Treatment of Magnetic Fields in Density-Functional Theory

Thesis submitted for the degree of *Philosophiae Doctor*

by

Sarah Reimann

Centre for Theoretical and Computational Chemistry/ Hylleraas Centre for Quantum Molecular Sciences
Department of Chemistry
Faculty of Mathematics and Natural Sciences
University of Oslo

Series of dissertations submitted to the
Faculty of Mathematics and Natural Sciences, University of Oslo
No. 2027

ISSN 1501-7710

Cover: Hanne Baadsgaard Utigard.
Print production: Reprosentralen, University of Oslo.

Abstract

Magnetic properties are an important application area in quantum chemistry. However, for the most widely used method in electronic structure calculations, density-functional theory, it is still not understood how to include the magnetic field into the exchange–correlation functional. With this work, we contribute to and further develop magnetic-field density-functional theory (BDFT), which is an alternative to, but less well-known, than current density-functional theory (CDFT). Using the four-way correspondence of convex analysis, we put BDFT into a common framework with CDFT and compare both approaches.

In order to improve upon current density functional approximations (DFAs), we extensively study their shortcomings when used for magnetic properties. We consider the field dependence of each Kohn–Sham energy component individually and study the self-consistent density of several frequently used DFAs. We make the important observation that the main error of present DFAs arises from their poor self-consistent densities. To fully benefit from a field dependence, it is therefore necessary to improve the density first.

Another main part of this work is to study and make use of adiabatic connection (AC) curves for systems including magnetic fields, both with highly accurate ab-initio methods and with DFAs. We show that the behaviour of AC curves depends on the choice of geometry. If the equilibrium geometry for zero field strength is kept for all fields, the curves bend more with increasing field strength, corresponding to an increase of static correlation. When each curve is calculated at the equilibrium geometry of that field strength, the magnetic field does not result in any significant changes regarding the shape of the AC curves. For AC curves computed with DFAs, the main error in the presence of a magnetic field is already there at zero field strength. To obtain more accurate results for magnetic properties with DFAs, it is therefore important not only to introduce the missing field dependence, but also to reduce the errors of present, field-independent functionals.

The investigations and results of this work provide valuable information both for the further theoretical development of BDFT and for the proper inclusion of magnetic field effects into DFAs, which will allow more accurate results for all kinds of density-functional calculations involving magnetic fields.

Preface

This thesis is the result of my work at the University of Oslo from August 2013 to April 2018 at the Centre for Theoretical and Computational Chemistry (CTCC), which was then taken over by the Hylleraas Centre for Quantum Molecular Sciences.

First, I would like to thank my main supervisor Trygve Helgaker for letting me work on this project. I am grateful for the support, the numerous discussions and all advice during those years, as well as to be given enough time and financial support to finish the project in the final stressful period. I furthermore appreciate having had the opportunity to attend several summer and winter schools, in addition to various conferences, allowing me to experience not only the work in the office but also the quantum chemistry community "in the big world".

I would further like to particularly thank Ulf Ekström, my first co-supervisor, who took really well care of me in the first years of the project. With my background from physics, he introduced me patiently to the language of quantum chemistry and guided me and the project into the right direction.

Moreover, I would very much like to thank Alex Borgoo, my second co-supervisor, who took over after Ulf finished working in the department. Changing the supervisor in the midst of a project is always challenging, but Alex took over really well, and was particularly helpful in the completion of the project, when the publication of papers and the writing of the thesis needed to be sped up in order to come to an end.

I would further like to thank all co-authors of the papers in my thesis. Since this thesis is centred around those papers, it would definitely not have been possible without all their work and effort regarding those papers, where we all complemented each other to get the best results. I would particularly like to thank Andy Teale, who his co-author in all of my papers and who answered the many questions that I had, especially concerning the use of the Dalton code, almost immediately, no matter at which time of the day (or night) I sent a mail. I am also grateful for the week he let me spend with him at the University of Nottingham, where I was impressed over how he so patiently and calmly debugged the completely convoluted code we had to work with.

I also want to thank all members of the CTCC for being a very pleasant group to work with. As daily user of the CTCC kitchen for making lunch, I hopefully did not interfere

with the coffee cookers too much, and especially apologize to Thomas Bondo Pedersen for the frequent cleaning of my pot exactly when he needed the water tap for the coffee. I also want to mention Jan Ingar Johnsen, our head of office, who was of incredible help with respect to any practical or administrative issues. In addition, it was nice to clear my mind now and then by having casual chats with him, or by watching the beautiful Norwegian sunset from his office window.

A special thanks goes to Lukas Wirz and Benjamin Knorr for reading through this thesis and helping to correct spelling and grammar, as well as pointing out potential ambiguities for people with another scientific background. Last, but not least, I would like to thank my family and friends who have supported and encouraged me throughout my studies and when working with this project.

It is important to realize that the results of this PhD go far beyond the numerical results presented here. What is presented here, are only the final outcomes. What is hidden and can only vaguely be imagined, is the process itself. Four years of intensive work on a project. Four years of daily challenges, four years with many reasons for frustration, worries and trouble. But also many opportunities to learn how to deal with them. Learning to accept that there are issues that we just have not understood and cannot understand with the data we have and can obtain by the end of the project. This means agreeing upon that we do not always find the final answers ourselves, but being content with having provided necessary preliminaries that other people can build upon. Learning how to face challenges as a group, realizing how far one can push oneself, but also seeing one's own limitations and knowing when to ask for help. All these are incredibly valuable skills that a PhD project gives many opportunities to work with, even if this does not become evident by reading a thesis. But finally these are the skills that can be taken along one's way. Many calculations fail, many of the numbers might get forgotten, but these experiences can be used and applied in whichever direction life evolves. And it evolves every day and leads us somewhere.

List of Papers

This thesis is based on the work presented in the following papers. A copy of each paper can be found at the end of the thesis.

I *The importance of current contributions to shielding constants in density-functional theory.*
S. Reimann, U. Ekström, S. Stopkowicz, A. M. Teale, A. Borgoo and T. Helgaker
Phys. Chem. Chem. Phys., Volume 17, Pages 18834-18842, 2015
DOI:10.1039/c5cp02682b

II *Magnetic-Field Density-Functional Theory (BDFT): Lessons from the Adiabatic Connection.*
S. Reimann, A. Borgoo, E. I. Tellgren, A. M. Teale and T. Helgaker
J. Chem. Theory Computat., Volume 13, Pages 4089-4100, 2017
DOI:10.1021/acs.jctc.7b00295

III *Kohn–Sham energy decomposition for molecules in magnetic fields.*
S. Reimann, J. Austad, A. Borgoo, E. I. Tellgren, A. M. Teale, T. Helgaker and S. Stopkowicz
Submitted to: *Molecular Physics*

Contents

Abbreviations

AC	Adiabatic Connection
BDFT	Magnetic-Field Density-Functional Theory
CCSD	Coupled-Cluster with Single and Double excitations
CCSD(T)	Coupled-Cluster with Single, Double and Perturbative Triple excitations
CDFT	Current Density-Functional Theory
DFA	Density Functional Approximation
DFT	Density-Functional Theory
FCI	Full Configuration Interaction
GGA	Generalized Gradient Approximation
LDA	Local Density Approximation
mGGA	Meta-Generalized Gradient Approximation
NMR	Nuclear Magnetic Resonance
OEP	Optimized Effective Potential

Chapter 1

Introduction

1.1 Motivation

Molecules in magnetic fields are an active area of research, studied as ultra-strong fields in astrophysics [1, 2, 3] and as weaker fields in quantum chemistry. In chemistry, magnetic properties like nuclear magnetic resonance (NMR) shielding constants and magnetizabilities play a major role. Over the last decades, quantum chemical calculations have reached a stage where they can strongly help to interpret experimental measurements, among others in magnetic-resonance spectroscopy [4, 5]. Using various levels of electronic structure theory, spectroscopic constants can gradually be brought closer to the exact, experimental value. But not only accuracy is progressively improved upon, in addition great effort is made to handle larger and larger systems [6, 7, 8]. The goal is that magnetic properties of systems comprised of hundreds of atoms can routinely be computed in the future.

Density functional theory (DFT) is the most widely used method in quantum chemistry. With the electron density as the basic variable, it is less expensive than methods working directly with the wave function. Moreover, it shows a faster basis-set convergence [9]. Since the exact universal density functional of DFT is not known, a large number of density functional approximations (DFAs) have been developed in the last decades.

However, those functionals have been developed without the consideration of magnetic fields. In the presence of a magnetic field, the electronic Hamiltonian changes, which needs to be reflected in the DFAs. How exactly the magnetic field enters the exchange–correlation functional is still unresolved. Therefore, with present DFAs, the results of DFT for molecular magnetic properties are not always satisfactory [10, 11, 12, 13]. The development of such functionals is an active field of research [13, 14].

1.2 Problems Addressed in this Work

This work aims at identifying the shortcomings of present DFAs regarding the computation of magnetic properties, since those errors point to the areas where the effects of the magnetic field are not properly understood and need to be improved upon when introducing a field-dependence into exchange–correlation functionals.

We particularly motivate the further development of magnetic-field density-functional theory (BDFT), which we demonstrate to be a worthwhile alternative to current density-functional theory (CDFT), and which has not been developed further after its introduction by Grayce and Harris in 1994 [15]. We present both BDFT (in Section 2.2) and CDFT (in Section 2.3) and put them into a common framework using the four-way correspondence of convex analysis (Section 2.4). We particularly focus on establishing and studying adiabatic connection (AC) curves in the framework of BDFT (Subsection 2.2.3 and content of Paper II).

We investigate magnetic properties in the context of DFT, since the prediction of those properties relies on an explicit dependence on the magnetic field. That way, magnetic properties allow us to apply the concepts of BDFT and serve as a test case of how well the field dependence is understood. Our study of magnetic properties is mainly presented in the papers attached to this work and summarized in Chapter 5. There, among other things, we quantify how large the field-dependent part of the NMR shielding constant and the magnetizability is, analyse the errors of existing DFAs, and study the field dependence of the Kohn–Sham energy components.

The results give insight into the behaviour of the exact energy components and our conclusions provide impetus for the development of improved DFAs.

Chapter 2

Exact Density-Functional Theory in a Magnetic Field

In this chapter, we discuss theoretical aspects of DFT in magnetic fields. We here consider the exact theory from a mathematical point of view, whereas approximations that are needed for practical calculations are dealt with in Chapters 3 and 4.

First, we present the basics of DFT without magnetic fields. This shows how the theory was historically developed and, outlining the fundamental theorems, it gives us the opportunity to introduce notation. We then expand on this first part by including magnetic fields. We establish and discuss first BDFT, then CDFT, and consider the adiabatic connection in those two frameworks. Finally, we put BDFT and CDFT into a common framework and relate them to each other.

2.1 Density-Functional Theory in the Absence of a Magnetic Field

Formally, DFT is an exact approach to determine the quantum mechanical ground state of a many-body system. DFT can be formulated using three different formalisms, the Hohenberg–Kohn formalism, the Levy–Lieb constrained-search formalism and Lieb's convex–conjugate theory. All three theories are discussed in the following, but first we introduce the Rayleigh–Ritz variation principle, which forms the basis of all three theories and DFT in general.

2.1.1 Rayleigh–Ritz Variation Principle

Considering a system of N interacting electrons, the Rayleigh–Ritz variation principle states that the ground-state energy is given by

$$E(v) = \inf_{\Psi \in \mathcal{W}_N} \langle \Psi | \hat{H}(v) | \Psi \rangle, \tag{2.1}$$

where $\hat{H}(v)$ is the molecular electronic Hamiltonian and $\mathcal{W}_N$ is the set of all normalized anti-symmetric N-electron wave functions Ψ with a finite Hamiltonian expectation value. The spin-free electronic Hamiltonian can be written as

$$\hat{H}(v) = -\frac{1}{2} \sum_{i=1}^{N} \nabla_i^2 + \sum_{i<j} \frac{1}{r_{ij}} + \sum_{i=1}^{N} \left(- \sum_{K} \frac{Z_K}{r_{iK}} \right) \tag{2.2}$$

$$= \hat{T} + \hat{V}_{ee} + \sum_{i=1}^{N} v(\mathbf{r}_i), \tag{2.3}$$

with interelectronic distance $r_{ij} = |\mathbf{r}_i - \mathbf{r}_j|$, kinetic energy operator $\hat{T}$, electron–electron interaction $\hat{V}_{\text{ee}}$, external potential $\sum_i v(\mathbf{r}_i)$ due to nuclear attraction, and where Z_K denotes the charge of nucleus K. To arrive at this Hamiltonian, the Born-Oppenheimer approximation has been adopted. This approximation is needed for all theory in this work. It is based on the fact that the nuclei are several orders of magnitude heavier than the electrons and therefore move much slower than the electrons. To a good approximation, the geometry of the nuclei can then be considered to be fixed, since the electrons will immediately adapt to each nuclear configuration. For the remainder of this work, we always work within the Born-Oppenheimer approximation, assuming fixed positions of the nuclei.

2.1.2 Hohenberg–Kohn Theory

For a given number of electrons, the specific form of the Hamiltonian (2.3) is determined by the nuclear potential v, which is a particular case of a more general multiplicative external potential. As a consequence of this form of the Hamiltonian, the wave function Ψ, as a solution of the Schrödinger equation (2.4),

$$\hat{H}(v)\Psi = E(v)\Psi, \tag{2.4}$$

is determined by this external potential, too. This wave function Ψ is used the construct the N-electron density $\rho(\mathbf{r})$, which depends only on three spatial coordinates $\mathbf{r}$,

$$\rho(\mathbf{r}) = N \int |\Psi(\mathbf{r}, s, \mathbf{x}_2, \ldots, \mathbf{x}_N)|^2 \, \mathrm{d}s \, \mathrm{d}\mathbf{x}_2 \ldots \mathrm{d}\mathbf{x}_N, \tag{2.5}$$

where the integration goes over $N-1$ space-spin coordinates $\mathbf{x}_i = (\mathbf{r}_i, s_i)$ and the spin index s, but not the spatial coordinate r, of the N-th particle. Since v determines Ψ and Ψ is used to construct $\rho(\mathbf{r})$, it is quite obvious that, for non-degenerate ground states, $\rho(\mathbf{r})$ then is a functional of v.

In 1964, Hohenberg and Kohn published a paper that states that the converse is also true [16]. This statement became later known as the Hohenberg–Kohn theorem and lay historically the theoretical foundation for DFT. In particular, it states the following:

First Hohenberg–Kohn Theorem. *The ground-state density ρ determines the external potential v up to a scalar.*

In other words, the theorem establishes that up to an arbitrary constant, there is a one-to-one correspondence between $v(\mathbf{r})$ and $\rho(\mathbf{r})$. A proof is given in Appendix A.

Representability restrictions It should be noted that Hohenberg–Kohn theory is only valid for the restricted sets of ρ-representable potentials $\mathcal{V}_N$ and v-representable densities $\mathcal{A}_N$ [17, 18],

$$\mathcal{V}_N = \{v|\ \hat{H}(v) \text{ has an } N\text{-electron ground state}\}, \tag{2.6}$$

$$\mathcal{A}_N = \{\rho|\ \rho \text{ comes from an } N\text{-electron ground state}\}. \tag{2.7}$$

A v-representable density determines all properties of its associated ground state, since through the one-to-one mapping, the potential $v(\mathbf{r})$ and with it the specific form of the Hamiltonian $\hat{H}(v)$, are uniquely defined. However, the conditions for a density ρ to be v-representable and for a potential v to be ρ-representable are still unknown. Neither $\mathcal{V}_N$ nor $\mathcal{A}_N$ are explicitly known and *a priori*, it is impossible to determine whether a given density ρ belongs to $\mathcal{A}_N$ or whether a given potential v belongs to $\mathcal{V}_N$.

However, this restriction of Hohenberg–Kohn theory is lifted in Subsection 2.1.3, where the set of densities and potentials is extended to a larger and known set.

Hohenberg–Kohn functional Let us denote the ρ-representable potential which supports a particular ground state with density ρ by v_ρ. The universal *Hohenberg–Kohn functional* $F_{\mathrm{HK}}(\rho)$ can then be defined as

$$F_{\mathrm{HK}}(\rho) = E(v_\rho) - (v_\rho|\rho), \qquad v \in \mathcal{V}_N, \rho \in \mathcal{A}_N, \tag{2.8}$$

with $(v_\rho|\rho) = \int v_\rho(\mathbf{r})\rho(\mathbf{r})\,\mathrm{d}\mathbf{r}$. With the Hamiltonian of Eq. (2.3), this is equivalent to

$$F_{\mathrm{HK}}(\rho) = T(\rho) + V_{\mathrm{ee}}(\rho), \quad \rho \in \mathcal{A}_N, \tag{2.9}$$

where the kinetic energy $T(\rho)$ and the electron-electron interaction term $V_{\mathrm{ee}}(\rho)$ are obtained as expectation values with the ground state wave function $\Psi(\rho)$ associated with the density ρ,

$$T(\rho) = \langle \Psi(\rho) | \hat{T} | \Psi(\rho) \rangle, \quad \rho \in \mathcal{A}_N, \tag{2.10}$$

$$V_{\mathrm{ee}}(\rho) = \langle \Psi(\rho) | \hat{V}_{ee} | \Psi(\rho) \rangle, \quad \rho \in \mathcal{A}_N. \tag{2.11}$$

This decomposition is unique in the absence of degeneracies.

Hohenberg–Kohn variation principle

In their landmark paper [16], Hohenberg and Kohn restated the Rayleigh-Ritz variational principle in terms of the density. This theorem is also known as the second Hohenberg–Kohn theorem and can be formulated the following way, valid for all $v \in \mathcal{V}_N$,

$$E(v) = \min_{\rho \in \mathcal{A}_N} \left(F_{\mathrm{HK}}(\rho) + (v|\rho) \right). \tag{2.12}$$

This result validates the use of the variational principle in DFT. It is much simpler than the Rayleigh–Ritz variation principle, since not the complicated wave function, but the ground state density ρ is used, which is a function of merely three variables, independent of the number N of electrons. In their original paper [16], Hohenberg and Kohn treated only non-degenerate ground states, but their results are also valid for degenerate systems. However, Hohenberg–Kohn theory has limitations related to representability, since it relies on the v-representability of the electron density ρ and the ρ-representability of the potential v and the sets $\mathcal{A}_N$ and $\mathcal{V}_N$ are not explicitly known. Moreover, reasonable non-v-representable densities have been shown to exist [19, 20].

2.1.3 Levy–Lieb Constrained-Search Formalism

The limitations of Hohenberg–Kohn theory were removed with the introduction of the Levy–Lieb constrained-search formalism. The Levy–Lieb constrained-search functional is defined as [21, 19]

$$F_{\mathrm{LL}}(\rho) = \min_{\Psi \mapsto \rho} \langle \Psi | \hat{T} + \hat{V}_{\mathrm{ee}} | \Psi \rangle. \tag{2.13}$$

Thus the Hohenberg–Kohn variation principle of Eq. (2.12) can in the Levy–Lieb constrained-search formalism be restated as

$$\begin{aligned} E(v) &= \inf_{\rho\in\mathcal{I}_N}\left(\inf_{\Psi\mapsto\rho}\langle\Psi|\hat{T}+\hat{V}_{\mathrm{ee}}+\sum_i v(\mathbf{r}_i)|\Psi\rangle\right) \\ &= \inf_{\rho\in\mathcal{I}_N}\left(\inf_{\Psi\mapsto\rho}\langle\Psi|\hat{T}+\hat{V}_{\mathrm{ee}}|\Psi\rangle+(v|\rho)\right) \\ &= \inf_{\rho\in\mathcal{I}_N}\left(F_{\mathrm{LL}}(\rho)+(v|\rho)\right), \quad v\in X^*. \end{aligned} \tag{2.14}$$

Compared to Hohenberg–Kohn theory, the set of allowed densities is extended from the set $\mathcal{A}_N$ of v-representable densities to the set $\mathcal{I}_N$ of N-representable densities [18],

$$\mathcal{I}_N = \{\rho|\ \rho \text{ can be obtained from some } \Psi\in\mathcal{W}_N\}. \tag{2.15}$$

The set $\mathcal{I}_N$ can be shown to be equivalent to the set of all non-negative functions that integrate to N and have a finite von Weizsäcker kinetic energy

$$T_W(\rho) = \frac{1}{2}\int\left|\nabla\rho^{1/2}(\mathbf{r})\right|^2\,\mathrm{d}\mathbf{r} = \frac{1}{8}\int|\nabla\rho(\mathbf{r})|^2\,\rho^{-1}(\mathbf{r})\,\mathrm{d}\mathbf{r}. \tag{2.16}$$

In other words, $\mathcal{I}_N$ can be defined as

$$\mathcal{I}_N = \{\rho\mid \rho\geq 0,\ \int\rho(\mathbf{r})\,\mathrm{d}\mathbf{r}=N,\ T_W(\rho)<\infty\}. \tag{2.17}$$

At the same time, the set of admissible external potentials is extended to $X^* = L^{3/2}+L^\infty$, where the space $L^{3/2}$ is a L^p-space with $p=3/2$. An L^p-space is a space of functions f on a set S, for which the p-th power of the absolute value is Lebesgue integrable,

$$\|f\|_p = \left(\int_S |f|^p\,\mathrm{d}\mu\right)^{1/p} < \infty. \tag{2.18}$$

For functions f in L^∞, there exists a positive constant $C<\infty$, such that $f(s)<C$ for almost all $s\in S$. The minimization in Eq. (2.13) is constrained in the way that one only searches in the space of trial wave functions Ψ that give the ground state density ρ. Lieb proved that, for the Levy–Lieb functional of Eq. (2.13), a minimizing wave function exists for all $\rho\in\mathcal{I}_N$ [20]. Thus F_{LL} can also be expressed in terms of this minimizing wave function Ψ_ρ,

$$F_{\mathrm{LL}}(\rho) = \min_{\Psi\mapsto\rho}\langle\Psi|\hat{T}+\hat{V}_{\mathrm{ee}}|\Psi\rangle = \langle\Psi_\rho|\hat{T}+\hat{V}_{\mathrm{ee}}|\Psi_\rho\rangle. \tag{2.19}$$

The Levy–Lieb functional F_{LL} is a continuation of the Hohenberg–Kohn functional F_{HK} from the unknown domain $\mathcal{A}_N$ to the known domain $\mathcal{I}_N$ [17], with

$$F_{\mathrm{LL}}(\rho) = F_{\mathrm{HK}}(\rho), \quad \rho \in \mathcal{A}_N. \tag{2.20}$$

The N-representability condition is weaker than the v-representability condition and is satisfied for any reasonable density.

2.1.4 Lieb's Convex–Conjugate Theory

Lieb formulated DFT in terms of convex conjugation, applying general results from convex analysis to DFT [20]. For a more general discussion of convex analysis, we refer to Ref. [22]. Lieb's convex-conjugate theory is based on the observation that $E(v)$ is concave and continuous in v. While continuity is difficult to show, concavity is simple to prove. It is instructive to demonstrate concavity explicitly, since this shows how the concavity of $E(v)$ depends on the linearity of the Hamiltonian $\hat{H}(v)$ in v and the Rayleigh–Ritz variation principle.

Proof of concavity of $E(v)$ To prove that $E(v)$ is concave in v, we have to show that for each $v = \lambda v_1 + (1-\lambda)v_2$, with $\lambda \in (0,1)$, it holds that $E(\lambda v_1 + (1-\lambda)v_2) \geq \lambda E(v_1) + (1-\lambda)E(v_2)$: Using the Rayleigh–Ritz variation principle, we have that

$$E(\lambda v_1 + (1-\lambda)v_2) = \inf_{\Psi} \langle \Psi | \hat{H}(\lambda v_1 + (1-\lambda)v_2) | \Psi \rangle. \tag{2.21}$$

The Hamiltonian is linear in v, which can easily be shown using Eq. (2.3),

$$\begin{aligned} \hat{H}(\lambda v_1 + (1-\lambda)v_2) &= \hat{T} + \hat{V}_{\mathrm{ee}} + \lambda v_1 + (1-\lambda)v_2 \\ &= \lambda \hat{T} + \lambda \hat{V}_{\mathrm{ee}} + \lambda v_1 + (1-\lambda)\hat{T} + (1-\lambda)\hat{V}_{\mathrm{ee}} + (1-\lambda)v_2 \\ &= \lambda \hat{H}(v_1) + (1-\lambda)\hat{H}(v_2), \end{aligned} \tag{2.22}$$

where we have used that $\lambda + (1-\lambda) = 1$, and where we, for simplicity, have omitted the summation over particles for the potential. Using the linearity of the Hamiltonian,

Eq. (2.21) then becomes

$$\begin{aligned} E(\lambda v_1 + (1-\lambda)v_2) &= \inf_{\Psi} \langle \Psi | \lambda \hat{H}(v_1) + (1-\lambda)\hat{H}(v_2) | \Psi \rangle \\ &= \inf_{\Psi} \Big(\langle \Psi | \lambda \hat{H}(v_1) | \Psi \rangle + \langle \Psi | (1-\lambda)\hat{H}(v_2) | \Psi \rangle \Big) \qquad (2.23) \\ &\geq \inf_{\Psi} \langle \Psi | \lambda \hat{H}(v_1) | \Psi \rangle + \inf_{\Psi} \langle \Psi | (1-\lambda)\hat{H}(v_2) | \Psi \rangle \qquad (2.24) \\ &= \lambda \inf_{\Psi} \langle \Psi | \hat{H}(v_1) | \Psi \rangle + (1-\lambda) \inf_{\Psi} \langle \Psi | \hat{H}(v_2) | \Psi \rangle \\ &= \lambda E(v_1) + (1-\lambda) E(v_2). \qquad (2.25) \end{aligned}$$

In the step from (2.23) to (2.24), we have applied the Rayleigh–Ritz variation principle. While in Eq. (2.24), the minimizing wave function is found separately for $\hat{H}(v_1)$ and $\hat{H}(v_2)$, this flexibility is not possible when the combined system is minimized, such that the energy is higher. With Eq. (2.25), the proof is completed.

With the two fundamental properties of continuity and concavity of $E(v)$, it follows from convex analysis that there exists a convex function $F(\rho)$, such that

$$F(\rho) = \sup_{v \in X^*} \left(E_0(v) - (v|\rho) \right), \quad \rho \in X \qquad (2.26)$$

$$E(v) = \inf_{\rho \in X} \left(F(\rho) + (v|\rho) \right), \quad v \in X^* \qquad (2.27)$$

where $X = L^3 \cap L^1$ and $X^* = L^{3/2} + L^\infty$. From a physical point of view, ρ can now be identified with the electron density. The domain X is the Banach space (complete normed vector space) of densities that integrate to N particles, $\int \rho(\mathbf{r})\, \mathrm{d}\mathbf{r} = N$, and give a finite kinetic energy. The dual space X^* is the set of potentials v for which $(v|\rho)$ is finite for all $\rho \in X$.

The functional $F(\rho)$ is convex and lower semi-continuous, but everywhere discontinuous. A function f is said to be lower semi-continuous if for all x and for any $\epsilon > 0$, there exists a real number $\delta > 0$, such that $\|x - y\| < \delta$ implies that $f(x) - \epsilon < f(y)$.

Equation (2.26) is called *Lieb variation principle* and it is the inverse of the Hohenberg–Kohn variation principle of Eq. (2.27). The functionals $E(v)$ and $F(\rho)$ are said to be *conjugate* functionals. Note that each of the functionals contains enough information to generate the other. Likewise, ρ and v are conjugate variables. The variation principles are examples of Legendre–Fenchel transforms, analogous to the transformation between Lagrangians and Hamiltonians in classical mechanics. While the Hohenberg–Kohn minimizers are precisely the ground-state densities of v, the Lieb maximizers are the potentials v with ground state density ρ.

The Lieb functional $F(\rho)$ is not the same as the Levy–Lieb constrained-search func-

tional F_{LL} of Eq. (2.13) unless F_{LL} is expressed in terms of ensembles, since

$$F(\rho) = \begin{cases} \min_{\gamma \mapsto \rho} \operatorname{tr} \gamma(\hat{T} + \hat{V}_{\mathrm{ee}}) & \rho \in \mathcal{I}_N \\ +\infty & \rho \notin \mathcal{I}_N. \end{cases} \tag{2.28}$$

The density matrix $\gamma \in \mathcal{D}_N$ is given by

$$\gamma = \sum_i c_i |\psi_i\rangle\langle\psi_i|, \quad c_i \geq 0, \quad \sum_i c_i = 1, \tag{2.29}$$

with antisymmetric wave functions $\psi_i \in \mathcal{W}_N$ and where $\mathcal{D}_N$ is the set of all normalized N-electron density matrices with a finite kinetic energy [23]. It should be noted that $F(\rho)$ is finite only for N-representable densities $\rho \in \mathcal{I}_N \subset X$. However, since $E(v)$ is obtained as an infimum, only densities $\rho \in \mathcal{I}_N$ contribute to the ground-state energy.

In Lieb's convex-conjugate theory, the domains are explicitly known and since the Lieb functional $F(\rho)$ is convex and $E(v)$ concave, each solution is a global solution. Furthermore, this theory reflects the fundamental symmetry between densities and potentials. Unless otherwise indicated, we will therefore work in Lieb's convex-conjugate framework, using the Lieb functional $F(\rho)$, in the remainder of this thesis.

2.1.5 The Adiabatic Connection

The same procedure as above can be applied to different interaction strengths. By scaling the electron-electron interaction with a coupling parameter $\lambda \in [0, 1]$, a series of partially interacting systems is generated, with $\lambda = 0$ corresponding to the noninteracting Kohn–Sham system and $\lambda = 1$ corresponding to the physically interacting system.

While, in principle, there exist many possible ways of including the coupling parameter [24], the simplest and most common choice is to scale the electron–electron interaction $\hat{V}_{\mathrm{ee}}$ linearly. The Hamiltonian of Eq. (2.3) then reads

$$\hat{H}_\lambda(v) = \hat{T} + \lambda \hat{V}_{\mathrm{ee}} + \sum_i v(\mathbf{r}_i). \tag{2.30}$$

The total energy is given by the Rayleigh–Ritz variation principle,

$$E_\lambda(v) = \inf_{\Psi \in \mathcal{W}_N} \langle \Psi | \hat{H}_\lambda(v) | \Psi \rangle. \tag{2.31}$$

Since $E_\lambda(v)$ is concave and continuous in v, we have as before

$$F_\lambda(\rho) = \sup_{v \in X^*} \left(E_\lambda(v) - (v|\rho) \right), \quad \rho \in X \tag{2.32}$$

$$E_\lambda(v) = \inf_{\rho \in X} \left(F_\lambda(\rho) + (v|\rho) \right), \quad v \in X^*. \tag{2.33}$$

In contrast to the Lieb and Hohenberg–Kohn variation principle of Eqs. (2.26) and (2.27), now both the ground state energy E_λ and the Lieb functional F_λ depend on λ. In Eq. (2.32), F_λ is expressed in the framework of Lieb's convex–conjugate theory. For N-representable densities, F_λ may also be expressed in the constrained-search form,

$$F_\lambda(\rho) = \min_{\gamma \mapsto \rho} \operatorname{tr} \gamma \left(\hat{T} + \hat{V}_{\mathrm{ee}} \right) = \operatorname{tr} \gamma_\lambda^\rho \left(\hat{T} + \hat{V}_{\mathrm{ee}} \right), \quad \rho \in \mathcal{I}_N, \tag{2.34}$$

where γ_λ^ρ is one of the minimizers. We may now relate F_λ at $\lambda = 0$ and $\lambda = 1$ using the integral representation,

$$\begin{aligned} F_\lambda(\rho) &= F_0(\rho) + \int_0^\lambda F'_\mu(\rho) \, \mathrm{d}\mu \\ &= F_0(\rho) + \lambda F'_0(\rho) + \int_0^\lambda \left(F'_\mu(\rho) - F'_0(\rho) \right) \, \mathrm{d}\mu, \end{aligned} \tag{2.35}$$

where the prime indicates differentiation with respect to μ. The functional $F_\lambda(\rho)$ is continuous and concave in λ. Since it is nearly linear in λ, it is useful to extract the linear part as done above. Although in principle, $F_\lambda(\rho)$ may not be everywhere differentiable with respect to λ, it is everywhere right and left differentiable and we may take the prime to indicate one of these one-sided derivatives. A function $f(x)$ is said to be right differentiable at a point x_0, if the derivative exits as x_0 is approached from the right (at values $x = x_0 + \epsilon$) and $f(x)$ is said to be left differentiable at x_0 if the derivative is approached from the left (at values $x = x_0 - \epsilon$), with $\epsilon > 0$.

Starting from this mathematical definition, we now identify the individual terms. For $\lambda = 0$, the constrained-search version of the Lieb functional, given by Eq. (2.34), reduces to

$$F_0(\lambda) = \min_{\gamma \mapsto \rho} \operatorname{tr} \gamma \hat{T} = T_{\mathrm{s}}(\rho), \tag{2.36}$$

where we have defined the noninteracting kinetic energy $T_{\mathrm{s}}(\rho)$. To obtain the derivative $F'_\lambda(\rho)$, we apply the Hellmann–Feynman theorem:

$$F'_\lambda(\rho) = \frac{\mathrm{d} \operatorname{tr} \gamma_\lambda^\rho \left(\hat{T} + \hat{V}_{\mathrm{ee}} \right)}{\mathrm{d}\lambda} = \operatorname{tr} \gamma_\lambda^\rho \hat{V}_{\mathrm{ee}} = \mathcal{W}_\lambda(\rho). \tag{2.37}$$

We may now rewrite the adiabatic connection in the form

$$F_\lambda(\rho) = T_\mathrm{s}(\rho) + \lambda \mathcal{W}_0(\rho) + \int_0^\lambda \left(W_\mu(\rho) - \mathcal{W}_0(\rho)\right)\, \mathrm{d}\mu. \tag{2.38}$$

We introduce the Hartree functional $J(\rho)$, also called Coulomb term, which describes the classical part of the electron-electron interaction,

$$J(\rho) = \iint \frac{\rho(\mathbf{r})\rho(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|}\, \mathrm{d}\mathbf{r}\mathrm{d}\mathbf{r}'. \tag{2.39}$$

It constitutes one part of $\mathcal{W}_0(\rho)$, while the rest gives the exchange functional $E_\mathrm{x}(\rho)$,

$$E_\mathrm{x}(\rho) = \mathcal{W}_0(\rho) - J(\rho). \tag{2.40}$$

The last term of Eq. (2.38), in general much smaller than the linear term, is the correlation functional $E_{\mathrm{c},\lambda}(\rho)$,

$$E_{\mathrm{c},\lambda}(\rho) = \int_0^\lambda \left(\mathcal{W}_\mu(\rho) - \mathcal{W}_0(\rho)\right)\, \mathrm{d}\mu. \tag{2.41}$$

Using these definitions of $J(\rho)$, $E_\mathrm{x}(\rho)$ and $E_{\mathrm{c},\lambda}(\rho)$ in Eq. (2.38), we obtain the Kohn–Sham decomposition of the universal density functional,

$$F_\lambda(\rho) = T_\mathrm{s}(\rho) + \lambda J(\rho) + \lambda E_\mathrm{x}(\rho) + E_{\mathrm{c},\lambda}(\rho). \tag{2.42}$$

Often the exchange and correlation functional are combined to the exchange–correlation functional:

$$E_{\mathrm{xc},\lambda}(\rho) = \lambda E_x(\rho) + E_{\mathrm{c},\lambda}(\rho). \tag{2.43}$$

The exchange–correlation functional is the only unknown in the Lieb functional. The strength of the adiabatic connection, as noted by Becke already in the early 1990s, is that an integration of the AC curve with respect to λ gives the exchange–correlation functional [25].

Decomposition of the AC integrand To study the exchange–correlation part only, the full AC integrand $\mathcal{W}_\lambda(\rho)$, given by Eq. (2.37), is often split into parts excluding the known Hartree term and exchange energy. We here follow the notation used in Ref. [26], where

$$E_\mathrm{typ}(\rho) = \int_0^1 \mathcal{W}_{\mathrm{typ},\lambda}(\rho)\, \mathrm{d}\lambda \tag{2.44}$$

and "typ" in the AC integrand $\mathcal{W}_{\mathrm{typ},\lambda}$ can be Hartree-exchange–correlation (Hxc), exchange–correlation (xc), or correlation only (c). The following relations are useful to relate the

different quantities [27],

$$\mathcal{W}_{\mathrm{xc},\lambda}(\rho) = \mathrm{tr}\,\gamma_\lambda^\rho \hat{V}_{\mathrm{ee}} - J(\rho) \tag{2.45}$$

$$\mathcal{W}_{\mathrm{c},\lambda}(\rho) = \mathcal{W}_{\mathrm{xc},\lambda}(\rho) - \mathcal{W}_{\mathrm{xc},0}(\rho). \tag{2.46}$$

Since only the correlation energy $E_{\mathrm{c},\lambda}$ of Eq.(2.41) depends on λ in a non-trivial way, the main interest is to study the dependence of $E_{\mathrm{c},\lambda}$ on λ. A number of models to describe $W_{\mathrm{c},\lambda}$ have been proposed, and this is still an ongoing topic of research [26, 28].

Properties of Adiabatic Connection Curves

All AC curves share some common properties. Two important features are that $\mathcal{W}_\lambda(\rho)$ is monotonically decreasing in λ, and that $E_{\mathrm{c},\lambda}(\rho) \leq 0$ is concave in λ. This follows from the fact that $F_\lambda(\rho) \geq 0$ is concave in λ and that the term $T_{\mathrm{s}}(\rho) + \lambda J(\rho) + \lambda E_{\mathrm{x}}(\rho) \geq 0$ is affine in λ [26]. Provided that the same ansatz $E_\lambda(v)$ is used for all values of λ, which is trivially satisfied, this holds for approximate theories, too.

The AC integrand $\mathcal{W}_{\mathrm{c},\lambda}$ gives the correlation contribution of $\hat{V}_{\mathrm{ee}}$ at interaction strength λ and therefore, we have that for the noninteracting system

$$\mathcal{W}_{\mathrm{c},0} = 0. \tag{2.47}$$

The relationships between $\mathcal{W}_{\mathrm{Hxc},\lambda}$, $\mathcal{W}_{\mathrm{xc},\lambda}$ and $\mathcal{W}_{\mathrm{c},\lambda}$ can be obtained using Eqs. (2.45) and (2.46),

$$\mathcal{W}_{\mathrm{xc},\lambda} = \lambda E_{\mathrm{x}} + \mathcal{W}_{\mathrm{c},\lambda} = \mathcal{W}_{\mathrm{Hxc}} - J, \tag{2.48}$$

which we want to point out since we mainly work with $\mathcal{W}_{\mathrm{c},\lambda}$ in Paper II of this work. Figure 2.1 illustrates a typical AC curve. The integral above the AC curve, up to $\mathcal{W}_{\mathrm{xc},0} = \mathrm{tr}\,\gamma_0^\rho \hat{V}_{\mathrm{ee}} - J(\rho)$, corresponds to $E_{\mathrm{c}}(\rho)$. The exchange energy is equal to the AC integrand at $\lambda = 0$, which means $E_{\mathrm{x}}(\rho) = \mathcal{W}_{\mathrm{xc},0}(\rho) = \mathrm{tr}\,\gamma_0^\rho \hat{V}_{\mathrm{ee}} - J(\rho)$. The Hartree contribution $J(\rho)$ is a constant, which means a shift of the AC curve along the y-axis.

Another quantity that can be obtained from the adiabatic connection is the correlation correction to the kinetic energy, T_{c}, which is defined as $T_{\mathrm{c}}(\rho) = T(\rho) - T_{\mathrm{s}}(\rho)$: The integral below the AC curve, down to $\mathcal{W}_{\mathrm{xc},1} = \mathrm{tr}\,\gamma_1^\rho \hat{V}_{\mathrm{ee}} - J(\rho)$, corresponds to $-T_{\mathrm{c}}(\rho)$, such that

$$\mathcal{W}_{\mathrm{c},1}(\rho) = -T_{\mathrm{c}}(\rho) + E_{\mathrm{c}}(\rho). \tag{2.49}$$

For each value of λ, $T_{\mathrm{c},\lambda}(\rho)$ can thus be obtained as [27]

$$T_{\mathrm{c},\lambda}(\rho) = \mathcal{W}_{\mathrm{c},\lambda}(\rho) - E_{\mathrm{c},\lambda}(\rho) = \int_0^\lambda (\mathcal{W}_{\mathrm{c},\mu}(\rho) - \mathcal{W}_{\mathrm{c},\lambda}(\rho))\,\mathrm{d}\mu. \tag{2.50}$$

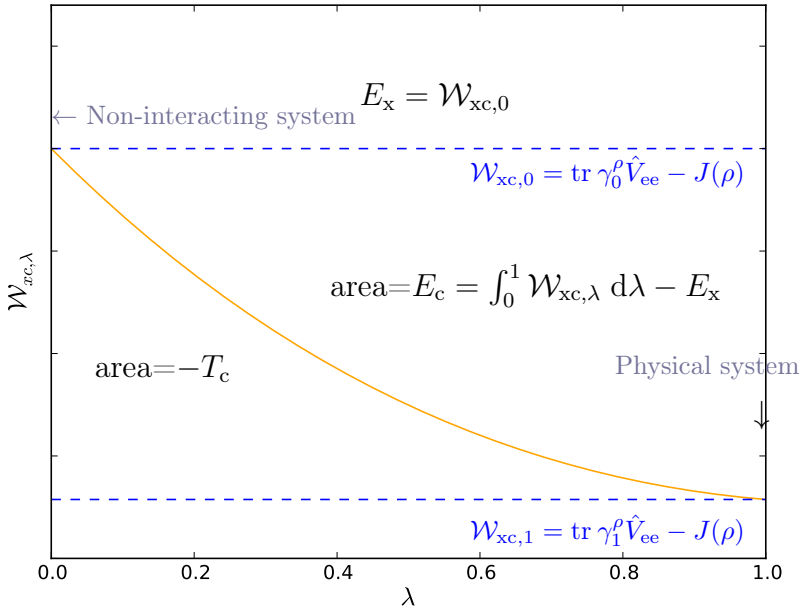


Figure 2.1: Typical adiabatic connection curve, demonstrating the relationship between E_c, T_c and E_x.

Dynamic and static correlation

AC curves are a useful tool to quantify the ratio between dynamic and static correlation in a system. Dynamic correlation refers to the instantaneous correlation between electrons, for example those ones occupying the same spatial orbital. Static correlation, on the other hand, is associated with electron pair that have a larger spatial separation. Static correlation becomes especially import in systems with near-degeneracies, when different orbitals have similar energies. A large amount of static correlation is problematic for methods that rely on a single Slater determinant as a reference, for example Hartree–Fock or coupled-cluster theories. [29]

For a molecular system with purely dynamic correlation, the AC curve is a nearly straight line, with the limit $E_\mathrm{c}(\rho) = -T_\mathrm{c}(\rho)$. A linear curve reflects a correlation energy that is proportional to λ^2. If the curve bends, dynamic correlation is mixed with static correlation. The curvature increases with the amount of static correlation. A larger curvature implies that $E_\mathrm{c} > -T_\mathrm{c}$. In the extreme case, where there is purely static correlation, the whole area between the $\mathcal{W}_{\mathrm{xc},0}$-line and the $\mathcal{W}_{\mathrm{xc},1}$-line corresponds to E_c and we have that $T_\mathrm{c} = 0$. In this case the AC curve has an infinitely negative curvature in the beginning, follows the y-axis downward and than makes a 90 degree angle to follow the $\mathcal{W}_{\mathrm{xc},1}$-line up to $\lambda = 1$. However, in practice there will always be some dynamic correlation left, unless there is a complete dissociation into uncorrelated parts.

A typical example for introducing static correlation is the stretching of a bond. For the hydrogen molecule, AC curves have been studied for increasing bond lengths and it has been shown that the curvature increases with the bond length, with the initial slope getting continuously steeper [26]. The increasing curvature as a measure for static correlation will be more discussed in the context of magnetic fields in Subsection 2.2.3.

2.1.6 Kohn–Sham Theory

Having introduced the Kohn–Sham decomposition of $F_\lambda(\rho)$ at various interaction strengths in Eq. (2.42), we may now set up the corresponding Hohenberg–Kohn variation principle:

$$E_\lambda(v) = \inf_{\rho\in X} \left(T_\mathrm{s}(\rho) + (v|\rho) + \lambda J(\rho) + E_{\mathrm{xc},\lambda}(\rho)\right). \tag{2.51}$$

In particular, we are interested in the fully interacting and noninteracting system,

$$E_1(v) = \inf_{\rho\in X} \left(T_\mathrm{s}(\rho) + (v|\rho) + J(\rho) + E_\mathrm{xc}(\rho)\right) \tag{2.52}$$

$$E_0(v_\mathrm{s}) = \inf_{\rho\in X} \left(T_\mathrm{s}(\rho) + (v_\mathrm{s}|\rho)\right), \tag{2.53}$$

where we use the notation $E_\mathrm{xc}(\rho) = E_{\mathrm{xc},1}(\rho)$. Furthermore, for the noninteracting system we have that $E_{\mathrm{xc},0}(\rho) = 0$, since the Hamiltonian reduces to $\hat{H}_0(v_\mathrm{s}) = \hat{T} + \sum_i v_\mathrm{s}(\mathbf{r}_i)$ and there is no contribution from the electron–electron interaction. It is much easier to find a minimizing density for the noninteracting system (2.53) than for the interacting system (2.52). To work with the noninteracting system instead of the interacting system, we must find a potential v_s, such that the noninteracting minimizing density becomes identical to the interacting minimizing density with external potential v. The Euler equations for the two systems, describing the stationary condition, are

$$\frac{\delta T_\mathrm{s}(\rho)}{\delta\rho(\mathbf{r})} = \mu - v(\mathbf{r}) - v_\mathrm{H}(\rho)(\mathbf{r}) - v_\mathrm{xc}(\rho)(\mathbf{r}) \tag{2.54}$$

$$\frac{\delta T_\mathrm{s}(\rho)}{\delta\rho(\mathbf{r})} = \mu_\mathrm{s} - v_\mathrm{s}(\mathbf{r}), \tag{2.55}$$

where the Hartree and exchange–correlation potential are given by

$$v_\mathrm{H}(\rho)(\mathbf{r}) = \frac{\delta J(\rho)}{\delta\rho(\mathbf{r})}, \qquad v_\mathrm{xc}(\rho)(\mathbf{r}) = \frac{\delta E_\mathrm{xc}(\rho)}{\delta\rho(\mathbf{r})}, \tag{2.56}$$

and where μ and μ_s are chemical potentials, arising from the constraint $\int \rho(\mathbf{r})\,\mathrm{d}\mathbf{r} = N$. Since both Eqs. (2.54) and (2.55) must hold simultaneously, we conclude that

$$\begin{aligned} v_\mathrm{s}(\mathbf{r}) &= v(\mathbf{r}) + v_\mathrm{H}(\rho)(\mathbf{r}) + v_\mathrm{xc}(\rho)(\mathbf{r}) \\ &= v(\mathbf{r}) + \int \frac{\rho(\mathbf{r}')}{|\mathbf{r}-\mathbf{r}'|}\,\mathrm{d}\mathbf{r}' + \frac{\delta E_\mathrm{xc}(\rho)}{\delta\rho(\mathbf{r})}. \end{aligned} \tag{2.57}$$

Thus the ground-state density of the interacting system can be determined by solving the noninteracting system with Hamiltonian

$$\hat{H}_0(v_\mathrm{s}) = -\sum_i \frac{1}{2}\nabla_i^2 + \sum_i v_\mathrm{s}(\mathbf{r}_i). \tag{2.58}$$

and corresponding Rayleigh–Ritz variation principle

$$E_\mathrm{s}(v_\mathrm{s}) = \inf_{\Phi\in\mathcal{W}_N} \left\langle \Phi|\hat{H}_0(v_\mathrm{s})|\Phi \right\rangle, \tag{2.59}$$

where the minimization is over Slater determinants only. This approach was introduced by Kohn and Sham [30], such that the noninteracting system with potential v_s is also referred to as *Kohn–Sham system*. This problem is separable and the Kohn–Sham orbitals ϕ_i are solutions to the Kohn–Sham equations

$$\left[-\frac{1}{2}\nabla^2 + v_\mathrm{s}(\mathbf{r})\right]\phi_i(\mathbf{r}) = \epsilon_i\phi_i(\mathbf{r}), \tag{2.60}$$

where ϵ_i are the orbital energies. The exact wave function is a single Slater determinant Φ, which consists of N occupied Kohn–Sham orbitals ϕ_i,

$$\Phi(\mathbf{x}_1,\mathbf{x}_2,\ldots,\mathbf{x}_N) = \frac{1}{\sqrt{N!}}\begin{vmatrix} \phi_1(\mathbf{x}_1) & \phi_2(\mathbf{x}_1) & \ldots & \phi_N(\mathbf{x}_1) \\ \phi_1(\mathbf{x}_2) & \phi_2(\mathbf{x}_2) & \ldots & \phi_N(\mathbf{x}_2) \\ \vdots & \vdots & \ddots & \vdots \\ \phi_1(\mathbf{x}_N) & \phi_2(\mathbf{x}_N) & \ldots & \phi_N(\mathbf{x}_N) \end{vmatrix} \tag{2.61}$$

with space-spin coordinates $\mathbf{x}_i = (\mathbf{r}_i, s_i)$ and electron density

$$\rho(\mathbf{r}) = \sum_i \sum_s |\phi_i(\mathbf{r},s)|^2. \tag{2.62}$$

Since v_s depends on the density, the Kohn–Sham equations (2.60) must be solved iteratively: One begins with a guess for the density $\rho(\mathbf{r})$ to construct $v_\mathrm{s}(\mathbf{r})$ according to Eq. (2.57). With the help of the Kohn–Sham equations (2.60) one then determines the new orbitals, uses them to construct the new density according to Eq. (2.62) and so on.

The Kohn–Sham method has the advantage that only a small fraction of the total

energy, the exchange–correlation energy E_{xc}, is unknown. In particular, the method provides a way of computing the dominant part of the kinetic energy, the noninteracting kinetic energy T_{s}, exactly. However, the drawback is that the introduction of orbitals increases the number of variables from three to $3N$ spatial variables.[1] Moreover, the functional F_λ is not known. Much effort has been put into understanding its properties and finding good approximations, which is discussed in Section 3.1.

2.2 Magnetic-Field Density-Functional Theory (BDFT)

In the following, we consider systems that include a magnetic field. We will set up DFT in the framework of BDFT. The formal basis of BDFT was first discussed by Grayce and Harris in 1994 [15]. We proceed as above, considering first the Hamiltonian and Rayleigh–Ritz variation principle, then setting up the Hohenberg–Kohn and Lieb variation principles, establishing the adiabatic connection and finally setting up the Kohn–Sham equations.

2.2.1 Hamiltonian and Rayleigh–Ritz Variation Principle in BDFT

Let us consider an external uniform magnetic field $\mathbf{B}$, which is represented by a vector potential $\mathbf{A}$, such that $\mathbf{B} = \nabla \times \mathbf{A}$. The electronic Hamiltonian is now given by

$$\hat{H}_\lambda(v, \mathbf{A}) = \hat{T}(\mathbf{A}) + \lambda \hat{V}_{\mathrm{ee}} + \sum_i v(\mathbf{r}_i) + \sum_i \mathbf{B}(r_i) \cdot \mathbf{S}, \tag{2.63}$$

where $\mathbf{S}$ is the spin operator. Since we in this work consider closed-shell systems only, the last term, the spin-Zeeman term, does not contribute, such that we in the following work with the Hamiltonian

$$\hat{H}_\lambda(v, \mathbf{A}) = \hat{T}(\mathbf{A}) + \lambda \hat{V}_{\mathrm{ee}} + \sum_i v(\mathbf{r}_i). \tag{2.64}$$

The kinetic energy operator now includes the vector potential,

$$\hat{T}(\mathbf{A}) = \frac{1}{2} \sum_j \big(-\mathrm{i}\boldsymbol{\nabla}_j + \mathbf{A}(\mathbf{r}_j) \big)^2. \tag{2.65}$$

The zero-field Hamiltonian $\hat{H}_\lambda(v)$ is thus expanded in the following manner,

$$\hat{H}_\lambda(v, \mathbf{A}) = \hat{H}_\lambda(v) - \mathrm{i} \sum_j \mathbf{A}(\mathbf{r}_j) \cdot \nabla_j + \frac{1}{2} \sum_j A^2(\mathbf{r}_j), \tag{2.66}$$

[1] Orbital free approaches to DFT exist, but at the moment they are in general less accurate than Kohn–Sham DFT models [31].

with $A = |\mathbf{A}|$ and assuming Coulomb gauge, $\nabla \cdot \mathbf{A} = 0$. We refer to the first additional term as *orbital paramagnetic* term, while the second one is called *diamagnetic* term. The ground-state energy is defined in terms of the Rayleigh–Ritz variation principle:

$$E_\lambda(v, \mathbf{B}) = \inf_{\gamma \in \mathcal{D}_N} \mathrm{tr}\gamma \, \hat{H}_\lambda(v, \mathbf{A}). \tag{2.67}$$

Gauge invariance of $E_\lambda(v, \mathbf{B})$ is ensured since a gauge transformation $\mathbf{A} \to \mathbf{A}'$ is precisely compensated for by a gauge transformation $\gamma \to \gamma'$ such that $\mathrm{tr}\gamma' \, \hat{H}_\lambda(v, \mathbf{A}') = \mathrm{tr}\gamma \, \hat{H}_\lambda(v, \mathbf{A})$ [32]. The topic of gauge-origin transformations is discussed in more detail in Subsection 4.1.1.

The theory of BDFT can be formulated in terms of either the vector potential $\mathbf{A}$ or in terms of the magnetic field $\mathbf{B}$. Both representations are equivalent, since after choosing the gauge, the magnetic field $\mathbf{B}$ determines the vector potential $\mathbf{A}$ via the relation $\mathbf{B} = \nabla \times \mathbf{A}$. Were are therefore free to work with either $E_\lambda(v, \mathbf{B})$ or $E_\lambda(v, \mathbf{A})$. In this work, we use consistently $\mathbf{B}$ as second variable in BDFT.

2.2.2 The BDFT Hohenberg–Kohn and Lieb Variation Principles

Since $E_\lambda(v, \mathbf{B})$ is concave and continuous in v for a fixed magnetic field $\mathbf{B}$ and a fixed interaction strength λ, we may introduce DFT in the usual manner

$$E_\lambda(v, \mathbf{B}) = \inf_{\rho \in X} \left[F_\lambda(\rho, \mathbf{B}) + (v|\rho) \right], \tag{2.68}$$

$$F_\lambda(\rho, \mathbf{B}) = \sup_{v \in X^*} \left[E_\lambda(v, \mathbf{B}) - (v|\rho) \right]. \tag{2.69}$$

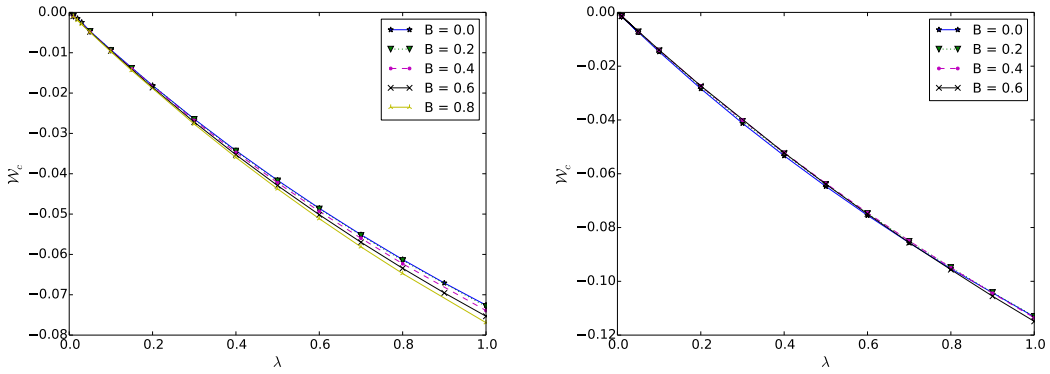
Both the Hohenberg–Kohn variation principle of Eq. (2.68) and the Lieb variation principle of Eq. (2.69) are simple extensions of their field-free counterparts. In both cases, $\mathbf{B}$ is treated as a parametric variable. The Lieb functional may also be expressed in constrained-search form,

$$F_\lambda(\rho, \mathbf{B}) = \min_{\gamma \mapsto \rho} \mathrm{tr} \, \gamma \hat{H}_\lambda(0, \mathbf{A}) = \mathrm{tr} \, \gamma^\rho_{\lambda, \mathbf{A}} \hat{H}_\lambda(0, \mathbf{A}), \tag{2.70}$$

where at least on minimizer $\gamma^\rho_{\lambda, \mathbf{A}}$ exists. The gauge invariance of $F_\lambda(\rho, \mathbf{B})$ follows from the gauge invariance of $E_\lambda(v, \mathbf{B})$, depending on $\mathbf{B}$ rather than $\mathbf{A}$. All arguments apply to all values of λ. Since $F_\lambda(\rho, \mathbf{B})$ explicitly depends on $\mathbf{B}$ in addition to ρ, it is said to be *semi-universal.*

2.2.3 The Adiabatic Connection in BDFT

The difference between the field-free Hamiltonian in Eq. (2.30) and the magnetic Hamiltonian in Eq. (2.64) is that the kinetic energy operator now explicitly depends on the mag-



(a) H_2, calculations at the FCI / aug-cc-pVTZ level.

(b) LiH, calculations at the CCD / aug-cc-pVTZ level.

Figure 2.2: AC curves of H_2 and LiH in a perpendicular magnetic field (atomic units, $\mathcal{W}_c$ in E_h). Each curve is calculated at the equilibrium bond distance in the applied field. The figure is taken from Paper II of this thesis, *"Magnetic-Field Density-Functional Theory (BDFT): Lessons from the Adiabatic Connection"*.

netic field. To set up the adiabatic connection, we therefore introduce the field-dependent noninteracting kinetic energy

$$T_s(\rho, \mathbf{B}) = F_0(\rho, \mathbf{B}) = \mathrm{tr}\, \gamma^{\rho}_{0,\mathbf{A}} \hat{T}(\mathbf{A}) \tag{2.71}$$

and the AC integrand

$$\mathcal{W}_\lambda(\rho, \mathbf{B}) = F'_\lambda(\rho, \mathbf{B}) = \mathrm{tr}\, \gamma^{\rho}_{\lambda,\mathbf{A}} \hat{V}_{ee}, \tag{2.72}$$

whose gauge invariance follows from the gauge invariance of $F_\lambda(\rho, \mathbf{B})$. We then obtain precisely as before

$$F_\lambda(\rho, \mathbf{B}) = T_s(\rho, \mathbf{B}) + \lambda J(\rho) + \lambda E_x(\rho, \mathbf{B}) + E_{c,\lambda}(\rho, \mathbf{B}), \tag{2.73}$$

where we have introduced the field-dependent BDFT exchange and correlation functionals,

$$E_x(\rho, \mathbf{B}) = \mathcal{W}_0(\rho, \mathbf{B}) - J(\rho), \tag{2.74}$$

$$E_{c,\lambda}(\rho, \mathbf{B}) = \int_0^\lambda \left(\mathcal{W}_\mu(\rho, \mathbf{B}) - \mathcal{W}_0(\rho, \mathbf{B})\right)\, \mathrm{d}\mu. \tag{2.75}$$

The field-dependence of the Hamiltonian affects the values of $\mathcal{W}_\lambda$ and particularly, the curvature of the AC curve. How the shape of the curve is influenced by the magnetic field is in detail studied in Paper II of this work.

Figure 2.2 is taken from Paper II and demonstrates our result that, as long as at each magnetic field strength the equilibrium geometry of the molecule at this particular field strength is used, the AC curves are almost indistinguishable and their shape is not affected

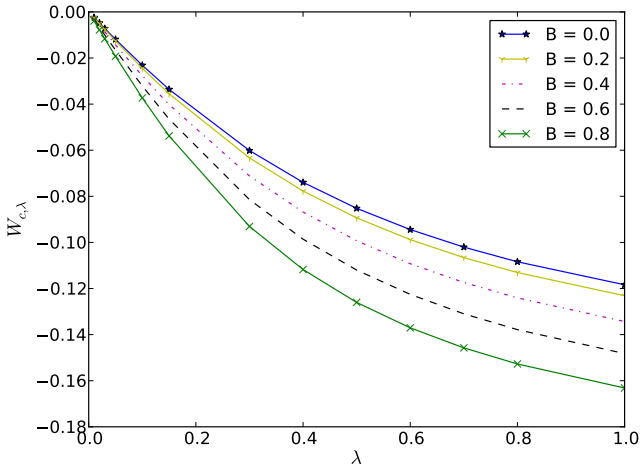


Figure 2.3: AC curves for H_2 at different magnetic field strengths with fixed bond length $3.0a_0$. Shown are accurate ab-initio AC integrands $\mathcal{W}_{c,\lambda}(\rho, \mathbf{A}) = \mathcal{W}_{xc,\lambda}(\rho, \mathbf{A}) - \mathcal{W}_{xc,0}(\rho, \mathbf{A})$ using the FCI density for each magnetic field strength (atomic units, $\mathcal{W}_{c,\lambda}$ in E_h).

by the field. However, Paper II also points out that this is not the case when the geometry is kept fixed at the one corresponding to zero field strength. In this case, we observe a greater curvature of the AC curve with increasing field strength, corresponding to an increase of static correlation. This is shown in Figure 2.3 and we refer to the discussion of Paper II in Section 5.2 for more details.

2.2.4 The BDFT Kohn–Sham Equations

The Kohn–Sham equations in BDFT are set up in the usual manner, by considering the interacting and noninteracting systems with the same minimizing density, but different potentials,

$$E_1(v, \mathbf{B}) = \inf_{\rho \in X} \left(T_s(\rho, \mathbf{B}) + (v|\rho) + J(\rho) + E_{xc}(\rho, \mathbf{B}) \right) \tag{2.76}$$

$$E_0(v_s, \mathbf{B}) = \inf_{\rho \in X} \left(T_s(\rho, \mathbf{B}) + (v_s|\rho) \right), \tag{2.77}$$

where we have

$$v_s = v + v_H(\rho) + v_{xc}(\rho, \mathbf{B}) \tag{2.78}$$

$$= v + \frac{\delta J(\rho)}{\delta \rho} + \frac{\delta E_{xc}(\rho, \mathbf{B})}{\delta \rho}. \tag{2.79}$$

The BDFT Kohn–Sham equations then read

$$\left[\frac{1}{2}\left(-\mathrm{i}\nabla + \mathbf{A}(\mathbf{r})\right)^2 + v_\mathrm{s}(\rho, \mathbf{B})(\mathbf{r})\right] \phi_n(\mathbf{r}, \mathbf{B}) = \epsilon_n(\mathbf{B})\phi_n(\mathbf{r}, \mathbf{B}). \tag{2.80}$$

When solving the Kohn–Sham equations (2.80), the additional difficulty in BDFT is that $E_\mathrm{xc}(\rho, \mathbf{B})$ must be modelled not only as a function of the electron density ρ, but also of the magnetic field $\mathbf{B}$. Until now, no field-dependent approximations of the exchange–correlation functional exist.

2.3 Current Density-Functional Theory (CDFT)

The foundations of CDFT were introduced in the 1980s by Vignale, Rasolt and Geldart [33, 34]. In contrast to BDFT, CDFT has since then been studied quite extensively in a number of works, for example by Pan and Sahni [35, 36], Tellgren et al. [23, 37], Lieb and Schrader [38], and Furness et al. [13].

2.3.1 Hamiltonian and Rayleigh–Ritz Variation Principle in CDFT

In standard DFT and BDFT, the concavity of $E(v)$ and $E(v, \mathbf{B})$ in v allows the conjugation of v to yield ρ, such that we can set up DFT with functionals $F(\rho)$ and $F(\rho, \mathbf{B})$. The natural question is whether we can treat the vector potential in the same manner, such that not only the scalar potential, but also the vector potential is conjugated. However, convex conjugation requires concavity of the respective variable, and $E(v, \mathbf{A})$ is not concave in $\mathbf{A}$ since the Hamiltonian does not depend linearly on $\mathbf{A}$, see Eq. (2.66).

To set up CDFT, we therefore use a different parametrization of the Hamiltonian. Introducing $u = v + \frac{1}{2}A^2$, variables can be changed from $(v, \mathbf{A})$ to $(u, \mathbf{A})$, yielding a Hamiltonian that is linear in $\mathbf{A}$,

$$\bar{H}_\lambda(u, \mathbf{A}) = \bar{H}_\lambda(u) - \mathrm{i}\sum_i \mathbf{A}(\mathbf{r}_i) \cdot \boldsymbol{\nabla}_i. \tag{2.81}$$

As a result, the ground-state energy

$$\mathcal{E}_\lambda(u, \mathbf{A}) = \inf_{\gamma \in \mathcal{D}_N} \mathrm{tr}\, \gamma \bar{H}_\lambda(u, \mathbf{A}) \tag{2.82}$$

is concave in both variables. The (v, A)- and (u, A)-representation are related through

the identities

$$\bar{H}_\lambda(v + \frac{1}{2}A^2, \mathbf{A}) = \hat{H}_\lambda(v, \mathbf{A}), \qquad \bar{H}_\lambda(u, \mathbf{A}) = \hat{H}_\lambda(u - \frac{1}{2}A^2, \mathbf{A}) \tag{2.83}$$

$$\mathcal{E}_\lambda(v + \frac{1}{2}A^2, \mathbf{A}) = E_\lambda(v, \mathbf{A}), \qquad \mathcal{E}_\lambda(u, \mathbf{A}) = E_\lambda(u - \frac{1}{2}A^2, \mathbf{A}). \tag{2.84}$$

2.3.2 The CDFT Hohenberg–Kohn and Lieb Variation Principles

The concavity of $\mathcal{E}_\lambda(u, \mathbf{A})$ in both u and $\mathbf{A}$ allows concave conjugation to be applied to both variables:

$$\mathcal{E}_\lambda(u, \mathbf{A}) = \inf_{\rho, \mathbf{j}_\mathrm{p}} [\mathcal{G}_\lambda(\rho, \mathbf{j}_\mathrm{p}) + (u|\rho) + (\mathbf{A}|\mathbf{j}_\mathrm{p})], \tag{2.85}$$

$$\mathcal{G}_\lambda(\rho, \mathbf{j}_\mathrm{p}) = \sup_{u, \mathbf{A}} [\mathcal{E}_\lambda(u, \mathbf{A}) - (u|\rho) - (\mathbf{A}|\mathbf{j}_\mathrm{p})], \tag{2.86}$$

where the universal density functional $\mathcal{G}_\lambda(\rho, \mathbf{j}_\mathrm{p})$ is convex in the electron density ρ and the paramagnetic current density $\mathbf{j}_\mathrm{p}$. It may also be expressed in constrained-search form,

$$\mathcal{G}_\lambda(\rho, \mathbf{j}_\mathrm{p}) = \min_{\gamma \mapsto (\rho, \mathbf{j}_\mathrm{p})} \mathrm{tr}\, \gamma \bar{H}_\lambda(0, \mathbf{0}) = \mathrm{tr}\, \gamma_\lambda^{\rho, \mathbf{j}_\mathrm{p}} \bar{H}_\lambda(0, \mathbf{0}). \tag{2.87}$$

2.3.3 The Adiabatic Connection in CDFT

To set up the adiabatic connection in CDFT, we introduce, as before, the noninteracting kinetic energy

$$\mathcal{K}_\mathrm{s}(\rho, \mathbf{j}_\mathrm{p}) = \mathcal{G}_0(\rho, \mathbf{j}_\mathrm{p}) = \mathrm{tr}\, \gamma_0^{\rho, \mathbf{j}_\mathrm{p}} \bar{H}_0(0, \mathbf{0}) \tag{2.88}$$

and the AC integrand

$$\mathcal{M}_\lambda(\rho, \mathbf{j}_\mathrm{p}) = \mathcal{G}'_\lambda(\rho, \mathbf{j}_\mathrm{p}) = \mathrm{tr}\, \gamma_\lambda^{\rho, \mathbf{j}_\mathrm{p}} \hat{V}_\mathrm{ee}. \tag{2.89}$$

As before, we then obtain the CDFT Kohn–Sham decomposition

$$\mathcal{G}_\lambda(\rho, \mathbf{j}_\mathrm{p}) = \mathcal{K}_\mathrm{s}(\rho, \mathbf{j}_\mathrm{p}) + \lambda J(\rho) + \lambda \mathcal{G}_\mathrm{x}(\rho, \mathbf{j}_\mathrm{p}) + \mathcal{G}_{\mathrm{c},\lambda}(\rho, \mathbf{j}_\mathrm{p}), \tag{2.90}$$

where we have introduced the CDFT exchange and correlation functionals

$$\mathcal{G}_\mathrm{x}(\rho, \mathbf{j}_\mathrm{p}) = \mathcal{M}_0(\rho, \mathbf{j}_p) - J(\rho), \tag{2.91}$$

$$\mathcal{G}_{\mathrm{c},\lambda}(\rho, \mathbf{j}_\mathrm{p}) = \int_0^\lambda (\mathcal{M}_\mu(\rho, \mathbf{j}_\mathrm{p}) - \mathcal{M}_0(\rho, \mathbf{j}_\mathrm{p}))\ \mathrm{d}\mu. \tag{2.92}$$

With the Kohn–Sham decomposition of Eq. (2.90), the CDFT Hohenberg–Kohn variation principle (2.85) becomes

$$\mathcal{E}_\lambda(u, \mathbf{A}) = \inf_{\rho, \mathbf{j}_\mathrm{p}} \left[\mathcal{K}_\mathrm{s}(\rho, \mathbf{j}_\mathrm{p}) + \lambda J(\rho) + \mathcal{G}_{\mathrm{xc},\lambda}(\rho, \mathbf{j}_\mathrm{p}) + (u|\rho) + (\mathbf{A}|\mathbf{j}_\mathrm{p})\right], \tag{2.93}$$

which is valid for all λ and where we have combined the CDFT exchange and correlation functionals to the CDFT exchange–correlation functional

$$\mathcal{G}_{\mathrm{xc},\lambda}(\rho,\mathbf{j}_\mathrm{p}) = \lambda\mathcal{G}_\mathrm{x}(\rho,\mathbf{j}_\mathrm{p}) + \mathcal{G}_{\mathrm{c},\lambda}(\rho,\mathbf{j}_\mathrm{p}). \tag{2.94}$$

2.3.4 The CDFT Kohn–Sham Equations

To set up Kohn–Sham theory in CDFT, we proceed in the usual manner. We consider the interacting and noninteracting systems

$$\mathcal{E}_1(u,\mathbf{A}) = \inf_{\rho,\mathbf{j}_\mathrm{p}} [\mathcal{K}_\mathrm{s}(\rho,\mathbf{j}_\mathrm{p}) + (u|\rho) + (\mathbf{A}|\mathbf{j}_\mathrm{p}) + J(\rho) + \mathcal{G}_\mathrm{xc}(\rho,\mathbf{j}_\mathrm{p})] \tag{2.95}$$

$$\mathcal{E}_0(u_\mathrm{s},\mathbf{A}_\mathrm{s}) = \inf_{\rho,\mathbf{j}_\mathrm{p}} [\mathcal{K}_\mathrm{s}(\rho,\mathbf{j}_\mathrm{p}) + (u_\mathrm{s}|\rho) + (\mathbf{A}_\mathrm{s}|\mathbf{j}_\mathrm{p})] \tag{2.96}$$

and require that they have identical minimizers. This means that not only the electron density ρ, but also the paramagnetic current density $\mathbf{j}_\mathrm{p}$ must be equal in the interacting and the noninteracting case, $(\rho,\mathbf{j}_\mathrm{p}) = (\rho_\mathrm{s},\mathbf{j}_{\mathrm{p,s}})$. For the potentials, this implies

$$u_\mathrm{s} = u + v_\mathrm{H}(\rho) + u_\mathrm{xc}(\rho,\mathbf{j}_\mathrm{p}) \tag{2.97}$$

$$\mathbf{A}_\mathrm{s} = \mathbf{A} + \mathbf{A}_\mathrm{xc}(\rho,\mathbf{j}_\mathrm{p}) \tag{2.98}$$

where the requirement on u_s follows from differentiation with respect to ρ and the requirement on $\mathbf{A}_\mathrm{s}$ follows from differentiation with respect to $\mathbf{j}_\mathrm{p}$. The scalar and vector exchange–correlation potentials are thus given by

$$u_\mathrm{xc}(\rho,\mathbf{j}_\mathrm{p})(\mathbf{r}) = \frac{\delta\mathcal{G}_\mathrm{xc}(\rho,\mathbf{j}_\mathrm{p})}{\delta\rho(\mathbf{r})} \tag{2.99}$$

$$\mathbf{A}_\mathrm{xc}(\rho,\mathbf{j}_\mathrm{p})(\mathbf{r}) = \frac{\delta\mathcal{G}_\mathrm{xc}(\rho,\mathbf{j}_\mathrm{p})}{\delta\mathbf{j}_\mathrm{p}(\mathbf{r})}. \tag{2.100}$$

As in usual DFT, Eq. (2.96) needs to be solved iteratively with potentials (2.99) and (2.100), such that we obtain the same ground-state density as in Eq. (2.95). Only a few CDFT exchange–correlation functionals $\mathcal{G}_\mathrm{xc}$ have been proposed, most of them based on the vorticity $\nu = \nabla\times(\mathbf{j}_\mathrm{p}/\rho)$, as suggested by Vignale, Rasolt and Geldart [39] and by Higuchi and Higuchi [40]. Another promising recent approach, which is found to be numerically more stable, is CDFT via the generalized kinetic-energy density as discussed by Furness et al. [13].

2.4 Comparison between BDFT and CDFT

In this section, we compare BDFT and CDFT and put both approaches into a common framework.

2.4.1 BDFT in the (u, A)-Representation

In Section 2.2, we have introduced BDFT in the (v, A)-representation. To make a direct comparison to CDFT, which relies on the (u, A)-representation to have $\mathcal{E}_\lambda(u, \mathbf{A})$ concave in both variables, we here present BDFT in the (u, A)-representation. As shown in Eqs. (2.83) and (2.84), the (v, A)- and (u, A)-representation can at any point easily be transformed into each other. To distinguish BDFT in the (u, A)-representation from BDFT in the (v, A)-representation, we use a different notation. In particular, we use the symbols $\mathcal{F}$, $\mathcal{E}$ and $\mathcal{T}_\mathrm{s}$, instead of F, E and T_s, for the Lieb functional, the energy and the noninteracting kinetic functional in the (u, A)-representation, respectively.

Using u and $\mathbf{A}$, the BDFT variation principles read

$$\mathcal{E}_\lambda(u, \mathbf{A}) = \inf_{\rho \in X} \left[\mathcal{F}_\lambda(\rho, \mathbf{A}) + (u|\rho) \right], \tag{2.101}$$

$$\mathcal{F}_\lambda(\rho, \mathbf{A}) = \sup_{u \in X*} \left[\mathcal{E}_\lambda(u, \mathbf{A}) - (u|\rho) \right]. \tag{2.102}$$

In the equivalent constrained-search form, the BDFT density functional of Eq. (2.70) becomes

$$\mathcal{F}_\lambda(\rho, \mathbf{A}) = \min_{\gamma \mapsto \rho} \operatorname{tr} \gamma \hat{H}_\lambda(0, \mathbf{A}) - \frac{1}{2}(A^2|\rho), \tag{2.103}$$

since the term $\frac{1}{2}(A^2|\rho)$ is included in $(u|\rho)$ and therefore needs to be subtracted from the density functional $\mathcal{F}_\lambda$. For the same reason, the noninteracting kinetic energy functional of Eq. (2.71) gets

$$\mathcal{T}_\mathrm{s}(\rho, \mathbf{A}) = \operatorname{tr} \gamma^\rho_{0,\mathbf{A}} \hat{T}(\mathbf{A}) - \frac{1}{2}(A^2|\rho). \tag{2.104}$$

With these changes for $\mathcal{F}_\lambda$ and $\mathcal{T}_\mathrm{s}$, the Kohn–Sham decomposition of the BDFT density functional in the (u, A)-representation is given as before by

$$\mathcal{F}_\lambda(\rho, \mathbf{A}) = \mathcal{T}_\mathrm{s}(\rho, \mathbf{A}) + \lambda J(\rho) + E_{\mathrm{xc},\lambda}(\rho, \mathbf{A}). \tag{2.105}$$

In terms of $\mathcal{F}_\lambda(\rho, \mathbf{A})$ and $\mathcal{T}_\mathrm{s}(\rho, \mathbf{A})$, the exact ground state energy $\mathcal{E}_\lambda(u, \mathbf{A})$ then is

$$\mathcal{E}_\lambda(u, \mathbf{A}) = \inf_{\rho \in X} \left(\mathcal{T}_\mathrm{s}(\rho, \mathbf{A}) + \lambda J(\rho) + \mathcal{F}_{\mathrm{xc},\lambda}(\rho, \mathbf{A}) + (u|\rho) \right). \tag{2.106}$$

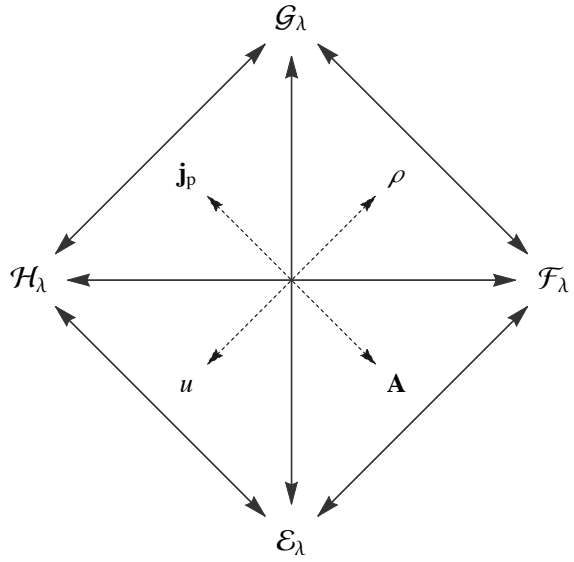


Figure 2.4: Four-way correspondence of DFT in a magnetic field. The figure is taken from Paper II of this work, which we also refer to for more details.

2.4.2 Four-Way Correspondence of DFT

The four-way correspondence, which in a general convex analysis context is discussed in Refs. [41] and [42], enables us to put BDFT and CDFT into a common framework, applying the theory of convex conjugation to DFT in two variables. This is illustrated in Fig. 2.4.

For both BDFT and CDFT, the starting point is the ground state energy $\mathcal{E}_\lambda(u, \mathbf{A})$, a concave function depending on the scalar potential u and the vector potential $\mathbf{A}$. In BDFT, the Lieb variation principle of Eq. (2.102) corresponds to the partial conjugation $u \to \rho$ from $\mathcal{E}_\lambda$ to $\mathcal{F}_\lambda$, where $\mathcal{F}_\lambda$ is a concave–convex saddle function. The upward conjugation $u \to \rho$ is a maximization with negative pairing $-(u|\rho)$. The BDFT Hohenberg–Kohn variation principle of Eq. (2.101) is the convex conjugate of the Lieb variation principle, corresponding to the partial conjugation $\rho \to u$ from $\mathcal{F}_\lambda$ to $\mathcal{E}_\lambda$, which is a minimization with positive pairing $+(u|\rho)$.

The CDFT Lieb variation principle of Eq. (2.86) involves the full conjugation $(u, \mathbf{A}) \to (\rho, \mathbf{j}_\mathrm{p})$ from $\mathcal{E}_\lambda$ to $\mathcal{G}_\lambda$, where $\mathcal{G}_\lambda$ is a convex function. In addition to the upward conjugation $u \to \rho$, it also requires the conjugation $\mathbf{A} \to \mathbf{j}_\mathrm{p}$, which is a maximization with negative pairing $-(\mathbf{A}|\mathbf{j}_\mathrm{p})$. The CDFT Hohenberg–Kohn variation principle of Eq. (2.85) corresponds to the full conjugation $(\rho, \mathbf{j}_\mathrm{p}) \to (u, \mathbf{A})$, where the downward conjugations $u \to \rho$ and $\mathbf{A} \to \mathbf{j}_\mathrm{p}$ are minimizations with positive pairings $+(u|\rho)$ and $+(\mathbf{A}|\mathbf{j}_\mathrm{p})$, respectively.

Fig. 2.4 also contains the concave–convex saddle function $\mathcal{H}_\lambda(v, \mathbf{j}_\mathrm{p})$, which corresponds to the partial conjugation $\mathbf{A} \to \mathbf{j}_\mathrm{p}$. Since it currently lacks any application in DFT, we will not discuss it further.

The BDFT and CDFT density functionals are related by

$$\mathcal{F}_\lambda(\rho, \mathbf{A}) = \inf_{\mathbf{j}_\mathrm{P}} \left[\mathcal{G}_\lambda(\rho, \mathbf{j}_\mathrm{P}) + (\mathbf{A}|\mathbf{j}_\mathrm{P})\right] \tag{2.107}$$

$$\mathcal{G}_\lambda(\rho, \mathbf{j}_\mathrm{P}) = \sup_{\mathbf{A}} \left[\mathcal{F}_\lambda(\rho, \mathbf{A}) - (\mathbf{A}|\mathbf{j}_\mathrm{P})\right], \tag{2.108}$$

which follows from the general four-way correspondence, irrespective of the application to DFT. Whether the semi-universal density functional of BDFT is easier or more difficult to model than the universal density functional of CDFT, is an open question.

Chapter 3

Approximate Density-Functional Theory

3.1 Density Functional Approximations

To use DFT in practical calculations, the exchange–correlation functional needs to be approximated. A wide range of density functional approximations (DFAs) has been developed, each with a specific functional form. The development of more accurate exchange–correlation functionals is still one of the main research topics in theoretical chemistry.

Typically, the functionals are separated into an exchange part and a correlation part,

$$E_{\mathrm{xc}}^{\mathrm{DFA}}(\rho) = E_{\mathrm{x}}^{\mathrm{DFA}}(\rho) + E_{\mathrm{c}}^{\mathrm{DFA}}(\rho). \tag{3.1}$$

In this section, we introduce the functionals used in this work. We follow the "Jacob's ladder of DFT" [43], which means starting with the simplest approximation and gradually adding more elements, increasing the complexity.

However, first we want to address one of the main problems of present DFAs, the problem of self-interaction, which is only partially solved for some selected DFAs, as the meta-generalized gradient approximations (mGGAs) discussed on Subsection 3.1.3.

Self interaction

The electron-electron interaction part $\hat{V}_{\mathrm{ee}}$ of the electronic Hamiltonian of Eq. (2.3) excludes the self-interaction term where $i = j$, since an electron does not interact with itself. When splitting $\hat{V}_{\mathrm{ee}}$ into the Hartree term J and the exchange–correlation energy E_{xc}, it is therefore necessary that the Coulomb energy of each occupied Kohn–Sham orbital is cancelled by its exchange–correlation energy. In Hartree–Fock theory, this is the case, and the self-exchange energy compensates for the self-Coulomb energy for each Hartree–Fock orbital. Approximate exchange–correlation functionals, however, often fail to satisfy this

condition.

Self-interaction leads to the wrong asymptotic behaviour of the exchange–correlation potential, which should satisfy

$$v_{\mathrm{xc}}(\mathbf{r}) \to -\frac{1}{r} \quad \text{for } r \to \infty, \tag{3.2}$$

and many of the most popular DFAs have a wrong long-range behaviour of v_{xc}. For a recent review of possible self-interaction corrections, we refer to Ref. [44].

3.1.1 Local Density Approximation

The *local density approximation* (LDA) is the oldest approximation in DFT and was proposed by Kohn and Sham [30]. It is assumed that the density can locally be treated as that of a uniform electron gas.

The exchange energy is then a known quantity, given by the Dirac formula

$$E_{\mathrm{x}}^{\mathrm{LDA}} = -C_x \int \rho^{4/3}(\mathbf{r})\mathrm{d}\mathbf{r}, \tag{3.3}$$

with $C_x = \frac{3}{4}\left(\frac{3}{\pi}\right)^{1/3}$. Since in this work, we only treat closed-shell atoms, we will restrict ourselves to the presentation of the functionals in their more basic form, without considering the case where the spin-up and spin-down density are not equal. For LDA for example, the alternative for a non-equal spin-up and down-down density would be the *Local Spin Density Approximation* (LSDA).

For the correlation energy, an exact expression exists only for the high and low density limit. However, very precise Monte Carlo simulation data [45] have been used by Vosko, Wilk and Nusair (VWN) [46] for parameterizations of $E_{\mathrm{x}}^{\mathrm{LDA}}$. To date, the most common form used for LDA is the so-called VWN5 correlation functional.

Although, for the application to molecules, the assumptions made in LDA are rather crude, the accuracy of the results is usually better than those obtained with the Hartree–Fock method. Ground state energies, molecular geometries and vibrational frequencies are reproduced within a few percent, although there is a systematic overbinding [47]. One of the main problems of LDA is its wrong asymptotic behaviour of the exchange–correlation potential due to self-interaction.

3.1.2 Generalized Gradient Approximation

As an improvement upon LDA, exchange–correlation functionals belonging to the class of *generalized gradient approximations* (GGAs) depend in addition to the density $\rho(\mathbf{r})$ on

the gradient of the density, $\nabla\rho(\mathbf{r})$,

$$E_{\mathrm{xc}}^{\mathrm{GGA}}(\rho) = \int \epsilon_{\mathrm{xc}}^{\mathrm{GGA}}(\rho(\mathbf{r}), \nabla\rho(\mathbf{r}))\, \mathrm{d}\mathbf{r}. \tag{3.4}$$

The specific form of $\epsilon_{\mathrm{xc}}^{\mathrm{GGA}}$ varies for the different functionals. Hundreds of GGA functionals have been developed over the years. The general idea has been to construct explicit expressions for $\epsilon_{\mathrm{xc}}^{\mathrm{GGA}}$ that satisfy as many of the known exact properties of ϵ_{xc} as possible, while remaining reasonably simple and giving good results. Some functionals have been specifically developed for the computation of certain chemical properties, as for example KT2 [48] for NMR shielding constants, which is discussed below. Typically, calculations with GGA functionals reproduce exchange–correlation energies and other integrated quantities quite well, giving much better results than corresponding LDA calculations. Exchange-correlation potentials, on the other hand, are often of poor quality.

BLYP

For the BLYP functional, the exchange energy is modelled by the Becke functional (B) [49]. It gives the correct $1/r$-asymptotic behaviour of the exchange potential and contains only one empirical parameter to fit the Hartree–Fock exchange energies for a wide range of atomic systems. The correlation part is given by the Lee-Yang-Parr functional (LYP) [50], which is based on the correlation-energy formula by Colle and Salvetti [51].

PBE

John Perdew and co-workers have developed a series of GGA functionals which are non-empirical. Instead of fitting parameters to empirical data, these functionals have been constructed to satisfy as many exact conditions as possible. They all consist of both an exchange and a correlation part. The Perdew-Burke-Ernzerhof functional (PBE) [52], which can be considered as the most refined one, employs the LDA functional and adds further terms and corrections to it.

KT2

In 2003, Keal and Tozer published two new functionals, developed to give accurate Kohn–Sham exchange–correlation potentials [48]. The idea was that these potentials determine directly the Kohn–Sham orbitals and orbital energies and that good orbital energies are important. In particular, Keal and Tozer demonstrated a correlation between the eigenvalue difference of the lowest unoccupied and highest occupied orbital and the accuracy of NMR shielding constants.

One of their functionals, the KT2 functional, contains four parameters and has the

form

$$E_{\mathrm{xc}}^{\mathrm{KT2}} = \alpha E_{\mathrm{x}}^{\mathrm{LDA}} + \beta E_{\mathrm{c}}^{\mathrm{LDA}} + \gamma \sum_{\sigma} \int \frac{|\nabla \rho_\sigma(\mathbf{r})|^2}{\rho_\sigma^{4/3}(\mathbf{r}) + \delta} \, \mathrm{d}\mathbf{r}, \tag{3.5}$$

where the parameters $\gamma = -0.006, \delta = 0.1$ were empirically fitted to reproduce NMR shielding constants for a set of molecules, while the parameters $\alpha = 1.07173$ and $\beta = 0.576727$ were empirically fitted for thermochemical and structural predictions. The KT2 functional typically gives shielding constants that are 2-3 times more accurate than those obtained using the most common GGAs. Atomization energies, ionization potentials and molecular bond lengths are comparable with those of most GGAs, while total energies are rather poorly reproduced.

3.1.3 Meta-Generalized Gradient Approximation

If, in addition to the density $\rho(\mathbf{r})$ and its gradient $\nabla\rho(\mathbf{r})$, the Laplacian of the density or the orbital kinetic energy density are also included, we arrive at the *meta-generalized gradient approximations* (mGGAs). The exchange–correlation functional is then of the form

$$E_{\mathrm{xc}}^{\mathrm{mGGA}}(\rho) = \int \epsilon_{\mathrm{xc}}^{\mathrm{mGGA}}(\rho(\mathbf{r}), \nabla\rho(\mathbf{r}), \nabla^2\rho(\mathbf{r})) \, \mathrm{d}\mathbf{r} \tag{3.6}$$

for the Laplacian or

$$E_{\mathrm{xc}}^{\mathrm{mGGA}}(\rho) = \int \epsilon_{\mathrm{xc}}^{\mathrm{mGGA}}(\rho(\mathbf{r}), \nabla\rho(\mathbf{r}), \tau(\mathbf{r})) \, \mathrm{d}\mathbf{r} \tag{3.7}$$

for the orbital kinetic energy density τ, with

$$\tau(\mathbf{r}) = \frac{1}{2} \sum_i |\nabla \phi_i(\mathbf{r})|^2. \tag{3.8}$$

In this thesis, we only consider mGGAs containing τ, but it should be noted that the information contained in the kinetic energy density is essentially equivalent to the one of the Laplacian of the density, since they are related via [29]

$$\tau(\mathbf{r}) = \frac{1}{2} \sum_i \epsilon_i |\phi_i(\mathbf{r})|^2 - v(\mathbf{r})\rho(\mathbf{r}) + \frac{1}{2}\nabla^2\rho(\mathbf{r}). \tag{3.9}$$

However, the use of τ is usually more stable than the use of $\nabla^2\rho(\mathbf{r})$. One complication that all functionals employing τ encounter is that $v_{\mathrm{xc}} = \delta E_{\mathrm{xc}}/\delta\rho$ cannot directly be computed, since τ is not an explicit functional of the density. One solution to this problem is the use of OEPs [53, 54], which, however, is computationally very expensive. Usually, mGGAs are therefore implemented such that the total energy is made stationary with respect to orbital variations. This, however, means that the local multiplicative Kohn–Sham potential is replaced with a differential operator [55]. Alternatively, GGA Kohn–Sham

orbitals can be used to evaluate the total energy, instead of the self-consistent orbitals.

TPSS

The Tao-Perdew-Staroverov-Scuseria (TPSS) functional [56, 57] is based on the PBE GGA and constructed to satisfy exact constraints without the use of empirical parameters. It employs the kinetic energy density τ, which for single real orbitals equals the von Weizsäcker kinetic energy τ_W,

$$\tau_W(\mathbf{r}) = \frac{|\nabla\rho(\mathbf{r})|^2}{8\rho(\mathbf{r})}. \tag{3.10}$$

Utilizing the iso-orbital indicator

$$z = \frac{\tau_W}{\tau} \leq 1, \tag{3.11}$$

the TPSS functional can identify one- or two-electron regions and is therefore partially free of self-interaction.

Extension to magnetic fields In magnetic fields, it is necessary to modify the kinetic-energy density to ensure independence of the particular choice of the gauge origin. One possible choice is the physical (mechanical) kinetic energy density proposed by Maximoff and Scuseria [58],

$$\tau_{\mathrm{MS}}(\mathbf{r}) = \frac{1}{2}\sum_l \left|(-\mathrm{i}\nabla + \mathbf{A}(\mathbf{r}))\phi_l(\mathbf{r})\right|^2. \tag{3.12}$$

This approach is well suited to BDFT, since τ_{MS} explicitly depends on the vector potential and therefore on the magnetic field.[1]

Another option is the gauge-invariant kinetic energy density proposed by Dobson [59] and used by Becke [60] and Tao [61],

$$\tau_{\mathrm{D}}(\mathbf{r}) = \tau(\mathbf{r}) - \frac{\mathbf{j}_{\mathrm{p}}^2(\mathbf{r})}{2\rho(\mathbf{r})}. \tag{3.13}$$

In contrast to τ_{MS}, the use of τ_{D} ensures that the iso-orbital indicator in the functional is extended in a rigorous way [62]. The performance of both generalizations is studied in the papers of this thesis. While we first retained the name TPSS and explicitly specified whether τ_{MS} or τ_{D} had been used, we later adopted different names for the two generalizations, as done in Paper III of this thesis. If working with τ_{D}, we refer to the functional as cTPSS, as done by Furness et al. [13], indicating that one uses the current to make τ gauge-invariant. Utilizing τ_{MS}, we call the functional aTPSS, indicating the use of the vector potential $\mathbf{A}$ to ensure gauge-invariance.

[1] We note that in Paper II of this thesis, τ_{MS} has been called τ_{phys} due to its characteristic of being the physical kinetic energy density. We later decided to go back to the notation of Paper I, where the sub-index 'MS' indicates who first proposed it. This is also the notation used in this thesis throughout.

SCAN

The SCAN functional [63] satisfies all known exact constraints that a mGGA functional can satisfy, specifically six for exchange, six for correlation and five for the sum of these two. Moreover, it is fitted to particular cases for which it is extremely accurate, in particular rare-gas atoms and non-bonded interactions. Since it is not fitted to any real bonded systems, it can be regarded as a non-empirical functional that can be applied to a wide range of molecular systems.

The accuracy of the SCAN functional is outstanding for systems where the exact exchange–correlation hole is localized near the electron. In general, SCAN is often comparable or even better than hybrid functionals, but at a lower computational cost.

3.1.4 Hybrid Density Functional Approximation

Another class of approximate functionals are the *hybrid density functional approximations*. They are based on the observation that the exchange energy is given by the first point of an AC curve, $E_{\mathrm{x}}(\rho) = \mathcal{W}_{\mathrm{xc},0}(\rho) = \langle \Psi_0^\rho | \hat{V}_{\mathrm{ee}} | \Psi_0^\rho \rangle - J(\rho)$. The expression for the exchange energy is the same as in Hartree–Fock theory, but in terms of Kohn–Sham orbitals,

$$E_{\mathrm{x}}(\rho) = -\frac{1}{2} \sum_{i,j} \int \frac{\phi_i^*(\mathbf{r}) \phi_j^*(\mathbf{r}') \phi_j(\mathbf{r}) \phi_i(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|} \, \mathrm{d}\mathbf{r} \mathrm{d}\mathbf{r}', \tag{3.14}$$

However, in general the Hartree–Fock orbitals differ from the Kohn–Sham orbitals.

Hybrid models include a certain percentage of exact exchange by mixing it with a standard LDA, GGA or mGGA exchange. A hybrid GGA, for example, is of the form

$$E_{\mathrm{xc}}^{\mathrm{hyb},GGA} = a E_{\mathrm{x}}^{\mathrm{HF}} + (1-a) E_{\mathrm{x}}^{\mathrm{GGA}} + E_{\mathrm{c}}^{\mathrm{GGA}}, \tag{3.15}$$

where the semi-empirical constant a is typically around 0.25. The first hybrid functional was introduced by Becke in 1993, where he mixed in an amount of 50% exact exchange in his *half-and-half* functional. However, already in this initial paper he suggested that even better results could be obtained by semi-empirical models where the amount of exact exchange is used as a parameter [25]. In general, approximate functionals rely on an error cancellation between exchange and correlation energies and this needs to be taken into account when determining the amount of exact exchange to be used.

B97

The B97 functional of Becke [64] is based on least-square fits to accurate thermochemical data. It is of the form

$$E_{\mathrm{xc}}^{\mathrm{B97}} = E_{\mathrm{x}}^{\mathrm{GGA}} + E_{\mathrm{c}}^{\mathrm{GGA}} + c_x E_{\mathrm{x}}^{\mathrm{exact}}, \tag{3.16}$$

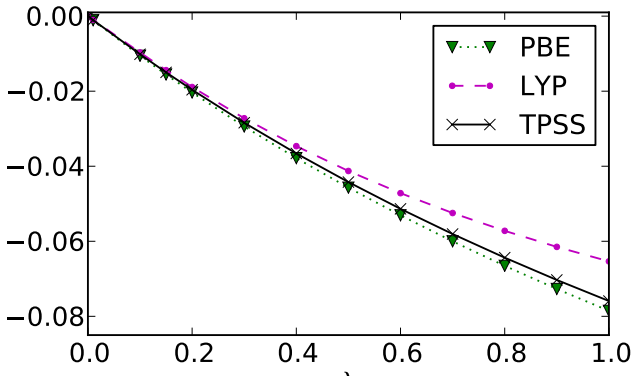


Figure 3.1: AC curves with several DFAs, for the H_2 molecule, bond length $1.4a_0$ (atomic units, $\mathcal{W}_c$ in E_h). All calculations have been performed with a contracted aug-cc-pVTZ basis.

where $c_x = 0.1943$ incorporates some exact exchange. The exchange and correlation energy functionals are of GGA type and contain three and six parameters, respectively, making a total of ten parameters. Becke claims that, with this large number of parameters, the limits have been attained regarding how much accuracy can be obtained within the GGA framework [64].

B3LYP

The Becke-3-parameter-Lee-Yang-Parr (B3LYP) functional is a hybrid GGA of the form

$$E_{xc}^{B3LYP} = (1-a)E_x^{LDA} + aE_x^{HF} + bE_x^{B88} + cE_c^{LYP} + (1-c)E_c^{LDA}, \tag{3.17}$$

with $a = 0.20$, $b = 0.72$ and $c = 0.81$ [65]. In the original paper by Becke [66], the PW91 correlation functional is used instead of LYP. The B3LYP functional is the most popular density functional and in the period of 1990-2006, it was used in 80% of all DFT calculations [67]. Close to equilibrium, it outperforms most other GGAs and mGGAs for calculations of structural and energetic properties.

3.2 Adiabatic Connection Curves for Approximate Functionals

Approximate density functionals have been developed for the fully interacting system, $\lambda = 1$. However, also the cases $\lambda \neq 1$ can be calculated, using scaling relations. Thus AC curves as illustrated in Figure 3.1 can be generated.

For any approximate functional, the AC integrand can be computed using [68]

$$\mathcal{W}_{\mathrm{c},\lambda}(\rho, \mathbf{A}) = \frac{\partial}{\partial\lambda}(\lambda^2 E_{\mathrm{c}}(\rho_{1/\lambda}, \mathbf{A})), \tag{3.18}$$

where the scaled density

$$\rho_{1/\lambda}(\mathbf{r}) = \lambda^{-3}\rho(\mathbf{r}/\lambda) = \lambda^{-3}\rho(\mathbf{r}') \tag{3.19}$$

is obtained by linear scaling of the coordinates, $\mathbf{r}' = \mathbf{r}/\lambda$. The scaling formula of Eq. (3.18) follows exactly as in standard DFT:[2]

$$\begin{aligned}
\mathcal{W}_{\mathrm{c},\lambda}(\rho, \mathbf{A}) &= \frac{\partial}{\partial\lambda} E_{\mathrm{c},\lambda}(\rho, \mathbf{A}) \\
&= \frac{\partial}{\partial\lambda}\left(F_\lambda(\rho, \mathbf{A}) - T_{\mathrm{s}}(\rho, \mathbf{A}) - \lambda J(\rho) - \lambda E_{\mathrm{x}}(\rho)\right) \\
&= \frac{\partial}{\partial\lambda}\lambda^2(F_\lambda(\rho_{1/\lambda}, \mathbf{A}) - T_{\mathrm{s}}(\rho_{1/\lambda}, \mathbf{A}) - J(\rho_{1/\lambda}) - E_{\mathrm{x}}(\rho_{1/\lambda}) \\
&= \frac{\partial}{\partial\lambda}\lambda^2 E_{\mathrm{c}}(\rho_{1/\lambda}, \mathbf{A}),
\end{aligned} \tag{3.20}$$

where we in the third step used the following coordinate scaling relations [68]

$$\begin{aligned}
F_\lambda(\rho, \mathbf{A}) &= \lambda^2 F_\lambda(\rho_{1/\lambda}, \mathbf{A}) \\
T_{\mathrm{s}}(\rho, \mathbf{A}) &= \lambda^2 T_{\mathrm{s}}(\rho_{1/\lambda}, \mathbf{A}) \\
J(\rho) &= \lambda J(\rho_{1/\lambda}) \\
E_{\mathrm{x}}(\rho) &= \lambda E_{\mathrm{x}}(\rho_{1/\lambda})
\end{aligned} \tag{3.21}$$

Note that we here consider the correlation part $\mathcal{W}_{\mathrm{c},\lambda}(\rho, \mathbf{A})$ only. The exchange part is trivially obtained, since it follows a simple linear scaling relation, see Eq. (2.48).

From relations (3.18) and (3.19), all quantities needed to evaluate a particular exchange–correlation functional can be computed. An overview over the most important equations is given in Paper II of this thesis. Here we discuss the relations and their derivation in more detail.

[2]We note that this derivation is presented in Paper II of this thesis in terms of $\mathbf{B}$ instead of $\mathbf{A}$. Both formulations are equivalent, and we here reformulate it in terms of $\mathbf{A}$ to be consistent with the notation of the remaining part of this thesis.

LDA

For a correlation functional that depends locally on the density only, the correlation energy is obtained by using the scaled density when evaluating the functional,

$$E_{\mathrm{c}}^{\mathrm{LDA}}(\rho_{1/\lambda}) = \int \epsilon_{\mathrm{c}}^{\mathrm{LDA}}(\rho_{1/\lambda}(\mathbf{r}))\,\mathrm{d}\mathbf{r} = \lambda^3 \int \epsilon_{\mathrm{c}}^{\mathrm{LDA}}(\lambda^{-3}\rho(\mathbf{r}'))\,\mathrm{d}\mathbf{r}'. \tag{3.22}$$

GGAs

GGAs depend in addition on the gradient, which must also be scaled,

$$\frac{\partial \rho_{1/\lambda}(\mathbf{r})}{\partial \mathbf{r}} = \lambda^{-3}\frac{\partial \rho(\mathbf{r}')}{\partial \mathbf{r}} = \lambda^{-3}\frac{\partial \rho(\mathbf{r}')}{\partial \mathbf{r}'}\frac{\partial \mathbf{r}'}{\partial \mathbf{r}} = \lambda^{-4}\frac{\partial \rho(\mathbf{r}')}{\partial \mathbf{r}'} \tag{3.23}$$

$$\implies \nabla_r^{1/\lambda}\rho_{1/\lambda}(\mathbf{r}) = \lambda^{-4}\nabla_{r'}\rho(\mathbf{r}'). \tag{3.24}$$

The expression for the correlation energy then becomes

$$\begin{aligned} E_{\mathrm{c}}^{\mathrm{GGA}}(\rho_{1/\lambda}) &= \int \epsilon_{\mathrm{c}}^{\mathrm{GGA}}(\rho_{1/\lambda}(\mathbf{r}), \nabla_r^{1/\lambda}\rho_{1/\lambda}(\mathbf{r}))\,\mathrm{d}\mathbf{r} \\ &= \lambda^3 \int \epsilon_{\mathrm{c}}^{\mathrm{GGA}}(\lambda^{-3}\rho_{1/\lambda}(\mathbf{r}), \lambda^{-4}\nabla_{r'}\rho(\mathbf{r}'))\,\mathrm{d}\mathbf{r}'. \end{aligned} \tag{3.25}$$

mGGAs

For mGGAs, furthermore the scaled kinetic energy density is needed. Using Eq. (3.19), it follows that

$$\sum_i |\phi_i^{1/\lambda}(\mathbf{r})|^2 = \rho_{1/\lambda}(\mathbf{r}) = \lambda^{-3}\rho(r/\lambda) = \lambda^{-3}\sum_i |\phi_i(\mathbf{r}/\lambda)|^2, \tag{3.26}$$

implying that the orbitals and their derivatives scale as

$$\phi_i^{1/\lambda}(\mathbf{r}) = \lambda^{-3/2}\phi_i(\mathbf{r}'), \tag{3.27}$$

$$\frac{\partial \phi_i^{1/\lambda}(\mathbf{r})}{\partial \mathbf{r}} = \lambda^{-\frac{3}{2}}\frac{\partial \mathbf{r}'}{\partial \mathbf{r}}\frac{\partial \phi_i(\mathbf{r}')}{\partial \mathbf{r}'} = \lambda^{-\frac{5}{2}}\frac{\partial \phi_i(\mathbf{r}')}{\partial \mathbf{r}'}. \tag{3.28}$$

The kinetic energy density τ, needed for mGGAs, then transforms as

$$\tau^{1/\lambda}(\mathbf{r}) = \frac{1}{2}\sum_i \left|\frac{\partial \phi_i^{1/\lambda}(\mathbf{r})}{\partial \mathbf{r}}\right|^2 = \frac{1}{2}\sum_i \left|\frac{\partial \phi_i^{1/\lambda}(\mathbf{r}')}{\partial \mathbf{r}'}\right|^2 = \lambda^{-5}\tau(\mathbf{r}'). \tag{3.29}$$

For the use in magnetic fields, the gauge invariant generalizations of τ, that were introduced in Subsection 3.1.3, need to be scaled. Both scale exactly as τ, as we show in the following.

For the physical kinetic energy density

$$\tau_{\rm MS}(\mathbf{r}) = \frac{1}{2}\sum_l \left|(-\mathrm{i}\nabla + \mathbf{A}(\mathbf{r}))\phi_l(\mathbf{r})\right|^2, \tag{3.30}$$

we use that

$$\frac{\partial}{\partial \mathbf{r}} = \frac{\partial \mathbf{r}'}{\partial \mathbf{r}}\frac{\partial}{\partial \mathbf{r}'} = \frac{1}{\lambda}\frac{\partial}{\partial \mathbf{r}'} \quad \Longrightarrow \nabla_r^{1/\lambda} = \lambda^{-1}\nabla_{r'} \tag{3.31}$$

and that $\mathbf{A}^{1/\lambda}(\mathbf{r}) = \lambda^{-1}\mathbf{A}(\mathbf{r}')$. We obtain

$$\begin{aligned}
\tau_{\rm MS}^{1/\lambda}(\mathbf{r}) &= \frac{1}{2}\sum_l \left|(-\mathrm{i}\nabla_r^{1/\lambda} + \mathbf{A}^{1/\lambda}(\mathbf{r}))\phi_l^{1/\lambda}(\mathbf{r})\right|^2 \\
&= \frac{1}{2}\sum_l \left|(-\mathrm{i}\lambda^{-1}\nabla_{r'} + \lambda^{-1}\mathbf{A}(\mathbf{r}'))\lambda^{-3/2}\phi_l(\mathbf{r}')\right|^2 \\
&= \frac{\lambda^{-5}}{2}\sum_l \left|(-\mathrm{i}\nabla_{r'} + \mathbf{A}(\mathbf{r}'))\phi_l(\mathbf{r}')\right|^2 \\
&= \lambda^{-5}\,\tau_{\rm MS}(\mathbf{r}'),
\end{aligned} \tag{3.32}$$

This shows that $\tau_{\rm MS}$ has the same scaling factor as τ. The same holds true for $\tau_{\rm D}$, which was defined in Subsection 3.1.3 and which we here repeat for convenience,

$$\tau_{\rm D}(\mathbf{r}) = \tau(\mathbf{r}) - \frac{\mathbf{j}_{\rm p}^2(\mathbf{r})}{2\rho(\mathbf{r})}. \tag{3.33}$$

From the definition $\mathbf{j}_{\rm p}(\mathbf{r}) = \frac{1}{2\mathrm{i}}\sum_k \Big(\phi_k^\dagger(\mathbf{r})\nabla\phi_k(\mathbf{r}) - (\nabla\phi_k^\dagger(\mathbf{r}))\phi_k(\mathbf{r})\Big)$, it follows that

$$\begin{aligned}
\mathbf{j}_{\rm p}^{1/\lambda}(\mathbf{r}) &= \frac{1}{2\mathrm{i}}\sum_k \Big(\phi_k^{(1/\lambda)\dagger}(\mathbf{r})\nabla_r^{1/\lambda}\phi_k^{1/\lambda}(\mathbf{r}) - (\nabla_r^{1/\lambda}\phi_k^{(1/\lambda)\dagger}(\mathbf{r}))\phi_k^{1/\lambda}(\mathbf{r})\Big) \\
&= \frac{1}{2\mathrm{i}}\sum_k \left(\frac{\phi_k^\dagger(\mathbf{r}')}{\lambda^{3/2}}\frac{\nabla_{r'}\phi_k(\mathbf{r}')}{\lambda^{5/2}} - \frac{\nabla_{r'}\phi_k^\dagger(\mathbf{r}')}{\lambda^{5/2}}\frac{\phi_k(\mathbf{r}')}{\lambda^{3/2}}\right) \\
&= \frac{1}{\lambda^4}\frac{1}{2\mathrm{i}}\sum_k \Big(\phi_k^\dagger(\mathbf{r}')\nabla_{r'}\phi_k(\mathbf{r}') - (\nabla_{r'}\phi_k^\dagger(\mathbf{r}'))\phi_k(\mathbf{r}')\Big) \\
&= \frac{1}{\lambda^4}\mathbf{j}_{\rm p}(\mathbf{r}'),
\end{aligned} \tag{3.34}$$

where we have utilized Eqs. (3.27) and (3.28). This relation (3.34) implies that

$$\frac{\mathbf{j}_{\rm p}^{(1/\lambda)^2}(\mathbf{r})}{2\rho_{1/\lambda}(\mathbf{r})} = \frac{\mathbf{j}_{\rm p}^2(\mathbf{r}')}{\lambda^8}\frac{\lambda^3}{2\rho(\mathbf{r}')} = \frac{1}{\lambda^5}\frac{\mathbf{j}_{\rm p}^2(\mathbf{r}')}{2\rho(\mathbf{r}')},$$

which scales in the same way as τ, and shows that

$$\tau_{\rm D}^{1/\lambda}(\mathbf{r}) = \frac{1}{\lambda^5}\tau_{\rm D}(\mathbf{r}'). \tag{3.35}$$

For mGGAs, one then uses the scaled density of Eq. (3.19), the scaled gradient of Eq. (3.24) and one of the scaled kinetic energy densities of Eqs. (3.29), (3.32) or (3.35) to compute the correlation functional,

$$\begin{aligned} E_{\mathrm{c}}^{\mathrm{mGGA}}(\rho_{1/\lambda}) &= \int \epsilon_{\mathrm{c}}^{\mathrm{mGGA}}(\rho_{1/\lambda}(\mathbf{r}), \nabla_r^{1/\lambda}\rho_{1/\lambda}(\mathbf{r}), \tau^{1/\lambda}(\mathbf{r}))\ \mathrm{d}\mathbf{r} \\ &= \lambda^3 \int \epsilon_{\mathrm{c}}^{\mathrm{mGGA}}(\lambda^{-3}\rho_{1/\lambda}(\mathbf{r}), \lambda^{-4}\nabla_{r'}\rho(\mathbf{r}'), \lambda^{-5}\tau(\mathbf{r}'))\ \mathrm{d}\mathbf{r}'. \end{aligned} \tag{3.36}$$

Chapter 4

Application of Density-Functional Theory in a Magnetic Field

While Chapter 2 deals with exact DFT in magnetic fields from a theoretical point of view, this chapter focuses more on the application of this theory and computational aspects of calculations in magnetic fields. First, we consider some additional complications that arise with the introduction of magnetic fields from a computational point of view. We then outline the difference between dia- and paramagnetic molecules as background for Paper III, where these molecules are studied separately. Next, we discuss magnetic properties, one of the key topics of this work, in particular the magnetizability and NMR shielding constants. We review two different ways to compute magnetic properties. Both methods, numerical and analytical differentiation, are used in our papers. Finally, we describe our implementation of the Lieb optimisation, whose theoretical aspects have been discussed in Chapter 2.

4.1 Computational Aspects in Magnetic Fields

The introduction of a magnetic field into the Hamiltonian results in additional computational challenges. One major complication is that the vector potential is dependent on the choice of the coordinate system, but the final results should not to be affected by this choice. This problem can be solved with the correct choice of orbitals, which is discussed in the following.

4.1.1 Gauge-Origin Transformations

The vector potential

$$\mathbf{A_O}(\mathbf{r}) = \frac{1}{2}\mathbf{B} \times \mathbf{r_O} \tag{4.1}$$

is dependent on the gauge origin, here denoted by $\mathbf{O}$, which is located where the vector potential vanishes,

$$\mathbf{A_O}(\mathbf{O}) = \mathbf{0}. \tag{4.2}$$

A gauge transformation of the vector potential means that the location of $\mathbf{O}$ is changed to $\mathbf{G}$. It may be expressed in terms of the gauge function f,

$$\mathbf{A_G}(\mathbf{r}) = \mathbf{A_O}(\mathbf{r}) + \nabla f, \qquad f(\mathbf{r}) = -\mathbf{A_O}(\mathbf{G}) \cdot \mathbf{r} \tag{4.3}$$

The corresponding changes in the Hamiltonian are compensated for by the exact wave function, such that observables are not affected [32]. However, there may not be invariance in the case of approximate calculations. Truncations in the wave function, for example, can mean that not all the terms that are needed for the compensation of the gauge transformation are included.

4.1.2 London Orbitals

To compensate for the gauge transformation of the Hamiltonian, the wave function must change as $\Psi' = \exp(-\mathrm{i}f)\Psi$ [32]. With the gauge transformation of Eq. (4.3), the exact wave function transforms in the following manner,

$$\Psi_{\mathbf{G}} = \exp(\mathrm{i}\mathbf{A_O}(\mathbf{G}) \cdot \mathbf{r})\Psi_{\mathbf{O}} = \exp\left(\mathrm{i}\frac{1}{2}\mathbf{B} \times (\mathbf{G} - \mathbf{O}) \cdot \mathbf{r}\right)\Psi_{\mathbf{O}}, \tag{4.4}$$

where $\mathbf{B}$ denotes the magnetic field and $\mathbf{G}$ the gauge origin. The exponential term is a position-dependent phase factor, introducing oscillations in the wave function. To model the phase factor accurately in approximate calculations, standard atomic orbitals do not have enough flexibility. As a consequence, with such orbitals, large basis sets have to be used to obtain accurate results in the presence of magnetic fields. In 1937, however, London introduced *London orbitals*, also known as *gauge-origin including atomic orbitals* [69], which incorporate the oscillatory behaviour directly into the orbitals,

$$\omega(\mathbf{r_K}, \mathbf{B}, \mathbf{G}) = \exp\left(\mathrm{i}\frac{1}{2}\mathbf{B} \times (\mathbf{G} - \mathbf{K}) \cdot \mathbf{r}\right)\chi(\mathbf{r_K}), \tag{4.5}$$

where $\chi(\mathbf{r_K})$ is a standard atomic orbital which is centred at $\mathbf{K}$, with corresponding position vector $\mathbf{r_K} = \mathbf{r} - \mathbf{K}$. When using London orbitals for atomic systems, the unperturbed wave function is an eigenfunction of the perturbed Hamiltonian to first order in the field [4]. London orbitals improve the quality and basis-set convergence of calculations involving magnetic fields. Importantly, the calculation of observables becomes gauge-origin independent.

4.2 Diamagnetic and Paramagnetic Molecules

If the energy of a molecule increases when exposed to an external magnetic field, this molecule is said to be *diamagnetic*, whereas when the energy decreases, it is said to be *paramagnetic*. All open-shell systems are paramagnetic. They have a permanent magnetic moment and in the field, an alignment of this magnetic moment with the field results in a lowering of the energy. Closed-shell molecules, which do not have a permanent magnetic moment, may either be dia- or paramagnetic. Most of them, however, are diamagnetic.

Let us consider the Hamiltonian in a magnetic field, Eq. (2.66), repeated here for convenience,

$$\hat{H}_\lambda(v, \mathbf{A}) = \hat{H}_\lambda(v) - \mathrm{i} \sum_j \mathbf{A}(\mathbf{r}_j) \cdot \nabla_j + \frac{1}{2} \sum_j A^2(\mathbf{r}_j). \tag{4.6}$$

The term quadratic in the field can only raise the energy and dominates over the linear operator for a diamagnetic molecule. The paramagnetic term, which is linear in the field, may either raise or lower the energy. For a paramagnetic molecule, it dominates over the quadratic term and lowers the energy.

For closed-shell paramagnetic molecules, a common feature is a low-lying excited state that couples with the ground state in the presence of a magnetic field. When the coupling is strong enough, the two states avoid each other, such that no crossing may occur. This mechanism leads to a lowering of the energy for sufficiently small fields. However, if the field gets strong enough, the diamagnetic term will always dominate, such that paramagnetic molecules become diamagnetic from a certain field strength on. [70]

4.3 Magnetic Properties

In this section, we introduce magnetic properties and, in particular, consider those two properties that are studied in the papers of this thesis, the magnetizability and the NMR shielding constant. Next, we discuss how the computation of magnetic properties depends on an accurate description of the field dependence, before we discuss how to perform these calculations in practice.

4.3.1 Energy Derivatives

Introducing a perturbation μ to a molecular electronic system, the energy can be written as a Taylor expansion in terms of this perturbation,

$$E(\mu) = E_0 + \frac{\mathrm{d}E}{\mathrm{d}\mu}\mu\bigg|_{\mu=0} + \frac{1}{2}\mu^T \frac{\mathrm{d}^2 E}{\mathrm{d}\mu^2}\mu\bigg|_{\mu=0} + \ldots, \tag{4.7}$$

where E_0 denotes the unperturbed energy. An external homogeneous magnetic field $\mathbf{B}$ and the magnetic moment

$$\mathbf{M}_K = \gamma_K \mathbf{I}_K, \tag{4.8}$$

associated with spin $\mathbf{I}_K$ and gyromagnetic ratio γ_K of nucleus K, can be considered as such perturbations to a field-free system. To second order in these two perturbations, the expression for the energy becomes

$$\begin{aligned} E(\mathbf{B},\mathbf{M}) =& E_0 + \frac{1}{2}\mathbf{B}^T \left.\frac{\mathrm{d}^2 E(\mathbf{B},\mathbf{M})}{\mathrm{d}\mathbf{B}^2}\right|_{\mathbf{B}=0} \mathbf{B} + \sum_K \mathbf{B}^T \left.\frac{\mathrm{d}^2 E(\mathbf{B},\mathbf{M})}{\mathrm{d}\mathbf{B}\mathrm{d}\mathbf{M}_K}\right|_{\mathbf{B},\mathbf{M}=0} \mathbf{M}_K \\ &+ \frac{1}{2}\sum_{K\neq L} \mathbf{M}_K^T \left.\frac{\mathrm{d}^2 E(\mathbf{B},\mathbf{M})}{\mathrm{d}\mathbf{M}_K \mathrm{d}\mathbf{M}_L}\right|_{\mathbf{M}=0} \mathbf{M}_L, \end{aligned} \tag{4.9}$$

where the magnetic moments are collectively referred to as $\mathbf{M} = \{\mathbf{M}_K\}$. The expansion in Eq. (4.9) does not include first-order terms since these terms vanish for closed-shell systems [4], which we exclusively consider in this work. The energy derivatives correspond to specific molecular properties, and the two relevant for this work, the magnetizability and the nuclear shielding constant, are be discussed in the following.

Magnetizability

The magnetizability is defined as the second-order response of a molecule to an external magnetic field,

$$\xi = -\left.\frac{\mathrm{d}^2 E(\mathbf{B})}{\mathrm{d}\mathbf{B}^2}\right|_{\mathbf{B}=0}. \tag{4.10}$$

For closed-shell systems, second-order perturbation theory yields the following explicit expression [71],

$$\xi = -\frac{1}{4}\left\langle 0 \left| \left(\mathbf{r}_\mathbf{O}^T \mathbf{r}_\mathbf{O}\mathbf{I} - \mathbf{r}_\mathbf{O}\mathbf{r}_\mathbf{O}^T\right) \right| 0 \right\rangle - \frac{1}{2}\sum_{n\neq 0} \frac{\left\langle 0 \left| \mathbf{l}_\mathbf{O}^T \right| n \right\rangle \left\langle n \left| \mathbf{l}_\mathbf{O} \right| 0 \right\rangle}{E_0 - E_n}, \tag{4.11}$$

where $\mathbf{I}$ is the 3×3 unit matrix, and $\mathbf{r}_\mathbf{O}$ and $\mathbf{l}_\mathbf{O}$ are the position vector and the electronic angular momentum vector around the gauge origin $\mathbf{O}$, respectively. We here follow the common notation of e.g. Ruud and Helgaker [71], where the summation over electrons is omitted for simplicity.

The first term, an expectation value depending on the ground state only, is called *diamagnetic* contribution, whereas the second term, a sum over states is called *paramagnetic* contribution. The average of the diagonal elements of the magnetizability tensor is referred to as *isotropic magnetizability*,

$$\xi^{\mathrm{iso}} = \frac{1}{3}\mathrm{Tr}\xi. \tag{4.12}$$

For atoms in a homogeneous magnetic field, the simple relationship

$$\xi^{\mathrm{dia}} = -\frac{1}{6}\left\langle r^2 \right\rangle \tag{4.13}$$

with $r = |\mathbf{r}|$ holds, as explicitly derived in Appendix B. Furthermore, for closed-shell atoms, the paramagnetic contribution vanishes since their ground-state is an eigenstate of the angular momentum operator with eigenvalue 0. Thus, only the diamagnetic part given in Eq. (4.13) contributes, such that the total magnetizability is directly proportional to the expectation value of r^2.

Nuclear Shielding Constant

Another second-order property is the nuclear shielding constant. It is used for molecular structure determination in NMR spectroscopy and therefore an important application area of quantum chemistry. The shielding constant describes the modification of an external magnetic field at the nucleus due to the magnetic field resulting from the motion of the electrons.

The shielding tensor σ_K associated with nucleus K describes the coupling between an external field and the nuclear magnetic moment of nucleus K. It is defined as the second-order derivative of the molecular electronic energy with respect to the magnetic field $\mathbf{B}$ and the magnetic moment $\mathbf{M}_K$ of that nucleus[1],

$$\sigma_K = \left.\frac{\mathrm{d}^2 E(\mathbf{B},\mathbf{M})}{\mathrm{d}\mathbf{B}\,\mathrm{d}\mathbf{M}_K}\right|_{\mathbf{B},\mathbf{M}=0} + \mathbf{I}. \tag{4.14}$$

The subdivision into dia- and paramagnetic contributions in approximate calculations is not unique and different conventions can be found in literature. We here follow the one of Helgaker et al. [74], where the diamagnetic contribution only contains terms that can be computed as an expectation value with the ground state wave function. All other terms are put into the paramagnetic part.

Again omitting the summation over electrons for simplicity, second-order perturbation theory yields Ramsey's expression for nuclear shielding constants [73]

$$\sigma_K = \frac{1}{2}\left\langle 0 \left| \frac{\mathbf{r}_{\mathbf{O}}^T \mathbf{r}_K - \mathbf{r}_{\mathbf{O}} \mathbf{r}_K^T}{r_K^3} \right| 0 \right\rangle - \sum_n \frac{\langle 0|\mathbf{l}_{\mathbf{O}}|n\rangle\langle n|\mathbf{l}_K^T \mathbf{r}_K^{-3}|0\rangle}{E_n - E_0} + \mathbf{I}, \tag{4.15}$$

where $\mathbf{r}_K, \mathbf{r}_{\mathbf{O}}$ and $\mathbf{l}_K, \mathbf{l}_{\mathbf{O}}$ are the position vectors and the orbital angular momentum vectors relative to nucleus K and gauge origin $\mathbf{O}$, respectively. We have added the unit matrix $\mathbf{I}$ to Ramsey's original expression to be consistent with our definition of the shielding tensor in Eq. (4.14).

[1] Note that sometimes the unit matrix is dropped in literature, see e.g. Refs. [72, 73].

The first term is the diamagnetic contribution, the second term the paramagnetic contribution. Since σ_K depends on the choice of the gauge origin, it is necessary to use London orbitals in order to obtain gauge-invariant results. As for the magnetizability, the paramagnetic term vanishes for closed-shell atoms, such that the shielding is given by the simple relationship

$$\sigma_K = \frac{1}{3}\left\langle 0 \left| \frac{1}{r_K} \right| 0 \right\rangle, \tag{4.16}$$

which is called *Lamb formula* [75]. The isotropic shielding constant is defined as

$$\sigma_K^{\text{iso}} = \frac{1}{3}\text{Tr}\sigma_K. \tag{4.17}$$

Comparing the expression for the shielding tensor, Eq. (4.15), with that of the magnetizability, Eq. (4.11), we note that the shielding tensor is more sensitive to regions close to the nuclei, whereas the magnetizability is more dependent on regions further away. This fact can be used to assess the quality of approximate exchange–correlation functionals in DFT. An accurate value for the magnetizability, but a poorer one for the shielding constant, for example, may indicate that the functional does not model regions close to the nuclei well enough.

4.3.2 Field-Dependence of Magnetic Properties

As mentioned above, for all magnetic properties there exists a (non-unique) subdivision into diamagnetic and paramagnetic parts. For the analysis of approximate methods, we further subdivide the paramagnetic part into field-independent and field-dependent parts. We use this subdivision for the NMR shielding constant in Paper I and for the magnetizability in Paper II .

There are thus three terms, here demonstrated for the magnetizability,

$$\xi = \xi^{\text{dia}} + \xi_{\rho}^{\text{para}} + \xi_{\mathbf{A}}^{\text{para}}, \tag{4.18}$$

where the first term refers to the diamagnetic part, the second term to the field-independent paramagnetic part, and the last term to the field-dependent paramagnetic part. The usefulness of this decomposition is that it allows us to quantify how large the error is if the field-dependent part is neglected. This is the case for all standard DFAs, where there is no explicit field-dependence in the exchange–correlation functional.

Computation of magnetic properties in practice From a practical point of view, the diamagnetic part is easily obtained as an expectation value, using the one-electron density matrix. For the paramagnetic contributions, first a calculation of the interacting system, at $\lambda = 1$, is performed. This can be done with any electronic structure method,

for example FCI, coupled-cluster or Hartree–Fock models. The obtained density matrix is used in a subsequent Lieb calculation at $\lambda = 0$, as discussed in Subsection 4.5. The obtained orbitals and orbital energies are then used to compute the property of interest, using the perturbative approach a response calculation, as discussed in Section 4.4.2. Within the limits of the method which is used to obtain the density, this response calculation with the Kohn–Sham orbitals gives the exact, but field-independent value for the property of interest. This is due to the fact that the response calculation does not contain any field dependence, but simply receives the orbitals that give the correct density, calculated in the Lieb calculation.

The field-dependent part is then obtained as the difference between the value of the, for example, coupled-cluster or FCI calculation, and the one from the Kohn–Sham response calculation. For the magnetizability, this gives $\xi_{\mathbf{A}}^{\text{para}} = \xi - \xi_{\text{KS}}$, and the same expression holds for the shielding constant with σ instead of ξ. The field-independent paramagnetic part is the difference between the Kohn–Sham value and the diamagnetic part, for the magnetizability in particular $\xi_{\rho}^{\text{para}} = \xi_{\text{KS}} - \xi^{\text{dia}}$. The density used in the Kohn–Sham calculation has to be the same one with which the diamagnetic part is computed. The different steps are illustrated in Figure 4.1.

4.4 Computation of Magnetic Properties

To compute magnetic properties, there exist two main approaches. One way is to compute the energy derivatives numerically. This is what we primarily do in the papers of this work, using the method of finite difference. Alternatively, the properties can be computed analytically, employing response theory. A short overview over both methods, along with their advantages and disadvantages, is given in the following two subsections.

4.4.1 Numerical Differentiation

As discussed in Subsection 4.3.1, magnetic properties are obtained as energy derivatives with respect to the magnetic field, computed at zero field strength. This means that any finite difference approach can be utilized, where the energy is computed for several sufficiently small field strengths. Afterwards, either the derivative is formed, or the points are fitted by a polynomial. For example, for the magnetizability, which is the negative second-order derivative of the energy with respect to the field, $\xi = -\mathrm{d}^2 E(\mathbf{B})/\mathrm{d}\mathbf{B}^2|_{\mathbf{B}=0}$, one needs at least three points. The challenge is to find a step length and an appropriate number of points, such that the derivative is computationally stable. A too large step length will make the method less accurate by the inclusion of higher-order terms. The derivative is needed at zero field strength, and one should therefore aim at step lengths as small as possible. However, too small steps give rise to numerical instabilities, such that

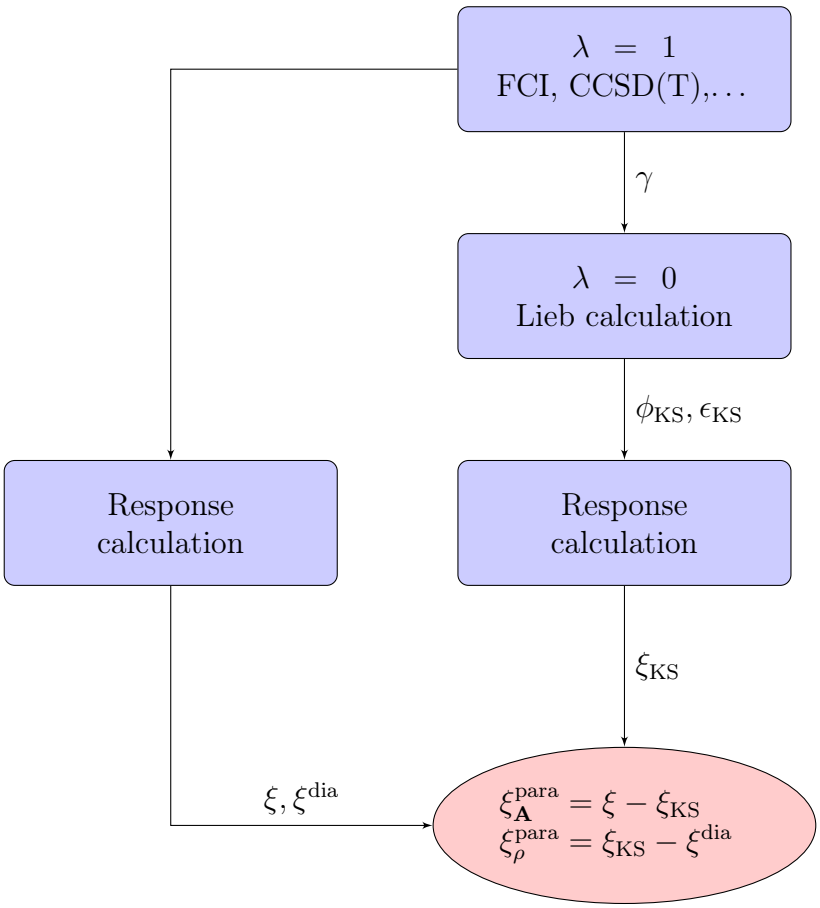


Figure 4.1: Demonstration of the different steps needed to compute all components of a magnetic property, here the magnetizability ξ. To have access to the field-independent part ξ_{KS}, we first perform a calculation at $\lambda = 1$. The obtained density matrix γ is used in a subsequent Lieb calculation at $\lambda = 0$, yielding Kohn–Sham orbitals ϕ_{KS} and Kohn–Sham energies ϵ_{KS}. Used in a response calculation, the Kohn–Sham orbitals and energies give the field-independent part of the magnetizability, ξ_{KS}, while the results from the $(\lambda = 1)$-calculation give the total and diamagnetic part, ξ and ξ^{dia}, respectively.

the right balance needs to be found.

One useful property of the molecular energy is that, due to symmetry, we have $E(\mathbf{B}) = E(-\mathbf{B})$, such that for each calculation of the energy at non-zero field strength, two points for the finite difference method are obtained.

The disadvantage of the numerical method is that calculations of the energy have to be done at several field strengths with tight convergence thresholds, which can be computationally costly, and that an optimal step length has to be determined to avoid numerical instabilities. The advantage, on the other hand, is that it is easy to implement.

4.4.2 Analytical Differentiation

To avoid having to perform calculations at several field strengths and to achieve greater speed and precision, magnetic properties can be computed analytically, using the method of response. We here restrict ourselves to time-independent, second-order properties, since only these are studied in our papers. A more general discussion of response theory is beyond the scope of this work. We will here follow the lectures of Trygve Helgaker at the Sostrup Summer School [76].

Consider the Taylor expanded energy of Eq. (4.9) up to second order in the perturbations $\mathbf{B}$ and $\mathbf{M}$. All energy derivatives,

$$\left.\frac{\mathrm{d}^2 E(\mathbf{B},\mathbf{M})}{\mathrm{d}\mathbf{B}^2}\right|_{\mathbf{B}=0}, \qquad \left.\frac{\mathrm{d}^2 E(\mathbf{B},\mathbf{M})}{\mathrm{d}\mathbf{B}\mathrm{d}\mathbf{M}_K}\right|_{\mathbf{B},\mathbf{M}=0}, \qquad \left.\frac{\mathrm{d}^2 E(\mathbf{B},\mathbf{M})}{\mathrm{d}\mathbf{M}_K\mathbf{M}_L}\right|_{\mathbf{M}=0} \tag{4.19}$$

are time-independent, second-order properties, to which the theory of linear response can be applied.

Perturbation theory yields explicit expressions for the energy derivatives, as given in Eq. (4.11) for the magnetizability and Eq. (4.15) for the NMR shielding constant. However, to compute these expressions explicitly, one would need the complete set of eigenstates $|n\rangle$ and associated eigenvalues E_n, which are only available at the FCI level. Therefore, these expressions are therefore not helpful for approximate methods like the Hartree–Fock and coupled-cluster methods.

For approximate methods, the energy functional $E(\mu, \mathbf{\Lambda})$ depends on two sets of parameters, the external perturbation parameter μ, as for example the magnetic field $\mathbf{B}$ or the magnetic moment $\mathbf{M}$, and the wave function parameters $\mathbf{\Lambda} = (\lambda_1, \lambda_2, \dots)$, like the molecular orbitals or cluster amplitudes. The parameters $\mathbf{\Lambda}(\mu)$ depend implicitly on the external parameter μ, while the Hamiltonian depends explicitly on μ. For each fixed value of μ, the parameters $\mathbf{\Lambda}$ in $E(\mu, \mathbf{\Lambda})$ are optimized to give the total energy $\mathcal{E}(\mu)$ with optimal values $\mathbf{\Lambda}^*(\mu)$,

$$\mathcal{E}(\mu) = E(\mu, \mathbf{\Lambda}^*). \tag{4.20}$$

Depending on the method, this optimization is not necessarily variational (e.g. for coupled-cluster theory). To compute the derivatives of $\mathcal{E}(\mu)$ with respect to μ, both the explicit and the implicit dependence must be taken into account.

For the first-order derivative, this means

$$\frac{\mathrm{d}\mathcal{E}(\mu)}{\mathrm{d}\mu} = \frac{\partial E(\mu, \mathbf{\Lambda}^*)}{\partial \mu} + \sum_i \left.\frac{\partial E(\mu, \mathbf{\Lambda})}{\partial \lambda_i}\right|_{\mathbf{\Lambda}=\mathbf{\Lambda}^*} \frac{\partial \lambda_i^*}{\partial \mu}, \tag{4.21}$$

where the first term considers the explicit dependence on μ and the second one the implicit dependence through the parameters $\mathbf{\Lambda}$. The quantity $\partial\lambda_i^*/\partial\mu$ is the linear response of the wave function, describing the change of the electronic structure when the system is perturbed.

Variational wave functions In the case of variational wave functions, the stationary condition

$$\left.\frac{\partial E(\mu, \mathbf{\Lambda})}{\partial \lambda_i}\right|_{\mathbf{\Lambda}=\mathbf{\Lambda}^*} = 0 \tag{4.22}$$

leads to a simplification of the molecular gradient to

$$\frac{\mathrm{d}\mathcal{E}(\mu)}{\mathrm{d}\mu} = \frac{\partial E(\mu, \mathbf{\Lambda}^*)}{\partial \mu}. \tag{4.23}$$

This means that the response of the wave function is not needed for the first derivative of the energy.

For the second-order derivative of $\mathcal{E}(\mu)$,

$$\begin{aligned}
\frac{\mathrm{d}^2\mathcal{E}(\mu)}{\mathrm{d}\mu^2} &= \frac{\mathrm{d}}{\mathrm{d}\mu}\frac{\partial E(\mu, \mathbf{\Lambda}^*)}{\partial \mu} \\
&= \left(\frac{\partial}{\partial \mu} + \sum_i \frac{\partial \lambda_i^*}{\partial \mu}\left.\frac{\partial}{\partial \lambda_i}\right|_{\mathbf{\Lambda}=\mathbf{\Lambda}^*}\right)\frac{\partial E(\mu, \mathbf{\Lambda}^*)}{\partial \mu} \\
&= \frac{\partial^2 E(\mu, \mathbf{\Lambda}^*)}{\partial \mu^2} + \sum_i \left.\frac{\partial^2 E}{\partial \mu \partial \lambda_i}\right|_{\mathbf{\Lambda}=\mathbf{\Lambda}^*} \frac{\partial \lambda_i^*}{\partial \mu},
\end{aligned} \tag{4.24}$$

however, the first-order response $\partial\lambda_i^*/\partial\mu$ is required. To compute the response, the stationary condition $\partial E(\mu, \mathbf{\Lambda})/\partial\lambda_i|_{\mathbf{\Lambda}=\mathbf{\Lambda}^*} = 0$, which holds for all perturbations μ, is differentiated with respect to μ,

$$\left.\frac{\mathrm{d}}{\mathrm{d}\mu}\frac{\partial E(\mu, \mathbf{\Lambda})}{\partial \lambda_i}\right|_{\mathbf{\Lambda}=\mathbf{\Lambda}^*} = \left.\frac{\partial^2 E(\mu, \mathbf{\Lambda})}{\partial \mu \partial \lambda_i}\right|_{\mathbf{\Lambda}=\mathbf{\Lambda}^*} + \sum_j \left.\frac{\partial^2 E(\mu, \mathbf{\Lambda})}{\partial \lambda_i \partial \lambda_j}\right|_{\mathbf{\Lambda}=\mathbf{\Lambda}^*} \frac{\partial \lambda_j^*}{\partial \mu} = 0. \tag{4.25}$$

This gives rise to the first-order time-independent response equations,

$$\sum_j \left.\frac{\partial^2 E(\mu,\mathbf{\Lambda})}{\partial\lambda_i\partial\lambda_j}\right|_{\mathbf{\Lambda}=\mathbf{\Lambda}^*} \frac{\partial\lambda_j^*}{\partial\mu} = -\left.\frac{\partial^2 E(\mu,\mathbf{\Lambda})}{\partial\mu\partial\lambda_i}\right|_{\mathbf{\Lambda}=\mathbf{\Lambda}^*} \tag{4.26}$$

The first factor $\partial^2 E(\mu,\mathbf{\Lambda})/\partial\lambda_i\partial\lambda_j$ is the electronic Hessian, a Hermitian matrix. The response equations (4.26) form a linear system of equations. Solving them yields the linear response vector $\partial\lambda_i^*/\partial\mu$. Usually, the response equations are solved iteratively, where the Hessian is multiplied with a trial vector. The Hessian matrix itself is generally too large to be explicitly constructed and stored.

It should be noted that only the first-order derivative of the wave function is needed to compute the second-order derivative of the energy. In general, we have the $2n+1$-rule, which says that for variational wave functions, the nth derivative of a wave function determines the energy derivatives to order $2n+1$.

Non-variational wave functions For non-variational wave functions, as for example coupled-cluster and the configuration interaction method, the energy can be made stationary by Lagrange's method of undetermined multipliers. For Lagrange multipliers the derivative of the wave function to order n determines the energy to order $2n+2$.

Advantages and disadvantages Compared to numerical differentiation, the advantages of the analytical differentiation are that the exact expressions allow a higher precision and are computationally more efficient. The main drawback is that the implementation is more difficult and requires more work. Moreover, for each property a different expression is needed while a numerical implementation can be used for any property.

4.5 Lieb Optimization in the Context of BDFT

In the BDFT framework, the Lieb functional is given by Eq. (2.69), which we repeat here for further use,

$$F_\lambda(\rho,\mathbf{A}) = \sup_{v\in X^*}\left[E_\lambda(v,\mathbf{A}) - (v|\rho)\right]. \tag{4.27}$$

In practice, the exact energy functional $E_\lambda(v,\mathbf{A})$ is not known, but approximate models can be used, as for example coupled-cluster or Hartree–Fock models. For the computation of AC curves in practice, we need to compute the AC integrand $\mathcal{W}_{\mathrm{xc},\lambda}(\rho,\mathbf{A}) = \langle\Psi_{\lambda,\mathbf{A}}|\hat{V}_{\mathrm{ee}}|\Psi_{\lambda,\mathbf{A}}\rangle - J(\rho)$. This requires wave functions with a fixed density ρ and at a fixed magnetic field, but with varying values for the interaction parameter λ.

In this work, we perform for each value of λ a Lieb optimization as given by Eq. (4.27). We follow the procedure by Wu and Yang [77], who originally applied this scheme in the

context of optimized effective potentials at $\lambda = 0$. This optimization method has later been applied to all interaction strengths $\lambda \in [0, 1]$, see for example Ref. [26].

The one-electron potential is constructed as a sum of the nuclear potential $v_{\text{nuc}}(\mathbf{r})$, a reference potential $v_{\text{ref}}(\mathbf{r})$ and a linear expansion in a basis set of Gaussians $\{g_t\}$ with expansion coefficients $\{b_t\}$,

$$v_{\text{eff}}(\mathbf{r}) = v_{\text{nuc}}(\mathbf{r}) + (1 - \lambda)v_{\text{ref}}(\mathbf{r}) + \sum_t b_t g_t(\mathbf{r}). \tag{4.28}$$

As a reference potential, we chose the Fermi-Amaldi potential [78],

$$v_{\text{ref}}(\mathbf{r}) = \frac{N-1}{N} \int \frac{\rho(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|} \, \mathrm{d}\mathbf{r}', \tag{4.29}$$

which ensures the correct long-range behaviour of the effective potential [79].

The aim in the Lieb maximization is to find those coefficients $\{b_t\}$ that maximize the quantity

$$G_{\lambda,b}(\rho, \mathbf{A}) = E_\lambda(v_{\lambda,b}, \mathbf{A}) - \int v_{\lambda,b}(\mathbf{r})\rho(\mathbf{r}) \, \mathrm{d}\mathbf{r}. \tag{4.30}$$

Depending on which optimization scheme is used, the gradient and the Hessian may be needed for the calculations, too. First-order methods, like the quasi-Newton method, require only the gradient, given by

$$\frac{\partial G_{\lambda,b}(\rho, \mathbf{A})}{\partial b_t} = \int [\rho_{\lambda,b}(\mathbf{r}) - \rho(\mathbf{r})] \, g_t(\mathbf{r}) \mathrm{d}\mathbf{r}, \tag{4.31}$$

whereas second-order methods, like the full Newton method, also require the Hessian,

$$\frac{\partial^2 G_{\lambda,b}(\rho, \mathbf{A})}{\partial b_t \partial b_u} = \iint g_t(\mathbf{r}) g_u(\mathbf{r}') \frac{\delta\rho(\mathbf{r})}{\delta v(\mathbf{r}')} \, \mathrm{d}\mathbf{r}\mathrm{d}\mathbf{r}'. \tag{4.32}$$

4.5.1 Our Procedure for Computing AC Curves with the Lieb Optimization

When generating an AC curve, we proceed in the following way:

1. First, we run a calculation at $\lambda = 1$ with a method of choice, for example Full Configuration Interaction (FCI), which gives a ground state density $\rho(\mathbf{r})$.

2. We then perform a Lieb optimization to find the optimal set of coefficients $\{b_t\}$, which determines the effective potential $v_{\text{eff}}(\mathbf{r})$ and with it the exact expression of the Hamiltonian $\hat{H}_\lambda(v, \mathbf{A})$.

3. With this set of coefficients $\{b_t\}$, we obtain the AC integrand $\mathcal{W}_{\mathrm{xc},\lambda}$ as

$$\begin{aligned}\mathcal{W}_{\mathrm{xc},\lambda}(\rho,\mathbf{A}) &= \frac{1}{\lambda}\langle\Psi_{1,\mathbf{A}}|\hat{V}_{\mathrm{ee}}|\Psi_{1,\mathbf{A}}\rangle - J(\rho) \\ &= \frac{1}{\lambda}\Big(\, E_\lambda(v,\mathbf{A}) - T(\rho,\mathbf{A}) - (v|\rho)\Big) - J(\rho),\end{aligned} \tag{4.33}$$

since V_{ee} scales linearly in λ [68],

$$\langle\Psi_{\lambda,\mathbf{A}}|\hat{V}_{\mathrm{ee}}|\Psi_{\lambda,\mathbf{A}}\rangle = \lambda\langle\Psi_{1,\mathbf{A}}|\hat{V}_{\mathrm{ee}}|\Psi_{1,\mathbf{A}}\rangle. \tag{4.34}$$

4.5.2 Our Implementation of the Lieb Optimization

To find the optimal coefficients $\{b_t\}$ for the potential in Eq. (4.28), we employ the Newton algorithm, which we, out of several optimization algorithms, found to require the smallest number of iterations until convergence. With Newton's method one approximates the function locally by a quadratic polynomial and uses the gradient and the Hessian matrix of second derivatives to find the optimum.

In practice, we start with a guess for the coefficients $\{b_t\}$. Then, with this set of coefficients, we perform an electronic structure calculation, e.g. on the FCI or coupled-cluster level, using the `LONDON` quantum-chemistry software [80, 81]. This code supplies us with the functional of Eq. (4.30) and the Lieb gradient of Eq. (4.31). Moreover, it provides us with the Hessian at the Hartree–Fock level, which we use as an approximation to the exact Hessian. In `LONDON`, this Hessian is computed as $H_{tu} = \sum_{ai}\langle i|g_t|a\rangle\langle a|g_u|i\rangle/(\epsilon_a - \epsilon_i)$, where a, i are the molecular orbital coefficients and ϵ_i, ϵ_a the molecular orbital energies. The functional $G_{\lambda,b}$, Lieb gradient and Hessian are used to determine iteratively the next set of coefficients $\{b_t\}$ until the norm of the gradient is smaller than a user-defined threshold.

Regularization to ensure a maximum

The Lieb functional of Eq. (4.27) does not necessarily have a maximum, in which case the optimizer fails to find a solution. Moreover, in the context of optimized effective potentials (OEPs), it has largely been discussed that the use of a finite basis set for the Kohn–Sham orbitals and the potential makes the problem ill-posed [82, 83, 84]. When the exchange–correlation functional is constructed in a finite basis set, its response to orbital perturbations is limited. The basis set for the potential can be inappropriate and there may be cases where changes in the potential cannot be properly reflected in the orbitals. Many different potentials, including non-physical ones, can numerically give the same total energy and orbitals.

One typical appearance, for example, are potentials with strong oscillations. Such

potentials are numerically possible if the oscillations occur in regions where the orbitals almost vanish, or if the potentials change so rapidly that the orbitals cannot adjust. Through regularization, however, it can be ensured that only physically meaningful potentials are generated. In the case of oscillations, it is possible to introduce a penalty term $\|v_{\lambda,b}(\mathbf{r})\|^2$, as done by Heaton-Burgess et al. [82] and Bulat et al. [84]. In this way, smooth potentials are favoured in the optimization procedure.

In our work, we use another regularization, which is inspired by the Moreau–Yosida regularization [85], easy to implement and at the same time effective to yield stable results in the optimization procedure: Instead of maximizing the quantity of Eq. (4.30), we are maximizing

$$G_{\lambda,b}(\rho, \mathbf{A}) - \frac{1}{2}\mu\|b\|^2 = E_\lambda(v_{\lambda,b}, \mathbf{A}) - \int v_{\lambda,b}(\mathbf{r})\rho(\mathbf{r})\ \mathrm{d}\mathbf{r} - \frac{1}{2}\mu\|b\|^2, \tag{4.35}$$

where $b = \sum_t b_t$ is the sum of the expansion coefficients of the potential and μ is a user-defined regularization constant. Accordingly, the gradient becomes

$$\frac{\partial G_{\lambda,b}(\rho, \mathbf{A})}{\partial b_t} - \mu\|b\| = \int [\rho_{\lambda,b}(\mathbf{r}) - \rho(\mathbf{r})]\ g_t(\mathbf{r})\mathrm{d}\mathbf{r} - \mu\|b\|. \tag{4.36}$$

Working with this regularization, one needs to choose μ large enough to achieve convergence, but at the same time one should try to keep it as small as possible in order not to change the Hamiltonian too much. Only for small values of μ, the regularized solution is a good approximation to the unregularized one.

Chapter 5

Discussion of Papers

In this chapter, the papers that form the foundation of this thesis are presented. For each paper, a short summary of the main results is given. Moreover, additional material may be included, mainly results that are closely linked to or have been important during the work on the paper, but that were eventually not included in the paper. Still, this material can be considered as valuable background information and is therefore presented in the context of this thesis.

5.1 Paper I

Title: *The importance of current contributions to shielding constants in density-functional theory*

Formal citation:
Reimann, S.; Ekström, U.; Stopkowicz, S.; Teale, A.M.; Borgoo, A.; Helgaker, T. The importance of current contributions to shielding constants in density-functional theory. *Phys. Chem. Chem. Phys.* **2015**, *17,* 18834-18842.

5.1.1 Main Results

In this paper, we studied the sources of error in the calculation of NMR shielding constants when utilizing standard DFAs. The aim is to quantify the error when the magnetic field and its corresponding current are neglected, which in turn serves as a motivation to develop BDFT to reduce this error. Moreover, errors originating from the electron density are analyzed, since the current (or field) corrections are in most cases relatively small. It is therefore important that errors in the underlying DFA are reduced as far as possible. Only in this way, the full benefits of a field-dependent functional can be realized and the advantages are not outweighed by much larger, field-independent errors.

Paper I is written from the perspective of CDFT and in terms of current corrections, instead of field corrections as done in the subsequent papers. We want to point out that all results presented here are equally valid for field corrections in the context of BDFT.

The main conclusions of this paper are given in the following paragraphs.

Missing current contribution is a leading source of error in shielding calculations using DFAs As explained in Subsection 4.3.2, we have subdivided the shielding tensor into a diamagnetic part, a current-independent paramagnetic part and a current-dependent paramagnetic part, and we analyzed those parts individually. Using the Lieb optimization outlined in Section 4.5, we were able to isolate the current-dependent paramagnetic part for a set of molecules. We found that, although the current contribution typically is one to two orders of magnitude smaller than the dia- and paramagnetic contributions, it may still be important since the dia- and paramagnetic contributions have opposite signs and in some cases nearly cancel. For the carbon atom in CO, for example, the current contribution is twice as large as the total shielding constant. A neglect of this contribution, which is the case for conventional DFAs, leads therefore to a large error in shielding calculations.

Assessment of current-independent functionals It is useful to judge existing, current-independent functionals on their ability to reproduce accurate *ab-initio* shielding constants, for example from coupled-cluster calculations, where the current contribution has been subtracted. If comparing the results of current-independent DFAs to *ab-initio* results including current effects, it is impossible to distinguish between the errors of a DFA and the missing current dependence. As a result, the functionals might not even benefit from accurate current corrections.

TPSS is a promising starting point Of the DFAs investigated, we found the gauge-invariant generalizations of TPSS to be the most promising candidates to apply a current correction to. Good current-independent shieldings require that the diamagnetic part is well described, which depends on a good density at zero field strength. For most DFAs, we observed large errors in the density around the nuclei, considerably larger than for Hartree–Fock. We consistently found TPSS to give the best densities. Also including the current-independent paramagnetic part, we found TPSS and its generalizations to be the approximate functional with the most balanced error.

5.2 Paper II

Title: *Magnetic-Field Density-Functional Theory (BDFT): Lessons from the Adiabatic Connection*

Formal citation:
Reimann, S.; Borgoo, A.; Tellgren, E. I.; Teale, A.M.; Helgaker, T. Magnetic-Field Density-Functional Theory (BDFT): Lessons from the Adiabatic Connection. *J. Chem. Theory Comput.* **2017**, *13,* 4089-4100.

5.2.1 Main Results

In this paper, the main question is whether for the accurate computation of magnetic properties using DFAs, we first need to fix the field-free functional or if the introduction of a field-dependence should be prioritized.

To address this question, we extend the study of AC curves to molecular systems involving magnetic fields, since by now, AC curves have only been studied for field-free systems. In particular, we take the approach from the perspective of BDFT. The intention is to consider both the weak-field and the strong-field regime, since an understanding of weak fields is necessary for an accurate modeling of magnetic properties, whereas strong fields are useful to improve our understanding of fundamental physical and chemical concepts.

Another aim of the paper is to compare the BDFT and CDFT correlation functionals, using the four-way correspondence of convex analysis, and to establish the relationships between those two approaches. We have presented most of this material in Section 2.3 of this thesis.

The main conclusions of the paper are presented in the following paragraphs.

Improvement of the magnetizability by using field-dependent density functionals We decompose the energy and its second derivative with respect to the magnetic field into its Kohn–Sham components, keeping in mind that the second derivative is proportional to the magnetizability, see Eq. (4.10). This decomposition allows us to study the individual contributions to the magnetizability in detail.

We show that the magnetizability is primarily determined by those energy components that are related to the density, namely $(v|\rho), J(\rho)$ and $T_\mathrm{s}(\rho)$. For the magnetizability, therefore the benefit of improving the density of existing DFAs will be larger than the benefit of introducing a magnetic-field dependence in the correlation functional. This is further motivated by our calculations which show how poor the self-consistent densities of present approximate functionals are.

At the same time, we show that once a good density has been attained, only an inclusion of magnetic-field effects into the functionals can lead to highly accurate results. This is illustrated by the fact that for some molecules, the field-dependent part of the magnetizability is comparable to the total correlation contribution. To get a good density,

however, we need a good correlation functional already at zero-field.

Behaviour of AC curves depends on the choice of geometry Studying AC curves in the regime of strong magnetic fields, we find that as long as each curve is calculated at the equilibrium geometry of that field strength, the curves are very close to each other and the magnetic field does not result in any significant changes.

However, if for all AC curves the equilibrium geometry of the field-free system is used, the curves bend more with increasing field strength. The reason is that static correlation is getting more important. Static correlation increases with the field, since a magnetic field results in a compression of the charge density. This means that keeping the molecular geometry fixed while increasing the magnetic field is equivalent to effectively stretching the bonds. We demonstrate that existing models for the AC integrand are able to capture this behaviour.

AC curves using approximate functionals Computing AC curves with approximate functionals, we show that the main error in the presence of a field is already present at zero field strength. The field does not introduce a major additional error, which justifies the use of DFAs for systems in strong fields without the need to treat additional correlation effects. For the DFAs studied in this paper, the most accurate AC integrands are obtained with TPSS using τ_{MS}, see Eq. (3.30), which we also refer to as aTPSS.

Since the major errors in the correlation functional are already present at zero-field, it is important to improve the functionals in the absence of a field. Only then can the benefits of field-dependent corrections be fully realized.

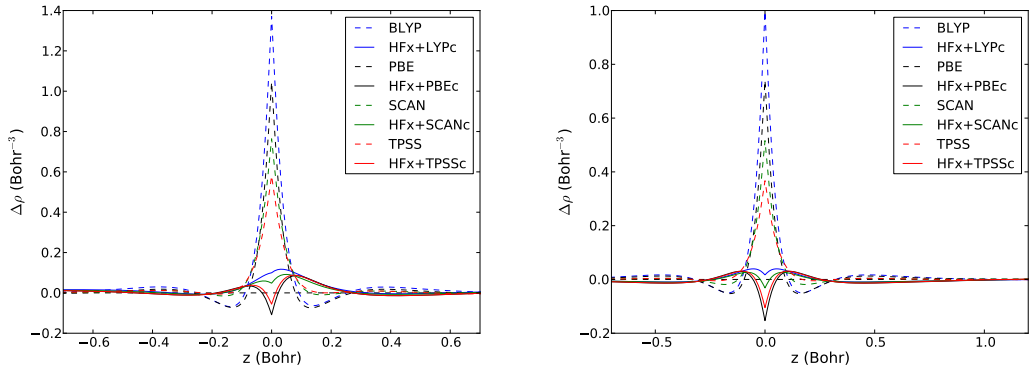
5.2.2 Additional Material

One of the main messages of this paper is that, when computing the magnetizability with DFAs, the main error comes from the poor self-consistent densities. We therefore thought it would be instructive to study to which a degree a better density could affect the results. From Paper I we know that, especially around the nuclei, the Hartree–Fock method has a much smaller error in the density than all DFAs we studied. Hence we wanted to investigate whether combining the correlation part of any DFA with 100% exact exchange would lead to an improvement of the density and the magnetizability. We are of course aware of the fact that this combination breaks the error cancellation between exchange and correlation part that most DFAs rely on for good energetics, apart from the fact that the internal parameters were optimized for the DFA and not the exact exchange. However, this analysis still gives an indication if, at least for the magnetizability, the importance of the density is so large that a slightly worse value for the exchange–correlation energy could be accepted and whether there are consistent trends when increasing the amount of

	He	Ne	HF	H_2O	NH_3	CH_4	CO	N_2
PBE	−0.02	−0.04	0.03	−0.04	−0.14	−0.01	−1.17	−1.22
HFx+PBEc	0.01	0.06	0.17	0.10	−0.02	0.09	−1.06	−1.18
BLYP	−0.02	−0.05	0.02	−0.06	−0.15	0.02	−1.19	−1.24
HFx+LYPc	0.01	0.04	0.14	0.07	−0.04	0.08	−1.09	−1.22
SCAN	−0.08	−0.25	−0.26	−0.46	−0.69	−0.53	−1.76	−1.81
HFx+SCANc	−0.02	−0.09	−0.03	−0.16	−0.32	−0.21	−1.40	−1.51
aTPSS	−0.05	−0.21	−0.20	−0.34	−0.49	−0.28	−1.57	−1.63
HFx+aTPSSc	−0.03	−0.10	−0.05	−0.18	−0.36	−0.25	−1.44	−1.56

Table 5.1: Error in the magnetizability compared to CCSD(T), $\Delta\xi = \xi - \xi_{\mathrm{CCSD(T)}}$, given in atomic units ($E_h B_0^{-2}$ with $B_0 = 2.35 \times 10^5$ T). For TPSS, we work with τ_{MS} of Eq. (3.30), and we therefore refer to it as aTPSS, see Subsection 3.1.3. For those lines in the table where exact exchange is combined with the correlation part of a DFA, we have, for emphasizing, indicated an improvement of the magnetizability with a blue number and a worsening with a red number.

exact exchange. This information can be valuable for the further development of DFAs.



(a) N_2 (b) CH_4

Figure 5.1: Error in the density compared to CCSD(T), $\Delta\rho = \rho - \rho_{\mathrm{CCSD(T)}}$. For N_2, the plot is along the molecular axis and shows only one nitrogen atom, which is located at the origin. For CH_4, the plot is along the C-H bond, where the carbon atom is located at the origin. All calculations have been done with a contracted aug-cc-pVQZ basis.

Figure 5.1 shows the error in the density compared to CCSD(T) for two molecules, both for the ordinary DFA and when the correlation part is combined with 100% exact exchange. When the DFA exchange is replaced by exact exchange, for all considered functionals, there is a considerable improvement of the density in the region of the nuclei. This leads to the question whether the magnetizability, too, benefits from this replacement.

Table 5.1 lists the error in the magnetizability compared to CCSD(T), $\Delta\xi = \xi - \xi_{\mathrm{CCSD(T)}}$ for several ordinary DFAs, as well as for the case where their correlation part is combined with 100% –Fock exchange. In most cases, the use of exact exchange leads to

an improvement of the result. It works particularly well for aTPSS and SCAN. For BLYP, however, the magnetizability gets in some cases worse, for PBE even in the majority of cases. This holds even for CH_4, where Fig. 5.1 explicitly shows a major improvement of the density for all DFAs. Consequently, an improvement of the density alone does not necessarily improve the magnetizability. Using 100% exact exchange apparently leads to a better density, but affects the exchange–correlation energy too much. This is confirmed by the fact that it works better for TPSS and SCAN, which already have a better density in the first place. Relatively, the change of the density and the effect on the exchange–correlation energy is therefore smaller, such that the magnetizability can benefit from the improvement of the density to a larger degree. Moreover, it may be that SCAN and TPSS, which are newer and more elaborate functionals than PBE and BLYP, rely less on the error cancellation between exchange and correlation energy, such that a replacement of the exchange with exact exchange is less problematic.

For the development of new approximate functionals, it is of course not necessary to use 100% exact exchange and most likely, it would be beneficial to optimize the amount, as it is done for hybrid functionals. Moreover, the test set of molecules would need to be increased, as well as the number of considered DFAs. We only chose a small selection here to get some first indication.

However, this preliminary study that arose in connection with Paper II, is enough to ascertain that a higher percentage of exact exchange usually results in a better density, but does not necessarily improve the magnetizability. This holds true despite of our finding that the major error that DFAs make in the magnetizability, arises from their poor self-consistent density. It is important to maintain the optimized exchange–correlation energy. For future work, the task therefore is to find a way to improve the density of approximate functionals without breaking their refined exchange–correlation energy.

5.3 Paper III

Title: *Kohn–Sham energy decomposition for molecules in magnetic fields*
Authors: S. Reimann, J. Austad, A. Borgoo, E.I. Tellgren, A.M. Teale, T. Helgaker and S. Stopkowicz

Submitted to *Molecular Physics*

5.3.1 Main Results

As in Paper II, we decompose the total energy into its Kohn–Sham components and investigate their behaviour in a magnetic field. The aim of this paper is to extent the study to a larger set of systems than in Paper II and to include both diamagnetic and

paramagnetic molecules. We analyze to which extent different approximate functionals are able to capture the correct field dependence obtained from Lieb optimizations based on magnetic-field dependent CCSD densities.

The main conclusions of the paper are presented in the following paragraphs.

Modeling of the Kohn–Sham components by approximate functionals We show that popular approximate exchange–correlation functionals at the GGA, mGGA and hybrid levels of theory model the Kohn–Sham components qualitatively well, although they neglect all field dependence in the correlation functional. This holds especially for diamagnetic molecules and can here be attributed to the fact that the exchange–correlation energy changes only slightly in a magnetic field.

For neutral closed-shell paramagnetic molecules, we find that not only the total energy behaves oppositely to (neutral) diamagnetic molecules, but all Kohn–Sham components do. An exception is the CH^+ molecule where the density becomes more compact rather than more diffuse with increasing field strength, corresponding to a decrease of the nuclear attraction energy, and at the same time a decrease in the noninteracting kinetic energy and an increase in the Hartree energy. In contrast to the diamagnetic molecules, the exchange–correlation energy increases with increasing field strength for all paramagnetic molecules, and the change in the field is larger than for diamagnetic molecules. It is therefore not surprising that approximate functionals and the MP2 method work less satisfactorily for paramagnetic molecules than for diamagnetic ones, since larger changes in the electronic structure have to be modelled. However, the qualitative behaviour of the Kohn–Sham components is nevertheless recovered by all approximate methods.

Performance of approximate functionals Utilizing the magnetizability as a measure for the performance in a magnetic field, we found that the errors approximate functionals make are one to two orders of magnitude larger for paramagnetic than for diamagnetic molecules. This means that the DFAs struggle much more to describe the correct field dependence of paramagnetic molecules than of diamagnetic molecules. While compared to CCSD(T), the LDA functional is least reliable, we found the cTPSS functional, with τ_D of Eq. (3.33), to perform particularly well. It significantly outperforms MP2 and gives similar errors for dia- and paramagnetic molecules.

Field dependence of the exchange energy To study the magnetic-field dependence of the universal density functional without interfering effects from the density, we have considered the dependence of the BDFT density functional $F_\lambda(\rho, \mathbf{B})$ for a fixed density ρ. We have shown that for atoms, where the Hamiltonian commutes with the angular-momentum operator in the field direction, the exchange energy is unaffected by the magnetic field. This holds also for single-orbital systems like the H_2 molecule. We observed that for

LiH and H_2O, too, the exchange-energy remains unaffected within the precision of our method. For these systems with constant exchange energy, this justifies the use of $E_{\mathrm{x}}(\rho)$ from field-free DFT without further modifications, and to put all field dependence into the correlation part $E_{\mathrm{c}}(\rho, \mathbf{B})$. By contrast, aromatic systems, like benzene and pyrrole, and paramagnetic molecules show a significant field dependence. The different qualitative behaviour of the molecules implies that the explicit dependence on the magnetic field is not trivial.

5.3.2 Additional Material

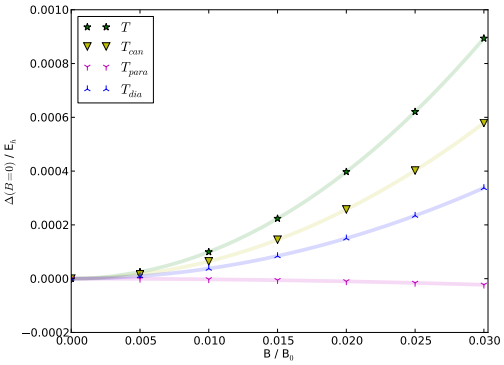
The major part of this paper was to decompose the total energy into its Kohn–Sham components and to analyze them separately. In principle, the kinetic energy can be even further subdivided, by splitting the kinetic energy operator of Eq. (2.65) into three contributions

$$\begin{aligned}\hat{T}(\mathbf{A}) &= \frac{1}{2}\sum_{i=1}^{N}\big(-\mathrm{i}\boldsymbol{\nabla}_i + \mathbf{A}(\mathbf{r}_i)\big)^2 \\ &= -\frac{1}{2}\sum_i \nabla_i^2 - \mathrm{i}\sum_{i=1}^{N}\mathbf{A}(\mathbf{r}_i)\cdot\nabla_i + \frac{1}{2}\sum_{i=1}^{N}A^2(\mathbf{r}_i). \end{aligned} \tag{5.1}$$

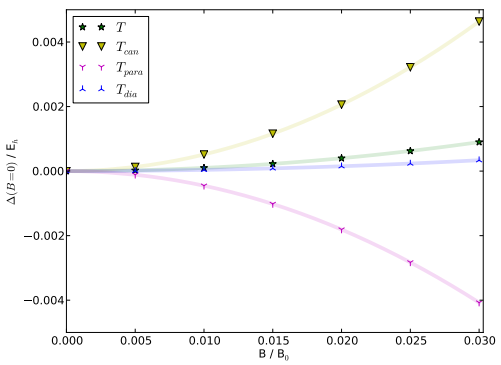
The total kinetic energy then consists of the the canonical kinetic energy $T_{\mathrm{can}} = -\frac{1}{2}\sum_i^N ||\nabla\phi_i(\mathbf{r})||^2$, the paramagnetic part $T_{\mathrm{para}} = -\mathrm{i}\sum_{i=1}^N \phi_i^\dagger(\mathbf{r})\mathbf{A}(\mathbf{r}_i)\cdot\nabla\phi_i(\mathbf{r})$ and the diamagnetic part $T_{\mathrm{dia}} = \frac{1}{2}\sum_{i=1}^N \rho(\mathbf{r}_i)A^2(\mathbf{r}_i)$. Those individual components, however, are not gauge invariant and their behaviour will be dependent on the exact location of the gauge origin. We were therefore interested whether this gauge dependence is only small or if there are significant qualitative differences in the shape of the curves.

Figure 5.2 shows for the H_2 all components of the kinetic energy as a function of the magnetic field, with different choices of the gauge origin. Both the the total kinetic energy T and T_{dia} are independent of the location of the gauge origin. However, the behaviour of the other two components, T_{can} and T_{para}, is strongly gauge-dependent. When the gauge-origin is located at the center of mass, $\mathbf{O} = (0,0,0)$, the upward curvature of T_{can} is smaller than of the total T, and the paramagnetic contribution T_{para} bends down just very slightly. Moving the origin to $\mathbf{O} = (3,3,0)$, the positive curvature of T_{can} and the negative one of T_{para} are significantly increasing, such that the plots become qualitatively different. In particular the curve of T_{can} now bends much more than the one of T. Moving the gauge origin even farther away from the center of mass, to $\mathbf{O} = (5,5,0)$, the qualitative behaviour of the curves stays as for $\mathbf{O} = (3,3,0)$, with the curvatures of T_{can} and T_{para} further increasing in absolute value.

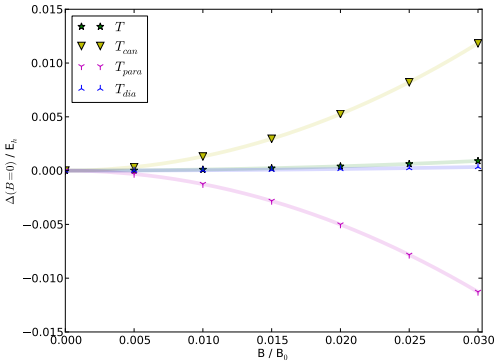
In conclusion, this shows that the behaviour of T_{can} and T_{para} depends very much on



(a) $\mathbf{O} = (0, 0, 0)$



(b) $\mathbf{O} = (3, 3, 0)$



(c) $\mathbf{O} = (5, 5, 0)$

Figure 5.2: For the H_2 molecule, all components of the kinetic energy are shown as a function of the magnetic field strength, with different locations of the gauge origin. The molecule is aligned with the z-axis, with the center of mass at (0,0,0), and the magnetic field is directed in x-direction. All calculations have been done at the FCI/aug-cc-pVTZ level.

the location of the gauge origin. It is valuable to know that the degree of this dependence is quite large, such that it generally makes little sense to study those components individually. The total kinetic energy T and T_{dia}, however, are both unaffected by the choice of origin.

Chapter 6

Summary and Outlook

6.1 Summary

In this work, we have investigated the importance to introduce a field dependence into approximate exchange–correlation functionals to improve the computation of magnetic properties. We have concluded that the missing field dependence is a leading source of error in NMR shielding calculations using DFAs.

For the magnetizability, however, we have found that the main error is due to the poor self-consistent density that present DFAs give. Therefore, for this property, the benefit of improving the density of existing DFAs will be larger than the benefit of introducing a magnetic-field dependence into the correlation functional. Since the main error that is made in the density in the presence of a magnetic field is already present at zero field strength, it is important to improve DFAs even in the absence of a field. Only then can we fully benefit of field-dependent corrections. Present DFAs struggle particularly to describe the magnetizability and its components for paramagnetic systems, where the changes of the electronic structure in the magnetic field are more dramatic than for diamagnetic systems. However, we have also shown that once a good density has been attained, highly accurate results can only be attained by an inclusion of magnetic-field effects into the approximate functionals and that a good density alone does not ensure the desired accuracy. This serves as a strong motivation for the development of BDFT, since until now, there do not exist any correlation functionals with an explicit field dependence.

We have presented BDFT in the context of Lieb's convex-conjugate theory, put it into a common framework with CDFT and compared both approaches. We have shown that BDFT is a promising alternative to CDFT, and the results of our calculations strongly suggest that an inclusion of magnetic field effects in the context of BDFT will allow a more accurate computation of magnetic properties. Presently, meta-GGA functionals that have been generalized to ensure independence of the gauge origin, are most suited for calculations in magnetic fields.

It is important to note that it is challenging to distinguish effects caused by an improvement of the density from effects due to an inclusion of the magnetic field into the exchange–correlation functional. Both effects are mutually dependent, since the inclusion of the magnetic field into the density functional also affects the self-consistent density obtained with this functional. We have shown that it is crucial to work on both aspects, since the neglect of one of them makes it impossible to fully optimize the other one.

6.2 Future Work

While this work provides many of the necessary preliminaries for the implementation of field-dependent functionals, the next step is to incorporate this information into actual exchange–correlation functionals. Having shown that a major error of present DFAs arises from their poor self-consistent density, we conclude that if the field-dependence is to be introduced as a correction to an already existing DFA, it is necessary to use one with a comparatively accurate density, as for example one of the gauge-invariant generalizations of TPSS. Alternatively, or even better, it would be to first find ways to improve the density of existing approximate functionals. Only this way the improvements due to a field-dependence will not be obscured by much larger errors introduced by poor densities.

Our work also points out that highly accurate results can only be obtained if modifications related to improved densities and field-dependencies do not break the otherwise optimized exchange–correlation energy of existing DFAs. This serves as an important reminder for future developers to take into account the limitations of current DFAs when improving upon them, since changes related to field-dependencies might impair otherwise well-balanced parameters that present approximate functionals rely on.

A central topic of this work have been Lieb optimizations at different interaction strengths, which we used for all three of our papers. While in theory, Lieb optimizations provide easily useful information about the Lieb functional, we found that in practice, the optimization is often difficult. Even with advanced optimisation algorithms and making use of different regularization parameters, the minimizing potential cannot always be found. For most molecules, if the optimization is performed at equilibrium density, we encounter no problems and the optimization converges smoothly within a few steps. However, if the Lieb optimization is performed at $\lambda > 0$ with a reference density that does not correspond to the equilibrium of the molecular system, for example utilizing a $(B = 0)$-density for a Hamiltonian with $B \neq 0$, the optimization often fails, which means that no potential can be found that minimizes the Lieb functional. Why exactly the optimizing potential is so difficult to find in those cases, is presently not clear to us and should be further investigated.

Bibliography

[1] D. T. Wickramasinghe and L. Ferrario, "Magnetism in Isolated and Binary White Dwarfs," *PASP*, vol. 112, no. 773, pp. 873–924, 2000.

[2] D. J. Price and S. Rosswog, "Producing Ultrastrong Magnetic Fields in Neutron Star Mergers," *Science*, vol. 312, no. 5774, pp. 719–722, 2006.

[3] K. N. Gourgouliatos, T. S. Wood, and R. Hollerbach, "Magnetic field evolution in magnetar crusts through three-dimensional simulations," *PNAS*, vol. 113, no. 15, pp. 3944–3949, 2016.

[4] T. Helgaker, M. Jaszuński, and K. Ruud, "Ab Initio Methods for the Calculation of NMR Shielding and Indirect Spin-Spin Coupling Constants," *Chem. Rev.*, vol. 99, no. 1, pp. 293–352, 1999.

[5] F. Neese, "Prediction of molecular properties and molecular spectroscopy with density functional theory: From fundamental theory to exchange-coupling," *Coord. Chem. Rev.*, vol. 253, no. 5, pp. 526 – 563, 2009.

[6] J. Kussmann and C. Ochsenfeld, "Linear-scaling method for calculating nuclear magnetic resonance chemical shifts using gauge-including atomic orbitals within Hartree-fock and density-functional theory," *J. Chem. Phys.*, vol. 127, no. 5, p. 054103, 2007.

[7] J. Kussmann and C. Ochsenfeld, "A density matrix-based method for the linear-scaling calculation of dynamic second- and third-order properties at the Hartree-fock and Kohn-Sham density functional theory levels," *J. Chem. Phys.*, vol. 127, no. 20, p. 204103, 2007.

[8] A. M. N. Niklasson and V. Weber, "Linear scaling density matrix perturbation theory for basis-set-dependent quantum response calculations: An orthogonal formulation," *J. Chem. Phys.*, vol. 127, no. 6, p. 064105, 2007.

[9] J. M. Martin, J. El-Yazal, and J.-P. François, "Basis set convergence and performance of density functional theory including exact exchange contributions for geometries and harmonic frequencies," *Mol. Phys.*, vol. 86, no. 6, pp. 1437–1450, 1995.

[10] O. B. Lutnæs, A. M. Teale, T. Helgaker, D. J. Tozer, K. Ruud, and J. Gauss, "Benchmarking density-functional-theory calculations of rotational g tensors and magnetizabilities using accurate coupled-cluster calculations," *J. Chem. Phys.*, vol. 131, p. 144104, 2009.

[11] A. M. Teale, O. B. Lutnæs, T. Helgaker, D. J. Tozer, and J. Gauss, "Benchmarking density-functional theory calculations of NMR shielding constants and spin–rotation constants using accurate coupled-cluster calculations," *J. Chem. Phys.*, vol. 138, no. 2, p. 024111, 2013.

[12] D. Flaig, M. Maurer, M. Hanni, K. Braunger, L. Kick, M. Thubauville, and O. C., "Benchmarking Hydrogen and Carbon NMR Chemical Shifts at HF, DFT, and MP2 Levels," *J. Chem. Theory Comput.*, vol. 10, pp. 572–578, 2014.

[13] J. W. Furness, J. Verbeke, E. I. Tellgren, S. Stopkowicz, U. Ekström, T. Helgaker, and A. M. Teale, "Current Density Functional Theory Using Meta-Generalized Gradient Exchange-Correlation Functionals," *J. Chem. Theory Comput.*, vol. 11, no. 9, pp. 4169–4181, 2015.

[14] W. Zhu, L. Zhang, and S. B. Trickey, "Comparative studies of density-functional approximations for light atoms in strong magnetic fields," *Phys. Rev. A*, vol. 90, p. 022504, Aug 2014.

[15] C. J. Grayce and R. A. Harris, "Magnetic-field density-functional theory," *Phys. Rev. A*, vol. 50, pp. 3089–3095, 1994.

[16] P. Hohenberg and W. Kohn, "Inhomogeneous Electron Gas," *Phys. Rev.*, vol. 136, pp. B864–B871, Nov 1964.

[17] R. Parr and W. Yang, *Density-Functional Theory of Atoms and Molecules*. International Series of Monographs on Chemistry, Oxford University Press, USA, 1994.

[18] T. Helgaker, P. Jørgensen, and J. Olsen, *Density-functional theory*. Book in preparation.

[19] M. Levy, "Electron densities in search of Hamiltonians," *Phys. Rev. A*, vol. 26, pp. 1200–1208, Sep 1982.

[20] E. H. Lieb, "Density functionals for coulomb systems," *Int. J. Quantum Chem.*, vol. 24, no. 3, pp. 243–277, 1983.

[21] M. Levy, "Universal variational functionals of electron densities, first-order density matrices, and natural spin-orbitals and solution of the v-representability problem," *PNAS*, vol. 76, no. 12, pp. 6062–6065, 1979.

[22] J. van Tiel, *Convex Analysis: An Introductory Text.* Wiley, 1984.

[23] E. I. Tellgren, S. Kvaal, E. Sagvolden, U. Ekström, A. M. Teale, and T. Helgaker, "Choice of basic variables in current-density-functional theory," *Phys. Rev. A*, vol. 86, p. 062506, Dec 2012.

[24] W. Yang, "Generalized adiabatic connection in density functional theory," *J. Chem. Phys.*, vol. 109, no. 23, pp. 10107–10110, 1998.

[25] A. D. Becke, "A new mixing of Hartree–fock and local density-functional theories," *J. Chem. Phys.*, vol. 98, no. 2, pp. 1372–1377, 1993.

[26] A. M. Teale, S. Coriani, and T. Helgaker, "Accurate calculation and modeling of the adiabatic connection in density functional theory," *The Journal of Chemical Physics*, vol. 132, no. 16, p. 164115, 2010.

[27] A. M. Teale, S. Coriani, and T. Helgaker, "The calculation of adiabatic-connection curves from full configuration-interaction densities: Two-electron systems," *J. Chem. Phys.*, vol. 130, no. 10, p. 104111, 2009.

[28] M. J. G. Peach, A. M. Miller, A. M. Teale, and D. J. Tozer, "Adiabatic connection forms in density functional theory: H2 and the He isoelectronic series," *J. Chem. Phys.*, vol. 129, no. 6, p. 064105, 2008.

[29] F. Jensen, *Introduction to Computational Chemistry.* Wiley, 2007.

[30] W. Kohn and L. J. Sham, "Self-Consistent Equations Including Exchange and Correlation Effects," *Phys. Rev.*, vol. 140, pp. 1133–1138, Nov 1965.

[31] S. S. Iyengar, M. Ernzerhof, S. N. Maximoff, and G. E. Scuseria, "Challenge of creating accurate and effective kinetic-energy functionals," *Phys. Rev. A*, vol. 63, p. 052508, Apr 2001.

[32] B. O. Roos, T. Helgaker, P. R. Taylor, and D. J. Tozer, *European Summerschool in Quantum Chemistry 2013, Book II.* 8 ed., 2013.

[33] G. Vignale and M. Rasolt, "Density-functional theory in strong magnetic fields," *Phys. Rev. Lett.*, vol. 59, pp. 2360–2363, 1987.

[34] G. Vignale and M. Rasolt, "Current- and spin-density-functional theory for inhomogeneous electronic systems in strong magnetic fields," *Phys. Rev. B*, vol. 37, pp. 10685–10696, 1988.

[35] X.-Y. Pan and V. Sahni, "Density and physical current density functional theory," *Int. J. Quantum Chem.*, vol. 110, no. 15, pp. 2833–2843, 2010.

[36] X.-Y. Pan and V. Sahni, "Generalization of the Hohenberg–Kohn theorem to the presence of a magnetostatic field," *J. Phys. Chem. Solids*, vol. 73, no. 5, pp. 630 – 634, 2012.

[37] E. I. Tellgren, A. M. Teale, J. W. Furness, K. K. Lange, U. Ekström, and T. Helgaker, "Non-perturbative calculation of molecular magnetic properties within current-density functional theory," *J. Chem. Phys.*, vol. 140, no. 3, p. 034101, 2014.

[38] E. H. Lieb and R. Schrader, "Current densities in density-functional theory," *Phys. Rev. A*, vol. 88, p. 032516, Sep 2013.

[39] G. Vignale, M. Rasolt, and D. J. W. Geldart, "Diamagnetic susceptibility of a dense electron gas," *Phys. Rev. B*, vol. 37, pp. 2502–2507, Feb 1988.

[40] K. Higuchi, M. Miyasita, M. Kodera, and M. Higuchi, "New practicable forms for exchange and correlation energy functionals of the cdft," *J. Magn. Magn. Mater.*, vol. 310, no. 2, pp. 1065 – 1066, 2007.

[41] R. T. Rockafellar, "A general correspondence between dual minimax problems and convex programs," *Pacific J. Math.*, vol. 25, pp. 597–611, 1968.

[42] V. Barbu and T. Precupano, *Convexity and Optimization in Banach Spaces*. Springer, 4 ed., 2014.

[43] J. P. Perdew and K. Schmidt, "Jacob's ladder of density functional approximations for the exchange-correlation energy," *AIP Conf. Proc.*, vol. 577, no. 1, pp. 1–20, 2001.

[44] T. Tsuneda and K. Hirao, "Self-interaction corrections in density functional theory," *J. Chem. Phys.*, vol. 140, no. 18, p. 18A513, 2014.

[45] D. M. Ceperley and B. J. Alder, "Ground State of the Electron Gas by a Stochastic Method," *Phys. Rev. Lett.*, vol. 45, pp. 566–569, Aug 1980.

[46] S. H. Vosko, L. Wilk, and M. Nusair, "Accurate spin-dependent electron liquid correlation energies for local spin density calculations: a critical analysis," *Can. J. Phys.*, vol. 58, no. 8, pp. 1200–1211, 1980.

[47] C. Ullrich, *Time-Dependent Density-Functional Theory: Concepts and Applications*. Oxford Graduate Texts, OUP Oxford, 2012.

[48] T. W. Keal and D. J. Tozer, "The exchange-correlation potential in Kohn-sham nuclear magnetic resonance shielding calculations," *J. Chem. Phys.*, vol. 119, no. 6, pp. 3015–3024, 2003.

[49] A. D. Becke, "Density-functional exchange-energy approximation with correct asymptotic behavior," *Phys. Rev. A*, vol. 38, pp. 3098–3100, Sep 1988.

[50] C. Lee, W. Yang, and R. G. Parr, "Development of the Colle-salvetti correlation-energy formula into a functional of the electron density," *Phys. Rev. B*, vol. 37, pp. 785–789, Jan 1988.

[51] R. Colle and O. Salvetti, "Approximate calculation of the correlation energy for the closed shells," *Theor. chim. acta*, vol. 37, no. 4, pp. 329–334, 1975.

[52] J. P. Perdew, K. Burke, and M. Ernzerhof, "Generalized Gradient Approximation Made Simple," *Phys. Rev. Lett.*, vol. 77, pp. 3865–3868, Oct 1996.

[53] W. Yang and Q. Wu, "Direct Method for Optimized Effective Potentials in Density-Functional Theory," *Phys. Rev. Lett.*, vol. 89, p. 143002, Sep 2002.

[54] S. Kümmel and L. Kronik, "Orbital-dependent density functionals: Theory and applications," *Rev. Mod. Phys.*, vol. 80, pp. 3–60, Jan 2008.

[55] F. Zahariev, S. S. Leang, and M. S. Gordon, "Functional derivatives of meta-generalized gradient approximation (meta-GGA) type exchange-correlation density functionals," *J. Chem. Phys.*, vol. 138, no. 24, p. 244108, 2013.

[56] J. Tao, J. P. Perdew, V. N. Staroverov, and G. E. Scuseria, "Climbing the Density Functional Ladder: Nonempirical Meta–Generalized Gradient Approximation Designed for Molecules and Solids," *Phys. Rev. Lett.*, vol. 91, p. 146401, Sep 2003.

[57] J. P. Perdew, J. Tao, V. N. Staroverov, and G. E. Scuseria, "Meta-generalized gradient approximation: Explanation of a realistic nonempirical density functional," *J. Chem. Phys.*, vol. 120, no. 15, pp. 6898–6911, 2004.

[58] S. N. Maximoff and G. E. Scuseria, "Nuclear magnetic resonance shielding tensors calculated with kinetic energy density-dependent exchange-correlation functionals," *Chem. Phys. Lett.*, vol. 390, no. 4-6, pp. 408–412, 2004.

[59] J. F. Dobson, "Alternative expressions for the Fermi hole curvature," *J. Chem. Phys.*, vol. 98, no. 11, pp. 8870–8872, 1993.

[60] A. D. Becke *J. Chem. Phys.*, vol. 117, p. 6935, 2002.

[61] J. Tao, "Explicit inclusion of paramagnetic current density in the exchange-correlation functionals of current-density functional theory," *Phys. Rev. B*, vol. 71, p. 205107, May 2005.

[62] E. Sagvolden, U. Ekström, and E. I. Tellgren, "Isoorbital indicators for current density functional theory," *Mol. Phys.*, vol. 111, no. 9-11, pp. 1295–1302, 2013.

[63] J. Sun, A. Ruzsinszky, and J. P. Perdew, "Strongly Constrained and Appropriately Normed Semilocal Density Functional," *Phys. Rev. Lett.*, vol. 115, p. 036402, Jul 2015.

[64] A. D. Becke, "Density-functional thermochemistry. V. Systematic optimization of exchange-correlation functionals," *J. Chem. Phys.*, vol. 107, no. 20, pp. 8554–8560, 1997.

[65] P. J. Stephens, F. J. Devlin, C. F. Chabalowski, and M. J. Frisch, "Ab Initio Calculation of Vibrational Absorption and Circular Dichroism Spectra Using Density Functional Force Fields," *J. Phys. Chem.*, vol. 98, no. 45, pp. 11623–11627, 1994.

[66] A. D. Becke, "Density-functional thermochemistry. iii. the role of exact exchange," *J. Chem. Phys.*, vol. 98, no. 7, pp. 5648–5652, 1993.

[67] S. F. Sousa, P. A. Fernandes, and M. J. a. Ramos, "General Performance of Density Functionals," *J. Phys. Chem. A*, vol. 111, no. 42, pp. 10439–10452, 2007. PMID: 17718548.

[68] M. Levy and J. P. Perdew, "Hellmann-feynman, virial, and scaling requisites for the exact universal density functionals. Shape of the correlation potential and diamagnetic susceptibility for atoms," *Phys. Rev. A*, vol. 32, pp. 2010–2021, 1985.

[69] F. London, "Théorie quantique des courants interatomiques dans les combinaisons aromatiques," *J. Phys. Radium*, vol. 8, no. 10, pp. 397–409, 1937.

[70] E. I. Tellgren, T. Helgaker, and A. Soncini, "Non-perturbative magnetic phenomena in closed-shell paramagnetic molecules," *Phys. Chem. Chem. Phys.*, vol. 11, pp. 5489–5498, 2009.

[71] K. Ruud and T. Helgaker, "The magnetizability, rotational g tensor, and quadrupole moment of PF3 revisited," *Chem. Phys. Lett.*, vol. 264, no. 1, pp. 17 – 23, 1997.

[72] A. M. Lee, N. C. Handy, and S. M. Colwell, "The density functional calculation of nuclear shielding constants using London atomic orbitals," *J. Chem. Phys.*, vol. 103, no. 23, pp. 10095–10109, 1995.

[73] N. F. Ramsey, "Magnetic Shielding of Nuclei in Molecules," *Phys. Rev.*, vol. 78, pp. 699–703, Jun 1950.

[74] T. Helgaker, P. J. Wilson, R. D. Amos, and N. C. Handy, "Nuclear shielding constants by density functional theory with gauge including atomic orbitals," *J. Chem. Phys.*, vol. 113, no. 8, pp. 2983–2989, 2000.

[75] P. Atkins and R. Friedman, *Molecular Quantum Mechanics*. OUP Oxford, 2011.

[76] T. Helgaker, "Time-Independent Molecular Properties." http://folk.uio.no/helgaker/talks/SostrupTI_14.pdf, 2014.

[77] Q. Wu and W. T. Yang, "A direct optimization method for calculating density functionals and exchange-correlation potentials from electron densities," *J. Chem. Phys.*, vol. 118, no. 6, pp. 2498–2509, 2003.

[78] E. Fermi and E. Amaldi, "Not known," *Mem. R. Acad. Italia*, vol. 6, no. 117, 1934.

[79] Q. Zhao, R. C. Morrison, and R. G. Parr, "From electron densities to Kohn-Sham kinetic energies, orbital energies, exchange-correlation potentials, and exchange-correlation energies," *Phys. Rev. A*, vol. 50, pp. 2138–2142, Sep 1994.

[80] E. I. Tellgren, A. Soncini, and T. Helgaker, "Nonperturbative ab initio calculations in strong magnetic fields using London orbitals," *J. Chem. Phys.*, vol. 129, no. 15, p. 154114, 2008.

[81] "LONDON, an ab-initio program package for calculations in finite magnetic fields, founded by E. I. Tellgren, A. Soncini and T. Helgaker. Programming by E. I. Tellgren (main programmer), K. K. Lange, A. M. Teale, U. E. Ekström, S. Stopkowicz and J. H. Austad.," 2016.

[82] T. Heaton-Burgess, F. A. Bulat, and W. Yang, "Optimized Effective Potentials in Finite Basis Sets," *Phys. Rev. Lett.*, vol. 98, p. 256401, Jun 2007.

[83] T. Heaton-Burgess and W. Yang, "Optimized effective potentials from arbitrary basis sets," *J. Chem. Phys.*, vol. 129, no. 19, p. 194102, 2008.

[84] F. A. Bulat, T. Heaton-Burgess, A. J. Cohen, and W. Yang, "Optimized effective potentials from electron densities in finite basis sets," *J. Chem. Phys.*, vol. 127, no. 17, p. 174101, 2007.

[85] S. Kvaal, U. Ekström, A. M. Teale, and T. Helgaker, "Differentiable but exact formulation of densiy-functional theory," *J. Chem. Phys.*, vol. 140, p. 18A518, 2014.

Appendix A

Proof of the Hohenberg–Kohn Theorem

The Hohenberg–Kohn theorem states that each N-electron density ρ is the ground state density of at most one external potential $v(\mathbf{r}) + c$, which is, up to an additive constant c, uniquely determined.

In other words, two different potentials cannot give rise to the same density. The proof is remarkably simple. Let us consider two different N-electron systems with potentials

$$v_1(\mathbf{r}) \neq v_2(\mathbf{r}) + c. \tag{A.1}$$

From Eq. (A.1), it follows that the associated Hamiltonians, $\hat{H}(v_1)$ and $\hat{H}(v_2)$, have two different ground state wave functions

$$\Psi_1 \neq \gamma\Psi_2. \tag{A.2}$$

Applying the Rayleigh-Ritz variation principle for the two ground states, we get

$$\begin{aligned} E(v_1) < \langle\Psi_2|\hat{H}_1|\Psi_2\rangle &= \langle\Psi_2|\hat{H}(v_2)|\Psi_2\rangle + \langle\Psi_2|\hat{H}(v_1) - \hat{H}(v_2)|\Psi_2\rangle \\ &= E(v_2) + (v_1 - v_2|\rho_2) \qquad \text{(A.3)} \\ E(v_2) < \langle\Psi_1|\hat{H}_2|\Psi_1\rangle &= \langle\Psi_1|\hat{H}(v_1)|\Psi_1\rangle - \langle\Psi_1|\hat{H}(v_1) - \hat{H}(v_2)|\Psi_1\rangle \\ &= E(v_1) - (v_1 - v_2|\rho_1). \qquad \text{(A.4)} \end{aligned}$$

Adding both inequalities leads to the strict inequality

$$E(v_1) + E(v_2) < E(v_1) + E(v_2) + (v_1 - v_2|\rho_2 - \rho_1). \tag{A.5}$$

If both potentials v_1 and v_2 gave rise to the same density, which means that $\rho_1 = \rho_2$, then Eq. (A.6) would simplify to

$$E(v_1) + E(v_2) < E(v_1) + E(v_2), \tag{A.6}$$

which is a contradiction. As a result, if two potentials v_1 and v_2 differ by more than an additive constant, then the associated densities ρ_1 and ρ_2 cannot be equal.

Appendix B

Derivation of the Diamagnetic Part of the Magnetizability for Closed-Shell Atoms

In the following, we will derive Eq. (4.13),

$$\xi^{\text{dia}} = -\frac{1}{6}\left\langle r^2 \right\rangle, \tag{B.1}$$

which is the diamagnetic part of the magnetizability for closed-shell atoms in a homogeneous magnetic field. From the Hamiltonian of Eq. (2.66), only the kinetic part depends on the magnetic field, such that we can write

$$\xi^{\text{dia}} = -\frac{1}{2}\,\frac{\mathrm{d}^2}{\mathrm{d}\mathbf{B}^2}(\rho|A^2)\bigg|_{\mathbf{B}=0} \tag{B.2}$$

With $\mathbf{A} = \frac{1}{2}\mathbf{B}\times\mathbf{r}$, we have that

$$A^2 = \frac{1}{4}|\mathbf{B}\times\mathbf{r}|^2 = \frac{1}{4}\mathbf{B}^T(r^2\mathbf{I} - \mathbf{r}\mathbf{r}^T)\mathbf{B}. \tag{B.3}$$

For a homogeneous magnetic field, $\mathbf{B} = (0, 0, B_z)$, this implies

$$A^2 = \frac{1}{4}B_z^2(x^2 + y^2). \tag{B.4}$$

Equation (B.2) thus becomes Using the argument that, due to symmetry, the magnetic field does not enter the density to first order, $\mathrm{d}\rho/dB_z = 0$, only two terms remain after the second derivation,

$$\xi^{\text{dia}} = -\frac{1}{8}\left\{\int \frac{\rho(\mathbf{r})}{\mathrm{d}B_z^2}B_z^2(x^2+y^2)\,\mathrm{d}\mathbf{r}\bigg|_{B_z=0} + 2\int \rho(\mathbf{r})(x^2+y^2)\,\mathrm{d}\mathbf{r}\bigg|_{B_z=0}\right\}. \tag{B.5}$$

From those two terms the first one vanishes due to the condition that the derivative is taken at $B_z = 0$. Hence

$$\xi^{\text{dia}} = -\frac{1}{4}\int \rho(\mathbf{r})(x^2 + y^2)\ \mathrm{d}\mathbf{r} \tag{B.6}$$

$$= -\frac{1}{6}\int \rho(\mathbf{r})r^2\ \mathrm{d}\mathbf{r}, \tag{B.7}$$

$$= -\frac{1}{6}\langle r^2 \rangle \tag{B.8}$$

where we have used that $\langle x^2 + y^2 \rangle = \frac{2}{3}\langle r^2 \rangle$.

Paper I

The importance of current contributions to shielding constants in density-functional theory.

S. Reimann, U. Ekström, S. Stopkowicz, A. M. Teale, A. Borgoo and T. Helgaker

Physical Chemistry Chemical Physics, Volume 17, Pages 18834-18842, 2015.

Paper II

Magnetic-Field Density-Functional Theory (BDFT): Lessons from the Adiabatic Connection

S. Reimann, A. Borgoo, E. I. Tellgren, A. M. Teale, A. Borgoo and T. Helgaker

Journal of Chemical Theory and Computation, Volume 13, Pages 4089-4100, 2017.

Paper III

Kohn–Sham energy decomposition for molecules in magnetic fields.
S. Reimann, J. Austad, A. Borgoo, E. I. Tellgren, A. M. Teale, T. Helgaker and S. Stopkowicz
Submitted to: *Molecular Physics*