Lower semi-continuity of universal functional in paramagnetic current-density functional theory

Simen Kvaal, Andre Laestadius, Erik Tellgren, and Trygve Helgaker
Hylleraas Centre for Quantum Molecular Sciences, Department of Chemistry, University of Oslo, P.O. Box 1033 Blindern, N-0315 Oslo, Norway
(Dated: Friday 3 July 2020)

A cornerstone of current-density functional theory (CDFT) in its paramagnetic formulation is proven. After a brief outline of the mathematical structure of CDFT, the lower semi-continuity and expectation valuedness of the CDFT constrained-search functional is proven, meaning that there is always a minimizing density matrix in the CDFT constrained-search universal density functional. These results place the mathematical framework of CDFT on the same footing as that of standard DFT.

INTRODUCTION

Density-functional theory (DFT) is at present the most widely used tool for first-principles electronic-structure calculations in solid-state physics and quantum chemistry. DFT was put on a solid mathematical ground by Lieb in a landmark paper [1] from 1983, where he introduced the universal density functional $F(\rho)$ as the convex conjugate to the concave ground-state energy $E(\rho)$ for an electronic system in the external scalar potential $v$.

For electronic systems under the influence of a classical external magnetic field $A$, current-density functional theory (CDFT) was introduced by Vignale and Rasolt in 1987 [2]. In addition to the density $\rho$, the paramagnetic current density $j$ becomes a basic variable. The mathematical foundation of CDFT was put in place by Tellgren et al. [3] and Laestadius [4,5] in the 2010s based on Lieb’s treatment of the field-free standard case. However, a central piece of the puzzle has been missing—namely, whether the CDFT constrained-search functional $F(\rho, j_p)$ is lower semi-continuous and expectation valued [6], i.e., that the infimum in its definition [see Eq. (3) below] is in fact attained.

In this letter, we provide proofs of these assertions. The CDFT constrained-search functional is indeed convex lower-semicontinuous, and can therefore be identified with the CDFT Lieb functional—that is, a Legendre–Fenchel transform of the energy. Without this fact, the ground-state energy functional $E(v, A)$ and the constrained-search functional $F(\rho, j_p)$ contain different information. If $F(\rho, j_p)$ were not expectation valued, one would lose the interpretation of the universal functional as intrinsic energy, which is very useful in standard DFT.

For an $N$-electron system in sufficiently regular external potentials $v$ and $A$, the ground-state energy is given by the Rayleigh–Ritz variation principle as

$$E(v, A) = \inf_{\Gamma} \text{Tr}(\Gamma H(v, A)), \quad (1)$$

where $H(v, A) = T(A) + W + \sum_{i=1}^{N} v(\mathbf{r}_i)$ is the electronic Hamiltonian with kinetic-energy operator $T(A) = \frac{1}{2} \sum_{i=1}^{N} (-i \nabla_i + A(\mathbf{r}_i))^2$ and two-electron repulsion operator $W$. The minimization is over all $N$-electron density matrices $\Gamma$ of finite kinetic energy, for which the one-electron density is $\rho \in X_L = L^1(\mathbb{R}^3) \cap L^2(\mathbb{R}^3)$, and $j_p \in X_p = L^1(\mathbb{R}^3) \cap L^3/2(\mathbb{R}^3)$ [7]. (The boldface notation indicates a space of vector fields.) The external potential energy $(v)\rho = \int_{\mathbb{R}^3} v(\mathbf{r}) \rho(\mathbf{r}) d\mathbf{r}$, the paramagnetic and diamagnetic terms $\frac{1}{2}(|A|^2)\rho$ and $(A)j_p = \int_{\mathbb{R}^3} A(\mathbf{r}) \cdot j_p(\mathbf{r}) d\mathbf{r}$, and thus the Hamiltonian $H(v, A)$, are well defined for any $v \in X_{L'} = L^{3/2}(\mathbb{R}^3) + L^\infty(\mathbb{R}^3)$ and $A \in X'_{p} = L^1(\mathbb{R}^3) + L^\infty(\mathbb{R}^3)$, where $X_{L'}$ and $X'_{p}$ are the dual spaces of $X_L$ and $X_p$, respectively. Examples of such potentials are the nuclear Coulomb potentials and uniform magnetic fields inside bounded domains. The symbol $X_L$ for the space of densities is so chosen to indicate it is the density space of Lieb’s analysis, while $X_p$ indicates “paramagnetic” current densities.

By a well-known reformulation of Eq. (I), we obtain the CDFT Hohenberg–Kohn variation principle

$$E(v, A) = \inf_{(\rho,j_p) \in X_L \times X_p} \left\{ F(\rho, j_p) \right. \right.$$  
\[ \left. + (v + \frac{1}{2}|A|^2)\rho + (A)j_p \right\}. \quad (2)

Here the Vignale–Rasolt constrained-search density functional $F: X_L \times X_p \to [0, +\infty]$ is defined by

$$F(\rho, j_p) = \inf_{\Gamma=(\rho,j_p)} \text{Tr}(\Gamma H_0), \quad (3)$$

For an $N$-electron system in sufficiently regular external potentials $v$ and $A$, the ground-state energy is given by the Rayleigh–Ritz variation principle as

$$E(v, A) = \inf_{\Gamma} \text{Tr}(\Gamma H(v, A)), \quad (1)$$

where $H(v, A) = T(A) + W + \sum_{i=1}^{N} v(\mathbf{r}_i)$ is the electronic Hamiltonian with kinetic-energy operator $T(A) = \frac{1}{2} \sum_{i=1}^{N} (-i \nabla_i + A(\mathbf{r}_i))^2$ and two-electron repulsion operator $W$. The minimization is over all $N$-electron density matrices $\Gamma$ of finite kinetic energy, for which the one-electron density is $\rho \in X_L = L^1(\mathbb{R}^3) \cap L^2(\mathbb{R}^3)$, and $j_p \in X_p = L^1(\mathbb{R}^3) \cap L^3/2(\mathbb{R}^3)$ [7]. (The boldface notation indicates a space of vector fields.) The external potential energy $(v)\rho = \int_{\mathbb{R}^3} v(\mathbf{r}) \rho(\mathbf{r}) d\mathbf{r}$, the paramagnetic and diamagnetic terms $\frac{1}{2}(|A|^2)\rho$ and $(A)j_p = \int_{\mathbb{R}^3} A(\mathbf{r}) \cdot j_p(\mathbf{r}) d\mathbf{r}$, and thus the Hamiltonian $H(v, A)$, are well defined for any $v \in X_{L'} = L^{3/2}(\mathbb{R}^3) + L^\infty(\mathbb{R}^3)$ and $A \in X'_{p} = L^1(\mathbb{R}^3) + L^\infty(\mathbb{R}^3)$, where $X_{L'}$ and $X'_{p}$ are the dual spaces of $X_L$ and $X_p$, respectively. Examples of such potentials are the nuclear Coulomb potentials and uniform magnetic fields inside bounded domains. The symbol $X_L$ for the space of densities is so chosen to indicate it is the density space of Lieb’s analysis, while $X_p$ indicates “paramagnetic” current densities.

By a well-known reformulation of Eq. (I), we obtain the CDFT Hohenberg–Kohn variation principle

$$E(v, A) = \inf_{(\rho,j_p) \in X_L \times X_p} \left\{ F(\rho, j_p) \right. \right.$$  
\[ \left. + (v + \frac{1}{2}|A|^2)\rho + (A)j_p \right\}. \quad (2)

Here the Vignale–Rasolt constrained-search density functional $F: X_L \times X_p \to [0, +\infty]$ is defined by

$$F(\rho, j_p) = \inf_{\Gamma=(\rho,j_p)} \text{Tr}(\Gamma H_0), \quad (3)$$
where \( H_0 = T(0) + W \) is the intrinsic electronic Hamiltonian, and \( \Gamma \mapsto (\rho, J_p) \) means that the infimum is taken over all \( N \)-electron density matrices \( \Gamma \) with density-current pair \( (\rho, J_p) \in L^1(\mathbb{R}^3) \times L^1(\mathbb{R}^3) \). Thus, if \( (\rho, J_p) \) is not \( N \)-representable, we have \( F(\rho, J_p) = +\infty \). The universal density functional \( F \) is the central quantity in any flavor of DFT, whose mathematical properties and approximation is of utmost importance to the field.

Although \( E \) in Eq. (2) is not concave, it is readily seen that the reparametrized energy

\[
\hat{E}(u, A) = E(u - \frac{1}{2}|A|^2, A)
\]

is concave. This reparametrization relies on a technical notion of compatibility of function spaces for the scalar and vector potentials \([7]\), satisfied for the potentials we consider here.

From the concavity and upper semi-continuity of the modified ground-state energy \( \hat{E} \), one deduces the existence of an alternative universal density functional \( \hat{F} : X_1 \times X_p \to [0, +\infty] \) related to the ground-state energy by Legendre–Fenchel transformations in the manner

\[
\hat{E}(u, A) = \inf_{(\rho, J_p)} \left\{ \hat{F}(\rho, J_p) + (u|\rho) + (A|J_p) \right\}, \quad (5)
\]

\[
\hat{F}(\rho, J_p) = \sup_{(u, A)} \left\{ \hat{E}(u, A) - (u|\rho) - (A|J_p) \right\}, \quad (6)
\]

where the optimizations are over the space \( X_1 \times X_p \) and its dual \( X'_1 \times X'_p \), respectively. As a Legendre–Fenchel transform, the functional \( \hat{F} \) is convex and lower semi-continuous. In this formulation of CDFT, the ground-state energy \( \hat{E} \) and the universal density functional \( \hat{F} \) contain precisely the same information: each functional can be obtained from the other and therefore contains all information about ground-state electronic systems in external scalar and vector fields.

From a comparison of the Hohenberg–Kohn variation principles in Eqs. (2) and (5), it is tempting to conclude that \( \hat{F} = F \) are the same functional, producing the same ground-state energy for each \( (v, A) \). However, there exist infinitely many functionals \( \hat{F} : X_1 \times X_p \to [0, +\infty] \) that give the correct ground-state energy \( E(v, A) \) (but not necessarily the same minimizing density, if any) for each \( (v, A) \) in the Hohenberg–Kohn variation principle. Each such \( \hat{F} \) is said to be an admissible density functional \([8]\). Among these, the functional \( \hat{F} \) stands out as being the only lower semi-continuous and convex universal density functional and a lower bound to all other admissible density functionals, \( \hat{F} \leq \hat{F} \). The functional \( \hat{F} \), often called the closed convex hull of all admissible density functionals, is thus the most well-behaved admissible density functional. Indeed, we may view it as a regularization of all admissible density functionals, known as the \( \Gamma \)-regularization in convex analysis. (This name is unrelated to our notation of density matrices.)

A fundamental result of Lieb’s analysis of DFT is the identification of the transparent constrained-search density functional with the mathematically well-behaved closed convex hull \( \hat{F} \). The identification follows since \( \hat{F} \) is convex and lower semi-continuous. Whereas convexity follows easily for the CDFT Vignale–Rasolt functional \( F \), the proof of lower semi-continuity is nontrivial. For standard DFT it is given in Ref. [1] and for CDFT in the present letter.

We simplify our analysis by merely assuming that the density–current pairs are \( (\rho, J_p) \in L^1(\mathbb{R}^3) \times L^1(\mathbb{R}^3) = [L^1(\mathbb{R}^3)]^4 \), which we denote as \( X \). With this topology, the potentials must be taken to be bounded functions, \( (v, A) \in X' = L^\infty(\mathbb{R}^3) \times L^\infty(\mathbb{R}^3) = [L^\infty(\mathbb{R}^3)]^4 \). This simplification is irrelevant in this context: if \( F \) can be shown to be lower semi-continuous in the \( [L^1(\mathbb{R}^3)]^4 \) topology, it will be lower semi-continuous in any stronger topology, as required if we enlarge the potential space to include more singular functions such as those in \( X_1' \times X_p' \). Indeed, the original proof of lower semi-continuity of the standard DFT Levy–Lieb functional \([3]\) was with respect to the \( L^1(\mathbb{R}^3) \) topology, from which the same property with respect to the \( X_1 \) topology immediately follows.

**THEOREM AND PROOF**

The intrinsic Hamiltonian \( H_0 = H(0, 0) \) is self-adjoint \( (H_0 = H_0^\dagger) \) over \( L_N^2 \), the Hilbert space of square-integrable \( N \)-electron wavefunctions (with spin and permutational antisymmetry built in). The expectation values of \( H_0 \) and \( H(v, A) \) are well-defined on the Sobolev space \( H_N \), the subset of \( L_N^2 \) with finite kinetic energy.

We denote by \( D_N \) the convex set of \( N \)-electron mixed states with finite kinetic energy. We have the mathematical characterization \([8]\)

\[
D_N = \left\{ \Gamma \in TC(L_N^2) \mid \Gamma^\dagger = \Gamma \geq 0, \text{Tr} \Gamma = 1, \nabla_1 \Gamma \nabla_1^\dagger \in TC(L_N^2) \right\},
\]

where \( TC(L_N^2) \) is the set of trace-class operators over \( L_N^2 \), the largest set of operators to which a basis-independent trace can be assigned. An operator \( A \) is trace class if and only if the positive
square root $|A| := \sqrt{AA^*}$ is trace class $\mathbb{H}$. A self-adjoint operator $A$ is trace-class if and only if it has a spectral decomposition of the form $A = \sum_{k=1}^{\infty} \lambda_k |\phi_k\rangle \langle \phi_k|$, where $\{\phi_k\}$ forms an orthonormal basis and where $\sum_k \lambda_k$ is absolutely convergent. Now $A = F \in \mathcal{D}_N$ if and only if $\lambda_k \geq 0$, $\sum_k \lambda_k = 1$, $\{\phi_k\} \subset H_N$, and if the total kinetic energy is finite, $\sum_k \lambda_k \langle \phi_k | T | \phi_k \rangle < +\infty$.

For any $\psi \in H_N^*$, the density-current pair $(\rho, j_\rho) \in L^1(\mathbb{R}^3) \times L^1(\mathbb{R}^3)$ is defined by

$$\rho(r_1) := N \int |\psi(r_1; \tau_{-1})|^2 d\tau_{-1}, \quad (8)$$

$$j_\rho(r_1) := N \text{Im} \int \overline{\psi^*(r_1; \tau_{-1})} \nabla_1 \psi(r_1; \tau_{-1}) d\tau_{-1}, \quad (9)$$

where we integrate over all spin variables and over $N - 1$ spatial coordinates, $\tau_{-1} = (\sigma_1, x_2, \cdots, x_N)$. For $A \in \mathcal{D}_N$, we can, for instance, compute $\rho = Tr A$ from $\sum_k \lambda_k \rho_k$ with $\rho_k$ obtained from Eq. (8) with $\psi = \phi_k$ (and similarly for $j_\rho$).

The theorem involves the weak topology on $X = L^1(\mathbb{R}^3) \times L^1(\mathbb{R}^3)$. Weak convergence of a sequence $\{x_n\} \subset X$, written $x_n \rightharpoonup x \in X$, means that, for any bounded linear functional $\omega \in X'$, we have $\omega(x_n) \rightarrow \omega(x)$ as a sequence of numbers—that is, weak convergence is the pointwise convergence of all bounded linear functionals. Recall that the dual space of $L^1(\mathbb{R}^3)$ is $L^\infty(\mathbb{R}^3)$, so that $\rho_n \rightharpoonup \rho \in L^1(\mathbb{R}^3)$ if and only if $(f|\rho_n) \rightarrow (f|\rho)$ for every $f \in L^\infty(\mathbb{R}^3)$. Likewise, $(\rho_n, j_{\rho_n}) \rightharpoonup (\rho, j_\rho) \in X$ if and only if $(f|\rho_n) \rightarrow (f|\rho)$ and $(a|j_{\rho_n}) \rightarrow (a|j_\rho)$ for every $(f, a) \in X'$.

The trace-class operators over a separable Hilbert space $\mathcal{H}$ are examples of compact operators, an infinite dimensional generalization of finite-rank operators. Indeed, the set $K(\mathcal{H})$ of compact operators is the closure of the finite-rank operators in the norm topology and thus a Banach space. The dual space of $K(\mathcal{H})$ is in fact $\text{TC}(\mathcal{H})$. For $B \in K(\mathcal{H})$ and $A \in TC(\mathcal{H})$, the dual pairing is $\text{Tr}(BA)$. Similar to the weak topology for a Banach space, the dual of a Banach space can be equipped with the weak* topology. A sequence of trace-class operators $\{A_n\}$ converges weak*-to $A \in TC(\mathcal{H})$ if, for each $B \in K(\mathcal{H})$, $\text{Tr}(B_n A) \rightarrow \text{Tr}(BA)$.

We now state and prove our main result, from which lower semi-continuity follows in Corollary 4.4 in Ref. [1].

Theorem 1. Suppose $(\rho, j_\rho) \in X$ and $\{\rho_n, j_{\rho_n}\} \subset X$ are such that $F(\rho, j_\rho) < +\infty$ and $F(\rho_n, j_{\rho_n}) < +\infty$ for each $n \in \mathbb{N}$ and suppose further that $(\rho_n, j_{\rho_n}) \rightarrow (\rho, j_\rho)$. Then there exists $\Gamma \in D_N$ such that $\Gamma \rightarrow (\rho, j_\rho)$ and $\text{Tr}(H_0 \Gamma) \leq \lim inf_n F(\rho_n, j_{\rho_n})$.

Proof of Theorem 1. The initial setup follows Ref. [1], which we here restate. Without loss of generality, we may replace $H_0 = T + W$ by $h^2 = T + W + 1$, which is self-adjoint and positive definite. The operator $h$ is taken to be the unique positive self-adjoint square root of $T + W + 1$.

Consider the sequence $\{g_n\}$ with elements $g_n := F(\rho_n, j_{\rho_n})$. If $g_n \rightarrow +\infty$, then the statement of the theorem is trivially true. Assume therefore that $\{g_n\}$ is bounded. Then there exists a subsequence such that $g := \lim_n g_n$ exists. Furthermore, for each $n$, there exists $\Gamma_n \in D_N$ such that $\Gamma_n \rightarrow (\rho_n, j_{\rho_n})$ and $\text{Tr}(h^2 \Gamma_n) \leq g + 1/n$. To see this, select for each $n$ a density matrix $\Gamma_n \rightarrow (\rho_n, j_{\rho_n})$ that satisfies $\text{Tr}(h^2 \Gamma_n) < g_n + 1/2n$ and choose $m$ such that $|g - g_n| < 1/2n$ for each $n > m$ (by taking a subsequence if necessary); for each $n > m$, we then have

$$0 \leq \text{Tr}(h^2 \Gamma_n) - g = |\text{Tr}(h^2 \Gamma_n) - g| \leq |\text{Tr}(h^2 \Gamma_n) - g_n| + |g_n - g| \leq 1/n.$$

(10)

Using the sequence $\{h^2 \Gamma_n\}$, we next establish a candidate limit density operator $\Gamma \in D_N$.

The dual-space sequence of (positive semi-definite) operators $y_n := h^2 \Gamma_n \in TC(L^2_N)$ is uniformly bounded in the trace norm: $\|y_n\|_{TC} \leq g + 1$. By the Banach–Alaoglu theorem, a norm-closed ball of finite radius in the dual space is compact in the weak*-topology. Thus, there exists $y \in TC(L^2_N)$ such that, for a subsequence, $\text{Tr}(B y_n) \rightarrow \text{Tr}(B y)$ for each $B \in K(L^2_N)$, meaning that $y$ is the (possibly nonunique) weak*-limit of a subsequence of $\{y_n\}$. The limit is positive definite, since the orthogonal projector $P_{\Phi}$ onto $\Phi \in L^2_N$ is a compact operator, which gives

$$\langle \Phi | y | \Phi \rangle = \text{Tr}(y P_{\Phi}) = \lim_n \text{Tr}(y_n P_{\Phi}) = \lim_n \langle \Phi | y_n | \Phi \rangle \geq 0.$$

(11)

We now define $\Gamma = h^{-1} y h^{-1}$, which fulfills all the criteria for being an element of $D_N$, except possibly $\text{Tr} \Gamma = 1$, although $\text{Tr} \Gamma \leq 1$ is already implied by the weak convergence. (Note that $\Gamma$ has finite kinetic energy since $\text{Tr}(h^2 \Gamma) < +\infty$.) If we can show that $\Gamma \rightarrow (\rho, j_\rho)$, then we are done with the complete proof, since $\Gamma \in D_N$ follows from $\text{Tr} \Gamma = N^{-1} \int_{\mathbb{R}^3} \rho(r) dr = 1$ and since

$$\text{Tr}(h^2 \Gamma) = \text{Tr} y \leq \lim inf_n \text{Tr} y_n$$

$$= \lim inf_n \text{Tr}(h^2 \Gamma_n)$$

$$\leq \lim inf_n \{F(\rho_n, j_{\rho_n}) + 1/2n\}$$

$$= \lim inf F(\rho_n, j_{\rho_n}).$$

(12)
Let \((\rho', j'_p) \leftarrow \Gamma\) be the density associated with \(\Gamma\). To demonstrate that \((\rho', j'_p) = (\rho, j_p)\), we recall that \((\rho_n, j_{pn}) \rightarrow (\rho, j_p)\) by assumption. Since weak limits are unique, our proof is complete if we can show that \((\rho_n, j_{pn}) \rightarrow (\rho, j_p)\) in \(L^1(\mathbb{R}^3) \times L^1(\mathbb{R}^3)\). The proof of \(\rho_n \rightarrow \rho\) is given in Ref. [2] and omitted here. We here demonstrate that \(j_p \rightarrow j'_p\) by showing that \(j_{pn} \rightarrow j'_p\) for each \(n \in \mathbb{L}^\infty(\mathbb{R}^3)\).

Let \(\Omega \subset \mathbb{R}^3\) be a bounded domain with characteristic function \(\chi\), equal to 1 on \(\Omega\) and 0 elsewhere. Since \(\rho, \rho' \in L^1(\mathbb{R}^3)\), we may, for a given \(\varepsilon > 0\), choose \(\Omega\) sufficiently large so that \(\int (1 - \chi) \rho \, d\tau < \varepsilon\) and \(\int (1 - \chi) \rho' \, d\tau < \varepsilon\). Since \(\rho_n \rightarrow \rho\), we also have \(\int (1 - \chi) (\rho_n - \rho) \, d\tau + \varepsilon\) for sufficiently large \(n\). From the triangle inequality, we obtain \(\int (1 - \chi) \rho_n \, d\tau \leq \int (1 - \chi) (\rho_n - \rho) \, d\tau + \int (1 - \chi) \rho' \, d\tau\), implying that \(\int (1 - \chi) \rho_n \, d\tau < 2\varepsilon\) for sufficiently larger \(n\).

In the notation \(\tau = (r_1, \tau_{-1}) = (x_1, x_2, \ldots, x_N)\) and \(\tau_{-1} = (\tau_{-1, x_1}, \ldots, \tau_{-1, x_N})\), let \(U_n = N \text{Im} \text{diag} \partial_{\alpha} \Gamma = N \text{Im} \sum_{\mu} \lambda_{\mu} \psi_{\mu}(\tau) \partial_{\alpha} \psi_{\mu}(r_1, \tau_{-1})\), where \(\alpha\) denotes a Cartesian component and where we have introduced the spectral decomposition \(\sum_{\mu} \lambda_{\mu} \psi_{\mu} \in D_N^{\mathbb{R}}\) with \(\psi_{\mu} \in \mathcal{H}_N^{\mathbb{R}}\). We note that, if \(\Gamma \rightarrow (\rho, j_p)\), then integration of \(U_n\) over \(\tau_{-1}\) gives \(j_{pn} = \int U_n (r, \tau_{-1}) \, d\tau_{-1}\).

Let now \(S = \prod_{i=1}^N \chi(r_i)\) be the characteristic function of \(\mathbb{L}^N \subset \mathbb{R}^{3N}\). By the definition of \(U_n\), we then have

\[
I(U_n) := \left| \int (1 - S) U_n \, d\tau \right| \leq N \int (1 - S) \sum_{\mu} \lambda_{\mu} \psi_{\mu} \, |\partial_{\alpha} \psi_{\mu}| \, d\tau.
\]

Applying the Cauchy–Schwarz inequality twice, we obtain

\[
I(U_n) \leq N \int (1 - S) \left( \sum_{\mu} |\lambda_{\mu} \psi_{\mu}|^2 \right)^{1/2} \left( \sum_{\mu} \lambda_{\mu} \partial_{\alpha} \psi_{\mu} \right)^{1/2} \, d\tau.
\]

Noting that \(1 - S \leq \sum_{i=1}^N (1 - \chi(r_i))\) and using the symmetry of \(|\psi_{\mu}|^2\), we obtain for the two factors

\[
\int (1 - S) \sum_{\mu} \lambda_{\mu} \psi_{\mu} \, d\tau \leq \int (1 - \chi) \rho' \, d\tau < \varepsilon,
\]

\[
\frac{N}{2} \int \sum_{\mu} \lambda_{\mu} \partial_{\alpha} \psi_{\mu} \, d\tau = \text{Tr}(TT) \leq 4|g|.
\]

We conclude that \(I(U_n)^2 \leq 2Ng\varepsilon\). Introducing \(U_{n,\alpha} = N \text{Im} \text{diag} \partial_{\alpha} \Gamma\) and proceeding in the same manner, we arrive at the bound \(I(U_{n,\alpha})^2 \leq 2Ng\varepsilon\), assuming that \(n\) has been chosen so large that \(\int (1 - \chi) \rho_n \, d\tau < 2\varepsilon\) holds.

We are now ready to consider the weak convergence \(j_{pn} \rightarrow j'_p\) in \(L^1(\mathbb{R}^3)\). For each \(a \in \mathbb{L}^\infty(\mathbb{R}^3)\) and for sufficiently larger \(n\), we obtain, using the Cauchy–Schwarz inequality and the H"older inequality in combination with the bounds \(I(U_n)^2 \leq 2Ng\varepsilon\) and \(I(U_{n,\alpha})^2 \leq 4Ng\varepsilon\), the inequality

\[
\left| \int (j_{pn} - j'_p) \cdot a \, d\tau \right| \leq \sum_{\alpha} \left| \int (j_{pn\alpha} - j'_{p\alpha}) a_{\alpha} \, d\tau \right| = \sum_{\alpha} \left| \int (U_{n\alpha} - U_{\alpha}) a_{\alpha}(r_1) \, d\tau \right|
\]

\[
\leq \sum_{\alpha} \left| \int (1 - S)(U_{n\alpha} - U_{\alpha}) a_{\alpha}(r_1) \, d\tau \right| + \sum_{\alpha} \left| \int S(U_{n\alpha} - U_{\alpha}) a_{\alpha}(r_1) \, d\tau \right|
\]

\[
\leq \sum_{\alpha} \left\| a_{\alpha} \right\|_\infty (6Ng\varepsilon)^{1/2} + \sum_{\alpha} \left| \int (U_{n\alpha} - U_{\alpha}) a_{\alpha}(r_1) S \, d\tau \right|.
\]

Since \(\varepsilon > 0\) is arbitrary, it only remains to show \(\int (U_{n\alpha} - U_{\alpha}) a_{\alpha}(r_1) S \, d\tau \rightarrow 0\) as \(n \rightarrow \infty\).

Let \(M\) be the compact multiplication operator associated with \(a_{\alpha}(r_1) S(\tau)\), a bounded function with compact support over \(\mathbb{R}^{3N}\). Let \(\Omega_n = \{1, \downarrow\}\) be the set consisting of the two spin states of the electrons. We note that

\[
\int U_{n\alpha} a_{\alpha}(r_1) S \, d\tau = \int_{\Omega_n \times 1} U_{n\alpha} a_{\alpha}(r_1) \, d\tau
\]

\[
= N \text{Im} \text{Tr}(\partial_{\alpha} \Gamma_n M)
\]

\[
= N \text{Im} \text{Tr}(h^{-1} M \partial_{\alpha} H^{-1} y_n),
\]

viewing \(\Gamma_n\) as an operator over \(L^2(\Omega \times \Omega_n)^N\) by domain restriction of the spectral decomposition.
elements—that is, \( \psi_\alpha \in H^1((\Omega \times \Omega_\sigma)^N) \), meaning that the \( 2^N \) spin components of \( \psi_\alpha \) are in \( H^1(\Omega^N) \). The spaces used here are not antisymmetrized, for simplicity.

Our next task is to demonstrate that \( B = h^{-1}M\partial_{\alpha}h^{-1} \) is compact over \( L^2((\Omega \times \Omega_\sigma)^N) \). We first show that \( h^{-1} \) is compact with range \( H^1((\Omega \times \Omega_\sigma)^N) \). We have \( h = \sqrt{T + W + 1} \) with domain \( H^1_N((\Omega \times \Omega_\sigma)^N) \). Now \( h^{-1} \) exists and is bounded since \( -1 \) is not in the spectrum of \( T + W \)—that is, \( h^{-1} : L^2((\Omega \times \Omega_\sigma)^N) \to H^1((\Omega \times \Omega_\sigma)^N) \) is bounded. By the Rellich–Kondrachov theorem, \( H^1((\Omega \times \Omega_\sigma)^N) \) (the standard Sobolev space without spin) is a compact subset of \( L^2(\Omega^N) \). It follows that \( H^1((\Omega \times \Omega_\sigma)^N) \) is a compact subset of \( L^2((\Omega \times \Omega_\sigma)^N) \), since the tensor product of compact sets is compact. Hence, \( h^{-1} \) is compact.

Next, the operator \( \partial_{\alpha} \) is, by the definition of the Sobolev space \( H^1(\Omega^N) \), bounded from \( H^1((\Omega \times \Omega_\sigma)^N) \) to \( L^2((\Omega \times \Omega_\sigma)^N) \). Thus \( \partial_{\alpha}h^{-1} \) is bounded over \( L^2((\Omega \times \Omega_\sigma)^N) \). It follows that \( B \in K(L^2((\Omega \times \Omega_\sigma)^N)) \) because it is a product of a compact operator \( h^{-1} \) with a bounded operator \( M\partial_{\alpha}h^{-1} \).

From compactness of \( B \), it follows that
\[
\int U_{\alpha\alpha}a_\alpha S \, d\tau = N \text{Im} \text{Tr}(B y_n) \\
\quad \to N \text{Im} \text{Tr}(B y) = \int U_{\alpha\alpha}a_\alpha S \, d\tau, \tag{15}
\]

by the weak-* convergence of \( y_n \) to \( y \). We conclude that \( J_{pn} \to J_p \) and hence that \( (\rho_n, J_{pn}) \to (\rho, J_p) \), completing the proof.

**Corollary 1.** \( F : L^1(\mathbb{R}^3) \times L^1(\mathbb{R}^3) \to [0, +\infty) \) is lower semi-continuous and also weakly lower semi-continuous.

**Proof.** Let \( (\rho_n, J_{pn}) \to (\rho, J_p) \) in \( L^1(\mathbb{R}^3) \times L^1(\mathbb{R}^3) \). From Theorem 1 we then obtain
\[
F(\rho, J_p) \leq \text{Tr}(H_0 \Gamma) \leq \liminf_n F(\rho_n, J_{pn}), \tag{16}
\]

where \( \Gamma \to (\rho, J_p) \). Hence, \( F \) is weakly lower semi-continuous. By Mazur’s Lemma [10], weak lower semi-continuity of a convex function implies strong lower semi-continuity.

**Corollary 2.** If \( F(\rho, J_p) < +\infty \), then the infimum in the CDFT constrained-search functional is a minimum:
\[
F(\rho, J_p) = \min_{\Gamma \to (\rho, J_p)} \text{Tr}(H_0 \Gamma). \tag{17}
\]

**Proof.** Simply take \( (\rho_n, J_{pn}) = (\rho, J_p) \) for all \( n \), and apply Theorem 1.

**CONCLUSION**

We have extended Theorem 4.4 of Ref.[3] to CDFT. As immediate corollaries, the constrained-search functional \( F(\rho, J_p) \) is lower semi-continuous and expectation valued, that is, if \( F(\rho, J_p) < +\infty \), then there exists a \( \Gamma \to (\rho, J_p) \) such that \( F(\rho, J_p) = \text{Tr}(H_0 \Gamma) \). These mathematical results are the final pieces in the puzzle of placing CDFT on a solid mathematical ground in a similar manner as done by Lieb for standard DFT.

**ACKNOWLEDGMENTS**

This work has received funding from the Research Council of Norway (RCN) under CoE Grant Nos. 287906 and 262695 (Hylleraas Centre for Quantum Molecular Sciences) and from ERC-STG-2014 under grant agreement no 639508.

*simen.kvaal@kjemi.uio.no*

[1] E.H. Lieb. Density Functionals for Coulomb Systems. *Int. J. Quant. Chem.*, 24(3):243–277, 1983.
[2] G. Vignale and M. Rasolt. Density-Functional Theory in Strong Magnetic Fields. *Phys. Rev. Lett.*, 59(20):2360–2363, 1987.
[3] E.I. Tellgren, S. Kvaal, E. Sagvolden, U. Ekström, A.M. Teale, and T. Helgaker. Choice of basic variables in current-density-functional theory. *Phys. Rev. A*, 86(6), 2012.
[4] A. Laestadius and M. Benedicks. Hohenberg–Kohn theorems in the presence of magnetic field. *Int. J. of Quant. Chem.*, 114(12):782–795, 2014.
[5] A. Laestadius. Density functionals in the presence of magnetic field. *Int. J. Quantum Chem.*, 114(21):1445–1456, 2014.
[6] S. Kvaal and T. Helgaker. Ground-state densities from the Rayleigh–Ritz variation principle and from density-functional theory. *J. Chem. Phys.*, 143(18):184106, 2015.
[7] A. Laestadius, E.I. Tellgren, M. Penz, M. Ruggenthaler, S. Kvaal, and T. Helgaker. Kohn–Sham Theory with Paramagnetic Currents: Compatibility and Functional Differentiability. *J. Chem. Theor. Comp.*, 15(7):4003–4020, 2019.
[8] E.I. Telgaren, S. Kvaal, and T. Helgaker. Fermion N-representability for prescribed density and paramagnetic current density. *Phys. Rev. A*, 89(1), 2014.
[9] M. Reed and B. Simon. *Methods of modern mathematical physics I: functional analysis*. Academic Press, San Diego, New York, Berkeley, Boston, London, Sydney, Tokyo, Toronto, 1980.
[10] I. Ekeland and B. Témam. *Convex Analysis and Variational Problems*. Society for Industrial and Applied Mathematics, Philadelphia, 1999.