(\mathbf{r})\ \nabla \cdot \mathbf{u}(x) - \mathbf{u}(x) \cdot \nabla n_0(\mathbf{r}) \ . \tag{79}$$

The first term above is the compressive component of the fluid motion, which comes from changes in the velocity distribution of the fluid. The second term originates from the rigid translation of the fluid, without any changes in the velocity distribution, such as occurs in the boost described in translational invariance. Thus the response to the compressive component should be frequency-dependent, but not that of the rigid component, i.e.,

$$v_{\mathrm{XC}}(\mathbf{r}, \omega) = f_\omega(n_0(\mathbf{r}))\ n_1^{\mathrm{comp}}(\mathbf{r}, \omega) + f_{\omega=0}(n_0(\mathbf{r}))\ n_1^{\mathrm{rigid}}(\mathbf{r}, \omega) \ . \tag{80}$$

Clearly, for the boosts used in translational invariance, $n_1^{\mathrm{comp}} = 0$, while this ansatz is also exact for the response of the uniform gas, where $n_1^{\mathrm{rigid}} = 0$.

5.4 Local current-density approximation – LCDA

Consider applying (54) to a slowly varying density, so that f_{XC} could be replaced by its uniform electron gas limit, which is known to be short-ranged. Then we find

$$f_{\mathrm{XC}}^{\mathrm{unif}}\ (q = 0, \omega, n(\mathbf{r}))\ \nabla n(\mathbf{r}) = \nabla v_{\mathrm{XC}}(\mathbf{r}) \ , \tag{81}$$

which is impossible, as the left-hand-side is frequency dependent, while the right is not. Thus f_{XC} for an inhomogeneous system is long-range in space and a nonlocal functional of the density [38].

Using the current density formalism, Vignale and Kohn [40] have shown how to overcome this difficulty, and produce a semilocal current-density approximation. Consider an almost uniform gas, where the deviations are both small and slowly-varying:

$$n_0(\mathbf{r}) = n\,[1 + 2\gamma\ \cos(\mathbf{q}_1 \cdot \mathbf{r})] , \qquad (q_1 \ll k_F, \omega/k_F; \gamma \ll 1) \ . \tag{82}$$

VK deduce an analytic expression for $\mathbf{f}_{\text{XC}}$ to first order in γ, by using (56) and (57) and the Ward identity. This contrasts with the scalar version, in which $f_{\text{XC}}(\mathbf{q}+\mathbf{q}_1, \mathbf{q}; \omega)$ has a singularity of the form $\mathbf{q}\cdot\mathbf{q}_1/q^2$ as $q \to 0$ at finite q_1, which causes the long-rangedness and non-locality of the scalar case. These effects do not show up here, as the tensor $\mathbf{f}_{\text{XC}}$ contains only differences and derivatives of $f_\omega(n)$ values.

Next, VK consider the problem of determining $\mathbf{a}_{\text{XC}}$ for a slowly-varying inhomogeneity, but which may have large global deviations from uniformity, and deduce the general form in terms of gradients of the density and current, up to two gradient operators. Then they equate these results to the perturbative ones mentioned above, in the same way as the gradient expansion coefficients have been found in the ground-state problem [81]. Some contributions of order γ^2 are not fixed by this comparison, but are fixed by requiring translational invariance (32). They give an explicit formula for the exchange-correlation vector potential in terms of $n(x)$ and $\mathbf{j}_1(x)$ and their gradients, with coefficients determined by f_ω and f_ω^{T} [40].

In the case where $|\nabla n|/n \ll q$, they recover the GK result, while in the case of a parabolic well and uniform perturbing electric field, they recover (80), thereby justifying that ansatz under those conditions.

Finally, note that this analysis applies only for $\omega \gg k_F q$, i.e., high frequency response, for which ALDA is not justified. A complete local current density response theory for all frequencies remains to be developed.

Very recently [82], the VK result has been simplified, as

$$-i\omega \mathbf{a}_{\text{XC}} = -\nabla v_{\text{XC}}^{\text{ALDA}} + \frac{1}{n_0(\mathbf{r})}\nabla \boldsymbol{\sigma}_{\text{XC}}(\mathbf{r},\omega) \tag{83}$$

where

$$\boldsymbol{\sigma} = \tilde{\eta}_{\text{XC}}\left(\nabla\otimes\mathbf{u} + \mathbf{u}\otimes\nabla^{\leftarrow} - \frac{2}{3}\nabla\cdot\mathbf{u}\mathbb{1}\right) + \tilde{\zeta}_{\text{XC}}\nabla\cdot\mathbf{u}\mathbb{1} \tag{84}$$

is a visco-elastic stress tensor with complex viscosity coefficients which are simply related to $f_\omega(n)$ and $f_\omega^{\text{T}}(n)$. For sufficiently slowly varying densities, these expressions yield the nonlinear response also [82].

5.5 Local-with-memory density approximation – LMDA

Some recent work goes way beyond the linear response regime, while being constrained to respect all the exact conditions discussed here. The ansatz suggested is [43]

$$\mathbf{F}_{\text{XC}}(x) = -\nabla v_{\text{XC}}(x) = -\frac{1}{n(x)}\nabla\int_{-\infty}^{\infty} dt'\ \Pi_{\text{XC}}\left(n\left(\mathbf{r}'_{t'}(x'), t-t'\right), t-t'\right) \tag{85}$$

where $\mathbf{F}_{\text{XC}}$ is the exchange-correlation force, and $\Pi_{\text{XC}}(n,\tau)$ is a pressure-like scalar local memory function of n and τ, where $\mathbf{r}'_{t'}(x)$ is the fluid displacement function. Thus locality is defined relative to the fluid element, rather than to a fixed position $\mathbf{r}$. This LMDA is the natural generalization of LDA from the static case.

To see explicitly that (85) satisfies translational invariance (32), note that under a rigid boost $\mathbf{X}(t)$, the many-body wavefunction also moves rigidly, so that $\mathbf{j}(x) = n_0(\mathbf{r}-\mathbf{X}(t))\dot{\mathbf{X}}$, and $\mathbf{u}$ is independent of $\mathbf{r}$. Thus $\mathbf{r}'_{t'} = \mathbf{r}+\mathbf{X}(t') - \mathbf{X}(t)$, and

$$\mathbf{F}_{\mathrm{XC}}(x) = -\frac{1}{n(x)}\nabla \Pi_{\mathrm{XC}}(n_0(\mathbf{r}-\mathbf{X}(t)), \omega = 0) = \mathbf{F}_{\mathrm{XC}}[n_0](\mathbf{r}-\mathbf{X}(t)) \quad , \tag{86}$$

satisfying (32).

Interestingly, the function $\Pi_{\mathrm{XC}}(n,\omega)$ is fully determined by requiring it to reproduce the scalar linear response of the uniform gas. Consider small motions at frequency ω in a uniform gas of density n. One finds the simple condition

$$\delta \Pi_{\mathrm{XC}}/\delta n = n \; f_\omega(n) \tag{87}$$

which implies

$$\Pi_{\mathrm{XC}}(n,\omega) = \int_0^n dn' \; n' \; f_\omega(n') \; . \tag{88}$$

Thus any parametrization of $f_{\mathrm{XC}}^{\mathrm{unif}}$ of section 4.5 fixes Π_{XC}.

Note that (85) defines the exchange-correlation force, rather than potential. For a one-dimensional inhomogeneity, one can simply integrate this force up to a point to determine the potential. Thus this scheme recovers the exact VK result for the slowly varying one-dimensional perturbation. In the more general three-dimensional case, work is continuing.

5.6 Expansion in powers of e^2

An alternative route to useful approximations is in powers of the Coulomb repulsion. To lowest order, one gets the time-dependent exchange-only density functional theory, which is most easily expressed in the OEP formalism, since the exchange action is an explicitly orbital-dependent functional:

$$u_{\mathrm{X},j}(x) = -\frac{1}{\phi_j^*(x)} \sum_{i=1}^{N} \int d^3r' \; \frac{\phi_i^*(\mathbf{r}',t) \; \phi_i(\mathbf{r}',t) \; \phi_i(\mathbf{r},t)}{|\mathbf{r}-\mathbf{r}'|} \tag{89}$$

This can be derived either by inserting the exchange action into the OEP equations [44], or directly from the Sham-Schlüter equation [33], expanded to first-order. These equations are computationally similar to the time-dependent Hartree-Fock equations, but their solution has one very important practical advantage over Hartree-Fock. The unoccupied orbitals in exchange-only density functional theory suffer no self-interaction error, whereas those in HF see N remaining occupied orbitals. This feature is not so important in the ground-state problem, but is very significant in time-dependent and excited-state problems.

One can also imagine going to higher order in e^2, via the Sham-Schlüter equation (23), to get ever more accurate results. Of course, for extended systems such as metals, one must include all orders within the random-phase approximation, to avoid blatantly unphysical results, just as in the ground-state problem.

For practical calculations, one can make a simple analytic approximation to the KLI exchange-correlation potential, which we denote PGG [26]:

$$v_{\text{XC}}^{\text{PGG}}(x) = \sum_{j=1}^{N} \frac{|\phi_j(x)|^2}{n(x)} \left[u_{\text{XC},j}(x) + c.c.\right] \tag{90}$$

which comes about by retaining only the dominant contribution. This leads to a simple analytic form for the exchange-correlation kernel:

$$f_{\text{XC}}^{\text{PGG}}(\mathbf{r},\mathbf{r}';\omega) = -2\frac{\left|\sum_{j=1}^{N} \phi_j(\mathbf{r})\, \phi_j^*(\mathbf{r}')\right|^2}{2\, n(\mathbf{r})\, |\mathbf{r}-\mathbf{r}'| n(\mathbf{r}')} \tag{91}$$

This approximate form is exact for exchange for two-electron spin-unpolarized systems. For more than two electrons, even at the exchange-only level, the exchange-correlation kernel has some frequency dependence.

6 Applications, Including Excitation Energies

6.1 Way beyond the linear response regime: Atoms in short laser pulses

Recently, experiments have been performed in which atoms are subjected to very intense, femto-second laser pulses . The external potential is then

$$v(x) = -\frac{Z}{r} + E_o\ f(t)\ z\ \sin(\omega_0 t), \tag{92}$$

where Z is the nuclear charge, E_o is the electric field of the laser, and $f(t)$ is the envelope function of the pulse.

In these experiments, the strength of the perturbation is comparable to the static field of the nucleus, so that linear response theory is simply inadequate. The full time-dependent Kohn-Sham equations are solved, and two quantities extracted from the self-consistent density. The induced dipole moment

$$d(t) = \int d^3r\ z\ n(x) \tag{93}$$

whose Fourier transform, $d(\omega)$, when squared, is proportional to the harmonic spectrum, i.e., the intensity distribution of emitted photons as a function of their frequency. Also, decreases in the norm of the single-particle Kohn-Sham orbitals in a finite volume

$$N_j(t) = \int_V d^3r \; |\phi_j(x)|^2 \tag{94}$$

where V is a finite volume, can be loosely equated with the probability of ionization of an electron from a given orbital. Calculations of these quantities have been performed for Be [44] and Ne [18] subjected to single frequency laser, and for He subjected to two-color lasers [18,20]. Most recently, the formalism has been extended to include the nuclear motion as well [20], which must be treated fully quantum mechanically when the nuclear probability densities do not stay in a tight classical distribution during the motion. It is also important to include the motion of *all* electrons in the outermost shell [20], which is not done in the single active electron (SAE) model in the traditional wavefunction approach [83].

6.2 Linear response and a little beyond

Many applications of time-dependent density functional theory have been in the linear response regime [16]. Calculations of photoresponse have been performed on atoms [21,84–88], molecules [89,90] and clusters [91–101] metallic surfaces [22–25,102–104] and semiconductor heterostructures [105–109] bulk semiconductors [110] and bulk metals [111–114].

For sufficiently long wavelengths, the electric field experienced by the system is spatially uniform, and the dipole approximation holds. For a finite system, the dynamic polarizability is

$$\alpha(\omega) = -\frac{2}{E} \int d^3r \; z \; n_1(\mathbf{r}, \omega) \tag{95}$$

and the photoabsorption cross section is

$$\sigma(\omega) = \frac{4\pi\omega}{c} \; \Im\alpha(\omega). \tag{96}$$

Most calculations of this kind have used the ALDA in the past, with reasonably good results. A similar formula defines the Feibelman d-parameter, $d_\perp(\omega)$, which characterizes the surface response in the long-wavelength limit and determines, e.g., the initial dispersion of the surface plasmon [115].

Another off-shoot of these calculations is the question of incorporating van der Waals forces into density functional approximations. Because these forces arise from *dynamic* fluctuations at separated points in the system, there is no (transparent) way to build them into a local or semilocal (GGA) ground-state exchange-correlation energy functional (but see [9,10]). A more natural language is that of dynamic linear response. Thus approximations for $\alpha(\omega)$ lead to predictions for C_6, the coefficient of the $1/R^6$ attractive potential between two widely separated neutral species [116,8], or the C_3 coefficient of the $1/R^3$ potential between an atom and a surface [117,118]. Furthermore, some approximation for $f_{\rm XC}$, inserted in the linear response formula and then inserted in the adiabatic connection formula for the exchange-correlation energy, will produce an energy estimate which includes van der Waals forces, at both small and large separations [11].

Senatore and Subbaswamy [119] have found explicit expressions for the response up to third order, while Gonze and Vigneron [120] have calculated the static case. Recently [7], it has been shown how to go to arbitrarily higher order response functions, order-by-order. Interestingly, the k-th order density response satisfies an integral equation whose kernel is independent of k, i.e., is the same for *any* order, but whose driving term depends on all $k-1$ lower order solutions:

$$n_k(x) = M_k(x) + \int d^4x' \int d^4x'' \; \chi_{\rm S}(x', x'') \; f_{\rm HXC}(x, x') \; n_k(x''). \tag{97}$$

(This corrects Eq. (185) of Ref. [7], which contains typographical errors.) Therefore once the first-order response is found, the next can be found by the same means, once the driving term is constructed.

6.3 Excitation Energies

The linear response of any system can be used to determine its excitation energies. Simply apply an oscillating potential and vary its frequency, and a resonance occurs whenever the frequency equals the difference of two energy eigenvalues of the system. This familiar physical statement can be used to great effect in time-dependent density functional theory, since the exact susceptibility χ must therefore contain poles at the exact excited state energy differences. To see this explicitly, we rewrite (42) and (43) as an integral equation for the first-order density response in terms of only the Kohn-Sham susceptibility and the exchange-correlation kernel:

$$\begin{aligned}\int d^3r' \quad & \left\{\delta(\mathbf{r}-\mathbf{r}') - \int d^3r'' \; \chi_{\rm S}(\mathbf{r},\mathbf{r}'';\omega) \; f_{\rm HXC}(\mathbf{r}',\mathbf{r}'';\omega)\right\} \; n_1(\mathbf{r}',\omega) \\ &= \int d^3r' \; \chi_{\rm S}(\mathbf{r},\mathbf{r}';\omega) \; v_1(\mathbf{r}',\omega)\end{aligned} \tag{98}$$

where the driving term contains the *external* potential. Since the true excitation energies Ω are generally not identical to the Kohn-Sham excitation energies, the right-hand-side remains finite as $\omega \to \Omega$, whereas n_1 has a pole at Ω. The operator acting on n_1 cannot be invertible here, so its eigenvalues must vanish, i.e., writing

$$\int d^3r' \int d^3r'' \; \chi_{\rm S}(\mathbf{r},\mathbf{r}';\omega) f_{\rm HXC}(\mathbf{r}',\mathbf{r}'';\omega) \; \zeta(\mathbf{r}'',\omega) = \lambda(\omega) \; \zeta(\mathbf{r},\omega) \tag{99}$$

we have $\lambda(\Omega) = 1$. This is an exact condition for determining the excitation energies of the system.

In practice, we usually don't want the full frequency-dependence of the functions in (99), which contain a vast amount of information on the response of the system. In the case where the poles are well-separated, we make a single-pole approximation (SPA), where we expand everything about a single Kohn-Sham energy difference, assuming the difference between the true excitation frequency and the Kohn-Sham one is smaller than the difference between energy levels [26]. We then find

$$\Omega = \omega_{jk} + 2\mathbb{R}(f_k - f_j) \int d^3r \int d^3r' \; \phi_j^*(\mathbf{r}) \; \phi_k^*(\mathbf{r}) \; f_{\text{HXC}}(\mathbf{r},\mathbf{r}';\omega) \; \phi_j(\mathbf{r}') \; \phi_k(\mathbf{r}') \tag{100}$$

This expression estimates the leading correction to the Kohn-Sham eigenvalues as excited state energies. It was used [30] to calculate the energies displayed in Fig. 2, starting from the exact Kohn-Sham orbitals, and using the self-interaction corrected adiabatic local density approximation for the exchange-correlation kernel.

These formulas have been applied to excited state calculations for atoms with considerable success [26], in the sense that the excitation energies are always improved, often significantly. The Kohn-Sham eigenvalues do not distinguish between the spin states of the excited states. When the excited state is degenerate at the Kohn-Sham level, the correction usually splits this degeneracy. In particular, the spin-dependent version of (100) yields the separate singlet and triplet levels, raising the singlet and lowering the triplet [28]. It is important to begin with a good approximation to $v_{\text{XC}}(\mathbf{r})$ for the ground-state, since the orbital energies, especially of the higher lying states, are very sensitive to the decay of the potential [27]. Beyond that, the time-dependent X-only approximation to f_{XC} is usually better than $f_{\text{XC}}^{\text{ALDA}}$, while the correlation correction within ALDA is usually in the right direction.

Other routes to excited state energies are also being explored, such as ensemble DFT [121–131], ΔSCF theory [132–134], and Görling-Levy perturbation theory [135,136]. It remains to be seen which will provide the best approach to the problem. Very likely, each of the various approaches will suggest its own approximations, etc., and be useful in its own way.

7 Outlook

The purpose of this chapter has been to review the current state of time-dependent density functional theory, especially for those who are more familiar with the traditional ground-state formulation and applications.

We believe that this subject is about to see an exponential growth in interest, as practical approximations are currently appearing and being developed, which should lead to applications in areas of chemistry [137], such as laser-induced photochemistry [138]. We anticipate that ultimately it will reach the same level of development and application as the ground-state theory currently enjoys, in both physics and chemistry.

Acknowledgments

This research has been supported by a NATO collaborative research grant 960759. Acknowledgment is made to the Donors of the Petroleum Research Fund, administered by the American Chemical Society, for partial support of this research. We thank Martin Petersilka for his unpublished results on the excitation spectrum of He. K.B. thanks Kin-Chung Lam, Federico G. Cruz, and Paul Hessler for technical assistance.

References

1. P. Hohenberg and W. Kohn, Phys. Rev. **136**, B 864 (1964).
2. W. Kohn and L.J. Sham, Phys. Rev. **140**, A 1133 (1965).
3. *Modern Density Functional Theory: A Tool for Chemistry*, edited by J. M. Seminario and P. Politzer (Elsevier, Amsterdam, 1995).
4. L.A. Curtiss, K. Raghavachari, P.C. Redfern, and J.A. Pople, J. Chem. Phys. **106**, 1067 (1997), and references therein.
5. W. Yang and Z. Zhou, in *Density Functional Theory of Molecules, Clusters, and Solids*, edited by D.E. Ellis (Kluwer Academic, Dordrecht, 1995).
6. K. Burke, J. P. Perdew, and M. Levy, in *Modern Density Functional Theory: A Tool for Chemistry*, edited by J. M. Seminario and P. Politzer (Elsevier, Amsterdam, 1995).
7. E.K.U. Gross, J.F. Dobson, M. Petersilka, in *Topics in Current Chemistry* **181**, pg. 81 (Springer, 1996).
8. Y. Andersson, D.C. Langreth, and B.I. Lunqvist, Phys. Rev. Lett. **76**, 102 (1996).
9. Y. Zhang, W. Pan, and W. Yang, submitted to J. Chem. Phys., June 1997.
10. D.C. Patton and M.R. Pederson, submitted to Int. J. Quantum Chem. 1997.
11. W. Kohn, Y. Meir. and D.E. Makarov, *Van der Waals energies in density functional theory*, preprint, July, 1997.
12. Y. Wang and J. P. Perdew, Phys. Rev. B **44**, 13298 (1991).
13. R. Colle and O. Salvetti, Theoret. Chim. Acta **37**, 329 (1975).
14. C. J. Umrigar and X. Gonze, in *High Performance Computing and its Application to the Physical Sciences*, Proceedings of the Mardi Gras 1993 Conference, edited by D. A. Browne et al. (World Scientific, Singapore, 1993).
15. *Virial energy density in density functional theory*, F.G. Cruz, K.C. Lam, and F.G. Cruz, to be submitted to J. Phys. Chem. A.
16. P. Mlynarski and D. R. Salahub, Phys. Rev. B **43**, 1399 (1990).
17. C.A. Ullrich, U.J. Gossman, and E.K.U. Gross, Ber. Bunsenges. Phys. Chem. **99**, 488 (1995).
18. C.A. Ullrich, S. Erhard, and E.K.U. Gross, in *Super Intense Laser Atom Physics IV*, edited by H.G. Muller (Kluwer, Dordrecht, 1996).
19. S. Erhard and E.K.U. Gross, in *Multiphoton Processes 1996*, ed. P. Lambropoulos and H. Walther (IOP Publishing, London, 1997).
20. C.A. Ullrich and E.K.U. Gross, Comments At. Mol. Phys., **33**, 211 (1997).
21. A. Zangwill and P Soven, Phys. Rev. A **21**, 1561 (1980).

22. M. Levy, in *Density Matrices and Density Functionals*, edited by R.M. Erdahl and V.H. Smith, Jr., (Reidel, Dordrecht, 1987), pg. 479.
23. J.F. Dobson and G.H. Harris, J. Phys. C **19**, 3971 (1986).
24. J.F. Dobson and G.H. Harris, J. Phys. C **20**, 6127 (1987).
25. J.F. Dobson and G.H. Harris, J. Phys. C **21**, L729 (1988).
26. M. Petersilka, U.J. Gossmann, and E.K.U. Gross, Phys. Rev. Lett. **76**, 1212 (1996).
27. M. Petersilka, U.J. Gossmann, and E.K.U. Gross, in *Electronic Density Functional Theory: Recent Progress and New Directions*, eds. J.F. Dobson, G. Vignale, and M.P. Das (Plenum, NY, 1997).
28. M. Petersilka and E.K.U. Gross, Int. J. Quantum Chem. S **30**, 1393 (1996).
29. C.J. Umrigar, A. Savin, and X. Gonze, in *Electronic Density Functional Theory: Recent Progress and New Directions*, eds. J.F. Dobson, G. Vignale, and M.P. Das (Plenum, NY, 1997).
30. M. Petersilka, private communication.
31. E. Runge and E.K.U. Gross, Phys. Rev. Lett. **52**, 997 (1984).
32. L. J. Sham and M. Schlüter, Phys. Rev. Lett. **51**, 1888 (1983).
33. R. van Leeuwen, Phys. Rev. Lett. **76**, 3610 (1996).
34. R.T. Sharp and G.K. Horton, Phys. Rev. **90**, 317 (1953).
35. J.D. Talman and W.F. Shadwick, Phys. Rev. A **14**, 36 (1976).
36. J. B. Krieger, Y. Li, and G. J. Iafrate, Phys. Rev. A **45**, 101 (1992).
37. G. Vignale, Phys. Rev. Lett. **74**, 3233 (1995).
38. G. Vignale, Phys. Lett. A **209**, 206 (1995).
39. J.F. Dobson, Phys. Rev. Lett. **73**, 2244 (1994).
40. G. Vignale and W. Kohn, Phys. Rev. Lett. **77**, 2037 (1996).
41. E.K.U. Gross and W. Kohn, Phys. Rev. Lett. **55**, 2850 (1985).
42. N. Iwamoto and E.K.U. Gross, Phys. Rev. B **35**, 3003 (1987).
43. J.F. Dobson, M. Bünner, and E.K.U. Gross, Phys. Rev. Lett. **79**, Sept 1 (1997).
44. C.A. Ullrich, U.J. Gossman, and E.K.U. Gross, Phys. Rev. Lett. **74**, 872 (1995).
45. P. Mlynarski and D.R. Salahub, J. Chem. Phys. **95**, 6050 (1991).
46. D. Tiszauer and K.C. Kulander, Phys. Rev. A **29**, 2909 (1984).
47. L.F.Errea, L. Méndez, A. Riera, M. Yáñez, J. Hanssen, C. Harel, and A. Salin, J. Physique **46**, 719 (1985).
48. I.L. Cooper, A.S. Dickinson, S.K. Sur, and C.T. Ta, J. Phys. B **20**, 2005 (1987).
49. A. Henne, H.J. Lüdde, A. Toepfer, and R.M. Dreizler, Phys. Lett. A **124**, 508 (1987).
50. W. Fritsch and C.D. Lin, Phys. Lett. A **123**, 128 (1987).
51. J.F. Reading and A.L. Ford, Phys. Rev. Lett. **58**, 543 (1987).
52. J.F. Reading and A.L Ford, Phys. Rev. B, **20**, 3747 (1987).
53. X. Tang, H. Rudolph, and P. Lambropoulos, Phys. Rev. A **44**, R6994 (1991).
54. G.D. Mahan, *Many-Particle Physics*, 2nd ed. (Plenum Press, New York, 1990).
55. L.V. Keldysh, Sov. Phys. JETP **20**, 1018 (1965).
56. R. van Leeuwen, *Causality and symmetry in time-dependent density functional theory*, preprint, September 7, 1997.
57. S. Erhard, Diplomarbeit, Universität Würzburg, 1996.
58. J.F. Dobson, in *Density Functional Theory*, Vol. 337 of *NATO ASI Series B*, edited by E.K.U. Gross and R.M. Dreizler (Plenum Press, New York, 1995), p. 393.

59. L. Brey, N.F. Johnson, and B.I. Halperin, Phys. Rev. B **40**, 10647 (1989).
60. T.K. Ng and K.S. Singwi, Phys. Rev. Lett. **59**, 2627 (1987).
61. W. Yang, Phys. Rev. A **38**, 5512 (1988).
62. N.W. Ashcroft, in *Density Functional Theory*, Vol. 337 of *NATO ASI Series B*, edited by E.K.U. Gross and R.M. Dreizler (Plenum Press, New York, 1995), p. 581.
63. A.A. Kugler, J. Stat. Phys. **12**, 35 (1975).
64. L.P Kadanoff and P.C. Martin, Ann. Phys. New York **24**, 419 (1963).
65. T.K. Ng, Phys. Rev. Lett. **62**, 2417 (1989).
66. S. Ichimaru, Rev. Mod. Phys. **54**, 1017 (1982).
67. V.D. Gorobchenko and E.G. Maksimov, Sov Phys. Usp. **23**, 35 (1980).
68. K.S. Singwi and M.P. Tosi, Solid State Physics **36**, 177 (1981).
69. Y. Wang and J. P. Perdew, Phys. Rev. B **43**, 8911 (1991).
70. M.K. Harbola and V. Sahni, Phys. Rev. Lett. **62**, 489 (1989).
71. A.J. Glick and W.F. Long, Phys. Rev. B **4**, 3455 (1971).
72. D.F Dubois and M.G. Kivelson, Phys. Rev. **186**, 409 (1969).
73. M. Hasegawa and M. Watabe, J. Phys. Soc. Jap. **27**, 1393 (1969).
74. K. Sturm, Phys. Rev. B **52**, 8028 (1995).
75. B. Dabrowski, Phys. Rev. B **34**, 4989 (1986).
76. K.L. Liu, Can. J. Phys. **69**, 573 (1991).
77. H.M. Böhm. S. Conti, and M.P. Tosi, J. Phys. C **8**, 781 (1996).
78. S. Conti, H.M. Böhm, and M.P. Tosi, Phys. Stat. Sol. B **193**, K11 (1996).
79. S. Conti, R. Nifosi, and M.P. Tosi, *Exchange-correlation potential for current density functional theory of frequency dependent response*, preprint cond-mat/9707062, July 1997.
80. K. Burke, M. Ernzerhof, and J.P. Perdew, *Why semi-local functionals work: Accuracy of the on-top hole density*, work in progress.
81. M. Rasolt and D.J.W. Geldart, Phys. Rev. B **34**, 1325 (1986).
82. G. Vignale, C.A. Ullrich, and S. Conti, *Time-dependent density functional theory beyond the ALDA*, preprint, July 1997.
83. K.C. Kulander and B.W. Shore, J. Opt. Soc. Am. B **7**, 502 (1990).
84. S.K. Ghosh and B.M. Deb, Chem. Phys. **71**, 295 (1982).
85. S.K. Ghosh and B.M. Deb, Theoret. Chim. Acta **62**, 209 (1983).
86. S.K. Ghosh and B.M. Deb, J. Mol. Struct. **103**, 163 (1983).
87. A. Zangwill and P Soven, Phys. Rev. Lett. **45**, 204 (1980).
88. A. Zangwill and P Soven, Phys. Rev. B **24**, 4121 (1981).
89. Z.H. Levine and P. Soven, Phys. Rev. Lett. **50**, 2074 (1983).
90. Z.H. Levine and P. Soven, Phys. Rev. A **29**, 625 (1984).
91. W. Ekardt, Phys. Rev. Lett. **52**, 1925 (1984).
92. M.J. Puska, R.M. Nieminen, and M. Manninen, Phys. Rev. B **31**, 3486 (1985).
93. W. Ekardt, Phys. Rev. B **31**, 6360 (1985).
94. D.E. Beck, Phys. Rev. B **35**, 7352 (1987).
95. A. Rubio, L.C. Balbás, and J.A. Alonso, Phys. Rev. B **46**, 4891 (1992).
96. W. Ekardt, Phys. Rev. B **32**, 1961 (1985).
97. W. Ekardt and Z. Penzar, Phys. Rev. B **43**, 1322 (1991).
98. J.A. Alonso, A. Rubio, and L.C. Balbás, Phil. Mag. B **69**, 1037 (1994).
99. J.M. Pacheco and W. Ekardt, Z. Phys. D **24**, 65 (1992).
100. J.M. Pacheco and W. Ekardt, Ann. Physik **1**, 254 (1992).

101. J.M. Pacheco and W. Ekardt, Phys. Rev. B **47**, 6667 (1993).
102. P. Gies and R.R. Gerhardts, Phys. Rev. B **36**, 4422 (1987).
103. P Gies and R.R. Gerhardts, Phys. Rev. B **37**, 10020 (1988).
104. K. Kempa and W.L. Schaich, Phys. Rev. B **37**, 6711 (1988).
105. T. Ando, Z. Phys. B. **26**, 263 (1977).
106. T. Ando, Solid State Commun. **21**, 133 (1977).
107. K.S. Yi and J.J. Quinn, Phys. Rev. B **27**, 2396 (1983).
108. L. Wendler and V.G. Grigoryan, Solid State Commun. **71**, 527 (1989).
109. J. M. Seminario, Chem. Phys. Lett. **206**, 547 (1993).
110. Z.H. Levine and D.C. Allan, Phys. Rev. Lett. **63**, 1719 (1989).
111. See, e.g., *Studies of Electron Distributions in Molecules and Crystals*, Transactions of the American Crystallographic Association, ed. R. Blessing (Polycrystal Book Service, Dayton, OH, 1990), Vol. 26.
112. A.A. Quong and A.G. Eguiluz, Phys. Rev. Lett. **70**, 3955 (1993).
113. N.E. Maddocks, R.W. Godby, and R.J. Needs, Europhys. Lett. **27**, 681 (1994).
114. A. Fleszar, A.A. Quong, and A.G. Eguiluz, Phys. Rev. Lett. **74**, 590 (1995).
115. P.J. Feibelman, Prog. Surf. Sci. **12**, 287 (1982).
116. J.F. Dobson and B.P. Dinte, Phys. Rev. Lett., in press.
117. Y. Andersson, B.I. Lunqvist, and D.C. Langreth, Phys. Rev. Lett. **77**, 2029 (1996).
118. Y. Andersson, E. Hult, H. Rydberg, P. Apell, B.I. Lundqvist, and D.C. Langreth, *Van der Waals interactions in density functional theory*, in *Electronic Density Functional Theory: Recent Progress and New Directions*, eds. J.F. Dobson, G. Vignale, and M.P. Das (Plenum, NY, 1997).
119. G. Senatore and K.R. Subbaswamy, Phys. Rev. A **35**, 2440 (1987).
120. X. Gonze and J.P. Vigneron, Phys. Rev. B **39**, 13120 (1989).
121. A. Theophilou, J. Phys. C **12**, 5419 (1979).
122. N. Hadjisavvas and A. Theophilou, Phys. Rev. A **32**, 720 (1985).
123. W. Kohn, Phys. Rev. A **34**, 5419 (1986).
124. E.K.U. Gross, L.N. Oliveira, and W. Kohn, Phys. Rev. A **37**, 2809 (1988).
125. E.K.U. Gross, L.N. Oliveira, and W. Kohn, Phys. Rev. A **37**, 2805 (1988).
126. L.N. Oliveira, E.K.U. Gross and W. Kohn, Phys. Rev. A **37**, 2821 (1988).
127. A. Nagy, Phys. Rev. A **42**, 4388 (1990).
128. A. Nagy, Phys. Rev. A **49**, 3074 (1994).
129. A. Nagy, Int. J. Quantum Chem. Symp. **29**, (1995).
130. M. Levy, in *Density Functional Theory*, eds. R. Dreizler and E. K. U. Gross, NATO ASI Series (Plenum, New York, 1995).
131. P.K. Chattaraj, S.K. Ghosh, S. Liu, and R.G. Parr, Int. J. Quant. Chem, in press.
132. O. Gunnarsson and B.I. Lundqvist, Phys. Rev. B **13**, 4274 (1976).
133. T. Ziegler, A. Rauk, and E. J. Baerends, Theoret. Chim. Acta **43**, 261 (1977).
134. U. von Barth, Phys. Rev. A **20**, 1693 (1979).
135. A. Görling, Phys. Rev. A **54**, 3912 (1996).
136. C. Filippi, C.J. Umrigar, and X. Gonze, *Excitation energies from density functional perturbation theory*, submitted to J. Chem. Phys., July 1997.
137. *Femtosecond Chemistry*, ed. by J. Manz and L. Wöste (VCH, Weinheim, 1995).
138. F. Zimmermann and W. Ho, Surf. Sci. Reps. **22**, 127 (1995).

Relativistic Density Functional Theory

Reiner M. Dreizler and Eberhard Engel

Institut für Theoretische Physik, J.W.Goethe-Universität Frankfurt,
Robert-Mayer-Str. 8-10, D-60054 Frankfurt/Main, Germany

1 Introduction

A few numbers will illustrate that it is necessary to develop a relativistic version of density functional theory (DFT) [1–6]. The relative relativistic correction

$$\frac{A_{rel} - A_{nrel}}{A_{nrel}}$$

is about 30% for the ionisation potential of the Gold atom, −13% for the bond length of the *AuH* molecule and about 50% for the dissociation energy of this system [7]. The error can even be larger for sensitive quantities like the electron affinity. For this quantity one finds

$$A_{nrel} = 0.10eV \qquad A_{rel} = 0.67eV$$

on the level of the Hartree Fock (HF) approximation [8], and

$$A_{nrel} = 1.02eV \qquad A_{rel} = 2.28eV$$

on the level of the configuration interaction approach [7,9]. Obviously, relativistic and correlation effects modify results on a comparable level.

While it is still possible to deal with relativistic effects in smaller systems in terms of traditional methods, the treatment of larger systems with heavier constituents will require a relativistic extension of DFT (RDFT). One of the questions that has to be answered in this context is: How much of the relativistic correction is due to kinetic effects (replace the nonrelativistic kinetic energy by its relativistic counterpart) and how much is due to exchange-correlation effects? The answer to this and other questions is the goal of our endeavours.

The proper frame for the discussion of RDFT is a suitable field theoretical formulation [10–12] of the problem at hand. For the case of Coulomb systems such a theory is quantum electrodynamics (QED). One might ask: Why not base the discussion on the many-body Dirac equation, as eg. in standard Dirac-Fock calculations?

Here is a more extensive answer. As one knows, Dirac theory leads to a positive definite one particle density

$$n(\underline{x}) = \overline{\psi}(x)\gamma^0\psi(x) = \psi^+(x)\psi(x) \ .$$

On the other hand, the energy spectrum contains a negative energy continuum. In order to prevent disaster, one has to invoke the concept of a filled Dirac sea. Holes in this sea can be associated with antiparticles.

One possibility to handle this situation in an economic fashion is the transition to a field theoretical formulation, which is, for present purposes, indicated for the case of the free particle problem. The free Dirac equation (with $\hbar = c = 1$, as usual) is

$$(-i\partial\!\!\!/ + m)\psi(x) = 0 \quad , \quad \partial\!\!\!/ = \gamma^\mu \partial_\mu = \gamma^\mu \frac{\partial}{\partial x^\mu} \ . \tag{1}$$

A general solution of this equation (a spinor wave packet) is

$$\psi(x) = \int d^3k \sum_{l=1}^{2} \left[b_l(k) u^{(l)}(k) e^{-ik\cdot x} + c_l(k) v^{(l)}(k) e^{ik\cdot x} \right] \ , \ k_0 = \sqrt{\underline{k}^2 + m^2} \ .$$

The notation is

- positive, negative energy spinors $u^{(l)}, v^{(l)}$
- energy-momentum four vector $k = (k_0, \underline{k})$
- Minkowski space scalar product $k \cdot x = k_0 t - \underline{k} \cdot \underline{x}$.

In the process of field quantisation the spinor wave functions are replaced by field operators,

$$\psi(x), \psi^+(x) \quad \longrightarrow \quad \hat{\psi}(x), \hat{\psi}^+(x) \ ,$$

by demanding that canonical, equal time anticommutation relations hold,

$$\{\hat{\psi}_\alpha(\underline{x}, t), \hat{\psi}^+_\beta(\underline{y}, t)\} = \delta_{\alpha\beta}\delta^{(3)}(\underline{x} - \underline{y}) \qquad (\alpha, \beta = 1, \ldots 4) \ . \tag{2}$$

This requires that the Fourier coefficients b_l, c_l be replaced by operators,

$$\begin{array}{ll} b_l(k) \to \hat{b}_l(k) & b_l^*(k) \to \hat{b}_l^+(k) \\ c_l(k) \to \hat{c}_l(k) & c_l^*(k) \to \hat{c}_l^+(k) \ . \end{array}$$

In order to incorporate the concept of the Dirac sea directly, one reinterprets the operators associated with the negative energy solutions as

$$\hat{c}_l(k) \to \hat{d}_l^+(k) \qquad \hat{c}_l^+(k) \to \hat{d}_l(k) \ ,$$

the interpretation being: The destruction of a particle with negative energy corresponds to the creation of an antiparticle with positive energy (and vice versa). This language refers to a vacuum state that is free of particles rather than the filled sea,

$$\hat{b}_l(k)|vac>= 0 \quad , \quad \hat{d}_l(k)|vac>= 0 \ .$$

The corresponding relativistic field operator

$$\hat{\psi}(x) = \int d^3k \sum_{l=1}^{2} \left[\hat{b}_l(k) u^{(l)}(k) e^{-ik\cdot x} + \hat{d}_l^+(k) v^{(l)}(k) e^{ik\cdot x}\right] \tag{3}$$

then describes the destruction of a particle or the creation of an antiparticle. $\hat{\psi}^+(x)$ describes the creation of a particle and the destruction of an antiparticle. Other operators can then be constructed like

(i) the charge operator

$$\hat{Q} = -e \int d^3k \sum_{l=1}^{2} \left[\hat{b}_l^+(k)\hat{b}_l(k) - \hat{d}_l^+(k)\hat{d}_l(k)\right] \tag{4}$$

$$= -e(\hat{N}_{part} - \hat{N}_{antipart}) \ , \tag{5}$$

where infinite but trivial vacuum expectation values have been subtracted. One then notes that $\hat{\psi}(x)$ raises the charge by one unit, while $\hat{\psi}^+(x)$ decreases the charge correspondingly.

(ii) the free Hamiltonian

$$\hat{H} = \int d^3x \, \bar{\hat{\psi}}(x)(-i\underline{\gamma} \cdot \underline{\nabla} + m)\hat{\psi}(x) \tag{6}$$

$$= \int d^3k \sum_{l=1}^{2} [\hat{b}_l^+(k)\hat{b}_l(k) - \hat{d}_l(k)\hat{d}_l^+(k)] \; k_0$$

$$= \int d^3k \sum_{l=1}^{2} [\hat{b}_l^+(k)\hat{b}_l(k) + \hat{d}_l^+(k)\hat{d}_l(k)] \; k_0 + \infty \ , \tag{7}$$

so that one can redefine the trivially renormalised Hamiltonian

$$\hat{H}_R \equiv \hat{H} - < vac|\hat{H}|vac > = \int d^3k \; k_0\{\hat{N}_{part}(k) + \hat{N}_{antipart}(k)\} \, . \tag{8}$$

Contrasting once more energy and charge in Dirac and field theory we note for a free particle (charge $-e$)/antiparticle:

	charge	energy
Dirac theory	negative definite	positive/negative
field theory	negative/positive	positive definite

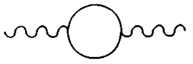
Obviously, the second line corresponds to the experimental solution.

In addition, the field theoretical formulation of the interacting problem (the coupling to the electromagnetic field can be introduced by invoking local $U(1)$ gauge invariance) involves for instance the features:

(i) The interaction between the fermions is described by the photon propagator rather than the Coulomb interaction. In lowest order this propagator describes the exchange of a free photon,

$D^{(0)}_{\mu\nu}(x,y) \longrightarrow$  .

It contains retardation and magnetic effects. One may even consider more exact forms like

$D_{\mu\nu}(x,y) =$  $+$  $+ \cdots$,

involving lowest order vacuum polarisation effects.

(ii) The full machinery of many-body physics [13, 14] can be activated directly, with fermion propagators, Dyson equations etc.

If one accepts these statements as an answer to the question posed, then one can start to look at a selection of the background material that is necessary for the discussion of RDFT. This is done in section 2, which is entitled

2 Assorted Remarks on Vacuum QED, the Relativistic Homogeneous Electron Gas and QED with External Potentials

A system of Dirac particles (charge $-e$) and antiparticles, which interact by the exchange of photons is characterised by the QED Lagrangian density

$$\mathcal{L}_{QED}(x) = \mathcal{L}_e(x) + \mathcal{L}_\gamma(x) + \mathcal{L}_{int}(x) \ . \tag{9}$$

The three terms represent the free Dirac Lagrangian of the fermions

$$\mathcal{L}_e(x) = \frac{i}{2}[\hat{\bar{\psi}}(x)\gamma^\mu\partial_\mu\hat{\psi}(x) - (\partial_\mu\hat{\bar{\psi}}(x))\gamma^\mu\hat{\psi}(x)] - m\hat{\bar{\psi}}(x)\hat{\psi}(x) \ , \tag{10}$$

the free photon Lagrangian

$$\mathcal{L}_\gamma(x) = -\frac{1}{16\pi}\hat{F}_{\mu\nu}(x)\hat{F}^{\mu\nu}(x) - \frac{\lambda}{8\pi}\left(\partial_\nu\hat{A}^\nu(x)\right)^2 \tag{11}$$

and the interaction term

$$\mathcal{L}_{int}(x) = -e\hat{j}^\nu(x)\hat{A}_\nu(x) \ . \tag{12}$$

The following explanatory remarks are necessary.

(i) $\hat{\psi}(x)$ and $\hat{A}_\mu(x)$ are the fermion and photon field operators, the actual electromagnetic field is characterised by the field tensor

$$\hat{F}_{\mu\nu}(x) = \partial_\mu\hat{A}_\nu(x) - \partial_\nu\hat{A}_\mu(x) \ . \tag{13}$$

$\hat{j}^\nu(x)$ is the fermion four current. It is most conveniently specified in the commutator form

$$\hat{j}^\nu(x) = \frac{1}{2}[\hat{\bar{\psi}}(x), \gamma^\nu\hat{\psi}(x)] \ , \tag{14}$$

which explicitly expresses charge conjugation invariance.

(ii) The quantisation of the free photon problem is more complicated due to the transversality of the photon field. The form of the Lagrangian specified refers to a covariant gauge (and thus involves the so-called gauge fixing term). Standard gauges are the Feynman gauge ($\lambda = 1$) and the Landau gauge ($\lambda = \infty$). The photon sector of the Fock space is supposed to be specified in terms of the Gupta-Bleuler indefinite metric quantisation. We will, however, not display any details concerning this point.

Essentially, all results available for (vacuum) QED systems are based on perturbation theory. The basic quantities used in this context are the fermion and the photon propagators (where T denotes time-ordering)

$$S_{F,vac}(x,y)_{\alpha\beta} = -i < vac|T\hat{\psi}_\alpha(x)\hat{\bar{\psi}}_\beta(y)|vac > \tag{15}$$

$$= \int \frac{d^4p}{(2\pi)^4}\, S_{F,vac}(p)_{\alpha\beta}\; e^{-ip\cdot(x-y)} \tag{16}$$

$$D_{\mu\nu,vac}(x,y) = -ie^2 < vac|T\hat{A}_\mu(x)\hat{A}_\nu(y)|vac > \tag{17}$$

$$= \int \frac{d^4q}{(2\pi)^4} D_{\mu\nu,vac}(q)\; e^{-iq\cdot(x-y)} \ . \tag{18}$$

The fact that both quantities depend only on the difference of the Minkowski coordinates is an expression of the translational invariance of the theory.

In lowest order, these propagators can be calculated directly. For the fermion propagator (describing the time development of a free fermion forward and of a free anti-fermion backward in time) one finds

$$S^{(0)}_{F,vac}(p)_{\alpha\beta} = \left(\frac{\not p + m}{p^2 - m^2 + i\epsilon}\right)_{\alpha\beta} \ . \tag{19}$$

In diagrammar this quantity will be denoted by

$$S^{(0)}_{F,vac}(p)_{\alpha\beta} = \alpha \xrightarrow{p} \beta \ . \tag{20}$$

The photon propagator depends on the choice of gauge. In Landau gauge one finds

$$D^{(0)}_{\mu\nu}(q) = -\frac{4\pi e^2}{q^2 + i\epsilon}\left\{g_{\mu\nu} - \frac{q_\mu q_\nu}{q^2}\right\} \ , \tag{21}$$

the second term in the bracket is not present in Feynman gauge. In diagrammar we represent this quantity by

$$iD^{(0)}_{\mu\nu}(q) = \mu \overset{q}{\sim\!\sim\!\sim\!\sim} \nu \ . \tag{22}$$

We shall later use the decomposition of this propagator into a nonretarded Coulomb (longitudinal) and a transverse contribution,

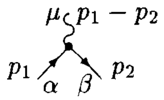
$$D^{(0)}_{\mu\nu}(q) = D^{(0)L}_{\mu\nu}(q) + D^{(0)T}_{\mu\nu}(q) \ , \tag{23}$$

with the longitudinal part

$$D^{(0)L}_{\mu\nu}(q) = g_{0\mu}g_{0\nu}\frac{4\pi e^2}{\underline{q}^2} \quad . \tag{24}$$

The link between the propagators is provided by the vertex, which describes the emission and the absorption of a photon by a fermion. In momentum space, the lowest order vertex just corresponds to a γ-matrix provided one implies four momentum conservation

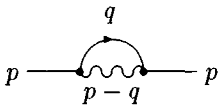
$$\Gamma^{(0)}_{\mu}(p_1,p_2)_{\alpha\beta} = \gamma_\mu \longrightarrow$$



$\cdot \quad (25)$

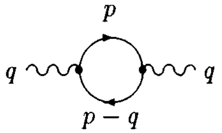
In the next order of perturbation theory (characterised by an additional free photon line), the contributions are

$$S^{(1)}_{F,vac}(p) =$$



(26)

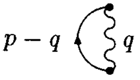
$$iD^{(1)}_{\mu\nu,vac}(q) =$$



$\cdot \quad (27)$

From these diagrams one extracts the lowest order contributions to the basic two point functions of the theory: the electron selfenergy,

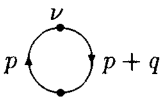
$$\Sigma^{(1)}_{vac}(p) =$$



$$= i\int\frac{d^4q}{(2\pi)^4}\, D^{(0)}_{\mu\nu}(q)\gamma^\mu S^{(0)}_{F,vac}(p-q)\gamma^\nu \quad , \tag{28}$$

and the vacuum polarisation,

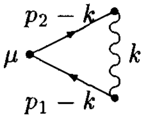
$$-i\Pi^{(0)}_{\mu\nu,vac}(q) =$$



$$= -\int\frac{d^4p}{(2\pi)^4}\, tr\left[\gamma_\mu S^{(0)}_{F,vac}(p)\gamma_\nu S^{(0)}_{F,vac}(p-q)\right] . \tag{29}$$

In addition, one finds for the next order contribution to the vertex function (a three point function)

$$\Gamma^{(1)}_{\mu,vac}(p_1,p_2) =$$



$$= i\int \frac{d^4k}{(2\pi)^4}\, D^{(0)}_{\rho\nu}(k)\gamma^\rho S^{(0)}_{F,vac}(p_1-k)\gamma_\mu S^{(0)}_{F,vac}(p_2-k)\gamma^\nu \ . \qquad (30)$$

If one evaluates the integrals over the loop momenta, one finds that neither of the three quantities $\Sigma^{(1)}, \Pi^{(0)}$ and $\Gamma^{(1)}$ is well defined. One encounters UV-divergencies (divergencies due to large values of the loop momenta). One way to deal with these divergencies and to endow the three quantities with a meaning is the counterterm technique based on dimensional regularisation [15]. In this method one first evaluates the integrals indicated in $(1, d-1)$-dimensional Minkowski space (rather than in the (1,3)-dimensional space). These results can then be analytically continued to noninteger d, so that the physically interesting limit $d \to 4$ can be taken. The results (involving an expansion about this limit) are

$$\Sigma^{(1)}_{vac}(p) = \frac{e^2}{16\pi^2}\Gamma(s)(-\not{p} + 4m) + \Sigma^{(1)}_{vac,finite}(p) \qquad (31)$$

$$\Pi^{(0)}_{\mu\nu,vac}(q) = (q^2 g_{\mu\nu} - q_\mu q_\nu)\left\{\frac{\Gamma(s)}{12\pi^2} + \omega^{(0)}_{vac,finite}(q^2)\right\} \qquad (32)$$

$$\Gamma^{(1)}_{\mu,vac}(p_1,p_2) = \frac{e^2}{16\pi^2}\Gamma(s)\gamma_\mu + \Gamma^{(1)}_{\mu,vac,finite}(p_1,p_2) \ . \qquad (33)$$

The divergent part of the loop integrals is isolated in each case in the Γ-function

$$\Gamma(s) \equiv \Gamma\left(\frac{1}{2}(4-d)\right) \xrightarrow[s\to 0]{} \frac{1}{s} + \ \dots \ . \qquad (34)$$

For the finite remainders, the limit $d \to 4$ can be taken directly. The rather lengthy expressions that result are not of interest for the moment.

Two additional features might be of interest though: The tensorial structure of the vacuum polarisation is a consequence of the general gauge invariance,

$$q^\mu\ \Pi_{\mu\nu}(q) = 0 \ . \qquad (35)$$

Another manifestation of gauge invariance is the Ward-Takahashi identity,

$$(p_\mu - p'_\mu)\Gamma^\mu_{vac}(p,p') = \Sigma_{vac}(p') - \Sigma_{vac}(p) \ , \ , \qquad (36)$$

which is satisfied by the divergent (as well as the finite) contributions to Σ and Γ.

The next step, the actual renormalisation, amounts to a redefinition of the physical constants and operators in the initial Lagrangian, leaving the physical implications of this Lagrangian unmodified. The renormalisation procedure can readily be demonstrated (on the one loop level) for the case of mass renormalisation. Using the Dyson equation

$$S_{F,vac}(p)^{-1} = S^{(0)}_{F,vac}(p)^{-1} - \Sigma_{vac}(p) = \not p - m - \Sigma_{vac}(p) \qquad (37)$$

one finds to first order

$$S^{(1)}_{F,vac}(p)^{-1} = \not p\left(1 + \frac{e^2}{16\pi^2}\Gamma(s)\right) - m\left(1 + \frac{e^2}{4\pi^2}\Gamma(s)\right) - \Sigma^{(1)}_{vac,finite}(p) \ . \qquad (38)$$

In order to compensate the terms that diverge in the limit $d \to 4$, one adds to the "unrenormalised Lagrangian" a "counterterm Lagrangian" of the same form.

The unrenormalised Lagrangian has the form (9) originally given with physical constants (m, e) and physical field operators. The counterterm Lagrangian for (fermion) mass renormalisation is

$$\mathcal{L}_{e,CT} = \hat{\bar\psi}(x)(iA\not\partial - B)\hat\psi(x) \ . \qquad (39)$$

The renormalised fermion Lagrangian is thus

$$\begin{aligned}\mathcal{L}_{e,R} &= \mathcal{L}_{e,unren} + \mathcal{L}_{e,CT} \\ &= \hat{\bar\psi}(x)\big(i(1+A)\not\partial - (m+B)\big)\hat\psi(x) \ . \end{aligned} \qquad (40)$$

Evaluating the electron propagator on the one loop level as before (with the same $\mathcal{L}_{int}$) one finds

$$\begin{aligned} S_{F,vac,R}(p)^{-1} &= \not p\left(1 + A + \frac{e^2}{16\pi^2}\Gamma(s)\right) \\ &\quad -m\left(1 + \frac{B}{m} + \frac{e^2}{4\pi^2}\Gamma(s)\right) - \Sigma^{(1)}_{vac,finite}(p) \ . \end{aligned} \qquad (41)$$

The obvious choice

$$A = -\frac{e^2}{16\pi^2}\Gamma(s) \quad ; \quad B = -\frac{e^2 m}{4\pi^2}\Gamma(s) \qquad (42)$$

leads to a finite result, in diagrammar,

The additional diagram represents the counterterm contribution

$$-\!\!\times\!\!- \;=\; S^{(0)}_{F,vac}(p)[-A\not p + B]S^{(0)}_{F,vac}(p) \ . \qquad (43)$$

The form invariance of the Lagrangian under this renormalisation procedure is then implemented by defining the wavefunction (actually field operator) renormalisation

$$\hat\psi_{bare}(x) = [1+A]^{1/2} \ \ \hat\psi(x) \qquad (44)$$

and the bare mass

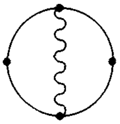
$$m_{bare} = m - \delta m \tag{45}$$

(to the given order), so that

$$\mathcal{L}_{e,R} = \hat{\bar{\psi}}_{bare}(x)(i\partial\!\!\!/ - m_{bare})\hat{\psi}_{bare}(x) \ , \tag{46}$$

that is the counterterms necessary to keep $\Sigma^{(1)}$ finite have been completely absorbed in a redefinition of the ingredients of the (free) fermion Lagrangian.

An equivalent procedure can be carried out for $\Pi^{(0)}_{\mu\nu,vac}$ addressing $\mathcal{L}_\gamma$, as well as for $\Gamma^{(1)}_{\mu,vac}$ addressing $\mathcal{L}_{int}$. The argumentation can be extended to any order, after a discussion of overlapping divergencies, that occur for instance in the diagram



We shall return to some aspects of the renormalisation problem when we talk about the derivation of explicit relativistic functionals.

We then leave vacuum QED and discuss briefly the simplest relativistic many fermion problem: The **Relativistic Homogeneous Electron Gas** (RHEG). This system is characterised by a ground state that contains N electrons per volume V,

$$n_0 = \frac{N}{V} \ . \tag{47}$$

The relativistic version is essentially the same model as the nonrelativistic equivalent with the difference that the kinetic energy is replaced by the relativistic form and the Coulomb interaction is replaced by the exchange of photons.

The fermion propagator reads

$$S_F(x,y) = -i \ < \phi_0|T(\hat{\psi}(x)\hat{\bar{\psi}}(y))|\phi_0> \ , \tag{48}$$

with $|\phi_0>$ denoting the ground state of the RHEG. This differs from the vacuum propagator $S_{F,vac}$ already in lowest order as $S_{F,vac}$ describes free electrons and positrons while S_F of the RHEG describes electrons and positrons embedded in a medium.

The result of an explicit, straightforward evaluation can be expressed in the form

$$S^{(0)}_F(p) = S^{(0)}_{F,vac}(p) + S^{(0)}_{F,D}(p) \tag{49}$$

with

$$S^{(0)}_{F,D}(p) = 2\pi i \ \delta(p^0 - E_p) \ \frac{p\!\!\!/ + m}{2E_p} \ \Theta(k_F - |\underline{p}|) \tag{50}$$

$$E_p = [\underline{p}^2 + m^2]^{1/2} \quad ; \quad k_F = [3\pi^2 n_0]^{1/3} \ . \tag{51}$$

The quantity in Eq.(50) is the contribution to the propagator due to finite density of real electrons. An alternative decomposition of the propagator in lowest order can be given,

$$S_F^{(0)}(p) = S_{F,-}^{(0)}(p) + S_{F,+}^{(0)}(p) \ . \tag{52}$$

This decomposition contains a contribution due to the electrons,

$$S_{F,+}^{(0)}(p) = \frac{(\not{p}_+ + m)}{2E_p}\left[\frac{\Theta(|\underline{p}| - k_F)}{p^0 - E_p + i\epsilon} + \frac{\Theta(k_F - |\underline{p}|)}{p^0 - E_p - i\epsilon}\right] , \tag{53}$$

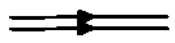
with

$$p_\pm^\mu = (\pm E_p, \underline{p}) \ , \tag{54}$$

(which has some resemblance to the corresponding nonrelativistic propagator) and a positron contribution,

$$S_{F,-}^{(0)}(p) = \frac{(\not{p}_- + m)}{2E_p}\left(\frac{-1}{p^0 + E_p - i\epsilon}\right) \ . \tag{55}$$

The lowest order fermion propagator of the RHEG will in the following be denoted by a double line

$S_F^{(0)}(p) =$  . (56)

The other two basic elements of perturbation theory, the free photon propagator and the simple vertex, remain unchanged. On the other hand, the full photon propagator, which in this case is given by

$$D_{\mu\nu}(x,y) = -ie^2 < \phi_0|T(\hat{A}_\mu(x)\hat{A}_\nu(y))|\phi_0 > \ , \tag{57}$$

and the full vertex differ from the vacuum QED result. The reason is that not only $e^+ - e^-$ pairs screen the bare interaction but also virtual electron-hole pairs.

The discussion of the ground state energy of the RHEG is the basis of the local density approximation (LDA), as in the nonrelativistic case. We shall discuss the necessary details in connection with the explicit consideration of this limit in section 4. We shall see there that the renormalisation procedure, which is also necessary for this system, concerns essentially only the vacuum aspects involved.

In the discussion of inhomogeneous systems (like atoms, molecules etc.) the fermions have to be subjected to an additional external field, which represents the fixed nuclei and/or any other external perturbation present. In this case the original QED Lagrangian has to be supplemented by an external interaction term,

$$\mathcal{L}(x) = \mathcal{L}_{QED}(x) + \mathcal{L}_{ext}(x) \tag{58}$$

$$\mathcal{L}_{ext}(x) = -e\hat{j}^\nu(x)V_\nu(x) \ . \tag{59}$$

For the discussion of stationary systems the external four vector is static,

$$V_\mu(x) \equiv V_\mu(\underline{x}) \ . \tag{60}$$

It will be denoted in diagrammar as

$$V \longrightarrow \quad . \tag{61}$$

One can then show (with the aid of Noether's theorem if necessary) that the corresponding energy is conserved, so that the Hamiltonian is given by

$$\begin{aligned}\hat{H} &= \int d^3x \, \hat{T}^{00}(x) & (62)\\ &= \int d^3x \, \bar{\hat{\psi}}(x) \, (-i\underline{\gamma}\cdot\underline{\nabla} + m)\hat{\psi}(x) & (63)\\ &- \frac{1}{8\pi}\int d^3x \Big\{\partial^0 \hat{A}_\nu(x)\partial^0\hat{A}^\nu(x) + \underline{\nabla}\hat{A}_\nu(x)\cdot\underline{\nabla}\hat{A}^\nu(x)\Big\}\\ &+ e\int d^3x \, \hat{j}^\mu(x)\hat{A}_\mu(x) + e\int d^3x \, \hat{j}^\mu(x)V_\mu(\underline{x}) \ .\end{aligned}$$

The renormalised Hamiltonian has the form

$$\hat{H}_R = \hat{H} - VEV + CT \ , \tag{64}$$

involving the subtraction of the vacuum expectation value and the addition of counterterms analoguous to the counterterms of the Langrangian.

3 Foundations

The first topic is the extension of the Hohenberg-Kohn (HK) theorem [16] to the case of relativistic QED systems.

3.1 The relativistic Hohenberg-Kohn theorem

This extension was first formulated by Rajagopal and Callaway [17,18] (and by MacDonald and Vosko [19]). As expected for a relativistically covariant situation, the theorem states that the ground state energy is a unique functional of the ground state four current,

$$E_0[j^\mu] = F[j^\mu] + \int d^3x \, j^\mu(\underline{x})V_\mu(\underline{x}) \ , \tag{65}$$

with F being a universal functional of j^μ. There are, nonetheless, a number of points that should be discussed more closely.

The arguments used in the original proof are based on QED, but the question of possible divergencies was ignored. As the proof relies (as in the nonrelativistic case) on the celebrated "reductio ad absurdum", one has to make sure that well defined, finite quantities for the ground state energy and the four current are involved [20],

$$E_0 = < \Phi_0|\hat{H}|\Phi_0 > - < vac\,|\hat{H}_e + \hat{H}_\gamma + \hat{H}_{int}|\,vac > + \Delta E_{0,CT} \quad (66)$$

$$j^\mu(\underline{x}) = < \Phi_0|\hat{j}^\mu(\underline{x})|\Phi_0 > + \Delta j^\mu(\underline{x}) \ , \quad (67)$$

that is quantities including all vacuum subtractions and counterterms (Δ). In addition, one has to make sure that the structure of the counterterms does not invalidate the proof. With these precautions, one can establish the usual chain of unique maps

$$\left\{V_\nu \middle| V_\nu + \partial_\nu \Lambda\right\} \Longleftrightarrow \left\{|\Phi_0> \,\middle|\, \text{with } |\Phi_0> \text{ from } V_\nu + \partial_\nu \Lambda\right\} \Longleftrightarrow j^\nu(\underline{x}) \ . \quad (68)$$

There exists a one-to-one correspondence between the class of external potentials differing by a static gauge transformation and the corresponding class of ground states. This class of physically equivalent ground states is uniquely determined by the ground state four current. As a consequence one can express all ground state observables (including the energy) as a unique functional of the ground state four current

$$O[j^\mu] =< \Phi_0[j^\mu] \mid \hat{O} \mid \Phi_0[j^\mu] > + \Delta O - VEV \ . \quad (69)$$

As a side remark one may note, that both the ground state energy and the four current are gauge invariant quantities.

For the case of a purely electrostatic, external potential,

$$V^\mu(\underline{x}) = \left\{V^0(\underline{x}), \underline{0}\right\} \ , \quad (70)$$

the proof can be repeated using just the zeroth component of the four current (i.e. the charge density) $j^0(\underline{x}) \equiv n(\underline{x})$ [21]. The ground state energy and all variables are then functionals of the density alone. It should be emphasised that this does not imply that the spatial components of the current vanish in this case. It only implies that the three current has to be interpreted as a functional of $n(\underline{x})$,

$$\underline{j}(\underline{x}) =< \Phi_0[n] \mid \hat{\underline{j}}(\underline{x}) \mid \Phi_0[n] > = \ \underline{j}([n], \underline{x}) \ . \quad (71)$$

With the Ritz principle, already involved in the details of the proof of the HK theorem, one may formulate the basic variational principle of RDFT as

$$\frac{\delta}{\delta j^\nu(\underline{x})} \left\{ E_0[j^\nu] - \mu \int d^3y \, j^0(\underline{y}) \right\} = 0 \ . \quad (72)$$

The subsidiary condition implies charge (rather than number) conservation.

This variational principle is utilised directly in relativistic extended Thomas-Fermi (RETF) models [22,23,20], in which an approximate density functional representation of $E_0[j^\nu]$ is specified. The mainstay of applications are, however, the

3.2 Relativistic Kohn-Sham equations

The first step necessary to set up this scheme [24] is the representation of the exact four current of the interacting system in terms of auxiliary spinors [20],

$$j^{\nu}(\underline{x}) = j^{\nu}_{vac}(\underline{x}) + j^{\nu}_{D}(\underline{x}) \ . \tag{73}$$

The vacuum polarisation current j^{ν}_{vac} is given as [25]

$$j^{\nu}_{vac}(\underline{x}) = \frac{1}{2}\left\{ \sum_{\epsilon_k \leq -m} \overline{\varphi}_k(\underline{x})\gamma^{\nu}\varphi_k(\underline{x}) - \sum_{-m<\epsilon_k} \overline{\varphi}_k(\underline{x})\gamma^{\nu}\varphi_k(\underline{x}) \right\} , \tag{74}$$

where we do not specify a necessary counterterm in this equation and in the following. The current due to occupied discrete orbitals is

$$j^{\nu}_{D}(\underline{x}) = \sum_{-m<\epsilon_k\leq\epsilon_F} \overline{\varphi}_k(\underline{x})\gamma^{\nu}\varphi_k(\underline{x}) \ . \tag{75}$$

This result follows directly from the charge conjugation invariant form of the four current operator (which has been specified beforehand) if one considers a system of noninteracting particles.

In the same vein one defines the noninteracting kinetic energy (including a mass term)

$$T_s[j^{\nu}] = T_{s,vac}[j^{\nu}] + T_{s,D}[j^{\nu}] \ , \tag{76}$$

$$T_{s,vac}[j^{\nu}] = \frac{1}{2}\int d^3x \left\{ \sum_{\epsilon_k \leq -m} \overline{\varphi}_k(\underline{x})\Big[-i\underline{\gamma}\cdot\underline{\nabla} + m\Big]\varphi_k(\underline{x}) \right. \tag{77}$$

$$\left. - \sum_{-m<\epsilon_k} \overline{\varphi}_k(\underline{x})\Big[-i\underline{\gamma}\cdot\underline{\nabla} + m\Big]\varphi_k(\underline{x}) \right\} ,$$

$$T_{s,D}[j^{\nu}] = \int d^3x \sum_{-m<\epsilon_k\leq\epsilon_F} \overline{\varphi}_k(\underline{x})\Big[-i\underline{\gamma}\cdot\underline{\nabla} + m\Big]\varphi_k(\underline{x}) \tag{78}$$

$T_{s,vac}$ is the kinetic contribution to the so-called Casimir energy [11] (again no counterterms displayed), while $T_{s,D}$ results from bound real electrons.

Adding and subtracting T_s as well as the covariant Hartree energy,

$$E_H[j^{\nu}] = \frac{1}{2}\int d^3x \int d^4y \; j^{\mu}(\underline{x}) \; D^{(0)}_{\mu\nu}(x-y) \; j^{\nu}(\underline{y}) \ , \tag{79}$$

for time independent currents

$$= \frac{e^2}{2}\int d^3x \int d^3y \; \frac{j^{\nu}(\underline{x}) \; j_{\nu}(\underline{y})}{|\underline{x}-\underline{y}|} \ , \tag{80}$$

to the total energy, one can rearrange the ground state energy as

$$E_0[j^\mu] = T_s[j^\mu] + E_{ext}[j^\mu] + E_H[j^\mu] + E_{xc}[j^\mu] \ , \tag{81}$$

with the exchange-correlation energy

$$E_{xc}[j^\mu] = F[j^\mu] - T_s[j^\mu] - E_H[j^\mu] \tag{82}$$

as usual.

Minimising E_0 with respect to the auxiliary orbitals leads to the most general relativistic Kohn-Sham (KS)-equations

$$\gamma^0 \Big\{ -i\underline{\gamma} \cdot \underline{\nabla} + m + e\not{V}(\underline{x}) + \not{v}_H(\underline{x}) + \not{v}_{xc}(\underline{x}) \Big\} \varphi_k(\underline{x}) = \epsilon_k \varphi_k(\underline{x}) \ , \tag{83}$$

with the effective potentials

$$v_H^\nu(\underline{x}) = e^2 \int d^3y \, \frac{j^\nu(\underline{y})}{|\underline{x} - \underline{y}|} \tag{84}$$

$$v_{xc}^\nu(\underline{x}) = \frac{\delta E_{xc}[j^\mu]}{\delta j_\nu(\underline{x})} \ . \tag{85}$$

This set of equations has to be solved self-consistently, leading in principle to the exact $j^\mu(\underline{x})$ and hence energies etc.

A glance at the ingredients of the general relativistic KS-scheme reveals a problem of considerable difficulty. For instance, the evaluation of the vacuum contributions (in j^μ and T_s) requires summation over all negative and positive energy solutions (as well as renormalisation) in each step of the self-consistency procedure. All effective potentials involved are endowed with a Minkowski space structure.

Fortunately, for practical electronic structure calculations two approximations are possible:

1. The most important simplification arises from the **no-sea approximation**, in which all radiative corrections are neglected

$$j^\mu_{vac} = 0 \quad T_{s,vac} = 0 \quad E_{xc,vac} = 0 \ . \tag{86}$$

   If necessary, one may evaluate these contributions perturbatively, that is after self-consistency has been achieved without these terms. This approximation should be useful for all systems of interest, with the exception possibly of super-heavy atoms.
2. For the case that the external potential is purely **electrostatic** (a situation commonly encountered in electronic structure calculations), the charge density is the only relevant variable, for instance via

$$\tilde{E}_H[n] = E_H[n, \underline{j}[n]] \quad , \quad \tilde{E}_{xc}[n] = E_{xc}[n, \underline{j}[n]] \ . \tag{87}$$

   As a consequence the Hartree and the xc–potentials only consist of a time-like component (rather than a four vector structure), so e.g.

$$v^{\mu}_{H}(\underline{x}) = (\tilde{v}_H(\underline{x}), \underline{0}) \tag{88}$$

with

$$\tilde{v}_H(\underline{x}) = \frac{\delta \tilde{E}_H[n]}{\delta n(\underline{x})} = \frac{\delta E_H[j^{\nu}]}{\delta n(\underline{x})} + \sum_k \int d^3x' \, \frac{\delta E_H[j^{\nu}]}{\delta j^k(\underline{x}')} \frac{\delta j^k(\underline{x}')}{\delta n(\underline{x})} \tag{89}$$

as there is an explicit functional dependence of $\underline{j}$.

The resulting **electrostatic, no-sea KS-equations** are (all ˜ are dropped)

$$\Big\{ - i\underline{\alpha}\cdot\underline{\nabla} + \beta m + V_{ext}(\underline{x}) + v_H(\underline{x}) + v_{xc}(\underline{x})\Big\}\varphi_k(\underline{x}) = \epsilon_k \varphi_k(\underline{x}) \, , \tag{90}$$

where the density and the three current are given by

$$n(\underline{r}) = \sum_{-m<\epsilon_k\leq\epsilon_F} \varphi_k^{+}(\underline{x})\varphi_k(\underline{x}) \tag{91}$$

$$\underline{j}(\underline{r}) = \sum_{-m<\epsilon_k\leq\epsilon_F} \varphi_k^{+}(\underline{x})\underline{\alpha}\,\varphi_k(\underline{x}) \ . \tag{92}$$

One should note that the exact current $\underline{j}[n]$ has been replaced by the KS current $\underline{j}(\underline{r})$. The KS current is not necessarily identical with the (unknown) functional $\underline{j}[n]$, but expected to be an acceptable approximation. In any case, possible differences that arise (eg. from inserting $\underline{j}[n]$ in the transverse Hartree energy) are absorbed in a redefinition of E_{xc}.

One may also decompose the interaction mediated by the free photon propagator into a Coulomb (longitudinal) and a transverse part (as discussed before)

$$D^{(0)}_{\mu\nu}(x-y) = g_{\mu 0} g_{\nu 0} \frac{e^2}{\mid \underline{x} - \underline{y} \mid} \, \delta(x^0 - y^0) + D^{T}_{\mu\nu}(x-y) \ . \tag{93}$$

If one neglects the transverse part, one obtains the longitudinal limit of RDFT, which corresponds to the use of the Dirac-Coulomb Hamiltonian, a kind of standard in quantum chemistry. Inclusion of the transverse term recovers the retardation and magnetic effects, which are usually included in a weakly relativistic limit in the form of the Dirac-Coulomb-Breit (DCB) Hamiltonian.

One of the problems often encountered in the application of DFT is to ensure that self-interaction effects contained in the Hartree term are properly cancelled by the x-energy functional. This problem can be handled if one defines the x-energy in terms of KS-orbitals. For this purpose one starts with the definition of the covariant exchange energy

$$E_x = \frac{1}{2} \int d^3x \int d^4y \; D^{(0)}_{\mu\nu}(x-y) \; tr\Big[S_F(x,y) \, \gamma^{\nu} \; S_F(y,x) \, \gamma^{\mu}\Big] \ . \tag{94}$$

If one evaluates the propagator in the KS-picture,

$$S_F \to S^{KS} = -i\Big\{ \Theta(x^0 - y^0) \sum_{\epsilon_n > \epsilon_F} \varphi_n(\underline{x})\, \overline{\varphi}_n(\underline{y}) exp[-i\epsilon_n(x^0 - y^0)] \qquad (95)$$

$$- \Theta(y^0 - x^0) \sum_{\epsilon_F \geq \epsilon_n} \varphi_n(\underline{x})\overline{\varphi}_n(\underline{y}) exp[-i\epsilon_n(x^0 - y^0)] \Big\} ,$$

and uses the Feynman gauge for the photon propagator, one obtains in the electrostatic, no-sea approximation the relativistic Fock term

$$E_x^{KS}[n] = -\frac{e^2}{2} \int d^3x \int d^3y \sum_{-m<\epsilon_k,\epsilon_l\leq\epsilon_F} \frac{\cos(\omega_{kl}|\underline{x} - \underline{y}|)}{|\underline{x} - \underline{y}|} \qquad (96)$$

$$\times \overline{\varphi}_k(\underline{x})\gamma_\mu \varphi_l(\underline{x})\, \overline{\varphi}_l(\underline{y})\gamma^\mu \varphi_k(\underline{y}) ,$$

with

$$\omega_{kl} = |\ \epsilon_k - \epsilon_l\ | . \qquad (97)$$

The functional dependence of E_x^{KS} on n arises via the functional dependence of the KS-orbitals on n. The correlation energy is then

$$E_c^{KS} = E_{xc} - E_x^{KS} . \qquad (98)$$

The x-potential corresponding to E_x^{KS} can not be evaluated directly via

$$v_x^{KS}(\underline{x}) = \frac{\delta E_x^{KS}[n]}{\delta n(\underline{x})} . \qquad (99)$$

In order to obtain this quantity, one has to activate the relativistic extension of the

3.3 Optimised potential method

The optimised potential method (OPM) [26–30,21] relies on the fact that the functional derivative of an energy expression with respect to the density can be evaluated with the aid of the chain rule

$$\frac{\delta E}{\delta n(\underline{x})} = \sum_k \int d^3x' d^3x'' \frac{\delta E}{\delta\varphi_i(\underline{x}')} \frac{\delta\varphi_i(\underline{x}')}{\delta v_{KS}(\underline{x}'')} \frac{\delta v_{KS}(\underline{x}'')}{\delta n(\underline{x})} + c.c. , \qquad (100)$$

if the dependence on the density is implicit via the orbitals. The quantity $\delta v_{KS}(\underline{x}')/\delta n(\underline{x})$ is the inverse KS response function. For the functional derivative of the orbitals with respect to the potential an explicit result can be derived from the KS equations (which also allows the direct evaluation of the KS response function)

$$\frac{\delta\varphi_k(\underline{x})}{\delta v_{KS}(\underline{y})} - G_k(\underline{x}, \underline{y})\varphi_k(\underline{y}) \qquad (101)$$

with

$$G_k(\underline{x},\underline{y}) = \sum_{-m<\epsilon_l<m,l\neq k} \frac{\varphi_l(\underline{x})\varphi_l^+(\underline{y})}{(\epsilon_l-\epsilon_k)} \; . \tag{102}$$

Finally, the functional derivative of the energy with respect to the orbitals is known if the energy is specified as $E = E[\varphi_i, \bar{\varphi}_i]$.

The standard procedure can readily be applied to the exchange term (analoguous to the original derivation given by Talman and Shadwick [27]) and leads to the ROPM integral equation for the x-potential

$$\int d^3x' K(\underline{x},\underline{x}')v_x(\underline{x}') = Q(\underline{x}) \; , \tag{103}$$

with

$$K(\underline{x},\underline{x}') = \sum_{-m<\epsilon_K\leq\epsilon_F} \varphi_k^+(\underline{x})G_k(\underline{x},\underline{x}')\varphi_k(\underline{x}') + c.c. \tag{104}$$

$$Q(\underline{x}) = \sum_{-m<\epsilon_k\leq\epsilon_F} \int d^3x' \; \varphi_k^+(\underline{x})G_k(\underline{x},\underline{x}')\frac{\delta E_x^{KS}[n]}{\delta\varphi_k^+(\underline{x}')} + c.c. \; , \tag{105}$$

which corresponds to the electrostatic, no-sea limit. A covariant extension can readily be derived.

This integral equation has to be solved self-consistently together with the KS-orbital equations. In this fashion one establishes the functional relation between v^{KS} and n, implicitly. One advantage of the OP-method is (as advertised) the fact that self-interaction effects are cancelled correctly. If one adjusts a trivial constant, so that

$$v_x(r\rightarrow\infty) = 0 \tag{106}$$

one finds for finite systems the asymptotic behaviour

$$v_x^{ROPM}(r\rightarrow\infty) = -\frac{1}{r} \; . \tag{107}$$

The OP-procedure can be applied for the full x-energy (longitudinal as well as transverse). It produces spinor solutions that do not depend on the gauge of the free photon propagator $D_{\mu\nu}^{(0)}$, justifying in retrospect the use of the Feynman gauge in the definition (By contrast gauge problems arise for the transverse, nonlocal, orbital-dependent Dirac-Fock exchange).

As the three current $\underline{j}(\underline{x})$ is a trivial functional of the orbitals and the orbitals are functionals of the density, the procedure establishes (indirectly) a functional dependence of the form

$$\underline{j}^{KS,OPM} = \underline{j}^{KS}[n] \; . \tag{108}$$

The self-consistent OPM-procedure is much more involved than the direct KS-scheme. For this reason a search for some shortcuts seems mandatory. This aspect as well as a valuation of numerical results will follow later.

As a final point of this section on fundamentals, we take a brief look at the weakly relativistic limit of the theory and the connection with nonrelativistic current density functional theory [31–33].

3.4 Weakly relativistic limit

The standard weakly relativistic limit of the QED Hamiltonian can be obtained with techniques as the low order Foldy-Wouthuysen transformation or by direct expansion. The results (all constants reinstated in this case) is the Pauli-type Hamiltonian

$$\hat{H}_P = \int d^3x\, \hat{\varphi}^+(\underline{x})\left\{ \frac{1}{2m}\left[(-i\hbar\underline{\nabla})^2 + 2i\hbar\,\frac{e}{c}\,\underline{V}(\underline{x})\cdot\underline{\nabla} + \frac{e^2}{c^2}\,\underline{V}(\underline{x})^2\right] - \frac{e\hbar}{2mc}\,\underline{\sigma}\cdot\left(\underline{\nabla}\times\underline{V}(\underline{x})\right) + eV_0(\underline{x})\right\}\hat{\varphi}(\underline{x}) + \hat{H}_{ee} \;. \tag{109}$$

The notation implies

- $\hat{\varphi}(x)$ is a nonrelativistic field operator with a two component structure
- $\underline{\sigma}$ are the Pauli matrices
- $\hat{H}_{ee}$ represents the standard Coulomb interaction as the limit of the relativistic $e^- - e^-$ interaction.

For the further discussion it is relevant to note that the gauge term $\frac{e^2}{c^2}\,\underline{V}^2$ is of order $1/c^2$, while the other terms are at most of order $1/c$.

The weakly relativistic limit of the three current operator can be extracted with the same techniques leading to

$$\hat{\underline{j}}(\underline{x}) = \hat{\underline{j}}_p(\underline{x}) - \frac{e}{c}\,\underline{\nabla}\times\hat{\underline{m}}(\underline{x}) - \frac{e}{mc}\,\underline{V}(\underline{x})\hat{n}(\underline{x}) \;. \tag{110}$$

It contains the paramagnetic current operator

$$\hat{\underline{j}}_p(\underline{x}) = -\frac{i\hbar}{2m}\left[\hat{\varphi}^+(\underline{x})\left(\underline{\nabla}\hat{\varphi}(\underline{x})\right) - \left(\underline{\nabla}\hat{\varphi}^+(\underline{x})\right)\hat{\varphi}(\underline{x})\right] \;, \tag{111}$$

the magnetisation-density operator

$$\hat{\underline{m}}(\underline{x}) = -\frac{e\hbar}{2mc}\,\hat{\varphi}^+(\underline{x})\underline{\sigma}\hat{\varphi}(\underline{x}) \;, \tag{112}$$

and the standard density operator

$$\hat{n}(\underline{x}) = \hat{\varphi}^+(\underline{x})\varphi(\underline{x}) \;. \tag{113}$$

In discussing the gauge structure of the problem at hand, some difficulties seem to arise. One first notes that the Pauli Hamiltonian is invariant under the gauge transformation

$$\hat{\varphi}'(\underline{x}) = e^{-ie\lambda(\underline{x})/\hbar}\hat{\varphi}(\underline{x}) \qquad ; \qquad \underline{V}'(\underline{x}) = \underline{V}(\underline{x}) - c\underline{\nabla}\,\lambda(\underline{x}) \;, \tag{114}$$

that is

$$\hat{H}_P(\hat{\varphi}',\underline{V}') = H_P(\hat{\varphi},\underline{V}) \;. \tag{115}$$

Concerning the currents one can state that the paramagnetic current $\hat{j}_p(\underline{x})$ is not invariant under this gauge transformation, but the combination

$$\hat{\underline{j}}_p(\underline{x}) - \frac{e}{mc}\,\underline{V}(\underline{x})\,\hat{n}(\underline{x}) \tag{116}$$

and hence the total nonrelativistic current $\hat{\underline{j}}(\underline{x})$ is invariant.

The problem arises if one reexpresses the Hamiltonian in terms of the (physical) current and density operators

$$\begin{aligned}\hat{H}_P = \hat{T} + &\hat{H}_{ee} \\ &- \int d^3x \Big\{ -\frac{e}{c}\underline{V}(\underline{x}) \cdot \hat{\underline{j}}(\underline{x}) + \Big[eV_0(\underline{x}) - \frac{e^2}{2mc^2}\hat{\underline{V}}^2(\underline{x})\Big]\hat{n}(\underline{x})\Big\}.\end{aligned} \tag{117}$$

This form suggests that n and $\underline{j}$ are the basic variables of the theory, but it is not possible to prove a HK-theorem with this Hamiltonian, establishing $E_0 = E_0[n, \underline{j}]$, which seems to contradict the statements of RDFT at first glance. The resolution of this dilemma is the fact that not all terms of order $1/c^2$ are included in $\hat{H}_P$, i.e. the Hamiltonian is not consistent with respect to an expansion in $1/c$. The contribution to $\hat{H}_P$ which does not allow the proof of a HK-theorem is the gauge term $\frac{e^2}{c^2}\,\underline{V}^2$. This means: If one neglects all terms of the order $1/c^2$ consistently, the proof of a HK-theorem with n and $\underline{j}$ as basic variables is possible. It remains to be investigated whether inclusion of all terms of order $1/c^2$ leads to a consistent gauge invariant result to that order.

4 Functionals

Applications of RDFT may, as in nonrelativistic DFT, either use the KS-scheme or RETF-methods. In the first instance knowledge of

$$E_{xc}\,[j^\mu] \quad \text{resp} \quad v_{xc}([j^\mu], \underline{x})$$

is required. For RETF-applications one needs, in addition, an explicit density functional representation of the noninteracting kinetic energy

$$T_s = T_s[j^\mu] \ .$$

We shall start the discussion by consideration of E_{xc}. The simplest approximation for this quantity is obtained in the

4.1 Relativistic local density approximation (E_{xc})

The procedure used to establish the relativistic LDA (RLDA) is exactly the same as in the nonrelativistic case [4]. One calculates the energy density of the relativistic homogeneous electron gas (RHEG) and replaces the constant density n_0 by a locally varying density,

$$E_{xc}^{RLDA}[n] = \int d^3x\; e_{xc}^{RHEG}(n_0 = n(\underline{x})) \ . \tag{118}$$

A dependence on the three current does not occur, as the spatial current vanishes for a homogeneous system

$$\underline{j}^{RHEG} = 0 \ . \tag{119}$$

In the nonrelativistic case rather accurate Monte Carlo results [34] are available for the correlation contribution (the x-contribution can be obtained analytically). No Monte Carlo results exist in the relativistic case. This means that one has to start from scratch with the evaluation of diagrammatic contributions, as far as this is possible.

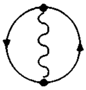
The difference compared to the nonrelativistic case is the fact that evaluation of the various contributions to e_{xc}^{RHEG} is more involved in the relativistic case. We shall demonstrate some of the details for the simplest case e_x^{RHEG}, which is given by (see eq.(94))

$$e_x^{RHEG} = \frac{1}{2}\int d^4y\, D_{\mu\nu}^{(0)}(x-y)\, tr[S_F^{(0)}(x-y)\gamma^\nu S_F^{(0)}(y-x)\gamma^\mu] + CT - VEV, \tag{120}$$

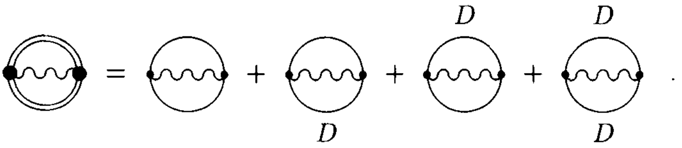
where the fermion propagators are the propagators of the RHEG. Going over to momentum space, one has

$$e_x^{RHEG} = \frac{1}{2}\int \frac{d^4q}{(2\pi)^4}\int \frac{d^4p}{(2\pi)^4}\, D_{\mu\nu}^{(0)}(q)\, tr\Big[S_F^{(0)}\,(q-p)\,\gamma^\nu\, S_F^{(0)}(p)\,\gamma^\mu\Big] \tag{121}$$
$$+\, CT \, - \, VEV \ .$$

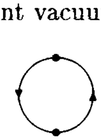
In diagrammar the loop integral and the VEV look like this,

$e_x^{RHEG} = \frac{i}{2}\Big[$  $-$  $\Big] + CT \ .$ (122)

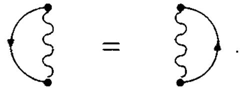
The electron propagators in the first term decompose into a vacuum and density part. Thus the x-bubble corresponds to



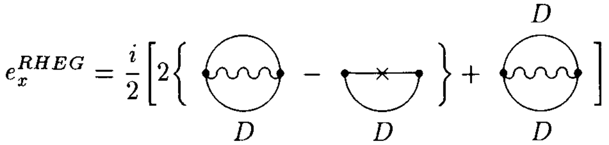
Only the density-density loop gives a finite contribution. In the three remaining terms we recognise the divergent vacuum polarisation



and the selfenergy insertions



The first term is cancelled by the VEV, the (identical) selfenergy subgraphs have to be renormalised by appropriate counterterms,



(123)

The first term is thus seen to contain the renormalised vacuum-selfenergy insertion,

$$e_{x,1}^{RHEG} = -i\int \frac{d^4p}{(2\pi)^4}\, tr\Big[S_D^{(0)}(p)\ \Sigma^{(1)}_{vac,ren}(p)\Big] \ . \tag{124}$$

This term vanishes for the following reason. The propagator $S_D(p)$ contains the factor $(\not{p}+m)$ and the renormalised selfenergy insertion satisfies the on-shell condition

$$\Big[(\not{p}+m)\Sigma^{(1)}_{vac,ren}(p)\Big]_{p^2=m^2} = 0 \ . \tag{125}$$

Thus only the finite density-density term remains. All vacuum corrections have been eliminated by the standard renormalisation scheme. The remaining term

$$e_{x,2}^{RHEG} \equiv e_x^{RHEG} = \frac{1}{2}\int \frac{d^4p}{(2\pi)^4}\ \frac{d^4q}{(2\pi)^4}\ D^{(0)}_{\mu\nu}(q)\ tr\Big\{S^{(0)}_{F,D}(q-p)\gamma^\mu S^{(0)}_{F,D}(p)\gamma^\mu\Big\} \tag{126}$$

can be evaluated in a straightforward manner [35–37], giving

$$e_x^{RHEG} = \Big[-\frac{e^2}{4\pi^3}\ k_F^4\Big]\Phi_x(\beta) \tag{127}$$

$$\Phi_x(\beta) = 1 - \frac{3}{2}\Big[\frac{\eta}{\beta} - \frac{1}{\beta^2}\ \mathrm{arsinh}(\beta)\Big]^2 \tag{128}$$

$$\beta = \frac{(3\pi^2 n_0)^{1/3}}{m} = \frac{k_F}{m}\ , \qquad \eta = (1+\beta^2)^{1/2} \tag{129}$$

Using the decomposition of the photon propagator into a longitudinal and a transverse part, one can split e_x^{RHEG} into corresponding contributions [38],

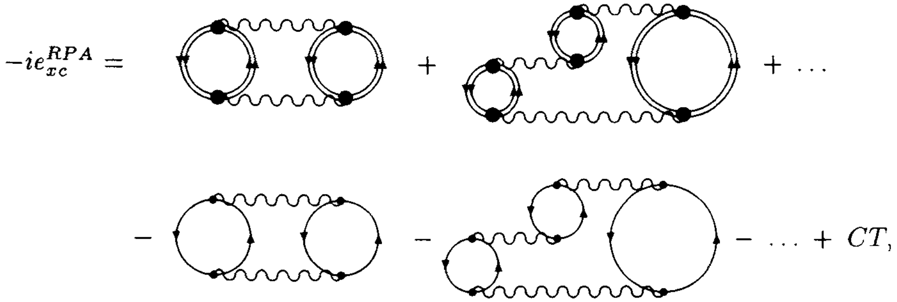
$$\Phi_x^L(\beta) = \frac{5}{6} + \frac{1}{3\beta^2} + \frac{2\eta}{3\beta}\mathrm{arsinh}(\beta) - \frac{2\eta^4}{3\beta^4}\ln(\eta) - \frac{1}{2}\Big(\frac{\eta}{\beta} - \frac{\mathrm{arsinh}(\beta)}{\beta^2}\Big)^2 \tag{130}$$

$$\Phi_x^T(\beta) = \frac{1}{6} - \frac{1}{3\beta^2} - \frac{2\eta}{3\beta}\mathrm{arsinh}(\beta) + \frac{2\eta^4}{3\beta^4}\ln(\eta) - \Big(\frac{\eta}{\beta} - \frac{\mathrm{arsinh}(\beta)}{\beta^2}\Big)^2 \ . \tag{131}$$

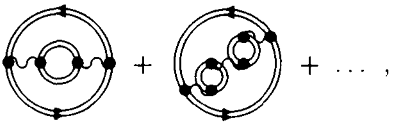
The variation of the relativistic correction factors is illustrated in Fig.4.1 of [6]. One finds that the longitudinal part does not differ very much from the nonrelativistic limit. The transverse correction factor is negative and small for low densities. It grows, however, sufficiently in magnitude so that the total x-energy density changes sign at about $\beta = 2.4$. As for instance the maximal density in Hg amounts to $\beta \approx 3$, one realizes that relativistic effects should be relevant for the inner shells of atoms.

One may look a bit more closely at the transverse part. It can either be decomposed into a magnetic and a retardation contribution or one may consider expansion in the weakly relativistic limit, giving the Breit contribution. One then finds that retardation and magnetic effects have opposite signs, the latter is dominant though. The Breit limit reproduces the exact transverse correction factor over the full range of density values of interest quite closely.

The calculation (including the renormalisation) becomes more involved if one addresses correlation contributions. As a matter of fact the only correlation contribution that has been evaluated is the relativistic random phase approximation (RPA) [35,37,39,40]. It corresponds to the following diagrams,

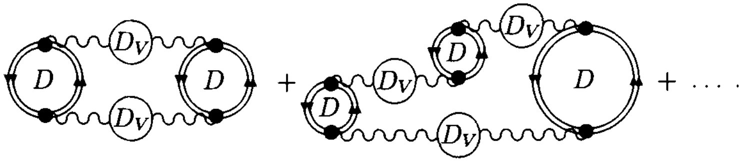


where the vacuum subtraction and the fact that counterterms are necessary has been indicated. Just to illustrate the diagrammatic games, we give an alternative representation of the diagrams

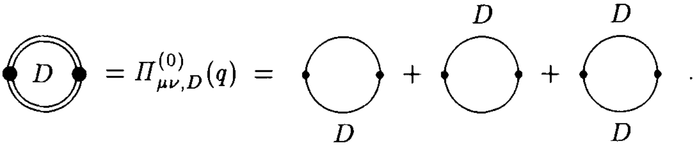


which indicates that the RPA corresponds essentially to an exchange type term, in which the free propagator (that is the free interaction) is replaced by a specifically screened interaction. We will not go through the messy details of further processing the corresponding equations, but rather look at an indication of the final result in diagrammar,

$-ie_{c,s}^{RPA} =$



Here the D in the electron loops indicate that only the electron gas part has to be inserted there



The wiggly line with D_V is the full vacuum photon propagator

$$D_{\mu\nu,V}(q)_{ren} = g_{\mu\nu}\ D_V(q)_{ren} \quad \text{(in Feynman gauge)} \tag{132}$$

with

$$D_V(q)_{ren} = \frac{D_V^{(0)}(q)}{1 - D_V^{(0)}(q)\Pi_{vac}(q)_{ren}} \ . \tag{133}$$

The series of RPA-subdiagrams can be resummed [35,37,41] leading to a structure of the form

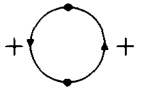
$$e_c^{RPA} \propto \int \frac{d^4p}{(2\pi)^4}\left\{ln[1 \ \pm \ D\Pi] \pm D\Pi\right\} \tag{134}$$

for both the longitudinal and transverse contributions. These integrals have only been evaluated in some approximations. In the **no-sea** approximation the full photon propagator is replaced by the free propagator,

$$D_{vac}(q)_{ren} \longrightarrow D^{(0)}(q) \ .$$

In the no-pair approximation (corresponding to the standard procedure in quantum chemistry) one also uses the free photon propagator and evaluates in addition the polarisation insertion as

$$\Pi^{(0)}(q)_{no-pair} = \tag{135}$$



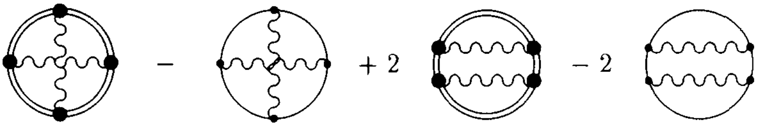
with the electron propagator on the basis of the decomposition (52). Thus in both cases one neglects the screening effects due to the vacuum. In addition, there is a (slight) difference in results due to the different evaluation of the polarisation insertion [42].

Even with these approximations the final result can only be obtained numerically. Writing the longitudinal as well as the transverse contribution of the no-sea result in the form

$$e_c^{RPA}(n_0) = \left\{ e_c^{RPA}(n_0) \right\}_{nonrel} \Phi_c^{RPA}(\beta) \ , \tag{136}$$

one finds that the correction factors can, as in the case of exchange, can be quite substantial for higher densities [6].

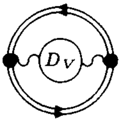
Further diagrams have not been evaluated for the relativistic homogeneous electron gas, but a number of high density limits are available. To second order (e^4) two additional correlation diagrams contribute,



Both diagrams require renormalisation beyond the vacuum subtraction indicated. In the high density limit one finds [43] for their total contribution,

$$e_c^{(2)}(n_0) \xrightarrow[\beta>>1]{} \frac{e^4}{12\pi^4} k_F^4 \Big(-3.18 \pm 0.12 \Big) \ . \tag{137}$$

In addition the two loop contribution to the screened exchange [43],



,

which is not contained in the no-sea approximation, can also be calculated in this limit,

$$e_x^{(2)}(n_0) \xrightarrow[\beta>>1]{} \frac{e^4}{12\pi^4} k_F^4 \Big(\ln(2\beta) - \frac{11}{6} \Big) \ . \tag{138}$$

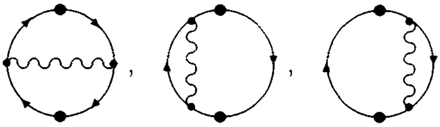
By comparison, for the no-sea RPA-result one has [35,37]

$$e_c^{RPA}(n_0) \xrightarrow[\beta>>1]{} \frac{e^4}{12\pi^4} k_F^4 \Big(-7.796 \Big) \ . \tag{139}$$

One notes that in the limit considered, the additional second order contributions amount to about 40% of the RPA value and that in the extreme high density limit the screened exchange contribution eventually dominates over all other known contributions. This occurs, however, only for $\beta \approx 10^3$, which is not relevant for electronic structure calculations.

4.2 Relativistic generalised gradient approximation (E_{xc})

As the x and RPA correlation contributions in LDA are known not to yield optimal results (for atoms and other systems) in the nonrelativistic case, one has to consider improvements. The next step (thinking of the history for nonrelativistic systems) would be direct gradient expansions [44–51]. The problem is, that the corresponding contributions as e.g.



of the homogeneous electron gas are difficult to evaluate (and have not been evaluated) for relativistic systems. In addition, they have not been found to be very accurate in the nonrelativistic regime. In order to make some headway, we have carried through the following scheme [52,53]:

For the x-part

Step 1: Solve the KS-OPM problem for a selection of atoms with closed subshells (17 atoms were chosen).

Step 2: Use the results to set up a semiempirical relativistic generalised gradient approximation (GGA), relying on the form

$$E_{x,rel}^{GGA}[n] = \int d^3x \; e_x^{LDA}(n) \left[\Phi_{x,0}(\beta) + g(\xi)\Phi_{x,2}(\beta)\right] \; . \tag{140}$$

For the function g of the dimensionless density gradient

$$\xi = (\underline{\nabla} n)^2 \; / \; \left[4n^2(3\pi^2 n)^{2/3}\right] \tag{141}$$

we choose nonrelativistic GGA forms. We used the Becke 88 [54], the Engel-Chevary-Macdonald-Vosko 92 [55] and the Perdew-Wang 91 [56] forms and found that final results for E_x only varied marginally with $g(\xi)$. The function $\Phi_{x,0}$ is the LDA relativistic correction factor indicated earlier. For the relativistic correction factor $\Phi_{x,2}$ we choose a reasonably flexible ansatz in the form of a [2/2] Padé-approximant,

$$\Phi_{x,2}(\beta) = \frac{a_0 + a_1\beta^2 + a_2\beta^4}{1 + b_1\beta^2 + \beta_2\beta^4} \; . \tag{142}$$

The form can be used for both the longitudinal and the transverse contributions, if one sets

$$a_0^L = 1 \qquad a_0^T = 0 \; , \tag{143}$$

which guarantees that the correct weakly relativistic limit is obtained. The fact that $\Phi_{x,2}$ must be an even function of β follows from the time reversal invariance.

Step 3: The coefficients have been fitted to the exact relativistic correction to the longitudinal exchange energy,

$$\Delta E_x^L = E_{x,rel}^L[n_{rel}] - E_{x,nrel}[n_{nrel}] \ , \tag{144}$$

and the exact transverse exchange energy $E_{x,rel}^T[n_{rel}]$, with all quantities being obtained by corresponding ROPM calculations. Explicit results, indicating their quality, will be shown in Section 5. For the moment we look at an illustration of the correction factors for the B88 and ECMV92 GGAs (Fig.1) and an abbreviated comparison of the constants in the Padé-ansatz for various GGA functionals and the longitudinal term:

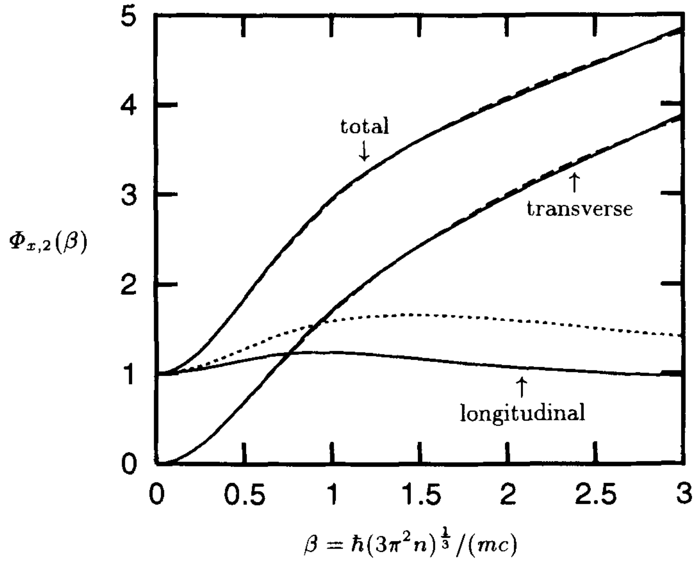


Fig. 1. Relativistic correction factors for the gradient contribution to the exchange energy density for both ECMV92 (solid line) and B88 (long dashes). Also shown is the relativistic correction factor for the second order gradient correction for $T_s[n]$ (short dashes).

Functional	a_1^L	a_2^L	b_1^L	b_2^L
B88 [54]	2.209	0.669	1.331	0.795
ECMV92 [55]	2.213	0.669	1.330	0.795
PW91 [56]	2.216	0.670	1.327	0.794

It remains to be stated that we also used other Padé-forms (eg. [3/3]) without finding significant improvements.

For the case of the **correlation term** basic data are hard to come by. In this case a more global form [53],

$$E_{c,rel}^{GGA}[n] = \int d^3x \; e_{c,nrel}^{GGA}(n, (\underline{\nabla} n)^2, \ldots) \; \Phi_c(\beta) \tag{145}$$

has been fitted to second order perturbation theory results (on the basis of a Dirac-Coulomb-Breit Hamiltonian) for the Neon isoelectronic series [57], as it seemed to be the most systematic set of quantum chemical data available. For the nonrelativistic GGA the Perdew-Wang 91 [56] and the Lee-Yang-Parr [58] forms have been used, Φ_c is again a [2/2] Padé-approximant. The fitting procedure used

$$\Delta E_c = E_{c,rel}[n_{rel}] - E_{c,nrel}[n_{nrel}] \tag{146}$$

as before, in order to suppress errors in the individual energy values as much as possible.

The last functional to be discussed is the

4.3 Relativistic gradient expansion for $T_s[n]$

As already indicated, this functional is of interest for RETF-applications. We discuss it for two reasons. First, it provides another example for the need of renormalisation. Second, the results exhibit a certain amount of physics [59, 22].

The starting point of the discussion is the definition of the exact kinetic energy and the exact current in terms of the exact fermion propagator

$$T[j^\mu] = -\, i \int d^3x \, \lim_{y \to x}{}_s \, tr\Big[\Big(- i\underline{\gamma} \cdot \underline{\nabla} + m \Big) S_F(x,y) \Big] - VEV + CT \tag{147}$$

$$j^\mu(\underline{x}) = -i \, \lim_{y \to x}{}_s \, tr\Big[S_F(x,y) \gamma^\nu \Big] - VEV + CT \; . \tag{148}$$

The symmetric limit indicated is defined as

$$\lim_{y \to x}{}_s \equiv \frac{1}{2} \left(\lim_{y \to x, y^0 > x^0} + \lim_{y \to x, y^0 < x^0} \right) \Bigg|_{(x-y)^2 \geq 0} \; . \tag{149}$$

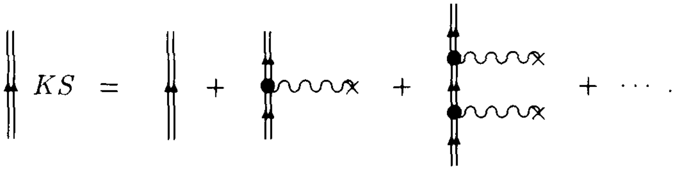
It is the relativistic equivalent of the nonrelativistic limit

$$\lim_{\underline{y} \to \underline{x}} \; \lim_{t_y \to t_x^+} \; .$$

The definitions given are quite general. In order to arrive at the noninteracting situation, one has to replace the exact fermion propagator by the KS propagator, which can be specified alternatively by the differential equation [60]

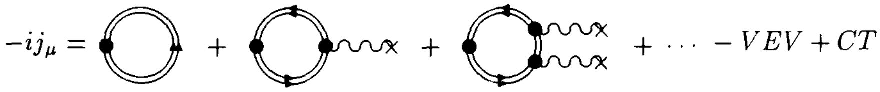
$$(i \partial\!\!\!/_x - m - v\!\!\!/_{KS}(\underline{x})) \; S_F^{KS}(x,y) \; = \; \delta^{(4)}(x-y) \; . \tag{150}$$

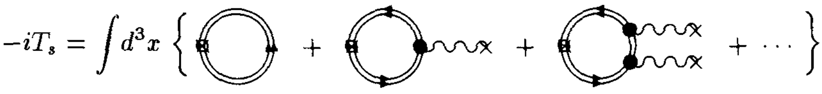
The standard perturbation expansion of this propagator in powers of the potential can be indicated in diagrammar as



We have seen the explicit form in terms of KS-orbitals before. Renormalisation is not necessary.

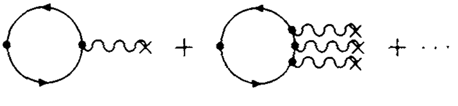
The symmetric limit required for the calculation of T_s and j_μ corresponds to closing the ends of the fermion lines, after supplying them with the required weight, that is





$-VEV + CT$.

We recognise outermost loops (integration), which introduce, as in standard QED without external fields, UV-divergencies. It does not make any difference whether the virtual electron-positron pairs involved in the loops are generated by the photon field or by an external potential. As a consequence the renormalisation procedure is the same as the procedure that one uses for the renormalisation of the propagators of vacuum QED. A quick inspection for j^μ shows that only the second diagram on the right-hand side is divergent. For the vacuum four current all contributions with an odd number of vertex points in the loop vanish due to Furry's theorem. From the remaining diagrams,



,

only the first one needs to be renormalised.

To see how this works, one has to evaluate T_S and j_μ explicitly. We do this using a semiclassical gradient expansion as a first step.

In order to obtain the semiclassical gradient expansion (an expansion in terms of derivatives of the effective KS-potential) one solves the differential equation for the propagator explicitly by iteration [60]. The general ansatz for the iteration is the expansion (dropping the index KS)

$$S_F(x,y) = e^{-i(x-y)\cdot v(\underline{x})} \int \frac{d^4p}{(2\pi)^4}\, e^{-ip(x-y)} \sum_{k=0}^{\infty} S_F^{[k]}(p_\mu, v_\mu(\underline{x})) \ . \tag{151}$$

The index $[k]$ denotes the order of the potential gradients involved. Insertion of the ansatz into the differential equation yields the recursion

$$(\not{p} - m) S_F^{[k]}(p_\mu, v_\mu(\underline{x})) = \left[(i\not{\partial} v_\mu(\underline{x}))\, \frac{\partial}{\partial p_\mu} - i\not{\partial} \right] S^{[k-1]}(p_\mu, v_\mu(\underline{x})) \ . \tag{152}$$

The starting point for the recursion is the solution of the differential equation for a constant potential,

$$S_F^{[0]}(p_\mu, v_\mu) = S_{F,vac}^{[0]}(p) - 2\pi i \frac{(\not{p} + m)}{2E}\, \delta(p^0 - E)\, \Theta(\epsilon_F - v^0 - p^0) \Big|_{v_0 = v_0(\underline{r})} \ . \tag{153}$$

As an example of the explicit results at this level, we look at the second order semiclassical gradient expansion of the density n and the kinetic energy density t_s for the case of a purely electrostatic external potential $v^\mu = (v^0, \underline{0})$,

$$\begin{aligned} \tilde{n}[v_0]_{reg} = {} & \frac{p^3}{3\pi^2} + \frac{1}{12\pi^2} \Gamma\left(2 - \frac{d}{2}\right) (\underline{\nabla}^2 v_0) \\ & - \frac{1}{12\pi^2} \left\{ \left[\frac{E}{p} + 2\mathrm{arsinh}\left(\frac{p}{m}\right) \right] (\underline{\nabla}^2 v_0) + \left[\frac{E^2}{p^2} - 3 \right] \frac{(\underline{\nabla} v_0)^2}{2p} \right\} \end{aligned} \tag{154}$$

$$\begin{aligned} \tilde{t}_s[v_0]_{reg} = {} & \frac{m^4}{16\pi^2} \Gamma\left(2 - \frac{d}{2}\right) + \frac{1}{8\pi^2} \left\{ pE^3 + p^3 E - \mathrm{arsinh}\left(\frac{p}{m}\right) \right\} \\ & + \frac{1}{24\pi^2} \Gamma\left(2 - \frac{d}{2}\right) (\underline{\nabla} v_0)^2 - \frac{1}{12\pi^2} \left[\frac{E^2}{p} + p \right] (\underline{\nabla}^2 v_0) \\ & - \frac{1}{12\pi^2} \left[\frac{E^3}{2p^3} - \frac{E}{p} + \mathrm{arsinh}\left(\frac{p}{m}\right) \right] (\underline{\nabla} v_0)^2 , \end{aligned} \tag{155}$$

where

$$E = \epsilon_F - v_0(\underline{x}) \qquad ; \qquad p = \sqrt{E^2 - m^2}\; \Theta(E^2 - m^2) \ . \tag{156}$$

The results have been obtained with dimensional regularisation. One recognises contributions which diverge in the limit $d \to 4$. There is a divergent contribution to the kinetic energy due to the Dirac-sea, which is removed by vacuum energy subtraction

$$< vac | \hat{H}_e | vac > = \frac{m^4}{16\pi^2}\, \Gamma\left(2 - \frac{d}{2}\right) \ . \tag{157}$$

The UV divergencies (proportional to Δv_0) are removed by the same counterterms which are the lowest order contribution to the vacuum polarisation insertion in standard QED without external fields.

After renormalisation has been carried out, one has finite (in the limit $d \to 4$) expressions of the form

$$j^{\nu}(\underline{x}) = j^{\nu}(v_{\mu}(\underline{x}), \underline{\nabla} v_{\mu}(\underline{x}), \ldots) \tag{158}$$

$$t_s(\underline{x}) = t_s(v_{\mu}(\underline{x}), \underline{\nabla} v_{\mu}(\underline{x}), \ldots) \ . \tag{159}$$

The last step is the derivation of a current (or density in the electrostatic case) gradient expansion. This is obtained by order by order inversion of the first relation

$$v_{\mu}(\underline{x}) = v_{\mu}(j^{\nu}(\underline{x}), \underline{\nabla} j^{\nu}(\underline{x}), \ldots) \tag{160}$$

and insertion into the second relation.

We first give the results for the case of an electrostatic external potential to the fourth order

$$T_s^{RGE0}[n] = \int d^3x \, \frac{(3\pi^2 n)^{5/3}}{10\pi^2 m} \, \frac{10}{\beta^5} \left[\frac{1}{8} \Big(\beta\eta^3 + \beta^3\eta - \operatorname{arsinh}(\beta) \Big) - \frac{1}{3}\beta^3 \right] \tag{161}$$

$$T_s^{RGE2}[n] = \frac{1}{72m} \int d^3x \, \frac{(\underline{\nabla} n)^2}{n} \, \frac{1}{\eta} \left[1 + 2\frac{\beta}{\eta} \operatorname{arsinh}(\beta) \right] \tag{162}$$

$$T_s^{RGE4}[n] = \int d^3x \, \left[t_{s,V}^{RGE4}[n] + t_{s,D}^{RGE4}[n] \right] \tag{163}$$

$$t_{s,V}^{RGE4} = \frac{1}{360\pi^2} \left\{ 3\frac{\beta^2}{\eta^2} (\underline{\nabla}^2 \beta)^2 + 6\frac{\beta}{\eta^4} (\underline{\nabla}^2 \beta)(\underline{\nabla}\beta)^2 + \frac{3 - \beta^4\eta^2}{\eta^6} (\underline{\nabla}\beta)^4 \right\}$$

$$\begin{aligned} t_{s,D}^{RGE4} = \frac{1}{5760\pi^2} \Bigg\{ & 4\frac{(\underline{\nabla}^2 \beta)^2}{\beta\eta} \left[3(1 - 4\beta^2) + 5\left(1 + 2\frac{\beta}{\eta} \operatorname{arsinh}(\beta) \right)^2 \right] \\ & + 2\frac{(\underline{\nabla}^2 \beta)(\underline{\nabla}\beta)^2}{\beta^2\eta^3} \left[-41\beta^2 + 20\left(1 + 2\frac{\beta}{\eta} \operatorname{arsinh}(\beta) \right) \right. \\ & \qquad \left. \times \left(\frac{\beta^2 + \eta^2}{2} + 2\frac{\beta}{\eta} \operatorname{arsinh}(\beta) \right) \right] \\ & + \frac{(\underline{\nabla}\beta)^4}{\beta^3\eta^5} \left[3 - 19\beta^2 - 8\beta^4 + 8\beta^6 + 16\beta^8 \right. \\ & \qquad \left. + 20\left(\frac{\beta^2 + \eta^2}{2} + 2\frac{\beta}{\eta} \operatorname{arsinh}(\beta) \right)^2 \right] \Bigg\} \ , \end{aligned}$$

and offer the following comments

i The zeroth order contribution to the kinetic energy density is the TF (relativistic homogeneous electron gas) result, already obtained by Vallarta and Rosen in 1932 [61].

ii The second and fourth order terms were only obtained in 1987 [22] and 1991 [62]. While the results for the second order term are quite compact, the result for the fourth order term looks somewhat messy. It contains explicit radiative corrections ($t^{(4)}_{s,vac}$), which correspond exactly to the Euler-Heisenberg energy (for the case of an electrostatic potential).

iii The relativistic results go, for small value of β, over into the corresponding nonrelativistic results [63,64].

iv An illustration of the relativistic correction can readily be given for the zeroth and the second order. In this case the ratio $t^{[k]}_{s,rel}/t^{[k]}_{s,nonrel}$ is a function of β alone. As in the case of exchange and correlation the relativistic corrections are noticeable, especially for the case of $t^{[2]}_s$ (see Fig.1).

v The same technique can be applied to the generation of a current gradient expansion for the case of a full four potential [65]. Obviously, there are no zeroth order current terms (they vanish in a homogeneous system). The current contribution to $t^{[2]}_s$ has the form ($\beta = (3\pi^2 n)^{1/3}/m$)

$$\begin{aligned} T_s^{RGE2}[n,\underline{j}] = \frac{3}{16}\int d^3x \, \frac{1}{\operatorname{arsinh}(\beta(\underline{x}))} \\ \times \int d^3y \int d^3z \, \sum_{k,l=1}^{3} \frac{\partial_k^y j^l(\underline{y})\partial_k^z j^l(\underline{z}) - \partial_k^y j^l(\underline{y})\partial_l^z j^k(\underline{z})}{|\underline{x}-\underline{y}|\,|\underline{x}-\underline{z}|} \; . \end{aligned} \tag{164}$$

As we will not discuss applications of the RETF model [23,66] in the following, we offer one brief remark at this point. The model is given by

$$E_0^{[2K]} = \sum_{i=0}^{K} T_S^{[2i]}[n] + E_{ext}[n] + E_H[n] + E_x^{LDA}[n] \tag{165}$$

and often combined with a spherical average of the system. Evaluation of the direct variational equations reproduces the gross features of atoms, but does not reproduce quantal effects like the shell structure. As the model does not involve correlation contributions comparison with Dirac-Fock-Slater results is adequate. The results show that the accuracy that can be obtained in the relativistic case is comparable to the accuracy in the nonrelativistic case.

5 Applications to atoms

A standard approach to relativistic Coulomb problems is the Dirac-Fock-Slater (DFS) approximation, in which the kinetic energy is treated fully in terms of relativity, while the nonrelativistic x-only LDA is used for the exchange-correlation energy. There are few investigations that use the relativistic LDA exchange functional and only a scatter addressing relativistic correlation effects (see eg.[19,67,38,68–70,21,71,72]). In view of this state of affairs we set ourselves as a first goal a more detailed investigation of the quality of the RLDA functionals that are available.

Table 1. Longitudinal ground state energies ($-E^L_{tot}$) and highest occupied eigenvalues ($-\epsilon^L_{mk}$) for closed subshell atoms from nonrelativistic OPM (NROPM [73]), relativistic OPM (ROPM [21]) and relativistic HF (RHF [8]) calculations [74] (all energies are in hartree).

Atom	$-E^L_{tot}$			$-\epsilon^L_{mk}$		
	NROPM	ROPM	RHF	NROPM	ROPM	RHF
He (1s1/2)	2.862	2.862	2.862	0.918	0.918	0.918
Be (2s1/2)	14.572	14.575	14.576	0.309	0.309	0.309
Ne (2p3/2)	128.545	128.690	128.692	0.851	0.848	0.848
Mg (3s1/2)	199.611	199.932	199.935	0.253	0.253	0.253
Ar (3p3/2)	526.812	528.678	528.684	0.591	0.587	0.588
Ca (4s1/2)	676.751	679.704	679.710	0.196	0.196	0.196
Zn (4s1/2)	1777.828	1794.598	1794.613	0.293	0.299	0.299
Kr (4p3/2)	2752.028	2788.848	2788.861	0.523	0.515	0.514
Sr (5s1/2)	3131.514	3178.067	3178.080	0.179	0.181	0.181
Pd (4d5/2)	4937.858	5044.384	5044.400	0.335	0.319	0.320
Cd (5s1/2)	5465.056	5593.299	5593.319	0.266	0.282	0.281
Xe (5p3/2)	7232.018	7446.876	7446.895	0.456	0.439	0.440
Ba (6s1/2)	7883.404	8135.625	8135.644	0.158	0.163	0.163
Yb (6s1/2)	13391.070	14067.621	14067.669	0.182	0.196	0.197
Hg (6s1/2)	18408.313	19648.826	19648.865	0.262	0.329	0.328
Rn (6p3/2)	21865.826	23601.969	23602.005	0.427	0.382	0.384
Ra (7s1/2)	23093.258	25028.027	25028.061	0.149	0.167	0.166
No (7s1/2)	32787.471	36740.625	36740.682	0.171	0.209	0.209

We begin, however, by looking at **ROPM-results** in the x-only approximation [21] in order to assess relativistic effects in a more global fashion. Table 1 shows ground state energies for neutral, spherical (that is closed subshell) atoms in the *no-sea/longitudinal* approximation. We concentrate on the first three columns, in which nonrelativistic OPM-, relativistic OPM- and relativistic HF-results can be compared. One notes that for heavier atoms

(a) relativistic effects are obviously important,
(b) ROPM- and RHF-results agree quite closely.

As a specific example for comparison we will use the Hg atom (here and in the following). For this atom the relativistic contribution to the total ground state energy (in the approximation specified) amounts to

$$\Delta E^L_{tot,1} = E^{L,ROPM}_{tot} - E^{NROPM}_{tot} = -1240.5 hartree \ ,$$

while the energy difference for the two relativistic theories is

$$\Delta E^L_{tot,2} = E^{L,RHF}_{tot} - E^{L,ROPM}_{tot} = -39 mhartree \ .$$

This clearly establishes the need for a relativistic treatment of heavier atoms and shows that the ROPM gives an adequate representation of exchange effects. The fact that ROPM results are always slightly higher than RHF-energies can be understood on the basis of the reduced variational freedom of the ROPM orbitals.

The trends indicated are also found for the orbital energies of the highest occupied orbitals (see Table 1). The relativistic $6s1/2$-orbital in Hg is more bound by

$$\Delta\epsilon_{6s1/2} = \epsilon_{6s1/2}^{L,ROPM} - \epsilon_{6s1/2}^{NROPM} = -67 mhartree = -1.8 eV \ ,$$

while there is little difference between the RHF and ROPM orbital energies. The last statement might imply that the orbital energies are the same for all ROPM and RHF orbitals. Table 2 (for Hg) demonstrates that this is not the case. Although the total energies agree quite closely for ROPM and RHF (and

Table 2. Single particle energies $(-\epsilon_{nlj}^{L})$ for Hg from NROPM-, ROPM- and RHF-calculations in comparison with DFS-, and RLDA-results (longitudinal limit, all energies are in *hartree*).

Level	NROPM	ROPM	RHF	DFS	RLDA
1S1/2	2756.925	3047.430	3074.228	3047.517	3044.410
2S1/2	461.647	540.056	550.251	539.713	539.250
2P1/2	444.015	518.061	526.855	518.164	517.746
2P3/2	444.015	446.682	455.157	446.671	446.399
3S1/2	108.762	128.272	133.113	128.001	127.905
3P1/2	100.430	118.350	122.639	118.228	118.148
3P3/2	100.430	102.537	106.545	102.397	102.346
3D3/2	84.914	86.201	89.437	86.085	86.060
3D5/2	84.914	82.807	86.020	82.690	82.668
4S1/2	23.522	28.427	30.648	28.067	28.046
4P1/2	19.895	24.161	26.124	23.871	23.854
4P3/2	19.895	20.363	22.189	20.039	20.030
4D3/2	13.222	13.411	14.797	13.148	13.146
4D5/2	13.222	12.700	14.053	12.434	12.432
4F5/2	4.250	3.756	4.473	3.556	3.559
4F7/2	4.250	3.602	4.312	3.402	3.404
5S1/2	3.501	4.403	5.103	4.290	4.286
5P1/2	2.344	3.012	3.538	2.898	2.896
5P3/2	2.344	2.363	2.842	2.219	2.218
5D3/2	0.538	0.505	0.650	0.363	0.363
5D5/2	0.538	0.439	0.575	0.296	0.296
6S1/2	0.262	0.329	0.328	0.222	0.222

we shall see in a moment that this statement also applies to the individual contributions to E_{tot}), one finds eg.

$$\epsilon_{1s1/2}^{L,RHF} - \epsilon_{1s1/2}^{L,ROPM} = -26.80 hartree \ .$$

This difference is (as expected) much smaller than the relativistic corrections to the inner orbital energies

$$\Delta\epsilon_{1s1/2} = \epsilon_{1s1/2}^{L,ROPM} - \epsilon_{1s1/2}^{NROPM} = -290.51 hartree \ ,$$

which corresponds to a decrease of about 10.5 %. The percentage change of the outer orbital is still very large (25.6% for the $6s_{1/2}$ orbital). These results demonstrate that it is dangerous to attach too close a physical interpretation to the orbitals and their energies. Table 3 shows the longitudinal x-contribution to the total energy in various approximations [21,6]. For this quantity the relativistic correction in Hg amounts to

$$\Delta E_{x,1}^{L} = E_x^{L,ROPM} - E_x^{NROPM} = -19.96 hartree \ ,$$

Table 3. Longitudinal (Coulomb) x-only energies ($-E_x^L$) for closed subshell atoms from NROPM-, ROPM-, RHF-, DFS-, and RLDA-calculations [21,74] (all energies are in *hartree*).

Atom	NROPM	ROPM	RHF	DFS	RLDA
He	1.026	1.026	1.026	0.853	0.853
Be	2.666	2.667	2.668	2.278	2.278
Ne	12.105	12.120	12.123	10.952	10.944
Mg	15.988	16.017	16.023	14.564	14.550
Ar	30.175	30.293	30.303	27.897	27.844
Ca	35.199	35.371	35.383	32.702	32.627
Zn	69.619	70.245	70.269	66.107	65.834
Kr	93.833	95.048	95.072	89.784	89.293
Sr	101.926	103.404	103.429	97.836	97.251
Pd	139.113	141.898	141.930	134.971	133.887
Cd	148.879	152.143	152.181	144.931	143.687
Xe	179.062	184.083	184.120	175.926	174.102
Ba	189.065	194.804	194.841	186.417	184.363
Yb	276.143	288.186	288.265	278.642	274.386
Hg	345.240	365.203	365.277	354.299	347.612
Rn	387.445	414.082	414.151	402.713	394.102
Ra	401.356	430.597	430.664	419.218	409.871
No	511.906	564.309	564.415	554.242	538.040

which is still quite substantial. Comparison with the difference in the total energy indicates, however, that kinetic and direct potential effects constitute about 98% of the total relativistic effect. The difference between ROPM- and RHF-results is again fairly small

$$\Delta E^{L}_{x,2} = E_x^{L,RHF} - E_x^{L,ROPM} = -74 mhartree \ .$$

Also included in Table 3 are DFS results. From the difference

$$\Delta E^{L}_{x,3} = E_x^{L,RHF} - E_x^{DFS} = -10.98 hartree$$

one can infer (in comparison with the difference between the relativistic and the nonrelativistic results), that insertion of a relativistic density into a nonrelativistic x-functional corrects the deviation from the full relativistic result somewhat.

We now look at **RLDA-results**, first again for the case of x-only (Table 4). In the longitudinal limit, the error of the RLDA for the total energy is only of the order of 0.1% for the heavier systems (17.20*hartree* for Hg). If one compares this with the error in the x-contribution, one finds that this error is solely due to exchange

Table 4. Longitudinal x-only ground state energies: Selfconsistent ROPM, RHF, RLDA and RGGA results for neutral atoms with closed subshells (in *hartree* [74]).

Atom	$-E^{L}_{tot}$	$E^{L}_{tot} - E^{L,ROPM}_{tot}$		
	ROPM	RHF	RLDA	RPW91
He	2.862	0.000	0.138	0.006
Be	14.575	−0.001	0.350	0.018
Ne	128.690	−0.002	1.062	−0.024
Mg	199.932	−0.003	1.376	−0.001
Ar	528.678	−0.005	2.341	0.041
Ca	679.704	−0.006	2.656	0.026
Zn	1794.598	−0.014	4.140	−0.262
Kr	2788.848	−0.013	5.565	−0.021
Sr	3178.067	−0.013	5.996	−0.008
Pd	5044.384	−0.016	7.707	−0.067
Cd	5593.299	−0.020	8.213	−0.033
Xe	7446.876	−0.019	9.800	0.085
Ba	8135.625	−0.019	10.289	0.059
Yb	14067.621	−0.048	13.272	−0.893
Hg	19648.826	−0.039	17.204	−0.250
Rn	23601.969	−0.035	19.677	0.004
Ra	25028.027	−0.034	20.460	−0.006

Table 5. Total relativistic x-only ground state energies: Selfconsistent ROPM, RLDA and (R)GGA results for neutral atoms with closed subshells in comparison with perturbative RHF data (in *hartree* [74]).

Atom	$-E_{tot}^{L+T}$	$E_{tot}^{L+T} - E_{tot}^{L+T,ROPM}$			
	ROPM	RHF(p)	RLDA	RPW91	PW91
He	2.862	0.000	0.138	0.006	0.006
Be	14.575	−0.001	0.351	0.018	0.017
Ne	128.674	−0.002	1.080	−0.024	−0.043
Mg	199.900	−0.003	1.408	−0.001	−0.037
Ar	528.546	−0.005	2.458	0.041	−0.111
Ca	679.513	−0.006	2.818	0.026	−0.195
Zn	1793.840	−0.014	4.702	−0.263	−1.146
Kr	2787.429	−0.012	6.543	−0.022	−1.683
Sr	3176.358	−0.012	7.149	−0.010	−2.014
Pd	5041.098	−0.013	9.765	−0.069	−3.953
Cd	5589.495	−0.016	10.556	−0.035	−4.538
Xe	7441.172	−0.012	13.161	0.083	−6.706
Ba	8129.160	−0.010	14.050	0.057	−7.653
Yb	14053.748	−0.023	20.886	−0.896	−17.662
Hg	19626.702	0.005	29.159	−0.260	−27.256
Rn	23573.351	0.026	35.203	−0.012	−35.149
Ra	24996.942	0.034	37.391	−0.026	−38.271

$$\Delta E_x^L = E_x^{L,ROPM} - E_x^{L,RLDA} = -17.59 hartree \ ,$$

which amounts to about 5%. Compared to nonrelativistic systems, this shows that the relative error of the longitudinal exchange energy is comparable (Be 14.5%, Kr 6.1%), so that the LDA-exchange contribution can as well not be considered to be sufficiently accurate in the relativistic case.

In Table 5 we look at results obtained for the full relativistic x-functional. We first note that inclusion of the transverse contribution leads to a higher ground state energy

$$\Delta E_{tot} = E_{tot}^{L,ROPM} - E_{tot}^{L+T,ROPM} = -22.12 hartree \ .$$

This is in accord with the fact, that the transverse term has an opposite sign with respect to the longitudinal term. The absolute error of the total RLDA-energy has, however, changed to $29.16 hartree$, an increase by $11.96 hartree$ with respect to the longitudinal limit. The error in the transverse exchange energy is (see Table 6)

$$\Delta E_x^T = E_x^{T,ROPM} - E_x^{T,RLDA} = +22.17 - 34.20 = -12.03 hartree \ ,$$

Table 6. Transverse x-only energies (E_x^T) for closed subshell atoms: ROPM results in comparison with the values obtained by insertion of ROPM densities into the relativistic LDA (RLDA) and two relativistic GGAs (RECMV92 and RB88) (all energies are in hartree, [74]).

Atom	ROPM	RLDA	RECMV92	RB88
He	0.000064	0.000159	0.000060	0.000061
Be	0.00070	0.00176	0.00071	0.00072
Ne	0.0167	0.0355	0.0166	0.0167
Mg	0.0319	0.0654	0.0319	0.0319
Ar	0.132	0.251	0.132	0.132
Ca	0.191	0.356	0.191	0.191
Zn	0.759	1.328	0.760	0.759
Kr	1.420	2.410	1.421	1.419
Sr	1.711	2.878	1.712	1.710
Pd	3.291	5.374	3.291	3.291
Cd	3.809	6.180	3.809	3.809
Xe	5.712	9.114	5.712	5.713
Ba	6.475	10.282	6.475	6.477
Yb	13.900	21.597	13.895	13.900
Hg	22.169	34.257	22.169	22.169
Rn	28.679	44.382	28.681	28.680
Ra	31.151	48.275	31.149	31.151

which corresponds to a relative error slightly larger than 50%. Obviously, there is substantial room for improvement.

We next look at the correlation contribution in the LDA. As the correlation contribution in heavier atoms in LDA amounts to about $10 hartree$, with a relativistic correction of the order of $0.5 hartree$, there is hardly any difference if one performs a variational x-only calculation and evaluates the correlation-contribution with the resulting density or if one performs a more complete variational calculation. As the RPA limit is known not to be an accurate approximation to the correlation energy, we suggest to use [21]

$$E_c^{RLDA}[n] = E_{c,rel}^{RPA}[n] - E_{c,nonrel}^{RPA}[n] + E_{c,nonrel}^{LDA}[n] \ . \qquad (166)$$

We use only the relativistic correction to the RPA, which is added to a complete nonrelativistic functional (eg. LDA from Monte Carlo [76]). For high densities the RPA contribution in the two nonrelativistic terms cancel, so that the correlation energy is given by the relativistic RPA plus the nonrelativistic second order exchange graph. For low densities the first two terms cancel, so that the correlation energy is given by the more adequate nonrelativistic result.

Table 7. Comparison of LDA [21], CI (estimated from nonrelativistic CI-calculations for the three innermost electrons and the experimental ionisation potentials of all other electrons [75]) and MBPT2 [57] correlation energies for neutral atoms: E_c^{NR} — nonrelativistic correlation energy, ΔE_c^L — relativistic contribution in the longitudinal correlation energy, E_c^T — transverse correlation energy (in the case of the MBPT2 only the dominating Breit contribution to E_c^T is given — all energies are in *mhartrees*).

Atom	$-E_c^{NR}$			$-\Delta E_c^L$		$-E_c^T$	
	MBPT2	CI	LDA	MBPT2	LDA	MBPT2	LDA
He	37.14	42.04	111.47	0.00	0.00	0.04	0.00
Be		94.34	224.44		0.02		0.02
Ne	383.19	390.47	743.38	0.20	0.38	1.87	0.32
Mg		438.28	891.42		0.75		0.57
Ar	697.28	722.16	1429.64	0.84	2.60	7.92	1.89
Zn	1650.61		2665.20	10.51	10.97	26.43	7.92
Kr	1835.43		3282.95	11.39	19.61	41.07	13.10
Cd	2618.11		4570.56	35.86	44.79	82.32	28.58
Xe	2921.13		5200.19	37.57	64.73	108.75	39.27
Hg	5086.24		8355.68	203.23	200.87	282.74	113.08
Rn	5392.07		9026.90	195.36	257.00	352.60	138.43

One problem that one encounters for heavier elements, is the fact that experimental total energies (and hence experimental correlation energies) are not available. (It is difficult to measure successive ionisation energies of all positive ions for heavier systems.) Thus we compare LDA-results with results obtained in second order many-body perturbation theory (MBPT) [57]. Table 7 illustrates the well-known fact that nonrelativistic LDA correlation energies overestimate the correct values by a factor of about two. As better density functionals are available for this quantity, one can concentrate on the relativistic corrections (here with respect to the LDA). One finds the following situation: While the longitudinal part agrees with the results of MBPT (at least within a factor of two, but mostly better), the differences for the transverse part are much larger (up to factors of 4). The comparison should not be taken as final, as the quality of the results of MBPT is difficult to assess, but in view of the large differences, it is obvious that also the relativistic correlation-corrections need to be improved upon.

The semiempirical **relativistic GGA exchange functional** gives very reasonable results. We first consider the total energies in the x-only limit for the case of the modified PW91 functional (similar results are obtained for the other GGA x-functionals that we have investigated). For both the longitudinal as well as the full exchange the deviation from the OPM-standard is less than 0.2% (for He), for the heavier systems less than 0.01% (see Tables 4,5).

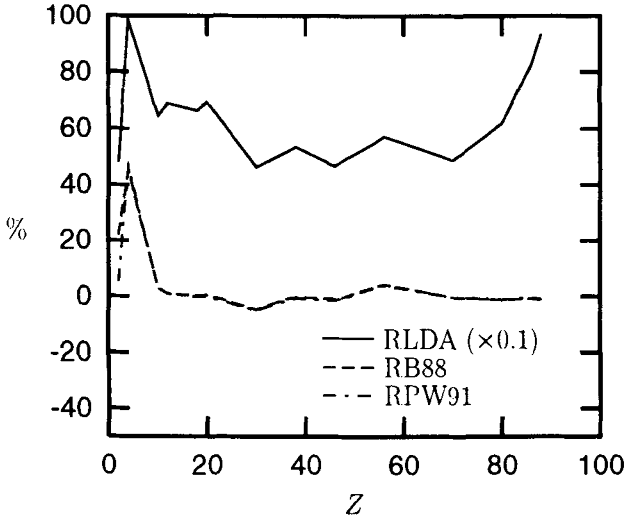


Fig. 2. Relativistic contribution ΔE_x to the x-only energy: Percentage deviation of self-consistent RLDA and RGGA results from ROPM-data for neutral atoms.

If one then looks at the x-contributions (Fig.2), one finds that these quantities are also reproduced very closely (with an absolute error of less than $100 mhartree$). The relativistic corrections themselves for both the longitudinal part as well as the transverse part agree very closely (which should not astonish as these quantities have been fitted).

The **relativistic GGA correlation functional** is not of the same quality (see Fig.3), still there is an order of magnitude improvement for the relativistic correlation contribution over the LDA (referred to MBPT as a standard). The fact that the situation for the correlation contribution is far from settled is illustrated for the case of neutral Xe. For this case the following results are available:

$$\begin{aligned}\Delta E_c &= 145 mhartree \quad RGGA\ [53] \\ &= 80 mhartree \quad RLDA\ [21] \\ &= 146 mhartree \quad DCB - MBPT2\ [57] \\ &= 105 mhartree \quad DCB - Coupled - Cluster\ [77]\ .\end{aligned}$$

6 Final Remarks

Here we offer some remarks on additional points and future problems: The point that we did not discuss (although we have a large number of case studies) is the question in how far the local quantities (rather than the integrated quantities) are reasonable. This can most easily be discussed by comparison

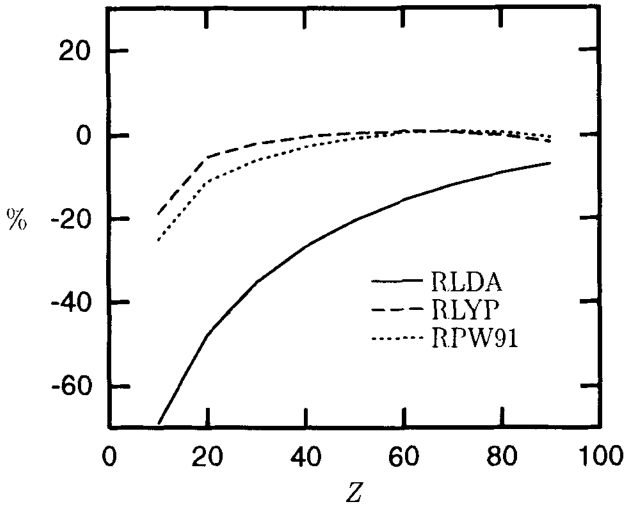


Fig. 3. Relativistic correction ΔE_c to the correlation energy: Percentage deviation of self-consistent RLDA- and RGGA-results from MBPT2-data for Ne isoelectronic series.

of the corresponding potentials, which show shell structure and finer effects more closely [21,6,52,53].

Obviously, there is much to be done: First the calculations for atoms have to be extended to the case of nonspherical systems (with the possibility of "spin polarisation", which in the relativistic case manifests itself in the appearance of current contributions). Of greater interest is, however, the investigation of relativistic effects in more complex systems, as for instance

i) Diatomic systems
   → changes in bond lengths, dissociation energies etc.
ii) Solids
   → changes in band structure features (eg. Fermi surfaces) and cohesive properties
iii) Pseudopotentials
   → as for heavy atoms even the outermost orbitals are affected by relativistic corrections, there is a modification of the pseudopotential [70]

Finally, some topics that have been addressed in the literature but have not been presented here (due to the usual lack of time), should at least be recorded.

- The discussion of RDFT has been extended to the case of strong, short range interactions on the basis of the field theoretical meson exchange model of nuclear physics, that is quantum hadrodynamics (QHD). Both ETF- [78,79] as well as KS-applications [80,81] have been given. In the latter instance it is of interest to note, that, due to the nature of the dominant interaction, results obtained with the x-only LDA agree quite closely with HF-results, which are available for a number of nuclei. The multiplicative character of the KS-exchange, however, allows also the consideration of superheavy nuclei [82], which, at the moment, are still not accessible via the HF-approach.

- Thermal RDFT has been discussed both on the basis of QED [83] as well as QHD [84]. Applications are restricted to thermal ETF-models [85].

Acknowledgements:
The rather extensive investigation of RDFT would not have been possible without several generations of Diploma and doctoral students. We take this opportunity to thank E. K. U. Groß, P. Malzacher, H. Müller, C. Speicher, T. Stroucken and W. Pohlner for their contributions in the earlier stages of this work. During the last year the "crew" included S. Keller, A. Facco Bonetti, R. N. Schmid, A. Höck and I. Andrejkovics.

References

1. Dreizler R M, da Providencia J (eds) (1985) Density Functional Methods in Physics. NATO ASI Series B, Vol 123, Plenum, New York
2. Parr R G, Yang W (1989) Density-Functional Theory of Atoms and Molecules. Oxford University Press, New York
3. Jones R O, Gunnarsson O (1989) Rev Mod Phys 61: 689
4. Dreizler R M, Gross E K U (1990) Density Functional Theory. Springer, Berlin
5. Gross E K U, Dreizler R M (eds) (1995) Density Functional Theory. NATO ASI Series B, Vol 337, Plenum, New York
6. Engel E, Dreizler R M (1996) Top Curr Chem 181: 1
7. Kaldor U, Heß B A (1994) Chem Phys Lett 230: 1
8. Dyall K G, Grant I P, Johnson C T, Parpia F A, Plummer E P (1989) Comp Phys Commun 55: 425
9. Pizlo A, Jansen G, Heß B A, von Niessen W (1993) J Chem Phys 98: 3945
10. Bjorken J D, Drell S D (1965) Relativistic Quantum Fields. McGraw-Hill, New York, chap 14.6
11. Itzykson C, Zuber J-B (1980) Quantum Field Theory. McGraw-Hill, New York
12. Weinberg S (1995) The Quantum Theory of Fields I/II. Cambridge University Press, London
13. Fetter A L, Walecka J D (1971) Quantum Theory of Many-Particle Systems. McGraw-Hill, New York
14. Inkson J C (1984) Many-Body Theory of Solids. Plenum, New York
15. t'Hooft G, Veltman M (1972) Nucl Phys B 44: 189
16. Hohenberg P, Kohn W (1964) Phys Rev 136 B: 864
17. Rajagopal A K, Callaway J (1973) Phys Rev B 7: 1912

18. Rajagopal A K (1978) J Phys C 11: L943
19. MacDonald A H, Vosko S H (1979) J Phys C 12: 2977
20. Engel E, Müller H, Speicher C, Dreizler R M (1995) In: Ref.[5], p 65
21. Engel E, Keller S, Facco Bonetti A, Müller H, Dreizler R M (1995) Phys Rev A 52: 2750
22. Engel E, Dreizler R M (1987) Phys Rev A 35: 3607
23. Engel E, Dreizler R M (1988) Phys Rev A 38: 3909
24. Kohn W, Sham L J (1965) Phys Rev 140 A: 1133
25. Greiner W, Müller B, Rafelski J (1985) Quantum Electrodynamics of Strong Fields. Springer, Berlin
26. Sharp R T, Horton G K (1953) Phys Rev 90: 317
27. Talman J D, Shadwick W F (1976) Phys Rev A 14: 36
28. Sahni V, Gruenebaum J, Perdew J P (1982) Phys Rev B 26: 4371
29. Sham L J (1985) Phys Rev B 32: 3876
30. Shadwick B A, Talman J D, Norman M R (1989) Comp Phys Commun 54: 95
31. Vignale G, Rasolt M (1987) Phys Rev Lett 59: 2360
32. Vignale G, Rasolt M (1988) Phys Rev B 37: 10685
33. Vignale G (1995) In: Ref.[5]
34. Ceperley D M, Alder B J (1980) Phys Rev Lett 45: 566
35. Akhiezer I A, Peletminskii S V (1960) Sov Phys JETP 11: 1316
36. Zapolsky H S (1960) Cornell University LNS Report. unpublished
37. Jancovici B (1962) Nuovo Cim XXV: 428
38. Ramana M V, Rajagopal A K, Johnson W R (1982) Phys Rev A 25: 96
39. Ramana M V, Rajagopal A K (1981) Phys Rev A 24: 1689
40. Müller H, Serot B D (1995) unpublished
41. Chin S A (1977) Ann Phys (NY) 108: 301
42. Andrejkovics I, Facco Bonetti A, Müller H, Engel E, Dreizler R M (1997) unpublished
43. Freedman B A, McLerran L D (1977) Phys Rev D 16: 1130, 1147, 1169
44. Ma S-K, Brueckner K A (1968) Phys Rev 165: 18
45. Herman F, Van Dyke J P, Ortenburger I B (1969) Phys Rev Lett 22: 807
46. Sham L J (1971) In: Marcus P M, Janak J F, Williams A R (eds) Computational Methods in Band Theory. Plenum, New York, p 458
47. Geldart D J W, Rasolt M (1976) Phys Rev B 13: 1477
48. Antoniewicz P R, Kleinman L (1985) Phys Rev B 31: 6779
49. Kleinman L, Lee S (1988) Phys Rev B 37: 4634
50. Chevary J A, Vosko S H (1990) Phys Rev B 42: 5320
51. Engel E, Vosko S H (1990) Phys Rev B 42: 4940; Phys Rev B 44: 1446(E)
52. Engel E, Keller S, Dreizler R M (1996) Phys Rev A 53: 1367
53. Engel E, Keller S, Dreizler R M (1997) In: Dobson J F, Vignale G, Das M P (eds) Electronic Density Functional Theory: Recent Progress and New Directions. Plenum, New York
54. Becke A D (1988) Phys Rev A 38: 3098
55. Engel E, Chevary J A, Macdonald L D, Vosko S H (1992) Z Phys D 23: 7
56. Perdew J P (1991) In: Ziesche P, Eschrig H (eds) Electronic Structure of Solids 1991. Akademie Verlag, Berlin
57. Ishikawa Y, Koc K (1994) Phys Rev A 50: 4733
58. Lee C, Yang W, Parr R G (1988) Phys Rev B 37: 785

59. Malzacher P, Dreizler R M (1986) Z Phys D 2: 37
60. Engel E, Müller H, Dreizler R M (1989) Phys Rev A 39: 4873
61. Vallarta M S, Rosen N (1932) Phys Rev 41: 708
62. Pohlner W F, Dreizler R M (1991) Phys Rev A 44: 7165
63. Kirzhnits D A (1967) Field Theoretical Methods in Many-Body Systems. Pergamon, London
64. Hodges C H (1973) Can J Phys 51: 1428
65. Müller H, Engel E, Dreizler R M (1989) Phys Rev A 40: 5542
66. Engel E, Dreizler R M (1989) J Phys B 22: 1901
67. Das M P, Ramana M V, Rajagopal A K (1980) Phys Rev A 22: 9
68. MacDonald A H, Daams J M, Vosko S H, Koelling D D (1981) Phys Rev B 23: 6377
69. MacDonald A H, Daams J M, Vosko S H, Koelling D D (1982) Phys Rev B 25: 713
70. Bachelet G B, Hamann D R, Schlüter M (1982) Phys Rev B 26: 4199
71. Rieger M M, Vogl P (1995) Phys Rev A 52: 282
72. Mayer M, Häberlein O D, Rösch N (1996) Phys Rev A 54: 4775
73. Engel E, Vosko S H (1993) Phys Rev A 47: 2800
74. In all our calculations the nuclei were represented by uniformly charged spheres with nuclear radii given by $R_{nucl} = 1.0793\ A^{1/3} + 0.73587\ fm$, A being the atomic mass (weighted by isotopic abundances) taken from Table III.7 of Ref.[86] unless explicitly stated otherwise. The speed of light has been set to $c = 137.0359895$.
75. Chakravorty S J, Gwaltney S R, Davidson E R, Parpia F A, Froese Fischer C (1993) Phys Rev A 47: 3649
76. Vosko S H, Wilk L, Nusair M (1980) Can J Phys 58: 1200
77. Eliav E, Kaldor U, Ishikawa Y (1994) Phys Rev A 49: 1724
78. Speicher C, Dreizler R M, Engel E (1992) Ann Phys (NY) 213: 312
79. Speicher C, Engel E, Dreizler R M (1993) Nucl Phys A562: 569
80. Schmid R N, Engel E, Dreizler R M (1995) Phys Rev C 52: 164
81. Schmid R N, Engel E, Dreizler R M (1995) Phys Rev C 52: 2804
82. Schmid R N, Engel E, Dreizler R M (1997) to be publ. in Found of Phys
83. Stroucken H J A C, Dreizler R M (1991) Phys Rev A 43: 3401
84. Müller H, Dreizler R M (1992) Z Phys A 341: 417
85. Müller H, Dreizler R M (1993) Nucl Phys A563: 649
86. Hisaka K, et al. (Particle Data Group) (1992) Phys Rev D 45: Number 11, Part II

Subject Index

Printing: Weihert-Druck GmbH, Darmstadt
Binding: Buchbinderei Schäffer, Grünstadt