

Kohn–Sham theory with paramagnetic currents: compatibility and functional differentiability

Andre Laestadius,^{*,†} Erik I. Tellgren,^{*,†} Markus Penz,[‡] Michael Ruggenthaler,[‡]
 Simen Kvaal,[†] and Trygve Helgaker[†]

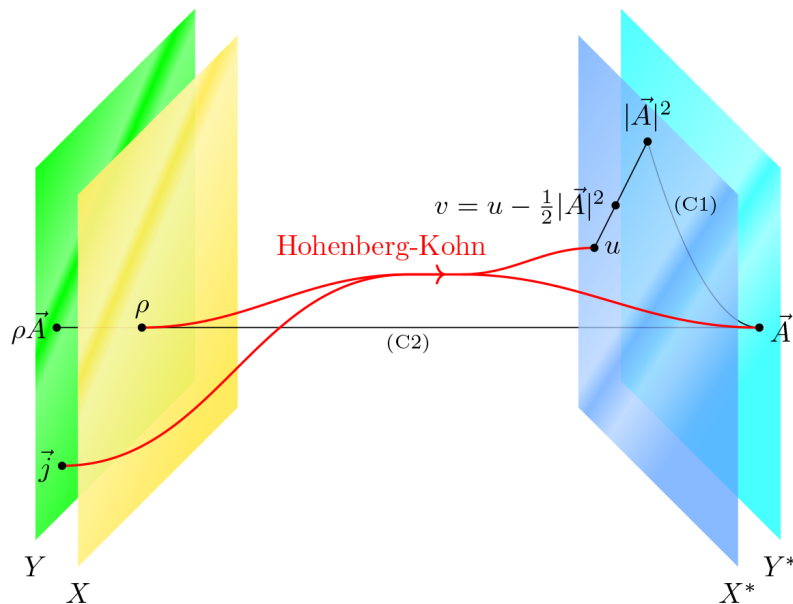
[†] *Hylleraas Centre for Quantum Molecular Sciences, University of Oslo, Norway*

[‡] *Max Planck Institute for the Structure and Dynamics of Matter, Hamburg, Germany*

E-mail: andre.laestadius@kjemi.uio.no; erik.tellgren@kjemi.uio.no

Abstract

Recent work has established Moreau–Yosida regularization as a mathematical tool to achieve rigorous functional differentiability in density-functional theory. In this article, we extend this tool to paramagnetic current-density-functional theory, the most common density-functional framework for magnetic field effects. The extension includes a well-defined Kohn–Sham iteration scheme with a partial convergence result. To this end, we rely on a formulation of Moreau–Yosida regularization for reflexive and strictly convex function spaces. The optimal L^p -characterization of the paramagnetic current density $L^1 \cap L^{3/2}$ is derived from the N -representability conditions. A crucial prerequisite for the convex formulation of paramagnetic current-density-functional theory, termed *compatibility* between function spaces for the particle density and the current density, is pointed out and analyzed. Several results about compatible function spaces are given, including their recursive construction. The regularized, exact functionals are calculated numerically for a Kohn–Sham iteration on a quantum ring, illustrating their performance for different regularization parameters.



1 Introduction

The theoretical foundation of density-functional theory (DFT) was established in a seminal paper by Hohenberg and Kohn.¹ There it was proven that two potentials that differ by more than a constant cannot share the same ground-state particle density ρ (see Eq. (3) for definition). This fact is referred to as the Hohenberg–Kohn (HK) theorem. Using this result, the Schrödinger equation was replaced by a minimization problem involving a universal density functional (HK variational principle). The work by Lieb² provided an abstract reformulation of DFT that eliminates some technical difficulties with the HK formulation and constitutes a more tractable framework for rigorous mathematical analysis. Lieb’s formulation relies on Legendre–Fenchel transformations between the ground-state energy and a universal density functional, analogous to the use of Legendre transformations in thermodynamics and classical mechanics. The HK theorem becomes recast into a fact about subgradients of convex functionals that are mapped one-to-one by Legendre–Fenchel transformations.^{3,4}

As far as practical purposes are concerned, DFT was first converted into a feasible algorithm for electronic structure calculations by Kohn and Sham.⁵ Here, both the unknown density of the full system and the effective Kohn–Sham (KS) potential for the non-interacting system are solved for in an iterative manner. Even though the important question of convergence of this procedure has been addressed in several works,^{6–10} it has only very recently been answered positively for finite-dimensional settings.¹¹

The motivation to include current densities and not just the particle density is to obtain a universal functional modelling the internal energy of magnetic systems. In terms of Lieb’s Legendre–Fenchel description, the current couples to the vector potential that now also enters the theory to account for the magnetic field. Recent work in current-density-functional theory (CDFT) has been devoted to the extension of the HK theorem, the HK variational principle, and the KS iteration scheme to include current densities,^{12–14} as well as to highlight the complexity of such a generalization.^{3,15,16} Other approaches are feasible as well, e.g., the magnetic-field density-functional theory (BDFT) of Grace and Harris,¹⁷ where a semi-universal functional is employed instead. There exists also a convexified formulation, in which BDFT and paramagnetic CDFT are related to each other by partial Legendre–Fenchel transformations.^{18,19} Furthermore, the physically important case of linear vector potentials (uniform magnetic fields) has been theoretically studied in linear-vector-potential density-functional theory (LDFT) without the need to include current densities.¹⁸ Works beyond the current density generalization exist too, e.g., spin-current density-functional theory, reduced-density-matrix-functional theory, Maxwell–Schrödinger density-functional theory (MDFT), and

quantum-electrodynamical density-functional theory (QEDFT).^{20–24} For more generalized density-functional theories see Ref. 25 and the references therein, e.g., the kinetic-density-functional theory of Sim *et al.*,²⁶ the work of Ayers²⁷ on k -density-functional theory, and Higuchi-Higuchi²⁸ who explored the use of different physical quantities as variables of the theory.

For mathematical reasons, CDFT is formulated in terms of the paramagnetic current density \vec{j} (see Eq. (4) for definition) rather than with the gauge-invariant total current density. A theoretical foundation in the sense of a HK theorem for the total current density has not yet been proven and its existence remains an open question in the general case.^{3,15} However, even if such a result could be shown, a HK variational principle does not exist for the total current density.¹⁶ Circumventing these problems may require the Maxwell–Schrödinger variational principle in place of the standard one.²³ For the CDFT that makes use of the paramagnetic current density, it is well-known that there are counterexamples that rule out any analogue of the HK theorem.^{3,14,15} Nevertheless, since the particle density and the paramagnetic current density determine the non-degenerate ground state (see Ref. 29 for results in the degenerate case), a universal Levy–Lieb^{2,30} constrained-search functional can be set up, as done by Vignale and Rasolt.¹² This functional can be extended to a Lieb functional that in this case also depends on the paramagnetic current density (for a first attempt see Ref. 31 with the choice of domain $(\rho, \vec{j}) \in (L^1 \cap L^3) \times \vec{L}^1$).

Since the Lieb functional within standard DFT suffers from non-differentiability,³² a property that CDFT inherits, we address this particular problem and formulate a regularized theory in a Banach space setting. We here apply our recent work⁹ that also extends the mathematical formalism of paramagnetic CDFT in Ref. 31. The need for differentiability—a fact that is usually overlooked in textbooks—is connected to the variational derivation and analysis of the Kohn–Sham scheme. This task, in the setting of paramagnetic CDFT, is the main aim of this work.

To set up a rigorous CDFT including the corresponding Kohn–Sham scheme, which is borrowed from our previous work⁹ and here baptized Moreau–Yosida–Kohn–Sham optimal damping algorithm (MYKSODA), we introduce and discuss the condition of *compatibility* between function spaces for the scalar and vector potentials on the one hand and for the paramagnetic current and the total physical current densities on the other. This condition is necessary both for the convex formulation of CDFT and the subsequent Moreau–Yosida regularization procedure. Moreover, to maintain compatibility the regularization procedure requires a Banach space formulation and we make use of our results employing *reflexive* Banach spaces.⁹ In this respect the approach presented here differs from that in standard DFT where a Hilbert space formulation has been previously considered,⁴ which does not allow the necessary compatibility in CDFT. However, to apply the Banach

space formulation outlined in Ref. 9, a suitable function space for the paramagnetic current density first needs to be identified. The choice \vec{L}^1 from Ref. 31 cannot be used for this purpose since it is not reflexive. It is therefore crucial to first prove that the paramagnetic current density is an element of L^p for some $1 < p < +\infty$. We prove in Corollary 2 that each component of \vec{j} is an element of $L^{3/2}$ under the assumption of finite kinetic energy.

This article is structured as follows. After introducing the basic quantum-mechanical model for paramagnetic CDFT in Sec. 2.1, we define suitable function spaces for particle and current densities in Sec. 2.2. In such a setting, the usual constrained-search functionals of DFT are defined and the energy functional and generalized Lieb functional are subsequently set up in Sec. 2.3. These functionals serve as the primary objects for a further study of the theory in terms of convex analysis. Here a first problem arises within CDFT: the lack of concavity of the energy functional. As a further ingredient of a well-defined Kohn–Sham iteration scheme, finiteness of the energy functional (and its concave version) is proven in Sec. 2.5. Like the authors recently showed,⁹ the variational Kohn–Sham construction can only be rigorously set up for a regularized theory. The respective form of Moreau–Yosida regularization is introduced and applied to the setting at hand in Sec. 3.1. Finally, the stage is set for a discussion of the Kohn–Sham iteration scheme in Sec. 3.2 and its precise formulation as MYKSODA in Sec. 3.3. We note possible convergence issues in the particular setting of a two-particle singlet state in Sec. 3.4. We conclude in Sec. 4 with a numerical study of the MYKSODA. For readers that are less familiar with Lebesgue spaces and functional analysis in general we recommend Refs. 33 and 34, while for the tools borrowed from convex analysis we point to Ref. 35.

2 Paramagnetic CDFT

2.1 Ground-state model

In what follows, we consider the Hamiltonian of an N -electron system to be specified by an external scalar potential $v : \mathbb{R}^3 \rightarrow \mathbb{R}$ and an external vector potential $\vec{A} : \mathbb{R}^3 \rightarrow \mathbb{R}^3$. The components of \vec{A} and other vectors are denoted A^k and are not to be confused with the Euclidean norm squared, $|\vec{A}|^2 = \vec{A} \cdot \vec{A}$. The physical kinetic energy operator and electron-electron repulsion are given by (in SI based atomic units),

$$T(\vec{A}) = \sum_{k=1}^N \frac{1}{2} \left(-i\nabla_{\vec{r}_k} + \vec{A}(\vec{r}_k) \right)^2, \\ W = \sum_{k=2}^N \sum_{l=1}^{k-1} \frac{1}{|\vec{r}_k - \vec{r}_l|}.$$

The full Hamiltonian then reads

$$H^\lambda(v, \vec{A}) = T(\vec{A}) + \lambda W + \sum_{k=1}^N v(\vec{r}_k), \quad (1)$$

where a scale factor $\lambda \geq 0$ is included in front of the electron–electron repulsion term. This means H^0 is the Hamiltonian of a non-interacting system while the usual interacting system is retrieved with H^1 . This extra parameter is motivated by its usefulness when addressing the KS theory and is standard in the literature.

We consider wavefunctions $\psi = \psi(\vec{x}_1, \dots, \vec{x}_N)$, where $\vec{x}_k = (\vec{r}_k, s_k)$ is the spatial and spin coordinate of the k -th particle. The wavefunctions are antisymmetric elements of the N -electron space $L^2((\mathbb{R}^3 \times \{\uparrow, \downarrow\})^N)$, i.e., the usual Hilbert space of quantum mechanics. We will be interested in ground-state CDFT, where several options are available to treat the spin degrees of freedom. Firstly, we could formulate a theory for the global ground state, obtained through minimization over all spin degrees of freedom, in which case a spin–Zeeman term could also be included in the Hamiltonian. Secondly, we could instead formulate a theory for the lowest singlet state ($S^2 = 0$) or the lowest state with some other prescribed value of the spin quantum number S^2 . For simplicity, we formulate a ground-state CDFT for the lowest singlet energy and adapt our notation accordingly. However, our analysis is mostly independent of this choice and applies equally well to a theory for global ground states. Thus without loss of generality the spin coordinate will in the sequel be omitted.

All wavefunctions are assumed to have finite kinetic energy,

$$\mathcal{K}(\psi) = \frac{1}{2} \sum_{k=1}^N \int_{\mathbb{R}^{3N}} |\nabla_{\vec{r}_k} \psi|^2 d\vec{r}_1 \dots d\vec{r}_N < +\infty.$$

We further assume L^2 normalization of ψ and denote the L^p norm by $\|\cdot\|_p$, $1 \leq p \leq \infty$. Henceforth the particle number N will be fixed and we define the set of admissible wavefunctions

$$\mathcal{W}_N = \{ \psi : \|\psi\|_2 = 1, \mathcal{K}(\psi) < +\infty, \\ \psi \text{ is antisymmetric} \}.$$

Moreover, $\gamma_\psi = |\psi\rangle\langle\psi|$ denotes the density matrix of a pure state ψ and $\mathcal{P}_N = \{ \gamma_\psi : \psi \in \mathcal{W}_N \}$ is the set of such states. The set of mixed states is given by (where the sum over n can be infinite)

$$\mathcal{D}_N = \left\{ \gamma = \sum_n p_n \gamma_n : \gamma_n \in \mathcal{P}_N, \right. \\ \left. \sum_n p_n = 1, p_n \geq 0, \text{Tr}(\gamma T(\vec{0})) < +\infty \right\}.$$

Note that in case $\gamma = \gamma_\psi \in \mathcal{P}_N$ we have $\text{Tr}(\gamma T(\vec{0})) = \mathcal{K}(\psi)$.

The energy functional for the ground-state energy can be written in the following alternative forms

$$\begin{aligned} E^\lambda(v, \vec{A}) &= \inf_{\gamma \in \mathcal{D}_N} \text{Tr}(\gamma H^\lambda(v, \vec{A})) \\ &= \inf_{\gamma \in \mathcal{P}_N} \text{Tr}(\gamma H^\lambda(v, \vec{A})) \\ &= \inf_{\psi \in \mathcal{W}_N} \langle \psi, H^\lambda(v, \vec{A}) \psi \rangle. \end{aligned} \quad (2)$$

Thus if a minimizer $\gamma \in \mathcal{D}_N$ exists one can always also obtain a pure ground state selected from one of the eigenvectors of γ .

2.2 Function spaces for densities

For any $\psi \in \mathcal{W}_N$, we define the particle density and the paramagnetic current density, respectively, according to

$$\rho_\psi(\vec{r}_1) = N \int_{\mathbb{R}^{3(N-1)}} |\psi|^2 d\vec{r}_2 \dots d\vec{r}_N, \quad (3)$$

$$\vec{j}_\psi(\vec{r}_1) = N \text{Im} \int_{\mathbb{R}^{3(N-1)}} \psi^* \nabla_{\vec{r}_1} \psi d\vec{r}_2 \dots d\vec{r}_N. \quad (4)$$

The aim of this section is to extract as much information as possible about the regularity of ρ_ψ and \vec{j}_ψ in terms of L^p spaces from the assumption that $\psi \in \mathcal{W}_N$. This will define the sets of admissible densities.

To avoid confusion, a word or two on our notation is appropriate at this point. Since the paramagnetic current density is the main current density of consideration we omit the usual superscript (or subscript) ‘‘p’’ for paramagnetic in \vec{j}^p . We write $\vec{r} = (\vec{r}_1, \dots, \vec{r}_N) \in \mathbb{R}^{3N}$, $\nabla = \nabla_{\vec{r}}$, and let \vec{r} denote any $\vec{r}_i \in \mathbb{R}^3$ but typically \vec{r}_1 . Further $\mathcal{H}^k(\mathbb{R}^n)$ denotes the Sobolev space that includes all functions in $L^2(\mathbb{R}^n)$ with weak derivatives up to k -th order in $L^2(\mathbb{R}^n)$. (\mathcal{H} should not be confused with the Hamiltonian H .) Finally, $\vec{X} = X \times X \times X$ is the triple copy of a Banach space X , here mostly used for L^p spaces as \vec{L}^p .

Hoffmann-Ostenhof and Hoffmann-Ostenhof (see Eq. (3.10) in Ref. 36) and Lieb (Theorem 1.1 in Ref. 2) have shown that the von Weizsäcker term involving ρ_ψ is bounded by the kinetic energy of ψ , i.e.,

$$\frac{1}{2} \int_{\mathbb{R}^3} |\nabla \rho_\psi^{1/2}|^2 d\vec{r} \leq \mathcal{K}(\psi), \quad (5)$$

and therefore $\rho_\psi \in \mathcal{I}_N$ if $\psi \in \mathcal{W}_N$, where

$$\mathcal{I}_N = \left\{ \rho \in L^1(\mathbb{R}^3) : \rho \geq 0, \|\rho\|_1 = N, \rho^{1/2} \in \mathcal{H}^1(\mathbb{R}^3) \right\}$$

denotes the set of N -representable particle densities. Even though the Hilbert space $\mathcal{H}^2(\mathbb{R}^{3N})$ is the natural domain of the kinetic energy operator, $\psi \in \mathcal{H}^1(\mathbb{R}^{3N})$ is sufficient to guarantee finite $\mathcal{K}(\psi)$. The Sobolev in-

equality in \mathbb{R}^3 (see, e.g., Theorem 8.3 in Ref. 33),

$$\left(\int_{\mathbb{R}^3} |f(\vec{r})|^6 d\vec{r} \right)^{1/3} \leq S \int_{\mathbb{R}^3} |\nabla f(\vec{r})|^2 d\vec{r}, \quad (6)$$

applied to $f = \rho_\psi^{1/2}$ yields

$$\|\rho_\psi\|_3 \leq S \int_{\mathbb{R}^3} |\nabla \rho_\psi^{1/2}|^2 d\vec{r}, \quad S = \frac{4}{3(16\pi^2)^{1/3}}. \quad (7)$$

Consequently $\mathcal{I}_N \subset L^1 \cap L^3$.

Remark 1. We make a brief comment concerning interpolation. For $1 \leq q \leq p \leq r \leq +\infty$, if $f \in L^q \cap L^r$ then by Hölder’s inequality

$$\|f\|_p^p \leq \|f\|_r^{p(r-q)/(r-q)} \|f\|_q^{q(r-p)/(r-q)}$$

and thus $f \in L^p$.

We proceed by summarizing criteria that follow from the works of Lieb and Kato for the space of particle densities in terms of L^p spaces.

Proposition 1 (Lieb² and Kato³⁷). *For $\psi \in \mathcal{W}_N$, $\rho_\psi \in L^p$ with $p \in [1, 3]$ and in particular ρ_ψ is an element of the Hilbert space L^2 . Moreover, $\psi \in \mathcal{H}^2$ implies $\rho_\psi \in L^p$ for all $p \in [1, \infty]$.*

Proof. The first part follows from Lieb² and Remark 1. By Lemma 3 in Ref. 37, the assumption $\psi \in \mathcal{H}^2$ gives $\rho_\psi \in L^\infty$. (Note that $A_i f(\vec{r}_i)$ in Ref. 37 corresponds to $(\rho(\vec{r}_i)/N)^{1/2}$ in our notation.) \square

An obvious limitation for the current density is that every component of \vec{j}_ψ is in L^1 .³¹ Here, by better exploiting the properties of $\psi \in \mathcal{W}_N$, we will be able to further characterize the set of current densities.

We start by giving some definitions. The kinetic-energy density $\tau_\psi : \mathbb{R}^3 \rightarrow \mathbb{R}_+$ is given by

$$\tau_\psi(\vec{r}) = N \int_{\mathbb{R}^{3(N-1)}} |\nabla_{\vec{r}} \psi|^2 d\vec{r}_2 \dots d\vec{r}_N,$$

and relates to the already defined kinetic energy by $\mathcal{K}(\psi) = \frac{1}{2} \|\nabla \psi\|_2^2 = \frac{1}{2} \|\tau_\psi\|_1$. We see that ψ being an element of \mathcal{W}_N guarantees finite \mathcal{H}^1 norm and thus $\mathcal{W}_N \subset \mathcal{H}^1$. Furthermore, let $\vec{r} \in \mathbb{R}^3$ be written $\vec{r} = (r^1, r^2, r^3)$ and define the component-wise kinetic energy density

$$\tau_\psi^k(\vec{r}) = N \int_{\mathbb{R}^{3(N-1)}} |\nabla_{r^k} \psi|^2 d\vec{r}_2 \dots d\vec{r}_N$$

and $\mathcal{K}^k(\psi) = \frac{1}{2} \|\tau_\psi^k\|_1$. A direct computation gives that the usual Sobolev norm satisfies

$$\|\psi\|_{\mathcal{H}^1}^2 = \int_{\mathbb{R}^3} (\rho_\psi(\vec{r})/N + \tau_\psi(\vec{r})) d\vec{r} = 1 + 2\mathcal{K}(\psi).$$

An important bound that will be used subsequently is the following one

$$\left| N \int_{\mathbb{R}^{3(N-1)}} \psi^* \nabla_{\vec{r}} \psi \, d\vec{r}_2 \dots d\vec{r}_N \right| \leq \tau_\psi(\vec{r})^{1/2} \rho_\psi(\vec{r})^{1/2}, \quad (8)$$

which is a consequence of the Cauchy–Schwarz inequality. Since both $\tau_\psi^{1/2}$ and $\rho_\psi^{1/2}$ are L^2 functions, integration and the Cauchy–Schwarz inequality give $\vec{j}_\psi \in \vec{L}^1$. Indeed

$$\int_{\mathbb{R}^3} |\vec{j}_\psi(\vec{r})| \, d\vec{r} \leq \frac{1}{2} (2N \mathcal{K}(\psi))^{1/2},$$

see Proposition 3 in Ref. 31. The idea is now to further extend such \vec{L}^p -characterizations. With $\vec{j}_\psi = (j_\psi^1, j_\psi^2, j_\psi^3)$ we now state and prove

Lemma 2. *Set $f_\psi^k = \partial_k \rho_\psi + 2ij_\psi^k$ for $k = 1, 2, 3$ and $\psi \in \mathcal{W}_N$. Then $f_\psi^k \in L^p(\mathbb{R}^3)$ for $1 \leq p \leq 3/2$.*

Proof. Since

$$f_\psi^k(\vec{r}) = N \int_{\mathbb{R}^{3(N-1)}} \psi^* \partial_{r^k} \psi \, d\vec{r}_2 \dots d\vec{r}_N,$$

we have similar to Eq. (8) for $1 \leq p \leq 3/2$

$$|f_\psi^k(\vec{r})|^p \leq (\rho_\psi(\vec{r}) \tau_\psi^k(\vec{r}))^{p/2} \quad (9)$$

by the Cauchy–Schwarz inequality. Next, we use Hölder’s inequality with q defined by $p/2 + 1/q = 1$ such that

$$\begin{aligned} & \left(\int_{\mathbb{R}^3} |f_\psi^k(\vec{r})|^p \, d\vec{r} \right)^{1/p} \\ & \leq \left[\left(\int_{\mathbb{R}^3} \rho_\psi(\vec{r})^{pq/2} \, d\vec{r} \right)^{1/q} \left(\int_{\mathbb{R}^3} \tau_\psi^k(\vec{r}) \, d\vec{r} \right)^{p/2} \right]^{1/p} \\ & = \|\rho_\psi\|_{pq/2}^{1/2} (2\mathcal{K}^k(\psi))^{1/2}. \end{aligned} \quad (10)$$

To conclude, we note that, by the assumption on p , we have $1 \leq pq/2 \leq 3$ and recall that ρ is in $L^1 \cap L^3$. \square

Note that \vec{f}_ψ in Lemma 2 can be seen as a generalized complex current, similar to the one considered by Tokatly³⁸ in a lattice version of time-dependent CDFT. It has also been considered before as “momentum density”.³⁹

As a direct consequence of Lemma 2, we have our main result about function spaces for the current density

Theorem 3. *For $\psi \in \mathcal{W}_N$, each component j_ψ^k of the paramagnetic current density \vec{j}_ψ is in L^p for any $1 \leq p \leq 3/2$ and we write $\vec{j}_\psi \in \vec{L}^p$. In particular, we have*

$$\|j_\psi^k\|_1 \leq (2N \mathcal{K}^k(\psi))^{1/2}, \quad \|j_\psi^k\|_{3/2} \leq S^{1/2} 2\mathcal{K}^k(\psi),$$

with the constant S given by Eq. (7).

Proof. Set $p = 1$ and apply Lemma 2. Since $pq/2 = 1$ and $\|\rho_\psi\|_1 = N$, Eq. (10) gives

$$\|j_\psi^k\|_1 \leq \int_{\mathbb{R}^3} |f_\psi^k(\vec{r})| \, d\vec{r} \leq (2N \mathcal{K}^k(\psi))^{1/2}.$$

With the choice $p = 3/2$ instead, we have $pq/2 = 3$. Eq. (10) then reduces to

$$\|j_\psi^k\|_{3/2} \leq S^{1/2} 2\mathcal{K}^k(\psi),$$

where we have also used the Sobolev inequality (6) and Eq. (5). Using interpolation (see Remark 1), it follows

$$\|j_\psi^k\|_p \leq [N^{3-2p} S^{3p-3} (2\mathcal{K}^k(\psi))^{4p-3}]^{\frac{1}{2p}}$$

and thus $j_\psi^k \in L^p$ for $1 \leq p \leq 3/2$. \square

From the proof of Lemma 2 (the Cauchy–Schwarz inequality applied to Eq. (9) with $p = 2$), we obtain the current correction to the von Weizsäcker kinetic energy (see Ref. 40 and note that this sharpens Theorem 14 in Ref. 31)

$$\frac{1}{2} \int_{\mathbb{R}^3} |\nabla \sqrt{\rho_\psi}|^2 \, d\vec{r} + \frac{1}{2} \int_{\mathbb{R}^3} \frac{|\vec{j}_\psi|^2}{\rho_\psi} \, d\vec{r} \leq \mathcal{K}(\psi). \quad (11)$$

The well-known inequality $\frac{1}{2} \int_{\mathbb{R}^3} |\vec{j}_\psi|^2 \rho_\psi^{-1} \, d\vec{r} \leq \mathcal{K}(\psi)$ is immediate from Eq. (11).

We next note that the space for the current density cannot be further restricted since $j_\psi^k \notin L^p$, $p > 3/2$, for some $\psi \in \mathcal{W}_N$. Before proving this fact, a further characterization of \vec{j}_ψ using $\psi \in \mathcal{H}^2$ is given.

Proposition 4. *$\psi \in \mathcal{H}^2$ implies $\vec{j}_\psi \in \vec{L}^2$.*

Proof. Suppose $\psi \in \mathcal{H}^2$. By Proposition 1 it follows that $\rho_\psi \in L^\infty$. Then from Eq. (8)

$$\begin{aligned} & \int_{\mathbb{R}^3} \left| N \int_{\mathbb{R}^{3(N-1)}} \psi^* \nabla_{\vec{r}} \psi \, d\vec{r}_2 \dots d\vec{r}_N \right|^2 \, d\vec{r} \\ & \leq \int_{\mathbb{R}^3} \tau_\psi(\vec{r}) \rho_\psi(\vec{r}) \, d\vec{r} \\ & \leq \|\rho_\psi\|_\infty 2\mathcal{K}(\psi) < +\infty, \end{aligned}$$

which gives $\vec{j}_\psi \in \vec{L}^2$. (Note that we could have argued by means of the Cauchy–Schwarz inequality to obtain $\int_{\mathbb{R}^3} |\vec{j}|^2 \, d\vec{r} \leq \|\rho\|_\infty \int_{\mathbb{R}^3} |\vec{j}|^2 \rho^{-1} \, d\vec{r}$, where the integral over $|\vec{j}|^2 \rho^{-1}$ is bounded in terms of $\mathcal{K}(\psi)$.) \square

Proposition 5. *For $N = 1$, there exists $\phi \in \mathcal{W}_N$ such that $\vec{j}_\phi \notin \vec{L}^p$, for every $p > 3/2$.*

Proof. Consider the function

$$\phi(\vec{r}) = r^{\alpha/2} e^{i\xi(r)}, \quad \xi(r) = r^\beta, \quad r = |\vec{r}|,$$

where $0 \leq r \leq 1$, $\alpha > -1$, and $2\beta > -\alpha - 1$. It then holds that $\phi \in \mathcal{H}^1(\mathbb{R}^3)$, since

$$\begin{aligned} \int_0^1 |\phi|^2 r^2 dr &= \int_0^1 r^{\alpha+2} dr = \left[\frac{r^{\alpha+3}}{\alpha+3} \right]_0^1 = \frac{1}{\alpha+3}, \\ \int_0^1 |\nabla\phi|^2 r^2 dr &= \int_0^1 |\nabla(r^{\alpha/2})|^2 r^2 dr + \int_0^1 |\nabla\xi|^2 r^{\alpha+2} dr \\ &= \frac{\alpha^2}{4} \int_0^1 r^\alpha dr + \beta^2 \int_0^1 r^{\alpha+2\beta} dr \\ &= \frac{\alpha^2}{4(\alpha+1)} + \frac{\beta^2}{\alpha+2\beta+1}. \end{aligned}$$

We wish to show that $\vec{j}_\phi \notin \vec{L}^p$, for all $p > 3/2$ for some choice of (α, β) in the set

$$\Theta = \{\alpha > -1, \beta > -(\alpha+1)/2\}.$$

Note that $\vec{j}_\phi = \text{Im} \phi^* \nabla \phi = r^\alpha \nabla r^\beta$. Let δ be an arbitrarily small positive number and set $\alpha = -1 + 3\delta$. Then with $\beta = -\delta > -3\delta/2$ we have $(\alpha, \beta) \in \Theta$. For all $p > 3/(2(1-\delta))$ we obtain

$$\begin{aligned} \int_0^\infty |\vec{j}_\phi|^p r^2 dr &\geq \delta^p \int_0^1 r^{2p(\delta-1)+2} dr \\ &= \delta^p \lim_{\varepsilon \rightarrow 0^+} \left[\frac{r^{2p(\delta-1)+3}}{2p(\delta-1)+3} \right]_\varepsilon^1 = +\infty. \end{aligned}$$

Since any $\delta > 0$ is allowed, we conclude $\vec{j}_\phi \notin \vec{L}^p$, $p > 3/2$. \square

Remark 2. Note that the same counterexample also shows that $\rho_\phi \notin L^p$, for $p > 3$.

From $\psi \in \mathcal{W}_N$, with Proposition 1 and Remark 2 we have thus arrived at the well-known optimal choice $L^1 \cap L^3$ of L^p spaces for the particle density.² Similarly, by Theorem 3 and Proposition 5 the optimal choice for the paramagnetic current density is $\vec{L}^1 \cap \vec{L}^{3/2}$. Note that densities and currents that are not from these spaces cannot be represented by admissible wavefunctions $\psi \in \mathcal{W}_N$. Later this choice will be further limited by the demands coming from *compatibility* (see Sec. 2.4), *reflexivity*, and *strict convexity* (in connection with the regularized KS iteration scheme, Sec. 3.3).

We summarize this section with some definitions and a concluding corollary. We also refer to Refs. 2,39,41–44 for further discussions on this topic (not only confined to CDFT).

Definition 6. We say that a density pair (ρ, \vec{j}) is *N-representable* if there exists a $\psi \in \mathcal{W}_N$ such that $\rho_\psi = \rho$ and $\vec{j}_\psi = \vec{j}$. If such a ψ is the ground state of some $H^\lambda(v, \vec{A})$, then (ρ, \vec{j}) is (λ) *v-representable*. Furthermore, we distinguish between *fully interacting* ($\lambda = 1$) and *non-interacting* ($\lambda = 0$) *v-representability*. If $\psi \in \mathcal{W}_N$ is replaced by $\gamma \in \mathcal{D}_N$ we call the above property *ensemble v-representability*.

Corollary 7. *The set of N-representable density pairs (ρ, \vec{j}) is a subset of $(L^1 \cap L^3) \times (\vec{L}^1 \cap \vec{L}^{3/2})$.*

The above set is given in terms of L^p spaces only. There are other well established constraints such as $\rho(\vec{r}) \geq 0$, $\int_{\mathbb{R}^3} \rho d\vec{r} = N$, and $\int_{\mathbb{R}^3} |\vec{j}|^2 \rho^{-1} d\vec{r} < +\infty$.^{2,31,42}

It will later be important to impose a reflexive (R) and strictly convex Banach space setting, and we therefore define such a space that includes the set of *N-representable density pairs*. Hanner’s inequality⁴⁵ shows that L^p , $1 < p < \infty$, is even a uniformly convex space which implies both strict convexity and reflexivity.

Definition 8. $X_{\text{R}} \times Y_{\text{R}} = L^3 \times \vec{L}^{3/2}$.

2.3 Constrained-search functionals

To formulate a rigorous CDFT, several requirements need to be placed on densities and potentials. Some of these requirements are related to *N-representability* and thus do not amount to any restriction, but merely exclude irrelevant densities that are invalid in the sense that they cannot arise from any quantum-mechanical state $\psi \in \mathcal{W}_N$. Other requirements amount to assumptions about the ground-state densities or restrictions on the external potentials that can be considered.

The universal part of the Hamiltonian $H^\lambda(v, \vec{A})$ in Eq. (1), independent of the potential pair (v, \vec{A}) , is $H^\lambda(0, \vec{0}) = T(\vec{0}) + \lambda W$. On a Banach space $X \times Y \subset X_{\text{R}} \times Y_{\text{R}}$ of measurable functions for particle densities (X) and current densities (Y), define the universal Levy–Lieb-type functionals F_{VR}^λ and $F_{\text{VR,DM}}^\lambda$ according to

$$\begin{aligned} F_{\text{VR}}^\lambda(\rho, \vec{j}) &= \inf_{\psi \in \mathcal{W}_N} \left\{ \langle \psi, H^\lambda(0, \vec{0}) \psi \rangle : \psi \mapsto (\rho, \vec{j}) \right\}, \\ F_{\text{VR,DM}}^\lambda(\rho, \vec{j}) &= \inf_{\gamma \in \mathcal{D}_N} \left\{ \text{Tr}(\gamma H^\lambda(0, \vec{0})) : \gamma \mapsto (\rho, \vec{j}) \right\}. \end{aligned}$$

These functionals are derived from the parts of the energy expressions in Eq. (2) that are independent of the potential pair (v, \vec{A}) . In analogy with Eq. (2) we generally have two possibilities, searching either over pure or mixed states. If a given density pair (ρ, \vec{j}) cannot be represented by a pure or mixed state, then the value of the functional will just be set to $+\infty$ by definition. Unlike the ground-state energy functional in Eq. (2), the pure and mixed search domains do not yield equivalent results, since the former is subject to more severe representability restrictions.³⁹

Remark 3. The density functionals are here denoted “VR” which stands for Vignale and Rasolt to credit their work.¹² We remark that we could just as well have chosen to credit Levy, Valone, and Lieb due to the obvious counterpart in DFT.^{2,30,46} “DM” refers to the use of density matrices for mixed states and was introduced for the DFT constrained-search functional by Valone in Ref. 46.

Remark 4. At this point we wish to keep the setting general and thus the space $X \times Y$ is not specified any closer. This setting includes the choice $X = L^1 \cap L^3$, $Y = L^1 \cap L^{3/2}$ such that all information from the previous section on function spaces is used. However, if a regularized theory is to be obtained, we need to have a reflexive Banach space setting and therefore we also consider the less restrictive choice $X = X_R$, $Y = Y_R$.

The Levy–Lieb-type functional $F_{VR}(\rho, \vec{j})$ is not convex, see Proposition 8 in Ref. 31. Yet by the linearity of the map $\gamma \mapsto (\rho, \vec{j})$ it follows that $F_{VR,DM}(\rho, \vec{j})$ is convex. Both functionals are *admissible*⁴⁷ in the sense that they can be used to compute the ground-state energy.

Since the energy expression will naturally include integrals over couplings of potentials with densities, it is helpful to introduce the notion of dual pairings (between elements of dual Banach spaces). For measurable functions f, g with domain \mathbb{R}^3 let

$$\langle f, g \rangle = \int_{\mathbb{R}^3} f(\vec{r})g(\vec{r}) \, d\vec{r},$$

whenever the integral is well-defined in $\mathbb{R} \cup \{\pm\infty\}$, and similarly for vector-valued functions \vec{f}, \vec{g} , but with the pointwise product replaced by $\vec{f} \cdot \vec{g}$. Then the energy expressions in Eq. (2) can be written as

$$E^\lambda(v, \vec{A}) = \inf_{(\rho, \vec{j}) \in X \times Y} \left\{ F_{VR}^\lambda(\rho, \vec{j}) + \langle \vec{A}, \vec{j} \rangle + \langle v + \frac{1}{2}|\vec{A}|^2, \rho \rangle \right\},$$

and equivalently by employing $F_{VR,DM}^\lambda$ defined with mixed states instead of F_{VR}^λ . In particular, $\lambda = 1$ corresponds to the fully interacting system, and $\lambda = 0$ to a non-interacting one.

At the outset, the formulation of paramagnetic CDFT relies on a decomposition of the total kinetic energy into canonical kinetic energy, the paramagnetic term $\langle \vec{A}, \vec{j} \rangle$, and the diamagnetic term $\langle |\vec{A}|^2, \rho \rangle$, with each of the terms separately finite.^{3,15,48} As in standard DFT, the electrostatic interaction with the external potential, $\langle v, \rho \rangle$, needs to be finite too. In the convexified form, a new potential variable is formed by absorbing the diamagnetic term into the scalar potential $u = v + \frac{1}{2}|\vec{A}|^2$.^{3,31} Minimally, then, the underlying function spaces should be such that

$$|\langle u, \rho \rangle| < +\infty \quad \text{and} \quad |\langle \vec{A}, \vec{j} \rangle| < +\infty.$$

A convex formulation achieves this automatically as it requires the stronger condition that densities and potentials are elements of dual Banach spaces:

Definition 9 (Density-potential duality). We say that there is *density-potential duality*, or just *duality*, when densities and potentials are confined to dual Banach

spaces

$$(D1) \quad \rho \in X \quad \text{and} \quad u \in X^*,$$

$$(D2) \quad \vec{j} \in Y \quad \text{and} \quad \vec{A} \in Y^*.$$

Remark 5. At this moment we do not assume any more specific properties for $X \subset X_R$ and $Y \subset Y_R$ besides duality. However, reflexivity and strict convexity of X and Y are additional assumptions that will become important in Sec. 3.

Remark 6. Note that $u \in X^*$, $X = L^1 \cap L^3$ and $\vec{A} \in Y^*$, $Y = L^1 \cap L^{3/2}$ imply $u \in L^{3/2} + L^\infty$ and $\vec{A} \in \vec{L}^3 + \vec{L}^\infty$. For the condition on the original scalar potential v see the next section on *compatibility*. As far as the vector potential is concerned, the restrictions on \vec{A} are stronger than the familiar setting of $\vec{A} \in L_{loc}^2$ (see e.g., Ref. 33), which is implied by $\vec{A} \in \vec{L}^3 + L^\infty$. Also the reflexive setting with $X = X_R$ and $Y = Y_R$, where $u \in L^{3/2}$ and $\vec{A} \in \vec{L}^3$, implies $\vec{A} \in L_{loc}^2$ again. Moreover, $u \in L_{loc}^{3/2}$ is a natural assumption.³³ We remark that our consideration of dual spaces is mathematically motivated and not a physical necessity. A truncated space domain is in many cases needed to cover the usual potentials of physical systems (see Ref. 4 for a discussion on this topic).

Since the potentials (v, \vec{A}) are *not* paired linearly with the densities (ρ, \vec{j}) , the functional E^λ defined in this way is *not* concave. The change of variables $u = v + \frac{1}{2}|\vec{A}|^2$ results in a convexification of paramagnetic CDFT, meaning that

$$\bar{E}^\lambda(u, \vec{A}) = E^\lambda(u - \frac{1}{2}|\vec{A}|^2, \vec{A}) \quad (12)$$

is a *jointly concave* functional.³ The consequences of this variable change for the choice of function spaces will be discussed in Sec. 2.4. The price to pay for concavity is a convoluted gauge symmetry. For all scalar fields χ with gradients in the same function space as \vec{A} one has

$$\begin{aligned} \bar{E}^\lambda(u, \vec{A}) &= E^\lambda(u - \frac{1}{2}|\vec{A}|^2, \vec{A}) \\ &= E^\lambda(u - \frac{1}{2}|\vec{A}|^2, \vec{A} + \nabla\chi) \\ &= \bar{E}^\lambda(u + \vec{A} \cdot \nabla\chi + \frac{1}{2}|\nabla\chi|^2, \vec{A} + \nabla\chi). \end{aligned}$$

But the benefit is much greater, making \bar{E}^λ jointly concave in both potentials and highlighting the linear structure of coupling between potentials and densities

$$\bar{E}^\lambda(u, \vec{A}) = \inf_{(\rho, \vec{j}) \in X \times Y} \left\{ F_{VR}^\lambda(\rho, \vec{j}) + \langle \vec{A}, \vec{j} \rangle + \langle u, \rho \rangle \right\}.$$

The convex formulation of paramagnetic CDFT can be outlined as follows. Let the dual space of $X \times Y$ be given by $X^* \times Y^*$. We define the generalized Lieb functional $F^\lambda(\rho, \vec{j})$ as the supremum over the energy plus the

linear coupling between densities and potentials, i.e.,

$$F^\lambda(\rho, \vec{j}) = \sup_{(u, \vec{A}) \in X^* \times Y^*} \left\{ \bar{E}^\lambda(u, \vec{A}) - \langle \vec{A}, \vec{j} \rangle - \langle u, \rho \rangle \right\}. \quad (13)$$

Such a functional is by construction convex. To extract more from Eq. (13) we first need

Definition 10. Let B be a Banach space with dual B^* , $f : B \rightarrow \mathbb{R} \cup \{\pm\infty\}$, and $g : B^* \rightarrow \mathbb{R} \cup \{\pm\infty\}$.

- (i) If f is convex, lower semi-continuous, has $f > -\infty$, and is not identically equal to $+\infty$, then it is called *closed convex* and we write $f \in \Gamma_0(B)$. Analogously, with weak-* lower semi-continuity we define $\Gamma_0^*(B^*)$. We also introduce the sets $\Gamma(B) = \Gamma_0(B) \cup \{\pm\infty\}$ and $\Gamma^*(B^*) = \Gamma_0^*(B^*) \cup \{\pm\infty\}$.
- (ii) Following Refs. 2,4, we define the (skew) conjugate functionals (Legendre–Fenchel transformations),

$$f^\wedge(b^*) = \inf_{b \in B} \{f(b) + \langle b^*, b \rangle\} \in -\Gamma^*(B^*),$$

$$g^\vee(b) = \sup_{b^* \in B^*} \{g(b^*) - \langle b^*, b \rangle\} \in \Gamma(B).$$

Definition 11. The standard norms for the intersection of two Banach spaces B, B' and their set sum are given by the following expressions⁴⁹

$$\|b\|_{B \cap B'} = \max\{\|b\|_B, \|b\|_{B'}\},$$

$$\|b\|_{B+B'} = \inf\{\|b\|_B + \|b'\|_{B'} : b \in B, b' \in B', b = b + b'\}.$$

Theorem 3.6 in Lieb² can be straightforwardly generalized to the statement that F^λ is lower semi-continuous on the space $X \times Y = (L^1 \cap L^3) \times (\bar{L}^1 \cap \bar{L}^{3/2})$, with topology defined by the norm from the above definition. (See also Proposition 12 in Ref. 31 where this was done for $(L^1 \cap L^3) \times \bar{L}^1$.) Also, by the same argument, we have for the reflexive setting

Lemma 12. $F^\lambda \in \Gamma_0(X_R \times Y_R)$.

Since F^λ is convex *and* lower semi-continuous, this clears the way for the application of powerful tools from convex analysis. Because Eq. (13) is already the Legendre–Fenchel transformation of the energy functional \bar{E}^λ we are able to switch back from F^λ to \bar{E}^λ with the inverse transformation (see Theorem 1 in Ref. 9)

$$\bar{E}^\lambda(u, \vec{A}) = \inf_{(\rho, \vec{j}) \in X \times Y} \left\{ F^\lambda(\rho, \vec{j}) + \langle \vec{A}, \vec{j} \rangle + \langle u, \rho \rangle \right\}. \quad (14)$$

Using the notation from Definition 10 (ii), we sum up the situation as

$$\bar{E}^\lambda = (F_{\text{VR}}^\lambda)^\wedge = (F_{\text{VR,DM}}^\lambda)^\wedge = (F^\lambda)^\wedge, \quad F^\lambda = (\bar{E}^\lambda)^\vee.$$

Moreover the closed convex F^λ is the smallest possible admissible functional,

$$F^\lambda \leq F_{\text{VR,DM}}^\lambda \leq F_{\text{VR}}^\lambda.$$

Solving the variational problem in Eq. (14) is the general task of CDFT.

Remark 7. It is to the best of our knowledge an open question whether $F_{\text{VR,DM}}^\lambda$ is lower semi-continuous and hence equal to F^λ in the context of paramagnetic CDFT.

2.4 Compatibility of function spaces

Duality as in Definition 9 is not strong enough to guarantee finiteness of the diamagnetic term. It also does not guarantee another natural condition on the function space for the diamagnetic contribution to the current density that we call *compatibility*.

Definition 13. The density function space X and the current density function space Y are said to be *compatible* if, for all $\rho \in X$ and all $\vec{A} \in Y^*$,

$$(C1) \quad |\vec{A}|^2 \in X^* \quad \text{and} \quad (C2) \quad \rho \vec{A} \in Y.$$

We emphasize both conditions in the definition above, as they have different physical interpretations, although it will be seen in Theorem 14 that (C1) and (C2) are equivalent. The first condition (C1) requires that the scalar potential v and $|\vec{A}|^2$ share the same function space, so that changes of variables between $v = u - \frac{1}{2}|\vec{A}|^2$ and $u = v + \frac{1}{2}|\vec{A}|^2$ stay within the space X^* . The second condition (C2) requires that the paramagnetic contribution, \vec{j} , and the diamagnetic contribution, $\rho \vec{A}$, to the total current density share the same function space Y .

Compatibility also ensures a sensible behavior under gauge transformations. Duality imposes a restriction on the gauge transformations that are allowed within the theory. Any gauge function χ that is used to transform \vec{A} to $\vec{A}' = \vec{A} + \nabla\chi$ must satisfy $\nabla\chi \in Y^*$. If (v, \vec{A}) has the ground-state density pair (ρ, \vec{j}) , the gauge transformed potential pair (v, \vec{A}') would be expected to have the ground-state density pair $(\rho, \vec{j}' = \vec{j} + \rho \nabla\chi)$. The second compatibility condition (C2) ensures that $\vec{j}' \in Y$, so that ground-state density pairs are never lost after an allowed gauge transformation.

The following theorem shows that the two compatibility conditions are in fact equivalent. First, we mention a fundamental result that will be used in the sequel. Suppose we are given a general Banach space B of measurable functions and a measurable function g . We can then check that g is contained in the dual B^* by verifying that the pairing $\langle g, f \rangle = \int g f \, d\vec{r}$ is finite for all $f \in B$, see Appendix A for a full proof of this statement.

Theorem 14. *If X, Y and their duals are Banach spaces of measurable functions, then (C1) \iff (C2).*

Proof. Part 1 (\Leftarrow): From (C2), we have that

$$\int_{\mathbb{R}^3} |\rho \vec{A} \cdot \vec{A}'| d\vec{r} < +\infty$$

for all $\rho \in X$ and all $\vec{A}, \vec{A}' \in Y^*$. Specialization to the case $\vec{A}' = \vec{A}$ yields (C1).

Part 2 (\Rightarrow): Suppose (C2) is false, i.e., $\rho \vec{A} \notin Y$. Then there exists an $\vec{A}' \in Y^*$ such that

$$\int_{\mathbb{R}^3} \rho \vec{A} \cdot \vec{A}' d\vec{r} = +\infty,$$

and we obtain

$$+\infty = |\langle \rho \vec{A}, \vec{A}' \rangle| \leq \frac{1}{2} (|\rho| |\vec{A}|^2 + |\vec{A}'|^2).$$

Thus $|\vec{A}|^2 + |\vec{A}'|^2 \notin X^*$ and either $|\vec{A}|^2$ or $|\vec{A}'|^2$ is not an element of X^* (or both). This contradicts (C1). \square

Compatible function spaces can be built up recursively, by combining different function spaces that are already known to be compatible.

Theorem 15 (Compatibility of intersections and sums). *Suppose X_1 and Y_1 are compatible, and the same holds for X_2 and Y_2 . Then (a) the intersections*

$$X = X_1 \cap X_2, \quad Y = Y_1 \cap Y_2, \quad (15)$$

are compatible. Moreover, (b) the sums

$$X = X_1 + X_2, \quad Y = Y_1 + Y_2, \quad (16)$$

are compatible. (Norms of intersections and sums are as given in Definition 11.)

Proof. By Theorem 14 it is sufficient to prove (C1).

Part (a): The dual spaces are $X^* = X_1^* + X_2^*$ and $Y^* = Y_1^* + Y_2^*$. Decompose the vector potential as $\vec{A} = \vec{A}_1 + \vec{A}_2$, with $\vec{A}_1 \in Y_1^*$ and $\vec{A}_2 \in Y_2^*$. Using the inequality

$$2|\vec{A}_1 \cdot \vec{A}_2| \leq |\vec{A}_1|^2 + |\vec{A}_2|^2$$

property (C1) follows immediately, since

$$|\langle \rho, |\vec{A}|^2 \rangle| \leq 2|\langle \rho, |\vec{A}_1|^2 \rangle| + 2|\langle \rho, |\vec{A}_2|^2 \rangle|$$

and each of the two terms is finite by hypothesis.

Part (b): The dual spaces in this case are $X^* = X_1^* \cap X_2^*$ and $Y^* = Y_1^* \cap Y_2^*$. Trivially, for any $\vec{A} \in Y^*$ we have $\vec{A} \in Y_i^*$ and therefore, by hypothesis, $|\vec{A}|^2 \in Y_i^*$ ($i = 1, 2$). Hence, $|\vec{A}|^2 \in Y_1^* \cap Y_2^*$. \square

As demonstrated in Ref. 31 (see also Theorem 3 above), the paramagnetic current density satisfies $\vec{j}_\psi \in \vec{L}^1$ for $\psi \in \mathcal{W}_N$. Combined with compatibility, this becomes a substantial condition on the vector potential space.

Theorem 16. *Let X and Y be compatible function spaces for the particle density and current density. Suppose furthermore that $Y \subseteq \vec{L}^1$. Then it follows that*

$$Y^* \subseteq \vec{X}^*$$

*and from that automatically $\vec{X} \subseteq Y \subseteq Y^{**}$.*

Proof. Let $\vec{A} = (f, g, h) \in Y^*$ be an arbitrary vector potential. From $Y \subseteq \vec{L}^1$ we immediately have $Y^* \supseteq \vec{L}^\infty$ and therefore we can add a constant to one of the components of \vec{A} without leaving the potential space Y^* ,

$$\vec{a} = (1 + f, g, h) \in Y^*.$$

Next, as we assume compatibility of X and Y , we have

$$\begin{aligned} |\vec{a}|^2 &= (1 + f)^2 + g^2 + h^2 \\ &= 1 + 2f + f^2 + g^2 + h^2 \in X^*. \end{aligned}$$

Furthermore, by assumption $f^2, g^2, h^2 \in X^*$. Compatibility combined with the fact that $1 \in L^\infty$ yields $1 \in X^*$. Thus, the only remaining term $2f$ must be an element of X^* too. Repeating the proof, with trivial changes for the other components, yields $f, g, h \in X^*$. \square

When the preconditions of the above theorem are satisfied and \vec{A} is an allowed vector potential, we thus have the peculiar situation that *both* $|\vec{A}|^2$ and $|\vec{A}|$ are contained in the function space of allowed scalar potentials.

Next, we turn to examples of reasonable choices of function spaces for paramagnetic CDFT that also illustrate the duality and compatibility conditions. Following Lieb,² we may choose the non-reflexive space $X = L^1 \cap L^3$ for the particle densities and its dual $X^* = L^{3/2} + L^\infty$ for the scalar potentials. For the current densities, we first consider the choice in the literature,³¹ where current densities were placed in the non-reflexive space $Y = \vec{L}^1$. Compatibility then follows trivially.

Proposition 17. *The choice $X = L^1 \cap L^3$ and $Y = \vec{L}^1$ is compatible.*

As a second example a choice of reflexive, compatible spaces should be given. Reflexivity is imperative for the construction of a well-defined KS scheme like in Sec. 3.3. To achieve this we just drop the non-reflexive L^1 from X in the example above and switch to $\vec{L}^{3/2}$ instead of \vec{L}^1 for Y .

Proposition 18. *The strictly convex and reflexive choice $X = L^3$, $Y = \vec{L}^{3/2}$ is compatible.*

Finally the admissible spaces $X = L^1 \cap L^3$ and $Y = \vec{L}^1 \cap \vec{L}^{3/2}$ that were derived from considerations regarding N -representability in Sec. 2.2 should be examined with respect to compatibility. It turns out that by applying Theorem 15 we can just build these spaces as intersections of the ones from Propositions 17 and 18.

Corollary 19. *The choice $X = L^1 \cap L^3$ and $Y = \vec{L}^1 \cap \vec{L}^{3/2}$ is compatible.*

Many other compatible function spaces can be constructed. The previous examples exclude the common case of uniform magnetic fields as these require linearly growing vector potentials, e.g., $\vec{A} = \frac{1}{2}\vec{B} \times \vec{r}$, which do not belong to any $\vec{L}^p(\mathbb{R}^3)$ space. We refer to Ref. 18 for a treatment of this situation. The next result provides function spaces that allow for inclusion of such vector potentials. It is a little detour to other possible choices for compatible Banach spaces, before we return to the discussion of energy functionals defined on them.

Let $w(\vec{r})$ be a suitable weight function. We use the notation $f \in L^p(w)$ for the weighted Lebesgue space defined by

$$\int_{\mathbb{R}^3} |f(\vec{r})|^p w(\vec{r}) d\vec{r} < +\infty.$$

Note that $wf \in L^p$ is equivalent to $f \in L^p(w^p)$. Some care is required however, as the two forms may produce inequivalent results when multiple weighted L^p -spaces are considered. For example, $f \in L^p(w') \cap L^q(w'')$ in general cannot be represented as $wf \in L^p \cap L^q$, for any weight function w . In the following examples the weight function is assumed to satisfy $w(\vec{r}) \geq 1$ for all $\vec{r} \in \mathbb{R}^3$.

Theorem 20. *Let Z be a normed space with dual Z^* . Then each of the following choices of function spaces is compatible,*

$$\begin{aligned} w^2 \rho &\in L^1 \cap Z, & w \vec{j} &\in \vec{L}^1, \\ w^{-2} v &\in L^\infty + Z^*, & w^{-1} \vec{A} &\in \vec{L}^\infty, \end{aligned} \quad (17)$$

or, with $1 < p < \infty$,

$$\begin{aligned} w^2 \rho &\in L^{p/(p-1)} \cap Z, & w \vec{j} &\in \vec{L}^{2p/(2p-1)}, \\ w^{-2} v &\in L^p + Z^*, & w^{-1} \vec{A} &\in \vec{L}^{2p}. \end{aligned} \quad (18)$$

Proof. In both cases, we exploit the equivalence of (C1) and (C2), and only prove one of them. First, for the choice in Eq. (17), condition (C1) is trivial, since $w^{-1} \vec{A} \in \vec{L}^\infty$ directly implies $w^{-2} |\vec{A}|^2 \in L^\infty$. Second, for Eq. (18), we similarly have that $w^{-1} \vec{A} \in \vec{L}^{2p}$ directly implies $w^{-2} |\vec{A}|^2 \in L^p$, establishing (C1). \square

In the first of the above examples, Eq. (17), we can make the trivial choice $w(\vec{r}) = 1$ and $Z = L^3$ to recover the function space analyzed in Ref. 31. Moreover, with suitable, non-trivial weight functions, unbounded vector potentials can be considered. For example, the choice $w(\vec{r}) = (1 + |\vec{r}|^2)^{1/2}$ was studied extensively in Ref. 18 as it allows for uniform magnetic fields and always ensures that the angular momentum is well-defined.

2.5 Finiteness of the energy functional

The following general property of the CDFT energy functional $E^\lambda(v, \vec{A})$ will be useful later. It says that for both interacting and non-interacting systems the energy is finite for all considered potentials. Furthermore, if the choice of function spaces is compatible, the same boundedness from below holds for $\bar{E}^\lambda(u, \vec{A})$.

Lemma 21. *$E^\lambda(v, \vec{A})$ is finite for*

$$(v, \vec{A}) \in (L^{3/2} + L^\infty) \times (\vec{L}^3 + \vec{L}^\infty),$$

corresponding to the density space $(\rho, \vec{j}) \in (L^1 \cap L^3) \times (\vec{L}^1 \cap \vec{L}^{3/2})$. By compatibility, $\bar{E}^\lambda(u, \vec{A})$ is also finite on the same domain.

Proof. To prove finiteness (of the infimum) it is enough to prove boundedness from below. By definition, we have for any $(v, \vec{A}) \in (L^{3/2} + L^\infty) \times (\vec{L}^3 + \vec{L}^\infty)$

$$\begin{aligned} E^\lambda(v, \vec{A}) = & \inf_{(\rho, \vec{j}) \in (L^1 \cap L^3) \times (\vec{L}^1 \cap \vec{L}^{3/2})} \left\{ F^\lambda(\rho, \vec{j}) + \int_{\mathbb{R}^3} \vec{A} \cdot \vec{j} d\vec{r} \right. \\ & \left. + \int_{\mathbb{R}^3} \left(v + \frac{1}{2} |\vec{A}|^2 \right) \rho d\vec{r} \right\}. \end{aligned} \quad (19)$$

Moreover, if (ρ, \vec{j}) is not N -representable then $F^\lambda(\rho, \vec{j}) = +\infty$ and we can consequently consider only N -representable density pairs. By definition $F^\lambda \geq F^0$ and furthermore $F^0(\rho, \vec{j}) \geq \inf_\psi \mathcal{K}(\psi)$ for all ψ with $\rho_\psi = \rho$ and $\vec{j}_\psi = \vec{j}$. By the von Weizsäcker bound in Eq. (11), we obtain

$$F^0(\rho, \vec{j}) \geq \frac{1}{2} \int_{\mathbb{R}^3} |\nabla \rho^{1/2}|^2 d\vec{r} + \frac{1}{2} \int_{\mathbb{R}^3} |\vec{j}|^2 \rho^{-1} d\vec{r},$$

for N -representable density pairs (ρ, \vec{j}) . The Sobolev inequality in Eq. (6) further allows for

$$F^0(\rho, \vec{j}) \geq \frac{1}{2S} \|\rho\|_3 + \frac{1}{2} \int_{\mathbb{R}^3} |\vec{j}|^2 \rho^{-1} d\vec{r}. \quad (20)$$

Combining Eqs. (19) and (20), it follows

$$\begin{aligned} E^\lambda(v, \vec{A}) \geq & \inf_{(\rho, \vec{j})} \left\{ \frac{1}{2S} \|\rho\|_3 + \frac{1}{2} \int_{\mathbb{R}^3} |\vec{j}|^2 \rho^{-1} d\vec{r} \right. \\ & \left. + \int_{\mathbb{R}^3} \vec{A} \cdot \vec{j} d\vec{r} + \int_{\mathbb{R}^3} \left(v + \frac{1}{2} |\vec{A}|^2 \right) \rho d\vec{r} \right\}, \end{aligned}$$

where the minimization is restricted to N -representable density pairs. Using the obvious inequality

$$|\vec{j}|/\rho^{1/2} + \rho^{1/2} \vec{A}^2 \geq 0$$

to replace the full square, we have

$$E^\lambda(v, \vec{A}) \geq \inf_{\rho \in \mathcal{I}_N} \left\{ \frac{1}{2S} \|\rho\|_3 + \int_{\mathbb{R}^3} v \rho d\vec{r} \right\}. \quad (21)$$

Next we bound the r.h.s. of Eq. (21) from below. Since \mathcal{C}_0^∞ is dense in $L^{3/2}$, also L^∞ is. This means that we can split $v = v_S + v_\infty$ with $v_S \in L^{3/2}$, $v_\infty \in L^\infty$ in such a way that $\|v_S\|_{3/2}$ is arbitrarily small. We choose the decomposition such that $\|v_S\|_{3/2} \leq (2S)^{-1}$ with the constant S from Eq. (7). Then using Hölder's inequality Eq. (21) can be estimated as

$$E^\lambda(v, \vec{A}) \geq \inf_{\rho \in \mathcal{L}_N} \left\{ \left(\frac{1}{2S} - \|v_S\|_{3/2} \right) \|\rho\|_3 + \int_{\mathbb{R}^3} v_\infty \rho \, d\vec{r} \right\} \geq -N \|v_\infty\|_\infty.$$

This proves the claim for $E^\lambda(v, \vec{A})$ and by compatibility also for $\bar{E}^\lambda(u, \vec{A})$. \square

Remark 8. The reflexive space $X_{\mathbb{R}}^* \times Y_{\mathbb{R}}^*$ is a subset of the domain given in Lemma 21 and corresponds to a compatible density space by Proposition 18. It follows by an adaptation of the proof of Lemma 21 that $E^\lambda(v, \vec{A})$ and $\bar{E}^\lambda(u, \vec{A})$ are finite on $X_{\mathbb{R}}^* \times Y_{\mathbb{R}}^*$. In this case when we just assume $(\rho, \vec{j}) \in L^3 \times \bar{L}^{3/2}$, however, it is crucial to exploit that the minimization can be restricted to N -representable density pairs, such that

$$\int_{\mathbb{R}^3} v \rho \, d\vec{r} = \int_{\mathbb{R}^3} (v - \varphi) \rho \, d\vec{r} + \int_{\mathbb{R}^3} \varphi \rho \, d\vec{r} \geq -\|v - \varphi\|_{3/2} \|\rho\|_3 - N \|\varphi\|_\infty,$$

where $\varphi \in \mathcal{C}_0^\infty$ has been chosen such that $v - \varphi$ has sufficiently small $L^{3/2}$ norm.

3 Regularization and the Kohn–Sham iteration scheme

In our previous work⁹ the general theory of a quantum system described by the (density) variable $b \in B$, where B is a reflexive ($B^{**} = B$) and strictly convex Banach space, was presented. In this theory the general problem

$$E(b^*) = \inf_{b \in B} \{f(b) + \langle b^*, b \rangle\}$$

(here $\langle b^*, b \rangle$ denotes the dual pairing that is not necessary given by an integral) with $f \in \Gamma_0(B)$ is studied. We here wish to apply this structure to paramagnetic CDFT, i.e., $b = (\rho, \vec{j})$, $b^* = (u, \vec{A})$, and $f(b) = F^\lambda(\rho, \vec{j})$. We choose the density space from Definition 8

$$X_{\mathbb{R}} \times Y_{\mathbb{R}} = L^3 \times \bar{L}^{3/2}$$

that was already discussed in Proposition 18 to meet the requirements of strict convexity, reflexivity, and compatibility. The dual potential space then is

$$X_{\mathbb{R}}^* \times Y_{\mathbb{R}}^* = L^{3/2} \times \bar{L}^3.$$

Another way to gain reflexivity is by limiting the spatial domain to a bounded set $\Omega \subsetneq \mathbb{R}^3$. Then $X =$

$L^1(\Omega) \cap L^3(\Omega)$, $Y = \bar{L}^1(\Omega) \cap \bar{L}^{3/2}(\Omega)$ automatically collapse to $X = L^3(\Omega)$, $Y = \bar{L}^{3/2}(\Omega)$, which are again reflexive. We also want to remark that in this case the usual Coulomb potential is included in $L^{3/2}(\Omega)$ since it is in $L_{\text{loc}}^{3/2}(\mathbb{R}^3)$.

Furthermore, compatibility of $X_{\mathbb{R}} \times Y_{\mathbb{R}}$ gives that the concave \bar{E}^λ (restricted to $X_{\mathbb{R}}^* \times Y_{\mathbb{R}}^*$) can be defined. By Lemma 21 and Remark 8, this energy is also bounded below. To connect the CDFT functionals $F_{\text{VR}}^\lambda(\rho, \vec{j})$ (or $F_{\text{VR,DM}}^\lambda(\rho, \vec{j})$) and $\bar{E}^\lambda(u, \vec{A})$, we have already introduced the Legendre–Fenchel transformation in Eqs. (13) and (14). It also relates the Lieb functional F^λ and the concave energy functional \bar{E}^λ vice versa as a conjugate pair. Also note that $F^\lambda \in \Gamma_0(X_{\mathbb{R}} \times Y_{\mathbb{R}})$ by Lemma 12. The next step lies in another type of transformation that makes F^λ functionally differentiable too, which is achieved by the Moreau–Yosida regularization.

3.1 Moreau–Yosida regularization

The original problem given in Eq. (14) of finding a ground-state density pair $(\rho, \vec{j}) \in X_{\mathbb{R}} \times Y_{\mathbb{R}}$ by minimizing the *convex* functional F^λ plus the potential energy can be restated using sub-/superdifferentials (Definition 2 in Ref. 9), both denoted ∂ . It means selecting (ρ, \vec{j}) from the superdifferential of \bar{E}^λ at $(u, \vec{A}) \in X_{\mathbb{R}}^* \times Y_{\mathbb{R}}^*$. Through the Legendre–Fenchel transformation Eq. (13), the same is possible for F^λ . With a minus sign in front, the potential pair (u, \vec{A}) yielding the ground state lies in the subdifferential of F^λ at (ρ, \vec{j}) (see Lemmas 3 and 4 in Ref. 9),

$$(\rho, \vec{j}) \in \partial \bar{E}^\lambda(u, \vec{A}) \iff -(u, \vec{A}) \in \partial F^\lambda(\rho, \vec{j}). \quad (22)$$

This statement can be seen as a more general reformulation of the HK theorem, but only with a (u, \vec{A}) potential pair, which is different from the physical (v, \vec{A}) setting.

The generalized notions of differentiability for convex/concave functionals involve the difficulty of non-existence or non-uniqueness. Sub- and superdifferentials are set-valued and can thus be empty or contain many elements. It is thus beneficial to “smooth out” the functional F^λ in such a way that it is differentiable, which implies that the subdifferential contains only one single element. (Note that in infinite dimensions only for a continuous functional a single element in the subdifferential implies differentiability.) This is achieved by the Moreau–Yosida regularization of $F^\lambda \in \Gamma_0(X_{\mathbb{R}} \times Y_{\mathbb{R}})$, for $\varepsilon > 0$ given by

$$F_\varepsilon^\lambda(\rho, \vec{j}) = \inf_{(\sigma, \vec{k}) \in X_{\mathbb{R}} \times Y_{\mathbb{R}}} \left\{ F^\lambda(\sigma, \vec{k}) + \frac{1}{2\varepsilon} \|(\rho, \vec{j}) - (\sigma, \vec{k})\|^2 \right\}. \quad (23)$$

Since F^λ is convex, the new functional F_ε^λ is convex as well and now also functionally differentiable by Theorem 9 in Ref. 9. This regularized functional then serves as the basis for defining a new energy functional $\bar{E}_\varepsilon^\lambda$

through the Legendre–Fenchel transformation Eq. (14) again

$$\bar{E}_\varepsilon^\lambda(u, \vec{A}) = \inf_{(\rho, \vec{j}) \in X_{\text{R}} \times Y_{\text{R}}} \left\{ F_\varepsilon^\lambda(\rho, \vec{j}) + \langle \vec{A}, \vec{j} \rangle + \langle u, \rho \rangle \right\}.$$

Note carefully that $\bar{E}_\varepsilon^\lambda$ is not the Moreau–Yosida regularization of some functional, but instead the Legendre–Fenchel conjugate of the regularized functional F_ε^λ . Then Theorem 10 in Ref. 9 can be used to relate the two energy functionals through

$$\bar{E}^\lambda(u, \vec{A}) = \bar{E}_\varepsilon^\lambda(u, \vec{A}) + \frac{\varepsilon}{2} \|(u, \vec{A})\|^2. \quad (24)$$

Since the infimum in the definition Eq. (23) is always uniquely attained at some $(\rho_\varepsilon, \vec{j}_\varepsilon) \in X_{\text{R}} \times Y_{\text{R}}$ (see Sec. 2.2.3 in Ref. 50), we can define a mapping $(\rho, \vec{j}) \mapsto (\rho_\varepsilon, \vec{j}_\varepsilon)$ that is called the proximal mapping, i.e.,

$$(\rho_\varepsilon, \vec{j}_\varepsilon) = \text{prox}_{\varepsilon F}(\rho, \vec{j}) \quad (25)$$

and

$$F_\varepsilon^\lambda(\rho, \vec{j}) = F^\lambda(\rho_\varepsilon, \vec{j}_\varepsilon) + \frac{1}{2\varepsilon} \|(\rho, \vec{j}) - (\rho_\varepsilon, \vec{j}_\varepsilon)\|^2.$$

The proximal mapping maps density pairs that are solutions of the regularized problem $(\rho, \vec{j}) \in \partial \bar{E}_\varepsilon^\lambda(u, \vec{A})$ back to solutions of the corresponding unregularized problem, $\text{prox}_{\varepsilon F}(\rho, \vec{j}) \in \partial \bar{E}^\lambda(u, \vec{A})$ by Corollary 11 in Ref. 9. Furthermore, the original functional F^λ is subdifferentiable at $(\rho_\varepsilon, \vec{j}_\varepsilon) \in X_{\text{R}} \times Y_{\text{R}}$, which means that the density pair $(\rho_\varepsilon, \vec{j}_\varepsilon)$ is v -representable. We note that by Theorem 9 in Ref. 9

$$\nabla F_\varepsilon^\lambda(\rho, \vec{j}) = \varepsilon^{-1} \mathcal{J}(\rho - \rho_\varepsilon, \vec{j} - \vec{j}_\varepsilon) \in \partial F^\lambda(\rho_\varepsilon, \vec{j}_\varepsilon),$$

where $\mathcal{J} : X_{\text{R}} \times Y_{\text{R}} \rightarrow X_{\text{R}}^* \times Y_{\text{R}}^*$ is the duality map that is always homogeneous (Definition 7 in Ref. 9). By Proposition 1.117 in Ref. 50, it is further bijective in the present setting of reflexive and strictly convex Banach spaces (including their duals). Letting $\mathcal{J} = (\mathcal{J}_{X_{\text{R}}}, \mathcal{J}_{Y_{\text{R}}})$, we get from $-(u, \vec{A}) = \nabla F_\varepsilon^\lambda(\rho, \vec{j})$ that

$$\begin{aligned} u &= -\varepsilon^{-1} \mathcal{J}_{X_{\text{R}}}(\rho - \rho_\varepsilon) \in X_{\text{R}}^*, \\ \vec{A} &= -\varepsilon^{-1} \mathcal{J}_{Y_{\text{R}}}(\vec{j} - \vec{j}_\varepsilon) \in Y_{\text{R}}^*, \end{aligned}$$

which straightforwardly transforms to

$$\begin{aligned} \rho &= \rho_\varepsilon - \varepsilon \mathcal{J}_{X_{\text{R}}}^{-1}(u) \in X_{\text{R}}, \\ \vec{j} &= \vec{j}_\varepsilon - \varepsilon \mathcal{J}_{Y_{\text{R}}}^{-1}(\vec{A}) \in Y_{\text{R}}. \end{aligned} \quad (26)$$

Here the compatibility of X_{R} as given by Proposition 18 again becomes important since it implies that we can decompose u as

$$u = \left(u - \frac{1}{2} |\vec{A}|^2 \right) + \frac{1}{2} |\vec{A}|^2 =: v + \frac{1}{2} |\vec{A}|^2. \quad (27)$$

Recall that for all $u \in L^{3/2}$ and $\vec{A} \in \vec{L}^3$ we have

$$u - \frac{1}{2} |\vec{A}|^2 \in L^{3/2}.$$

We conclude this section by discussing a Hilbert (H) space formulation. A direct adaptation of the approach taken in Ref. 4 is to choose

$$X_{\text{H}} \times Y_{\text{H}} = L^2 \times \vec{L}^2,$$

i.e., the Hilbert space built up from four copies of L^2 . The regularization presented in Ref. 4 for standard DFT can then be directly applied to the four-vector (ρ, \vec{j}) instead of just the particle density ρ . We note that $\psi \in \mathcal{H}^2$ is sufficient to obtain $j^k \in L^2$, see Proposition 4. However, when $X \times Y = X_{\text{H}} \times Y_{\text{H}}$ the density and potential spaces are not compatible in the meaning of Definition 13. This causes a problem for the regularization procedure of the Legendre–Fenchel pair F^λ and \bar{E}^λ , because we cannot decompose u as in Eq. (27) any more. To see this, note that $|\vec{A}|^2$ cannot in general be assumed to satisfy $|\vec{A}|^2 \in X_{\text{H}}$, which in turn would yield the desired $v \in X_{\text{H}}$. Consequently, we cannot obtain the physical setting of $E^\lambda(v, \vec{A})$ from the Moreau–Yosida setting of $\bar{E}^\lambda(u, \vec{A})$. This again highlights the usefulness of the more general reflexive Banach space formulation that allows a compatible choice of function spaces and makes a regularized paramagnetic CDFT possible.

3.2 Regularized Kohn–Sham iteration scheme in CDFT

We now revisit the KS iteration scheme that we previously analyzed for generic Banach spaces.⁹ Due to the similarity to the Optimal Damping Algorithm^{51,52} constructed for an unregularized setting, we baptize this iteration scheme the Moreau–Yosida Kohn–Sham Optimal Damping Algorithm (MYKSODA).

Again, let $X \times Y = X_{\text{R}} \times Y_{\text{R}}$ so that the space of densities is compatible, reflexive, and strictly convex. The latter two properties are also fulfilled by the dual $X^* \times Y^* = X_{\text{R}}^* \times Y_{\text{R}}^*$. From Sec. 3.1, the ground-state problem Eq. (14) can be reformulated in terms of sub- and superdifferentials

$$(\rho, \vec{j}) \in \partial \bar{E}^\lambda(u, \vec{A}) \iff -(u, \vec{A}) \in \partial F^\lambda(\rho, \vec{j}).$$

The regularized functionals F_ε^λ and $\bar{E}_\varepsilon^\lambda$ then allow the same relation with the benefit that F_ε^λ is now differentiable. This means we can switch from the subdifferential ∂ to the gradient ∇ of F_ε^λ , yet this is not permitted for $\bar{E}_\varepsilon^\lambda$.

We set up two problems side by side, the interacting problem with $\lambda = 1$ and the non-interacting reference

problem corresponding to $\lambda = 0$, i.e.,

$$\begin{aligned} (\rho_{\text{reg}}, \vec{j}_{\text{reg}}) \in \partial \bar{E}_\varepsilon^1(u_{\text{ext}}, \vec{A}_{\text{ext}}) \\ \iff -(u_{\text{ext}}, \vec{A}_{\text{ext}}) = \nabla F_\varepsilon^1(\rho_{\text{reg}}, \vec{j}_{\text{reg}}), \end{aligned} \quad (28)$$

$$\begin{aligned} (\rho_{\text{reg}}, \vec{j}_{\text{reg}}) \in \partial \bar{E}_\varepsilon^0(u_{\text{KS}}, \vec{A}_{\text{KS}}) \\ \iff -(u_{\text{KS}}, \vec{A}_{\text{KS}}) = \nabla F_\varepsilon^0(\rho_{\text{reg}}, \vec{j}_{\text{reg}}). \end{aligned} \quad (29)$$

In the setting of (regularized) KS theory the external potential pair $(u_{\text{ext}}, \vec{A}_{\text{ext}})$ is fixed and $(u_{\text{KS}}, \vec{A}_{\text{KS}})$ is to be determined under the assumption that both problems give the same (regularized) ground-state density pair $(\rho_{\text{reg}}, \vec{j}_{\text{reg}})$. We have highlighted the λ dependence by using “ext” for $\lambda = 1$ and by “KS” for $\lambda = 0$ (but for the functionals \bar{E} and F we keep 1 and 0). By combining Eqs. (28) and (29) we arrive at

$$\begin{aligned} (u_{\text{KS}}, \vec{A}_{\text{KS}}) &= (u_{\text{ext}}, \vec{A}_{\text{ext}}) \\ &+ \nabla F_\varepsilon^1(\rho_{\text{reg}}, \vec{j}_{\text{reg}}) - \nabla F_\varepsilon^0(\rho_{\text{reg}}, \vec{j}_{\text{reg}}) \end{aligned}$$

from which the iteration scheme will be derived by replacing the unknown variables by sequences. Let (ρ_i, \vec{j}_i) be the element of a sequence towards the (regularized) ground-state density pair $(\rho_{\text{reg}}, \vec{j}_{\text{reg}})$, and thus the next step towards the KS potential pair $(u_{\text{KS}}, \vec{A}_{\text{KS}})$ follows by

$$\begin{aligned} (u_{i+1}, \vec{A}_{i+1}) &= (u_{\text{ext}}, \vec{A}_{\text{ext}}) \\ &+ \nabla F_\varepsilon^1(\rho_i, \vec{j}_i) - \nabla F_\varepsilon^0(\rho_i, \vec{j}_i). \end{aligned} \quad (30)$$

The expression $\nabla(F_\varepsilon^1 - F_\varepsilon^0)$, which can be identified as a “Hartree exchange-correlation potential”, is where the usual approximation techniques of DFT enter. Yet in the domain of CDFT the variety of tried and tested functionals is meager^{53–55} compared to the wealth of options in conventional DFT.⁵⁶

The second major step in the iteration scheme is then the solution of the non-interacting reference system (instead of the computationally difficult interacting problem) which selects a ground-state density pair $(\rho_{i+1}, \vec{j}_{i+1})$ corresponding to the approximated KS potential pair (u_{i+1}, \vec{A}_{i+1}) that then serves as the next input in Eq. (30)

$$(\rho_{i+1}, \vec{j}_{i+1}) \in \partial \bar{E}_\varepsilon^0(u_{i+1}, \vec{A}_{i+1}). \quad (31)$$

That this (super)differential is indeed always non-empty, meaning that the associated ground-state problem has at least one solution, can be shown by using the result of finiteness of \bar{E}^λ from Lemma 21 (see the proof of Theorem 12 in Ref. 9).

The iteration stops when

$$(u_{\text{ext}}, \vec{A}_{\text{ext}}) = -\nabla F_\varepsilon^1(\rho_i, \vec{j}_i),$$

which means that (ρ_i, \vec{j}_i) solves the original interacting ground-state problem with external potential pair

$(u_{\text{ext}}, \vec{A}_{\text{ext}})$. This is so because this condition is equivalent to

$$(u_{i+1}, \vec{A}_{i+1}) = -\nabla F_\varepsilon^0(\rho_i, \vec{j}_i)$$

by Eq. (30), and the next step given by Eq. (31) would just yield the same density pair $(\rho_{i+1}, \vec{j}_{i+1}) = (\rho_i, \vec{j}_i)$ again. In such a case, or if we decide we have converged close enough to the supposed ground-state density pair $(\rho_{\text{reg}}, \vec{j}_{\text{reg}})$ of the regularized problem, a fixed relation between this solution and the solution of the unregularized problem is established by Eq. (26).

The question of convergence of the sequences $\{(\rho_i, \vec{j}_i)\}$ and $\{(u_i, \vec{A}_i)\}$ with respect to the Banach space topologies of $X_{\text{R}} \times Y_{\text{R}}$ and $X_{\text{R}}^* \times Y_{\text{R}}^*$ is immediately raised. The authors have answered this in Ref. 9, but only in a weak sense. More precisely, the associated energy sequence

$$F_\varepsilon^1(\rho_i, \vec{j}_i) + \langle \vec{A}_{\text{ext}}, \vec{j}_i \rangle + \langle u_{\text{ext}}, \rho_i \rangle$$

can be guaranteed to converge to *some* value larger or equal the correct value of the regularized energy functional $\bar{E}_\varepsilon^1(u_{\text{ext}}, \vec{A}_{\text{ext}})$. Arguably this is not what one expects from a well-formed KS iteration, where convergence to the correct ground-state density pair is the obvious aim. Further, such convergence in terms of energy can only be guaranteed if an additional step is inserted into the scheme consisting of Eqs. (30) and (31), coined “optimal damping”,^{6,51,57} that limits the step of the new density in such a way that the energy value assuredly decreases. It should be noted that a very recent development¹¹ treating the finite dimensional case finally proves full convergence of the exact KS iteration.

3.3 Weak-type convergence of MYK-SODA

We now have all the ingredients we need for an application of Theorem 12 in Ref. 9, which constitutes the main theoretical result of this work.

Theorem 22. *For the density spaces $X_{\text{R}} \times Y_{\text{R}} = L^3 \times \bar{L}^{3/2}$, and the corresponding dual for potential pairs $X_{\text{R}}^* \times Y_{\text{R}}^* = L^{3/2} \times \bar{L}^3$, a well-defined KS iteration can be set up for the energy functional \bar{E}^λ from Eq. (2). It starts with a fixed potential pair $(u_{\text{ext}}, \vec{A}_{\text{ext}}) \in X_{\text{R}}^* \times Y_{\text{R}}^*$ by setting $(u_1, \vec{A}_1) = (u_{\text{ext}}, \vec{A}_{\text{ext}})$ and selecting $(\rho_1, \vec{j}_1) \in \partial \bar{E}_\varepsilon^0(u_{\text{ext}}, \vec{A}_{\text{ext}})$. Then iterate $i = 1, 2, \dots$ according to:*

(a) Set

$$\begin{aligned} (u_{i+1}, \vec{A}_{i+1}) &= (u_{\text{ext}}, \vec{A}_{\text{ext}}) \\ &+ \nabla F_\varepsilon^1(\rho_i, \vec{j}_i) - \nabla F_\varepsilon^0(\rho_i, \vec{j}_i) \end{aligned}$$

and stop if

$$(u_{i+1}, \vec{A}_{i+1}) = -\nabla F_\varepsilon^0(\rho_i, \vec{j}_i) = (u_{\text{KS}}, \vec{A}_{\text{KS}}).$$

(b) Select $(\rho'_{i+1}, \vec{j}'_{i+1}) \in \partial \bar{E}_\varepsilon^0(u_{i+1}, \vec{A}_{i+1})$.

(c) Choose $t_i \in (0, 1]$ maximally such that for

$$(\rho_{i+1}, \vec{j}_{i+1}) = (\rho_i, \vec{j}_i) + t_i((\rho'_{i+1}, \vec{j}'_{i+1}) - (\rho_i, \vec{j}_i))$$

one still has

$$\begin{aligned} \frac{d}{dt_i} \left[F_\varepsilon^1(\rho_{i+1}, \vec{j}_{i+1}) + \langle (u_{\text{ext}}, \vec{A}_{\text{ext}}), (\rho_{i+1}, \vec{j}_{i+1}) \rangle \right] \\ = \langle \nabla F_\varepsilon^1(\rho_{i+1}, \vec{j}_{i+1}) + (u_{\text{ext}}, \vec{A}_{\text{ext}}), \\ (\rho'_{i+1}, \vec{j}'_{i+1}) - (\rho_i, \vec{j}_i) \rangle \leq 0. \end{aligned} \quad (32)$$

Then the strictly descending sequence

$$\{F_\varepsilon^1(\rho_i, \vec{j}_i) + \langle \vec{A}_{\text{ext}}, \vec{j}_i \rangle + \langle u_{\text{ext}}, \rho_i \rangle\}_i$$

converges as a sequence of real numbers to

$$\begin{aligned} e_\varepsilon(u_{\text{ext}}, \vec{A}_{\text{ext}}) &= \inf_i \{F_\varepsilon^1(\rho_i, \vec{j}_i) + \langle \vec{A}_{\text{ext}}, \vec{j}_i \rangle + \langle u_{\text{ext}}, \rho_i \rangle\} \\ &\geq \bar{E}_\varepsilon^1(u_{\text{ext}}, \vec{A}_{\text{ext}}). \end{aligned}$$

Thus,

$$e_\varepsilon(u_{\text{ext}}, \vec{A}_{\text{ext}}) + \frac{\varepsilon}{2} \|(u_{\text{ext}}, \vec{A}_{\text{ext}})\|^2$$

is an upper bound for the ground-state energy $\bar{E}_\varepsilon^1(u_{\text{ext}}, \vec{A}_{\text{ext}})$.

Proof. To be able to apply Theorem 12 from Ref. 9 we have to make sure that $X_{\text{R}} \times Y_{\text{R}}$ and $X_{\text{R}}^* \times Y_{\text{R}}^*$ are reflexive and strictly convex, the non-interacting energy functional \bar{E}^0 needs to be finite on all of $X_{\text{R}}^* \times Y_{\text{R}}^*$, and $F^\lambda, \bar{E}^\lambda$ must form a convex-concave pair linked by the Legendre–Fenchel transformation. Now, Proposition 18 shows that the chosen density spaces are indeed reflexive, all L^p with $1 < p < \infty$ are strictly convex anyway, and also that they are compatible. Compatibility gives that the energy functional can be transformed to a concave \bar{E}^λ by Eq. (12) that links to a convex Lieb functional F^λ by Eq. (13) (Legendre–Fenchel transformation). Then Lemma 21 and Remark 8 prove that \bar{E}^0 is indeed finite on $X_{\text{R}}^* \times Y_{\text{R}}^*$. With the results from Theorem 12 in Ref. 9 we get a strictly decreasing and converging sequence

$$\{F_\varepsilon^1(\rho_i, \vec{j}_i) + \langle \vec{A}_{\text{ext}}, \vec{j}_i \rangle + \langle u_{\text{ext}}, \rho_i \rangle\}_i$$

with the given lower bound. The transformed energy bound follows directly from Eq. (24). \square

Remark 9. Any candidate for a possible ground-state density pair from the iteration defined in Theorem 22 can be transformed to a solution of the corresponding “physical” unregularized problem with the help of

Eq. (26). But it has not been proven that the iteration actually converges in terms of densities and potentials as elements of the given Banach spaces and dual spaces *or* that if it converges, it actually reaches the ground-state density pair $(\rho_{\text{reg}}, \vec{j}_{\text{reg}})$ and the associated KS potential pair $(u_{\text{KS}}, \vec{A}_{\text{KS}})$.

Remark 10. We already discussed $(u_{\text{ext}}, \vec{A}_{\text{ext}}) = -\nabla F_\varepsilon^1(\rho_i, \vec{j}_i)$ as a stopping condition for the iteration before. To have such a stopping condition is important for at least a possible convergence to the correct ground-state density pair that is then a fixed point. Note that we still have the appearance of a whole set of ground-state density pairs in step (b), signifying that degeneracy is admitted. Thus, it is more beneficial to look at the sequence of potential pairs (u_i, \vec{A}_i) that, if the stopping condition is eventually reached, gives the correct KS potential for *some* ground-state densities. This is a reason why it was important to switch to differentiable functionals F_ε^λ through regularization: To be able to define a unique sequence of potentials that can converge to the KS potential. See also Ref. 10 where the traditional iteration in density space is supplemented with a *bivariate* formalism.

Remark 11. Note that step (c) in the KS iteration scheme above corresponds to a line-search between the points (ρ_i, \vec{j}_i) and $(\rho'_{i+1}, \vec{j}'_{i+1})$. If more solutions of the regularized, non-interacting reference system from step (b) are taken into account (degeneracy), then step (c) gets generalized to a search over a convex polytope.

3.4 Kohn–Sham iteration scheme for two-electron systems

As only a partial convergence result is presently available for the KS algorithm—leaving open the possibility that it does not always converge to the right energy and potential—it is interesting to consider a case where non-interacting N - and v -representability issues pose a challenge for the algorithm. Consider a formulation of paramagnetic CDFT for singlet ground states and restrict attention to a two-electron system. Then the (unregularized) non-interacting KS system is represented by a single orbital and its vorticity,

$$\vec{v}_{\text{KS}} = \nabla \times \frac{\vec{j}_{\text{KS}}}{\rho_{\text{KS}}},$$

vanishes if differentiability is assumed. For the small set of KS potentials that yield ground-state degeneracies, this can be circumvented by allowing the KS system to be represented by a mixed state. However, for most KS potentials, the KS ground state is unique and has trivial vorticity. On the other hand, for most external potentials, the correlated ground state of the interacting system has a nontrivial vorticity. Hence, most of the ground-state densities are not non-interacting ($\lambda = 0$) N -representable.

This situation poses an interesting challenge for the KS iteration scheme as the vorticity of the KS system cannot develop gradually. Until the algorithm has constructed potentials that yield an exact ground-state degeneracy, the vorticity is trivial, and it is not clear how “visible” the corresponding degrees of freedom are to the optimization algorithm. Moreau–Yosida regularization alleviates the challenge somewhat, since the relevant densities have contributions from the potentials, complicating the non-interacting N -representability conditions. With regularization, the relevant vorticity that should reproduce the interacting system using Eq. (26) is

$$\vec{v}_{\text{KS,reg}} = \nabla \times \frac{\vec{j}_{\text{KS}} - \varepsilon \mathcal{J}_{\text{YR}}^{-1}(\vec{A}^{\text{KS}})}{\rho_{\text{KS}} - \varepsilon \mathcal{J}_{\text{XR}}^{-1}(u^{\text{KS}})}.$$

If counterexamples that prevent a full convergence proof exist at all, the type of system sketched above is a promising candidate for further analysis.

4 Numerical application to quantum ring

The theory of regularized CDFT can be directly applied to a one-dimensional quantum ring. Although this is a toy model, it is sufficiently rich to contain simple formal analogues of many aspects of CDFT for a three-dimensional spatial domain. We limit attention to two-electron systems in singlet spin states. The Hamiltonian is given by

$$H^\lambda(v, A) = \frac{1}{2} \sum_{j=1}^2 \left(-\frac{\hbar}{R} \frac{\partial}{\partial \theta_j} + A(\theta_j) \right)^2 + \sum_{j=1}^2 v(\theta_j) + \lambda W(\theta_1, \theta_2),$$

where R is the radius of the ring and the potentials v and A as well as the electron-electron repulsion W are considered functions of the angular position along the ring. Note that gradients and the vector potential only have tangential components and may therefore effectively be treated as scalars.

Because of the limitation to singlet states, the spatial wave function $\psi(\theta_1, \theta_2) = \psi(\theta_2, \theta_1)$ must be symmetric. Any uncorrelated state, e.g., a KS state, takes the form $\psi(\theta_1, \theta_2) = \phi(\theta_1)\phi(\theta_2)$ and is defined by a single orbital ϕ . The densities that arise from such an uncorrelated, single-orbital $\phi(\theta) = \sqrt{\rho(\theta)} e^{i\chi(\theta)}$ state must satisfy

$$R \int_0^{2\pi} \frac{j(\theta)}{\rho(\theta)} d\theta = \int_0^{2\pi} \nabla \chi(\theta) d\theta = 2\pi m,$$

with m an integer if $\rho > 0$ everywhere. Note that $\chi(\theta + 2\pi) = \chi(\theta) + 2\pi m$ is in general a multivalued phase function. By contrast, a correlated state can well

give rise to a fractional value of m . This is the quantum ring analogue of the fact that vorticity is trivial for single-orbital systems in a three-dimensional spatial domain.

Next, in order to study regularized CDFT numerically, we discretize the quantum ring into N_G uniformly spaced grid points. The approach described below is implemented in a Matlab program named MYRING.⁵⁸ We replace the Laplacian by the standard second-order finite difference expression

$$\frac{\partial^2}{\partial \theta^2} \phi(\theta) \leftrightarrow \frac{\phi(\theta_{j+1}) - 2\phi(\theta_j) + \phi(\theta_{j-1}))}{h^2},$$

where $h = 2\pi R/N_G$ is the grid spacing. The paramagnetic term is discretized using the symmetric first-order expression,

$$\frac{\partial}{\partial \theta} \phi(\theta) \leftrightarrow \frac{\phi(\theta_{j+1}) - \phi(\theta_{j-1}))}{2h}.$$

Defining the particle density and current density at grid point k by

$$\rho_k = h \sum_l |\psi(\theta_k, \theta_l)|^2, \\ j_k = -\frac{\hbar}{2} h \sum_l \psi(\theta_k, \theta_l)^* \frac{\psi(\theta_{k+1}, \theta_l) - \psi(\theta_{k-1}, \theta_l)}{2h} + \text{c.c.},$$

we can define a constrained-search functional as well as linear pairings between ρ and $u = v + A^2/2$ as well as j and A ,

$$\langle u, \rho \rangle = h \sum_k u_k \rho_k, \\ \langle A, j \rangle = h \sum_k A_k j_k.$$

The densities and potentials may all be regarded as vectors in \mathbb{R}^{N_G} . Because all norms in finite dimensions are mathematically equivalent, we can choose to endow all function spaces with the same Euclidean $l^2(N_G)$ norm without losing compatibility. However, the norms may not be *numerically* equivalent. Moreover, to connect to the continuum limit when $N_G \rightarrow \infty$, it is likely that one needs more carefully chosen norms. This is left for future studies.

The grid discretization makes it trivial to construct compatible finite-dimensional function spaces. However, this is not true in arbitrary basis expansions of density pairs and potential pairs. Unless the respective basis sets have special properties, compatibility is in general lost.

By solving the discretized Schrödinger equation, we then obtain the ground-state energy and the regularized energy

$$\bar{E}_\varepsilon^\lambda(u, A) = \bar{E}^\lambda(u, A) - \frac{\varepsilon}{2} \|u\|_2^2 - \frac{\varepsilon}{2} \|A\|_2^2.$$

The universal density functional can be computed from the Lieb variational principle

$$F_\varepsilon^\lambda(\rho, j) = \sup_{u, A} G_\varepsilon^\lambda(u, A; \rho, j), \quad (33)$$

with $G_\varepsilon^\lambda(u, A; \rho, j) = \bar{E}_\varepsilon^\lambda(u, A) - \langle u, \rho \rangle - \langle A, j \rangle$. We have found that a cutting-plane bundle method for convex optimization^{59,60} provides robust, though occasionally very slow convergence to the maximum value. In more detail, our implemented method maintains a “bundle” of data $(g_l, u_l, A_l, \sigma_l, k_l)$ from previous iterations. The bundle contains the function value $g_l = G_\varepsilon^\lambda(u_l, A_l; \rho, j)$ and a supergradient (σ_l, k_l) evaluated at (u_l, A_l) . Then a model function is defined by all the tangent planes encoded in the bundle,

$$Q_l(u, A) = g_l + \langle u - u_l, \sigma_l \rangle + \langle A - A_l, k_l \rangle. \quad (34)$$

The next sample point (u_{l+1}, A_{l+1}) is determined by maximizing the model function subject to a trust region constraint (to guard against Q_l being unbounded, as may happen in the first iterations).

The stopping criterion requires care. As mentioned above, because the Moreau–Yosida regularization is only applied to F^λ , the energy functional $\bar{E}_\varepsilon^\lambda$ is not more differentiable than the original \bar{E}^λ . Hence, there is no guarantee that G_ε^λ is differentiable with respect to the potentials at the maximum. This is particularly true for the KS potentials at $\lambda = 0$, where N -representability constraints become more visible. Hence, it is not feasible to rely on vanishing (super)gradients as a stopping criterion for the optimization. This is connected to ground-state degeneracy and can be diagnosed by computing the energy gap to the first excited state.

4.1 Kohn–Sham potentials from the Lieb variational principle

We consider a discretization with $N_G = 30$ grid points and set the electron-electron interaction to

$$W(\theta_1, \theta_2) = 3\sqrt{1 + \cos(\theta_1 - \theta_2)}.$$

We choose the external potentials

$$\begin{aligned} v_{\text{ext}}(\theta) &= \cos(\theta), \\ A_{\text{ext}}(\theta) &= 0.6, \end{aligned}$$

in order to obtain a non-trivial example that is nonetheless simple to specify. The external potentials are visualized in Fig. 1. The resulting Hamiltonian $H^1(v_{\text{ext}}, A_{\text{ext}})$ has a highly correlated ground state with densities (ρ, j) displayed in Fig. 2. Performing maximization in the Lieb variation principle (Eq. (13) or (33)) defining $F_{\varepsilon=0}^\lambda(\rho, j)$ yields KS potentials $(u_{\text{KS}}, A_{\text{KS}})$ as a by-product, visualized in Fig. 1. The Hamiltonian $H^0(v_{\text{KS}}, A_{\text{KS}})$, with $v_{\text{KS}} = u_{\text{KS}} - (A_{\text{KS}})^2/2$, has a two-fold ground-state degeneracy and one of these ground-

state densities is shown in Fig. 2. The vanishing gap is seen in Fig. 3 and results in a non-differentiable kink in the ground-state energy $\bar{E}_{\varepsilon=0}^0(u_{\text{KS}}, A_{\text{KS}})$. The interacting density pair (ρ, j) is a supergradient at this non-differentiable point, but it is neither a left- nor a right-derivative. Due to the limitation that our implementation is limited to pure states, and furthermore that the choice of degenerate eigenvector basis is not optimized, it is seen in Fig. 2 that the interacting ground-state density pair (ρ, j) is not reproduced exactly by the KS ground state. In general, exact reproduction requires mixed states.

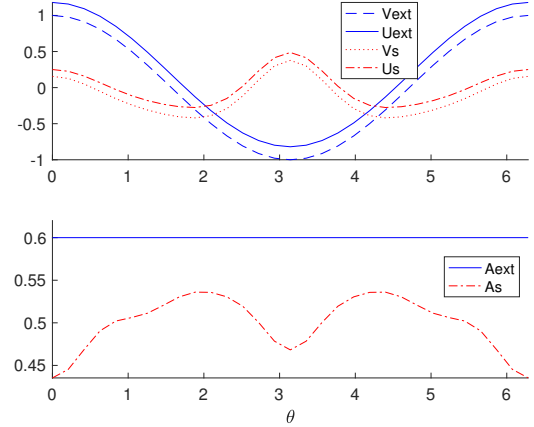


Figure 1: External potentials and KS potentials in the unregularized case ($\varepsilon = 0$).

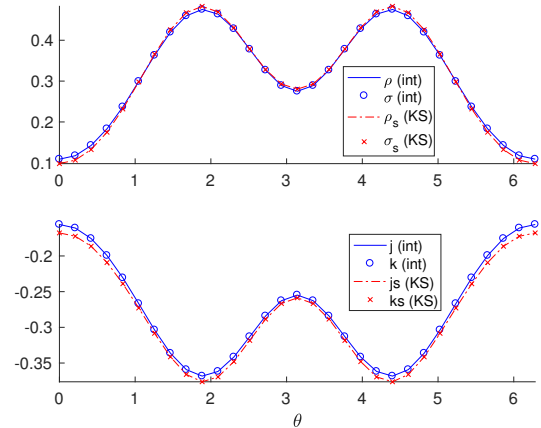


Figure 2: The unregularized ground-state density pair (ρ, j) for the correlated, interacting system subject to the external potentials together with the density pair $(\rho_{\text{KS}}, j_{\text{KS}})$ for the uncorrelated KS system. Because $\varepsilon = 0$, the regularized density pair $(\sigma = \rho - \varepsilon u_{\text{ext}}, k = j - \varepsilon A_{\text{ext}})$ trivially coincides with unregularized density pair. Note that (ρ, j) are nearly reproduced by the KS density pair $(\rho_{\text{KS}}, j_{\text{KS}})$, but failure of non-interacting N -representability prevents an exact match.

Next we illustrate the regularized setting by taking $\varepsilon = 0.1$. This relatively large regularization parameter is used to make the effects of regularization noticeable. It is now the pair $(\sigma = \rho - \varepsilon u_{\text{ext}}, k = j - \varepsilon A_{\text{ext}})$ that takes

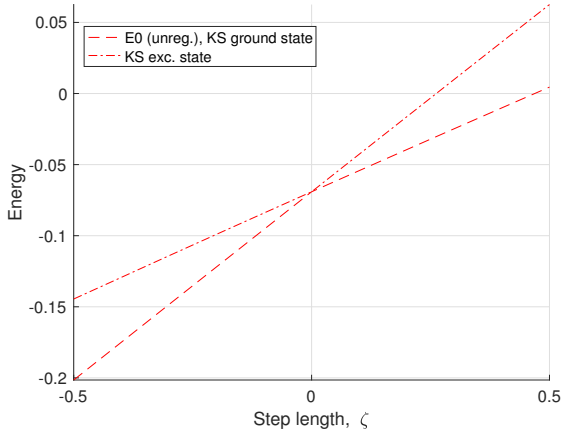


Figure 3: The unregularized KS ground-state energy $\bar{E}_\varepsilon^0(u_{\text{KS}} + \zeta h\sigma, A_{\text{KS}} + \zeta hk)$ and the first excited state as a function of the step length ζ . The non-differentiable kink at $\zeta = 0$ arises from a level crossing and the reference density pair from the interacting system corresponds to a particular supergradient at this kink.

over the role played by the density pair in the unregularized setting. In particular, the Lieb variation principle now yields a KS potential pair $(u_{\text{KS}}, A_{\text{KS}})$ such that $(\sigma_{\text{KS}} = \rho_{\text{KS}} - \varepsilon u_{\text{KS}}, k_{\text{KS}} = j_{\text{KS}} - \varepsilon A_{\text{KS}})$ coincides with the density pair (σ, k) , but $(\rho_{\text{KS}}, j_{\text{KS}}) \neq (\rho, j)$. Hence, the KS potentials shown in Fig. 4 are different from those in the unregularized setting (Fig. 1). The resulting densities are shown in Fig. 5.

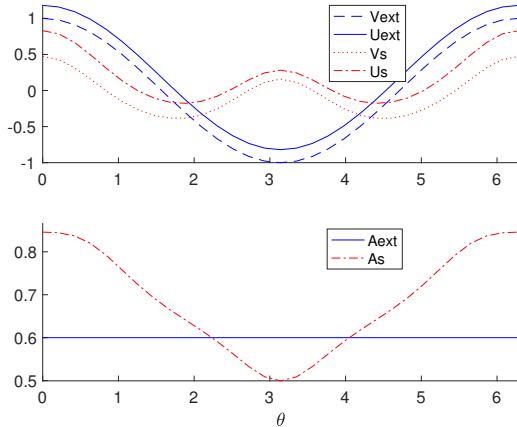


Figure 4: External potentials and KS potentials in the case of Moreau–Yosida regularization with parameter value $\varepsilon = 0.1$.

4.2 Kohn–Sham potentials from the iterative algorithm

In the previous section, the KS potentials were determined by first solving for correlated ground-state wave function of the interacting system, then constructing its densities, and finally plugging these densities into the Lieb variation principle. The iterative KS algorithm discussed in Sec. 3.2 above provides an alternative that

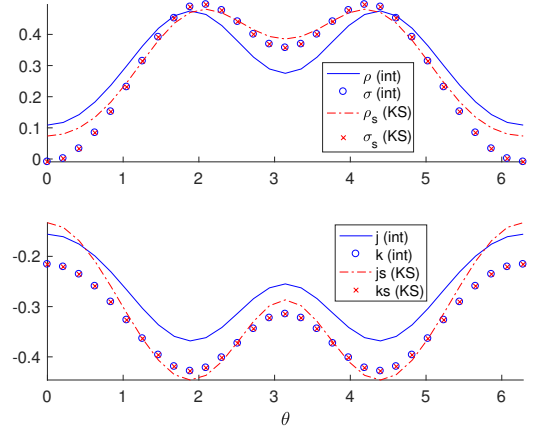


Figure 5: The ground-state density pair (ρ, j) for the correlated, interacting system subject to the external potentials together with the density pair $(\rho_{\text{KS}}, j_{\text{KS}})$ for the uncorrelated KS system. The regularized density pair $(\sigma = \rho - \varepsilon u_{\text{ext}}, k = j - \varepsilon A_{\text{ext}})$ is very nearly reproduced by the regularized KS density pair $(\sigma_{\text{KS}} = \rho_{\text{KS}} - \varepsilon u_{\text{KS}}, k_{\text{KS}} = j_{\text{KS}} - \varepsilon A_{\text{KS}})$, but failure of non-interacting N -representability prevents an exact match.

does not require any *a priori* information of the correlated ground state or its associated density. For simplicity, we have implemented the pure-state version of this algorithm, enabling us to see the consequences when a density pair is not representable by a pure ground state. The linesearch for the interpolation parameter t was implemented in the following way:

- (i) Successively try $t = 1, \frac{1}{2}, \frac{1}{4}, \frac{1}{8}, \dots$, until the criterion from the optimal damping step Eq. (32) is fulfilled.
- (ii) If already $t = 1$ fulfills Eq. (32), then use this value. Otherwise, let $t = 2^{-k}$ be the first parameter value such that Eq. (32) holds and estimate the critical t value by linear interpolation between $t = 2^{-k+1}$ and 2^{-k} . If the criterion is still not fulfilled at this t , perform another linear interpolation and choose the best of the sampled values.

The computation of gradients $\nabla F_\varepsilon^1(\rho, j)$ is done using the Lieb variation principle, with a maximum of 300 bundle optimization iterations and a convergence criterion of 10^{-5} for stopping earlier. When there is a degenerate ground state, the gradient criterion does not apply and we instead test for a small gap and stagnated bundle iterations. In cases of numerically very small, but non-zero gap between the ground state and first excited state, our implementation may fail to obtain an adequate solution from the (pure-state) Lieb variation principle. As the algorithm was not formulated to account for such failures, the energy seen in the KS iterations may not be bounded from below by the true energy $\bar{E}^1(u_{\text{ext}}, A_{\text{ext}})$, unless we override the reference density pair and instead use the actual density pair returned from the Lieb optimization.

Continuing with the same numerical example as in the previous section, we ran the KS iteration for different values of the regularization parameter. In the unregularized case, the consequences of failure of pure-state representability, both for the KS system and interacting systems corresponding to trial densities encountered in the course of the iterations, prevented a meaningful result. With Moreau–Yosida regularization, we were able to converge within the expected accuracy, given the finite precision of our implementation of the Lieb variation principle. In Fig. 6 the convergence of the energy difference,

$$\Delta E_i = F_\varepsilon^1(\rho_i, j_i) + \langle u_{\text{ext}}, \rho_i \rangle + \langle A_{\text{ext}}, j_i \rangle - \bar{E}_\varepsilon^1(u_{\text{ext}}, A_{\text{ext}}),$$

is shown for four different values, $\varepsilon = 0.05, 0.1, 0.2, 0.3$, of the regularization parameter. Fig. 7 shows the convergence of the gradient norm,

$$\|(u_{\text{ext}}, A_{\text{ext}}) + \nabla F_\varepsilon^1(\rho_i, j_i)\|_2,$$

which vanishes when the ground-state density of the interacting system has been reproduced. Although not encountered in the example studied here, small numerical inaccuracies especially in the Lieb variation principle lead to occasional small increases of the energy. The convergence is slow compared to experience with standard algorithms, such as Pulay’s DIIS,⁶¹ and approximate density functionals, as these result in quadratic convergence in favorable cases. However, most standard algorithms also lack formal convergence guarantees and have, for practical reasons, never been tested with the exact functional. An exception is the work by Wagner *et al.* that did explore convergence of the exact functional using an algorithm applied to one-dimensional systems.^{8,44} In Ref. 44 an adaptive choice of the damping (mixing) parameter was investigated, including discussions on line search and Hermite spline fit to the energy as a function of the damping parameter. (It is interesting to note that they use the curvature of the energy as information. In the regularized setting where derivatives are guaranteed to exist, the curvature is a key ingredient in the convergence proof of Ref. 11.) Furthermore, their study of an optimal damping parameter demonstrated numerically that convergence is more difficult for strongly correlated systems.

The present work is the first time a KS vector potential, corresponding to an exact CDFT functional, is calculated using a KS iteration scheme. As expected, the convergence in Figs. 6 and 7 is faster for larger values of the regularization parameter. This is partly due to the fact that the unregularized case features a KS system with vanishing gap and partly due to the increased regularity of the problem for larger ε .

Finally, Fig. 8 shows the calculated values of the damping parameter t as a function of iteration number. The parameter values vary substantially between

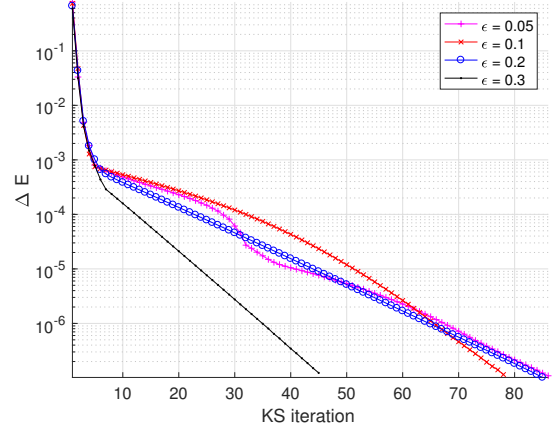


Figure 6: Energy convergence of KS iterations for different values of the regularization parameter.

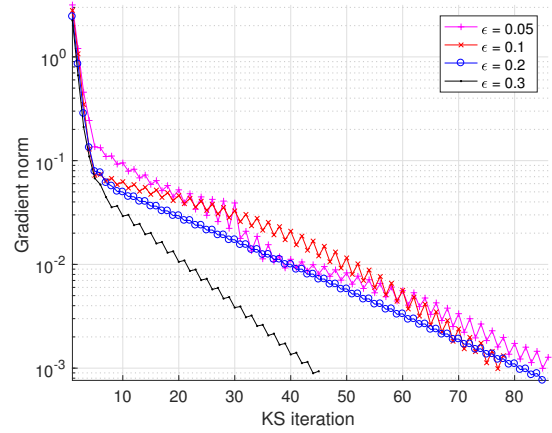


Figure 7: Convergence of the gradient norm in the KS iterations for different values of the regularization parameter.

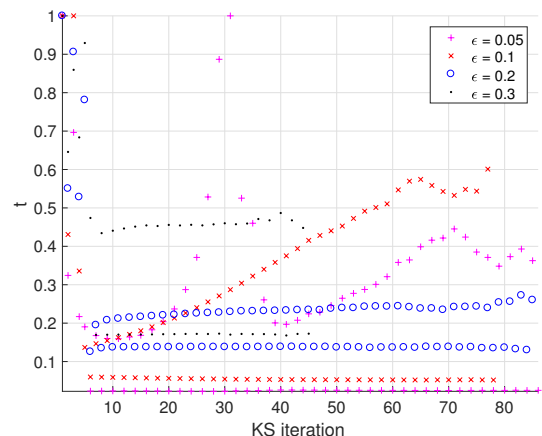


Figure 8: Calculated linesearch step t in the KS iterations.

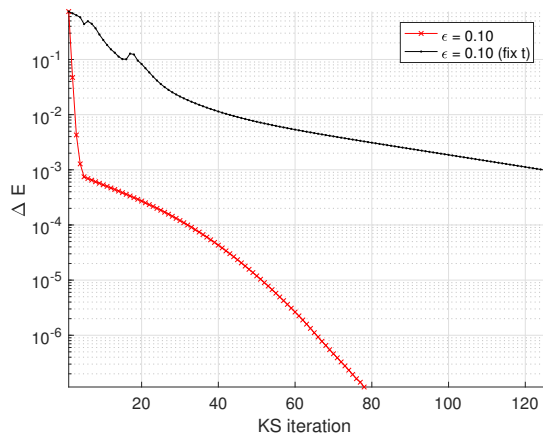


Figure 9: Comparison of KS iterations with maximal t determined in each iteration and with fixed $t = 0.05$.

the examples with different ε and also from one iteration to the next. In particular, for $\varepsilon = 0.1$ the MYKSODA iterations alternate between smaller $t \sim 0.05$ and larger values in the range $0.1 < t < 0.65$. A simpler iterative algorithm could use a fixed t in all iterations, as was done in Ref. 8. To explore this possibility, we fixed a conservative value $t = 0.05$ for the damping parameter. As seen in Fig. 9 this yields dramatically slower convergence, showing that t in general needs to be chosen adaptively.

5 Conclusions

We have given a comprehensive account of the rigorous formulation of Kohn–Sham theory for CDFT. An important point is that textbook treatments of DFT rely on ill-defined functional derivatives.³² However, recent work has demonstrated that functional derivatives can be made well-defined and rigorous using Moreau–Yosida regularization.^{4,9} We have extended that approach to functional differentiation in CDFT, enabling us to obtain well-defined Kohn–Sham potentials as well as an iteration scheme (MYKSODA). The presented MYKSODA is an algorithm for practical calculations in the setting of ground-state CDFT within a regularized framework. A toy model in the form of a quantum ring is solved numerically, and allowed a study of MYKSODA for the exact universal density functional. The calculations illustrate the performance of the algorithm and highlight the difference to iteration schemes with a constant damping factor. It is also the first implementation of a Moreau–Yosida regularized Kohn–Sham approach.

While our model was solved numerically with the exact functional, this is of course not feasible for more realistic settings where we must resort to density-functional approximations. This raises the question of how to develop such approximations for the Moreau–Yosida regularized setting, or alternatively, of how to compute the Moreau–Yosida regularization of well-established

density-functional approximations. This is an interesting topic for future investigation.

Central to the theory developed here was the concept of compatibility of spaces of densities and current densities. It allows a fully convex formulation of the theory and demands the use of Banach spaces for the basic variables. The respective \bar{L}^p constraints for current densities were determined optimally in order to complement knowledge from traditional DFT and previous work on CDFT. This article sets the stage for further inquiries into the field, such as the possible full convergence of the iteration scheme and the study of approximate (regularized) functionals for CDFT.

Acknowledgments

We thank an anonymous referee for improvements on our proof of Lemma 21. This work was supported by the Norwegian Research Council through the CoE Hylleraas Centre for Quantum Molecular Sciences Grant No. 262695. AL is grateful for the hospitality received at the Max Planck Institute for the Structure and Dynamics of Matter in Hamburg, while visiting MP and MR. MP acknowledges support by the Erwin Schrödinger Fellowship J 4107-N27 of the FWF (Austrian Science Fund) and is thankful for an invitation to the Hylleraas Centre just taking place writing this. AL and SK were supported by ERC-STG-2014 under grant agreement No. 639508. EIT was supported by the Norwegian Research Council through Grant No. 240674.

A A theorem on everywhere defined functionals on spaces of measurable functions

On any infinite-dimensional Banach space (assuming the axiom of choice) there exist everywhere defined linear maps that are unbounded. The following theorem shows that this cannot happen for linear functionals on spaces of measurable functions that are defined as integrals. The proof is based on a construction by D. Fischer.⁶²

Theorem 23. *Let B be a Banach space consisting of measurable functions $f : \mathbb{R}^n \rightarrow \mathbb{R}$. Let g be a measurable function. Then the functional $T : f \mapsto \int gf \, d\mu$ is in B^* if and only if for all $f \in B$,*

$$\left| \int fg \, d\mu \right| < +\infty.$$

Proof. Since a bounded linear functional must be everywhere defined, the only if part is trivial. Suppose g is measurable and that the integral $\int fg \, d\mu$ exists for all $f \in B$. For $n \in \mathbb{N}$, define a sequence of bounded

functions with bounded support,

$$g_n(x) = \begin{cases} 0, & \|x\| > n, \\ g(x), & \|x\| \leq n \text{ and } |g(x)| \leq n, \\ \frac{n}{|g(x)|} \cdot g(x), & \|x\| \leq n < |g(x)|. \end{cases}$$

Then g_n is measurable for all n , and $h_n(x) = g_n(x)f(x) \rightarrow h(x) = g(x)f(x)$ for all x . Moreover $|g_n(x)f(x)| \leq |h(x)|$ for all x , the latter function being integrable by assumption. By the dominated convergence theorem,

$$\int_{\mathbb{R}^n} f(x)g_n(x) \, d\mu \rightarrow \int_{\mathbb{R}^n} f(x)g(x) \, d\mu,$$

as $n \rightarrow +\infty$. Thus, the family of continuous linear functionals $T_n: f \mapsto \int f g_n \, d\mu$ is pointwise bounded.

The uniform boundedness principle states that a family $\{T_n\}$ of pointwise bounded linear functionals is in fact uniformly bounded. Thus, $\sup_n \|T_n\|_{B^*} < +\infty$. It then follows that

$$\|T\|_{B^*} \leq \sup_n \|T_n\|_{B^*} < +\infty.$$

Hence, $T \in B^*$. \square

References

- (1) Hohenberg, P.; Kohn, W. *Phys. Rev.* **1964**, *136*, B864–B871.
- (2) Lieb, E. H. *Int. J. Quantum Chem.* **1983**, *24*, 243–277.
- (3) Tellgren, E. I.; Kvaal, S.; Sagvolden, E.; Ekström, U.; Teale, A. M.; Helgaker, T. *Phys. Rev. A* **2012**, *86*, 062506.
- (4) Kvaal, S.; Ekström, U.; Teale, A. M.; Helgaker, T. *J. Chem. Phys.* **2014**, *140*, 18A518.
- (5) Kohn, W.; Sham, L. J. *Phys. Rev.* **1965**, *140*, A1133–A1138.
- (6) Cancès, E. *J. Chem. Phys.* **2001**, *114*, 10616–10622.
- (7) Cancès, E.; Kudin, K. N.; Scuseria, G. E.; Turinici, G. *J. Chem. Phys.* **2003**, *118*, 5364–5368.
- (8) Wagner, L. O.; Stoudenmire, E. M.; Burke, K.; White, S. R. *Phys. Rev. Lett.* **2013**, *111*, 093003.
- (9) Laestadius, A.; Penz, M.; Tellgren, E. I.; Ruggenthaler, M.; Kvaal, S.; Helgaker, T. *J. Chem. Phys.* **2018**, *149*, 164103.
- (10) Lammert, P. E. **2018**, preprint arXiv:1807.06125.
- (11) Penz, M.; Laestadius, A.; Tellgren, E. I.; Ruggenthaler, M. **2019**, preprint arXiv:1903.09579.
- (12) Vignale, G.; Rasolt, M. *Phys. Rev. Lett.* **1987**, *59*, 2360–2363.
- (13) Diener, G. *J. Phys.: Condens. Matter* **1991**, *3*, 9417–9428.
- (14) Capelle, K.; Vignale, G. *Phys. Rev. B* **2002**, *65*, 113106.
- (15) Laestadius, A.; Benedicks, M. *Int. J. Quantum Chem.* **2014**, *114*, 782–795.
- (16) Laestadius, A.; Benedicks, M. *Phys. Rev. A* **2015**, *91*, 032508.
- (17) Grayce, C. J.; Harris, R. A. *Phys. Rev. A* **1994**, *50*, 3089–3095.
- (18) Tellgren, E. I.; Laestadius, A.; Helgaker, T.; Kvaal, S.; Teale, A. M. *The Journal of Chemical Physics* **2018**, *148*, 024101.
- (19) Reimann, S.; Borgoo, A.; Tellgren, E. I.; Teale, A. M.; Helgaker, T. *J. Chem. Theory Comput.* **2017**, *13*, 4089–4100.
- (20) Pittalis, S.; Vignale, G.; Eich, F. G. *Phys. Rev. B* **2017**, *96*, 035141.
- (21) Ayers, P. W.; Golden, S.; Levy, M. *The Journal of Chemical Physics* **2006**, *124*, 054101.
- (22) Giesbertz, K. J.; Ruggenthaler, M. *Physics Reports* **2019**,
- (23) Tellgren, E. I. *Phys. Rev. A* **2018**, *97*, 012504.
- (24) Ruggenthaler, M. **2017**, preprint arXiv:1509.01417.
- (25) Ayers, P. W.; Fuentealba, P. *Phys. Rev. A* **2009**, *80*, 032510.
- (26) Sim, E.; Larkin, J.; Burke, K.; Bock, C. W. *The Journal of Chemical Physics* **2003**, *118*, 8140–8148.
- (27) Ayers, P. W. *Journal of Mathematical Physics* **2005**, *46*, 062107.
- (28) Higuchi, M.; Higuchi, K. *Phys. Rev. B* **2004**, *69*, 035113.
- (29) Laestadius, A.; Tellgren, E. I. *Phys. Rev. A* **2018**, *97*, 022514.
- (30) Levy, M. *Proc. Natl. Acad. Sci. USA* **1979**, *76*, 6062–6065.
- (31) Laestadius, A. *Int. J. Quantum Chem.* **2014**, *114*, 1445–1456.
- (32) Lammert, P. E. *Int. J. Quantum Chem.* **2007**, *107*, 1943–1953.
- (33) Lieb, E. H.; Loss, M. *Analysis*; American Mathematical Society, 2001.
- (34) Teschl, G. *Mathematical Methods in Quantum Mechanics*; American Mathematical Society, 2006.
- (35) van Tiel, J. *Convex analysis: an introductory text*; Wiley, 1984.
- (36) Hoffmann-Ostenhof, M.; Hoffmann-Ostenhof, T. *Phys. Rev. A* **1977**, *16*, 1782–1785.
- (37) Kato, T. *Trans. Amer. Math. Soc.* **1951**, *70*, 195–211.
- (38) Tokatly, I. V. *Phys. Rev. B* **2011**, *83*, 035127.
- (39) Tellgren, E. I.; Kvaal, S.; Helgaker, T. *Phys. Rev. A* **2014**, *89*, 012515.
- (40) Bates, J. E.; Furche, F. *J. Chem. Phys.* **2012**, *137*, 164105.
- (41) Englisch, H.; Englisch, R. *Physica A: Statistical Mechanics and its Applications* **1983**, *121*, 253 – 268.
- (42) Lieb, E. H.; Schrader, R. *Phys. Rev. A* **2013**, *88*, 032516.
- (43) Laestadius, A. *J. Math. Chem.* **2014**, *52*, 2581–

2595.

- (44) Wagner, L. O.; Baker, T. E.; Stoudenmire, E. M.; Burke, K.; White, S. R. *Phys. Rev. B* **2014**, *90*, 045109.
- (45) Hanner, O. *Arkiv för Matematik* **1956**, *3*, 239–244.
- (46) Valone, S. M. *The Journal of Chemical Physics* **1980**, *73*, 1344–1349.
- (47) Kvaal, S.; Helgaker, T. *J. Chem. Phys.* **2015**, *143*, 184106.
- (48) Kvaal, S.; Helgaker, T. *Unpublished manuscript*
- (49) Liu, T.-S.; Wang, J.-K. *Mathematica Scandinavica* **1969**, *23*, 241–251.
- (50) Barbu, V.; Precupanu, T. *Convexity and optimization in Banach spaces*, 4th ed.; Springer, 2012.
- (51) Cancès, E.; Le Bris, C. *Int. J. Quantum Chem.* **2000**, *79*, 82–90.
- (52) Cancès, E. *J. Chem. Phys.* **2001**, *114*, 10616–10622.
- (53) Vignale, G. *Adv. Quantum Chem.* **1990**, *21*, 235.
- (54) Tellgren, E. I.; Teale, A. M.; Furness, J. W.; Lange, K.; Ekström, U.; Helgaker, T. *The Journal of chemical physics* **2014**, *140*, 034101.
- (55) Furness, J. W.; Verbeke, J.; Tellgren, E. I.; Stopkowitz, S.; Ekström, U.; Helgaker, T.; Teale, A. M. *Journal of chemical theory and computation* **2015**, *11*, 4169–4181.
- (56) Burke, K. *The Journal of chemical physics* **2012**, *136*, 150901.
- (57) Cancès, E. In *Mathematical Models and Methods for Ab Initio Quantum Chemistry*; De Franceschi, M., Le Bris, C., Eds.; Lecture Notes in Chemistry; Springer, 2000; Vol. 74; pp 17–43.
- (58) MYring, a program for Moreau–Yosida regularization of a one-dimensional quantum ring. Available at <https://gitlab.com/et/myring>.
- (59) Cheney, E. W.; Goldstein, A. A. *Numer. Math.* **1959**, *1*, 253–268.
- (60) Kelley Jr., J. E. *J. SIAM* **1960**, *8*, 703–712.
- (61) Pulay, P. *J. Comput. Chem.* **1982**, *3*, 556–560.
- (62) Fischer, D. Discontinuous functionals on L^p . 2014; <https://math.stackexchange.com/questions/1008990>.