The Physics and Mathematics of Elliott Lieb

The 90th Anniversary Volume I

Edited by
Rupert L. Frank
Ari Laptev
Mathieu Lewin
Robert Seiringer

EMS PRESS

Editors:

Rupert L. Frank
Mathematisches Institut
Ludwig-Maximilians-Universität München
Theresienstraße 39
80333 München, Germany
Email: r.frank@lmu.de

Ari Laptev
Department of Mathematics
Imperial College London
Huxley Building, 180 Queen’s Gate
London SW7 2AZ, UK
Email: a.laptev@imperial.ac.uk

Mathieu Lewin
CEREMADE
Université Paris Dauphine
Place de Lattre de Tassigny
75016 Paris, France
Email: mathieu.lewin@math.cnrs.fr

Robert Seiringer
Institute of Science and Technology Austria
Am Campus 1
3400 Klosterneuburg, Austria
Email: robert.seiringer@ist.ac.at

2020 Mathematics Subject Classification: 00B30; 46-02, 49-02, 81-02, 82-02

Keywords: quantum many-body systems, statistical mechanics, entropy inequalities, functional inequalities, sharp constants

ISBN Vol. I 978-3-98547-021-1
ISBN Vol. II 978-3-98547-022-8
ISBN Set 978-3-98547-019-8 (set of both volumes)

Bibliographic information published by the Deutsche Nationalbibliothek
The Deutsche Nationalbibliothek lists this publication in the Deutsche Nationalbibliografie; detailed bibliographic data are available on the Internet at http://dnb.dnb.de.

Published by EMS Press, an imprint of the

European Mathematical Society – EMS – Publishing House GmbH
Institut für Mathematik
Technische Universität Berlin
Straße des 17. Juni 136
10623 Berlin, Germany

https://ems.press

© 2022 EMS Press

The cover drawing, based on a photograph of Elliott Lieb, was done by António B. Araújo.
The photograph of Elliott Lieb on page v is by Theodore H. Lewis III.

Typesetting: Ashry Abdalla, Roisin, Belgium; Alison Durham, Manchester, UK; WisSat Publishing + Consulting GmbH, Fürstenwalde, Germany
Printing and binding: Beltz Bad Langensalza GmbH, Bad Langensalza, Germany
♾ Printed on acid free paper

Chapter 22

Lieb variation principle in density-functional theory

Trygve Helgaker and Andrew M. Teale

Lieb's convex formulation of density-functional theory is presented in a pedagogical manner, emphasizing its connection to Hohenberg–Kohn theory and to Levy's constrained-search theory. The Hohenberg–Kohn and Lieb variation principles are discussed, highlighting the dual relationship between the ground-state energy and the universal density functional. Applications of the Lieb variation principle are reviewed, demonstrating how it may be utilized to calculate the Kohn–Sham potential of atoms and molecules, to study the exchange–correlation functional and the adiabatic connection by high-precision many-body methods, and to calculate the exchange–correlation hole and energy densities of atoms and molecules.

Density-functional theory (DFT) is a hugely successful and powerful theory, underlying nearly all electronic-structure calculations of molecules and materials in chemistry and physics today. It is also a very beautiful theory – an elegant application of convex analysis to electronic-structure theory, predicated on the simple observation that the ground-state electronic energy of an atom, molecule, or material is continuous and concave in the external potential, as realized in 1983 by Lieb in his seminal work "Density functionals for Coulomb systems" [21]. Unfortunately, nearly forty years later, Lieb's convex formulation of DFT is still not widely appreciated among the practitioners of DFT or even among many of its developers. Likewise, the teaching of DFT is still mostly based on the Hohenberg–Kohn theorem [14] and Levy's constrained-search theory [20]. This is a pity since Lieb's convex, unifying formulation of DFT is both accessible and transparent – and, in our experience, likely to intrigue students.

Lieb's theory does not only give valuable insight into DFT. It also furnishes us with a practical tool for computation: the Lieb variation principle, according to which we may calculate to high accuracy, for any given electron density, the corresponding external potential and the universal density functional, thereby providing us with a practical realization of the Hohenberg–Kohn mapping from densities to potentials. Furthermore, it allows us to study the adiabatic connection and to calcu-

---

2020 Mathematics Subject Classification. 81V55.
Keywords. Density-functional theory, Hohenberg–Kohn variation principle, Lieb variation principle, universal density functional theory, Hohenberg–Kohn theorem.

late the universal density functional of DFT using high-precision quantum-chemical methods, providing both insight and valuable benchmark data for the development of approximate density functionals.

The bulk of this chapter consists of two parts. We begin in Section 1 by reviewing Lieb's convex formulation of DFT, emphasizing its relationship to Hohenberg–Kohn theory and to Levy's constrained-search theory. Next, in Section 2, we review and illustrate the Lieb variation principle as a computational tool in DFT, with applications to both Kohn–Sham theory and orbital-free DFT. Section 3 contains some concluding remarks.

1 Lieb's convex formulation of density-functional theory

The following presentation of Lieb's formulation of DFT is intended to be an informal introduction rather than a rigorous exposition of the theory. Concepts and elements of convex analysis are introduced as needed, in a hopefully self-contained manner. For an accessible, rigorous treatment convex analysis, we highly recommend the introductory text of van Tiel [38].

1.1 Hohenberg–Kohn theorem

The Hamiltonian of an atomic or molecular system of N electrons in an external potential v can be written in the form

$$H(v) = T + W + \sum_{i=1}^{N} v(\mathbf{r}_i). \tag{1}$$

Whereas the external potential v varies from system to system, the kinetic operator

$$T = -\frac{\hbar^2}{2m_\mathrm{e}} \sum_{i=1}^{N} \nabla^2_{\mathbf{r}_i}$$

and the electron-repulsion operator

$$W = \frac{e^2}{4\pi\varepsilon_0} \sum_{i=1}^{N} \sum_{j=1}^{i-1} |\mathbf{r}_i - \mathbf{r}_j|^{-1}$$

are the same for all N-electron systems. We are particularly interested in those potentials v that support an electronic ground state and denote the set of these potentials by $\mathcal{V}_N$.

For a given $v \in \mathcal{V}_N$, the ground-state energy $E(v)$ is obtained by solving the Schrödinger equation,

$$H(v)\Psi(\mathbf{x}_i) = E(v)\Psi(\mathbf{x}_i).$$

Its solution is a complicated many-body problem, the wave function $\Psi(\mathbf{x}_i)$ depending on the spatial and spin coordinates $\mathbf{x}_i = (\mathbf{r}_i, \sigma_i)$ of all electrons. By contrast, the associated ground-state density is a much simpler quantity, depending only on three spatial coordinates:

$$\rho(\mathbf{r}) = \int |\Psi(\mathbf{r}, \sigma, \mathbf{x}_2, \dots, \mathbf{x}_N)|^2 \, \mathrm{d}\sigma \mathrm{d}\mathbf{x}_2 \cdots \mathrm{d}\mathbf{x}_N ,$$

where we interpret the integration over spin coordinates as a summation over α and β spins. According to the *Hohenberg–Kohn theorem*, we may use the density as a fundamental variable in place of the wave function when studying atomic and molecular systems [14]:

> *“Thus $v(\mathbf{r})$ is (to within a constant) a unique functional of $\rho(\mathbf{r})$; since, in turn, $v(\mathbf{r})$ fixes H we see that the full many-particle ground state is a unique functional of $\rho(\mathbf{r})$.”*

Equipped with this theorem, we can set up the *Hohenberg–Kohn mapping* from ground-state densities to ground-state wave functions via the external potential and the Hamiltonian:

$$\rho \mapsto v_\rho \mapsto H(v_\rho) \mapsto \Psi_\rho .$$

A density that is the ground-state density for some potential $v \in \mathcal{V}_N$ is said to be *v-representable* and the set of all v-representable densities is denoted by $\mathcal{A}_N$.

The Hohenberg–Kohn theorem (1964) has played an enormously important role in theoretical chemistry – indeed, it is typically viewed as the cornerstone of DFT, taught in all introductions to DFT along with Levy’s constrained-search theory (1979). At the same time, the Hohenberg–Kohn theorem may appear almost mysterious and difficult to understand intuitively. Although Levy’s constrained-search formulation brings clarity, the underlying structure and beauty of DFT is best appreciated from the point of view Lieb’s convex formulation of the theory (1983).

It is interesting to note that many of the elements or concepts of Lieb’s theory were already present in the early work of Hohenberg and Kohn but were not picked up on and developed further at the time. We will here highlight these connections, beginning with the subgradient inequality of the ground-state energy.

1.2 Subgradient inequality of the ground-state energy

As part of the proof of the Hohenberg–Kohn theorem [14], the following *subgradient inequality* of the ground-state energy at $v \in \mathcal{V}_N$ was established:

$$\begin{aligned} E(u) &< E(v) + (u - v \mid \rho_v) \quad \text{for all } u \notin v + \mathbb{R}, \\ E(u) &= E(v) + (u - v \mid \rho_v) \quad \text{for all } u \in v + \mathbb{R}. \end{aligned} \tag{2}$$

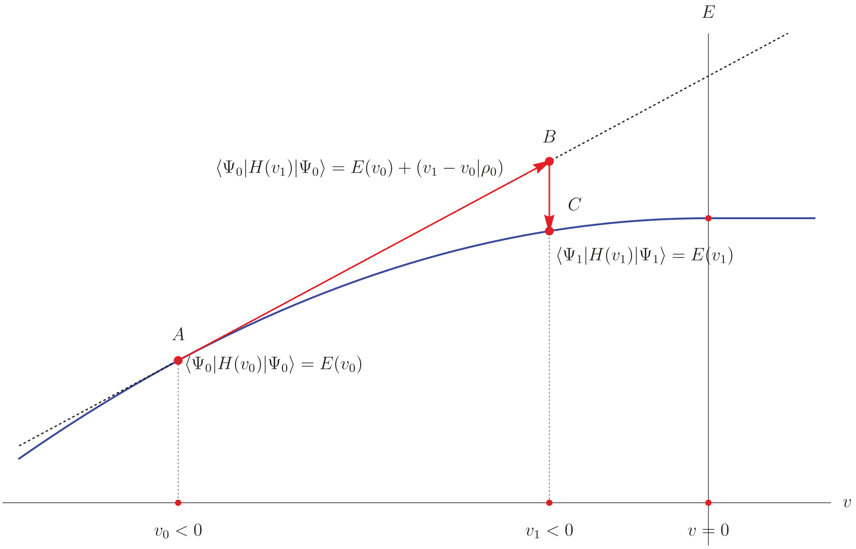
E
B
$\langle\Psi_0|H(v_1)|\Psi_0\rangle = E(v_0) + (v_1 - v_0|\rho_0)$
C
$\langle\Psi_1|H(v_1)|\Psi_1\rangle = E(v_1)$
A
$\langle\Psi_0|H(v_0)|\Psi_0\rangle = E(v_0)$
v
$v_0 < 0$
$v_1 < 0$
$v = 0$

Figure 1. The concavity of the ground-state energy $v \mapsto E(v)$ in the potential v (blue line), assuming that the potentials $v_0, v_1 \in \mathcal{V}_N$ have ground-state wave functions Ψ_0 and Ψ_1, respectively. As the potential changes from v_0 to v_1 with the ground-state wave function Ψ_0 fixed along the red line, the energy changes linearly. When the wave function relaxes from Ψ_0 to the ground-state wave function Ψ_1 at v_1, the ground-state energy is lowered, generating a concave curve $v \mapsto E(v)$. The slope of the tangent to E at v_0 is the density of the ground-state wave function Ψ_0 at v_0. In the terminology of convex analysis, the ground-state density ρ_0 is the "subgradient" of the ground-state at the potential v_0. The "subgradient inequality" in equation (2) therefore expresses the notion that the straight red line in the plot is a tangent to the concave blue curve.

Here $\rho_v \in \mathcal{A}_N$ is a ground-state density associated with $H(v)$ and we use the shorthand notation

$$(u|\rho) = \int u(\mathbf{r})\rho(\mathbf{r})\,\mathrm{d}\mathbf{r}$$

to denote the interaction of u with ρ. An intuitive understanding of the subgradient inequality can be obtained from Figure 1, which shows how the inequality arises from the Rayleigh–Ritz variation principle and the linearity of the Hamiltonian in the potential. As seen in the figure, the subgradient inequality is a consequence of the *concavity of the ground-state energy*. Geometrically speaking, a function is said to be concave if every chord connecting two points on its graph lies on or below the graph. As we shall see later, concavity of the ground-state energy is the key to DFT.

The subgradient inequality gives the Hohenberg–Kohn theorem directly. Let the potentials $u, v \in \mathcal{V}_N$ differ by more than an additive constant and let $\rho_u, \rho_v \in \mathcal{A}_N$

be the corresponding ground-state densities. From (2), we then obtain the strict subgradient inequalities $E(u) < E(v) + (u - v \mid \rho_v)$ and $E(v) < E(u) + (v - u \mid \rho_u)$. Adding these together, we conclude that $(v - u \mid \rho_v - \rho_u) < 0$ and hence that $\rho_u \neq \rho_v$ in accordance with the Hohenberg–Kohn theorem. The Hohenberg–Kohn theorem is thus a simple consequence of the (strict) concavity of the ground-state energy in the external potential, which in turn follows from the Rayleigh–Ritz variation principle and the linearity of the Hamiltonian in the potential.

1.3 Hohenberg–Kohn and Lieb variation principles

Note that the subgradient inequality of the ground-state energy is also the key to the Hohenberg–Kohn variation principle. For a v-representable density $\rho \mapsto v_\rho$, the *Hohenberg–Kohn universal density functional* [14] is defined as

$$F_{\mathrm{HK}}(\rho) = E(v_\rho) - (v_\rho|\rho),$$

where $F_{\mathrm{HK}}\colon \mathcal{A}_N \to \mathbb{R}$ and $E\colon \mathcal{V}_N \to \mathbb{R}$. Combining this definition of F_{HK} with the subgradient inequality of equation (2) expressed as $E(v_\rho) - (v_\rho|\rho) \geq E(v) - (v \mid \rho)$, we obtain

$$F_{\mathrm{HK}}(\rho) \geq E(v) - (v \mid \rho),$$

valid for every $\rho \in \mathcal{A}_N$ and for every $v \in \mathcal{V}_N$. From this inequality, the *Hohenberg–Kohn and Lieb variation principles*, respectively, follow directly:

$$E(v) = \min_{\rho \in \mathcal{A}_N} \left(F_{\mathrm{HK}}(\rho) + (v \mid \rho)\right) \quad \text{for all } v \in \mathcal{V}_N, \tag{3}$$

$$F_{\mathrm{HK}}(\rho) = \max_{v \in \mathcal{V}_N} \left(E(v) - (v \mid \rho)\right) \quad \text{for all } \rho \in \mathcal{A}_N. \tag{4}$$

We may thus calculate the ground-state energy from the universal density functional by the Hohenberg–Kohn variation principle [14]; conversely, the universal density functional is obtained from the ground-state energy by the Lieb variation principle. Together, these variation principles highlight a symmetry or duality between potentials and ground-state densities and also between the ground-state energy and the universal density functional. However, this duality was not emphasized in the work by Hohenberg and Kohn, who considered only the first variation principle equation (3). Lieb was the first to study both variation principles rigorously, within the framework of convex analysis [21].

We note the *v-representability problem* of Hohenberg–Kohn theory and the variation principles established above: the sets $\mathcal{A}_N$ and $\mathcal{V}_N$ do not form vector spaces and are not explicitly known. Also, we have no optimality conditions except by the Hohenberg–Kohn mapping $\rho \mapsto v_\rho$. These restrictions will be lifted with the development of the constrained-search formalism of Levy [20] and the convex formulation by Lieb [21].

1.4 Vector spaces of densities and potentials

To proceed, we will need to work with vector spaces that contain all densities and all potentials of interest to us. Lieb introduced the complete vector spaces

$$X = L^3(\mathbb{R}^3) \cap L^1(\mathbb{R}^3), \quad X^* = L^{3/2}(\mathbb{R}^3) + L^\infty(\mathbb{R}^3),$$

where the space of potentials X^* is dual to the space of densities X, thereby ensuring that all interactions $(v \mid \rho)$ with $\rho \in X$ and $v \in X^*$ are finite. While X^* is sufficiently large to include all Coulomb potentials, X contains all normalized densities arising from N-electron states of a finite kinetic energy – such densities are said to be *N-representable*. The set of N-representable densities $\mathcal{I}_N$ can be characterized in a simple manner,

$$\mathcal{I}_N = \left\{ \rho \in X \;\middle|\; \rho \geq 0, \int \rho(\mathbf{r})\, \mathrm{d}\mathbf{r} = N,\, T_\mathrm{W}(\rho) < +\infty \right\},$$

where $T_\mathrm{W}\colon \mathcal{I}_N \to [0, +\infty]$ is the *von Weizsäcker kinetic energy functional*

$$T_\mathrm{W}(\rho) = \frac{1}{2} \int |\nabla \rho^{1/2}(\mathbf{r})|^2\, \mathrm{d}\mathbf{r}. \tag{5}$$

It can be shown that $\mathcal{I}_N$ is a *convex set* and that T_W is a *convex function*, meaning that, for each pair of different densities $\rho_1, \rho_2 \in \mathcal{I}_N$ and every $\lambda \in (0, 1)$, the *convex combination* $\lambda\rho_1 + (1 - \lambda)\rho_2$ also belongs to $\mathcal{I}_N$ and that T_W satisfies the *convexity characterization* of a convex function:

$$T_\mathrm{W}(\lambda\rho_1 + (1 - \lambda)\rho_2) \leq \lambda T_\mathrm{W}(\rho_1) + (1 - \lambda) T_\mathrm{W}(\rho_2).$$

We note that $\mathcal{A}_N \subsetneq \mathcal{I}_N \subsetneq X$.

Even though X^* contains all Coulomb potentials, it does not contain all potentials that may be of interest to us – harmonic potentials, for example, support an electronic ground state but are not contained in X^*. Indeed, this limitation of the theory was already indicated in the title of Lieb's 1983 paper [21]: "Density functionals for Coulomb systems". The restriction is fairly mild, however, and does not reduce the usefulness of the theory significantly. In keeping with this limitation, we redefine $\mathcal{V}_N$ to be the set of potentials in X^* that support a ground state: $\mathcal{V}_N \subsetneq X^*$.

1.5 Levy–Lieb constrained-search functional

While Hohenberg–Kohn theory is based on the Schrödinger equation, the constrained-search theory is based on the *Rayleigh–Ritz variation principle* for the ground-state energy $E\colon X^* \to \mathbb{R}$, which for pure states may be written as

$$E(v) = \inf_{\Psi \mapsto N} \langle \Psi | H(v) | \Psi \rangle. \tag{6}$$

The ground-state energy $E(v)$ is well defined (finite) for each $v \in X^*$ provided the minimization is over all antisymmetric N-electron wave functions of a finite kinetic energy, as indicated by the notation $\Psi \mapsto N$. We note, however, that the infimum is not always achieved – for the oxygen atom, for example, a minimum is obtained for atoms containing up to nine electrons, but no minimum exists for more than nine electrons. In other words, the potential $v_Z(\mathbf{r}) = -Ze^2/(4\pi\varepsilon_0 r)$ with $Z = 8$ belongs to $\mathcal{V}_N$ when $N \leq 9$ but not when $N > 9$. Nevertheless, $E(v)$ as given by equation (6) is well defined in all cases and we use the term 'ground-state energy' even when the infimum is not achieved and no ground state exists.

Following the seminal work of Levy [20], we now rewrite the Rayleigh–Ritz variation principle of equation (6) in the manner

$$E(v) = \inf_{\rho \in \mathcal{I}_N} \inf_{\Psi \mapsto \rho} \langle \Psi | H(v) | \Psi \rangle,$$

where the short-hand notation $\Psi \mapsto \rho$ indicates that the minimization is restricted to wave functions with density ρ. Introducing the *Levy–Lieb constrained-search functional* $F_{\mathrm{LL}}\colon \mathcal{I}_N \to [0, +\infty]$ on the set of N-representable densities by

$$F_{\mathrm{LL}}(\rho) = \inf_{\Psi \mapsto \rho} \langle \Psi | H(0) | \Psi \rangle,$$

where $H(0) = T + W$, we arrive at the following *Hohenberg–Kohn variation principle* for $v \in X^*$ (see [20]):

$$E(v) = \inf_{\rho \in \mathcal{I}_N} (F_{\mathrm{LL}}(\rho) + (v \mid \rho)).$$

Since $\mathcal{I}_N$ and X^* are explicitly known, the v-representability problem of Hohenberg–Kohn theory has been solved.

1.6 Discontinuity of the universal density functional

Although the constrained-search theory solves the v-representability problem, the question regarding the optimality conditions in the Hohenberg–Kohn variation principle remains. Since F_{LL} is defined on the vector space X, it is tempting to assume that we can use the Euler equations for this purpose, since we no longer need to worry about whether or not the functions in the immediate neighbourhood of optimal functions are v-representable:

$$\frac{\delta F_{\mathrm{LL}}(\rho)}{\delta \rho(\mathbf{r})} = -v(\mathbf{r}) + c, \quad c \in \mathbb{R}. \tag{7}$$

However, use of the Euler equations presumes that the density functional is differentiable. Unfortunately, differentiability is precluded by the fact that the universal density-functional is discontinuous [19].

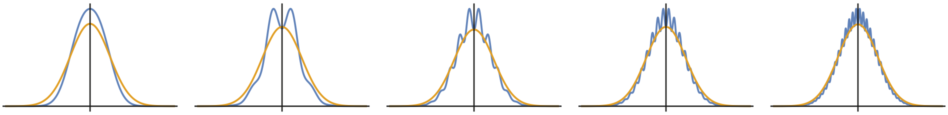

Figure 2. A convergent sequence of densities with a divergent von Weizsäcker kinetic energy.

To illustrate, consider a one-electron system, for which the universal density functional has an explicit form, being equal to the von Weizsäcker functional of equation (5). A one-electron Gaussian density of unit exponent has a finite kinetic energy:

$$\rho(\mathbf{r}) = \pi^{-\frac{3}{2}} \exp(-r^2), \quad T_{\mathrm{W}}(\rho) = \frac{3}{4} E_{\mathrm{h}}.$$

Let now $\{\rho_n\}_{n=1}^{\infty}$ be a sequence that approaches ρ in the norm of X:

$$\lim_{n \to +\infty} \|\rho - \rho_n\|_X = 0,$$

while developing increasingly rapid oscillations of increasingly small amplitude, as illustrated in Figure 2. The kinetic energy $T_{\mathrm{W}}(\rho_n)$ is then driven arbitrarily high in the sequence, implying that T_{W} is not continuous:

$$\lim_{n \to \infty} T_{\mathrm{W}}(\rho_n) = +\infty \neq T_{\mathrm{W}}\Big(\lim_{n \to \infty} \rho_n\Big) = \frac{3}{4} E_{\mathrm{h}}.$$

Since it can be shown that $T_{\mathrm{W}} \leq F_{\mathrm{LL}}$ for each $N \geq 1$, this result also means that F_{LL} is discontinuous and hence not differentiable. Consequently, the Euler equations in equation (7) are not well defined. We will return to the problem of identifying minimizers in the Hohenberg–Kohn variation principle later, when the proper tools have been developed.

1.7 Rayleigh–Ritz variation principle for ensembles

Our next step is to generalize the constrained-search theory to ensembles, taking as our starting point the *canonical-ensemble Rayleigh–Ritz variation principle*:

$$E(v) = \inf_{\gamma \mapsto N} \operatorname{tr}(\gamma H(v)). \tag{8}$$

Here each density matrix γ is a normalized self-adjoint trace-class operator on the N-electron Hilbert space, which may be given the spectral decomposition

$$\gamma = \sum_{i=1}^{\infty} \lambda_i |\Psi_i\rangle\langle\Psi_i|, \quad \lambda_i \geq 0, \quad \sum_{i=1}^{\infty} \lambda_i = 1,$$

in terms of an orthonormal basis $\{\Psi_i\}_{i=1}^{\infty}$. The expectation value in equation (8) may then be written as

$$\operatorname{tr}(\gamma H(v)) = \sum_{i=1}^{\infty} \lambda_i \langle \Psi_i | H(v) | \Psi_i \rangle.$$

Each M-degenerate ensemble ground-state density matrix obtained from (8) using $v \in \mathcal{V}_N$ is a (finite) convex combination of M pure ground-state density matrices,

$$\gamma_0 = \sum_{i=1}^{M} \lambda_{0i} |\Psi_{0i}\rangle\langle\Psi_{0i}|, \quad \lambda_{0i} \geq 0, \quad \sum_{i=1}^{M} \lambda_{0i} = 1,$$

where each Ψ_{0i} is a ground-state wave function obtained from equation (6) with the same potential v. By the same token, each ensemble ground-state density is a convex combination of pure ground-state densities with the same potential. Generalizing the concept of v-representability to ensembles, we say that a density $\rho \in \mathcal{I}_N$ is *ensemble* v*-representable* if it is the ensemble ground-state density for some external potential $v \in \mathcal{V}_N$. The set of all ensemble v-representable densities is denoted by $\mathcal{B}_N$. Clearly, $\mathcal{A}_N \subsetneq \mathcal{B}_N$.

It should be emphasized that the Rayleigh–Ritz variation principle for pure states and the Rayleigh–Ritz variation principle for ensembles give the same ground-state energy for every $v \in X^*$ – the only difference is that more solutions (minimizing density matrices) are obtained with the ensemble variation principle.

1.8 Lieb constrained-search functional

Proceeding as for pure states, we obtain from the ensemble Rayleigh–Ritz variation principle in equation (8) the *ensemble Hohenberg–Kohn variation principle*

$$E(v) = \inf_{\rho \in \mathcal{I}_N} (F_{\mathrm{DM}}(\rho) + (v \mid \rho)),$$

where the *Lieb density-matrix constrained-search functional* [21] is given by

$$F_{\mathrm{DM}}(\rho) = \inf_{\gamma \mapsto \rho} \operatorname{tr}(\gamma H(0)).$$

A nontrivial result due to Lieb [21] is that the infimum is always achieved,

$$F_{\mathrm{DM}}(\rho) = \min_{\gamma \mapsto \rho} \operatorname{tr}(\gamma H(0)) = \operatorname{tr}(\gamma_\rho H(0)), \tag{9}$$

with important implications for the theory. A similar result holds for F_{LL}.

Since the constrained search over ensembles in F_{DM} includes the constrained search over pure states in F_{LL}, we conclude that

$$F_{\mathrm{DM}} \leq F_{\mathrm{LL}}.$$

In fact, $F_{\mathrm{DM}} \neq F_{\mathrm{LL}}$, since there exist densities $\rho \in \mathcal{I}_N$ for which $F_{\mathrm{DM}}(\rho) < F_{\mathrm{LL}}(\rho)$. Nevertheless, both functionals are *admissible*, meaning that both give the correct ground-state energy for every potential in the Hohenberg–Kohn variation principle. They differ only in that F_{DM} gives more minimizing solutions – specifically, each minimizing density obtained with F_{DM} is a convex combination of those obtained with F_{LL}. The reason for this behavior is that F_{DM} is the *convex hull* of F_{LL} – that is, its greatest convex lower bound [21].

Convexity is a powerful property of F_{DM}, with far-reaching consequences for the theory. The proof is simple. Let $\rho_1, \rho_2 \in \mathcal{I}_N$ be different and let $\lambda \in (0, 1)$. There then exist by equation (9) minimizing density matrices $\gamma_1 \mapsto \rho_1$ and $\gamma_2 \mapsto \rho_2$ in F_{DM} such that $F_{\mathrm{DM}}(\rho_1) = \mathrm{tr}(\gamma_1 H(0))$ and $F_{\mathrm{DM}}(\rho_2) = \mathrm{tr}(\gamma_2 H(0))$. By the definition of F_{DM}, we then have

$$\begin{aligned} F_{\mathrm{DM}}(\lambda\rho_1 + (1-\lambda)\rho_2) &= \min_{\gamma \mapsto \lambda\rho_1 + (1-\lambda)\rho_2} \mathrm{tr}(\gamma H(0)) \\ &\leq \mathrm{tr}((\lambda\gamma_1 + (1-\lambda)\gamma_2) H(0)), \end{aligned} \tag{10}$$

where the inequality holds since $\lambda\gamma_1 + (1-\lambda)\gamma_2 \mapsto \lambda\rho_1 + (1-\lambda)\rho_2$ is an allowed (but not necessarily minimizing) density matrix in the constrained search. Reintroducing F_{DM} on the right-hand side of equation (10), we obtain

$$F_{\mathrm{DM}}(\lambda\rho_1 + (1-\lambda)\rho_2) \leq \lambda F_{\mathrm{DM}}(\rho_1) + (1-\lambda) F_{\mathrm{DM}}(\rho_2)$$

verifying convexity.

Another important property of F_{DM} is *lower semi-continuity* [21], meaning that, for each $\rho \in X$ and each sequence $\{\rho_n\}_{n=1}^{\infty}$ converging to ρ, we have

$$\liminf_{n\to\infty} F_{\mathrm{DM}}(\rho_n) \geq F_{\mathrm{DM}}(\rho).$$

Roughly speaking, lower semi-continuity at ρ means that F_{DM} may jump up but not down as we move away from ρ. Functions that are both convex and lower semi-continuous are said to be *closed convex*, an important, well-behaved class of convex functions to which F_{DM} belongs.

1.9 Minimizing densities

A useful property of convex functions is that every local minimum is a global minimum, greatly simplifying their minimization. Since F_{DM} is convex, it follows that $\rho \mapsto F_{\mathrm{DM}} + (v \mid \rho)$ is also convex, implying that *every local minimizing density in the (ensemble) Hohenberg–Kohn variation principle is a global minimizer*. Moreover, for a given potential $v \in X^*$, the set of minimizers constitute a convex set, which is nonempty if and only if $v \in \mathcal{V}_N$.

Let us consider the relationship of the minimizing densities in the Hohenberg–Kohn variation principle to the ground-state densities in the Rayleigh–Ritz variation principle. Since $F_{\mathrm{DM}}(\rho) + (v \mid \rho) = \min_{\gamma\mapsto\rho} \operatorname{tr}(\gamma H(v))$, we may write

$$E(v) = \begin{cases} \inf_{\gamma} \operatorname{tr}(\gamma H(v)), & \text{Rayleigh–Ritz variation principle,} \\ \inf_{\rho\in\mathcal{I}_N} \min_{\gamma\mapsto\rho} \operatorname{tr}(\gamma H(v)), & \text{Hohenberg–Kohn variation principle,} \end{cases} \tag{11}$$

where the searches in the two variation principles are over the same complete set of density matrices. Therefore, if $v \in \mathcal{V}_N$, then the infimum is achieved by the same density matrices in both cases:

$$E(v) = \min_{\gamma} \operatorname{tr}(\gamma H(v)) = \min_{\rho\in\mathcal{I}_N} \min_{\gamma\mapsto\rho} \operatorname{tr}(\gamma H(v)), \quad v \in \mathcal{V}_N.$$

It follows that *the minimizing densities in the Hohenberg–Kohn variation principle are precisely the ground-state densities in the Rayleigh–Ritz variation principle*. We note how this result depends critically on the existence of a minimizing density matrix $\gamma \mapsto \rho$ in the constrained-search functional in equation (9) and hence in equation (11), precluding the existence of minimizing densities in the Hohenberg–Kohn variation principle that are not ground-state densities.

In conclusion, the ensemble Hohenberg–Kohn variation principle is faithful in the sense that it gives not only the same ground-state energy as does the ensemble Rayleigh–Ritz variation principle, but also the same ground-state densities, when these exist. A similar result holds for pure states, with the Levy–Lieb functional.

1.10 The Lieb functional and variation principle

Let us now develop DFT within the framework of convex analysis. The key to this development is the *concavity* of the ground-state energy $E\colon X^* \to \mathbb{R}$ (alluded to in Section 1.2), meaning that the interpolation characterization

$$E(\lambda v_1 + (1-\lambda)v_2) \geq \lambda E(v_1) + (1-\lambda)E(v_2)$$

holds for each pair $v_1, v_2 \in X^*$ and each $\lambda \in (0,1)$. To demonstrate concavity, we first note that $H(\lambda v_1 + (1-\lambda)v_2) = \lambda H(v_1) + (1-\lambda)H(v_2)$ holds by the linearity of $v \mapsto H(v)$. We may therefore rewrite the Rayleigh–Ritz variation principle (here in a pure-state formulation) as

$$\begin{aligned} E(\lambda v_1 + (1-\lambda)v_2) &= \inf_{\Psi} \langle \Psi \mid H(\lambda v_1 + (1-\lambda)v_2) \mid \Psi\rangle \\ &= \inf_{\Psi} (\lambda\langle \Psi \mid H(v_1) \mid \Psi\rangle + (1-\lambda)\langle \Psi \mid H(v_2) \mid \Psi\rangle). \end{aligned}$$

Minimizing the two expectation values separately, we obtain

$$\begin{aligned} E(\lambda v_1 + (1-\lambda)v_2) &\geq \inf_{\Psi} \lambda\langle \Psi \mid H(v_1) \mid \Psi\rangle + \inf_{\Psi} (1-\lambda)\langle \Psi \mid H(v_2) \mid \Psi\rangle \\ &= \lambda \inf_{\Psi} \langle \Psi \mid H(v_1) \mid \Psi\rangle + (1-\lambda) \inf_{\Psi} \langle \Psi \mid H(v_2) \mid \Psi\rangle, \end{aligned} \tag{12}$$

where the last step follows since λ and $1 - \lambda$ are both nonnegative. Concavity of the ground-state energy is thus an immediate consequence of the Rayleigh–Ritz variation principle and of the linearity of the Hamiltonian in the external potential. We note that the proof of concavity does not carry through for excited states (except for the lowest state of each symmetry) since the Rayleigh–Ritz minimization is then subject to orthogonality constraints that depend on the external potential, precluding the inequality in equation (12) from being established.

Since the ground-state energy is concave and everywhere finite, it is also continuous and hence *closed concave*. We may then apply the *Fenchel–Moreau biconjugation theorem* of convex analysis to deduce the existence of a closed convex function $F: X \to \overline{\mathbb{R}}$ such that

$$E(v) = \inf_{\rho \in X} (F(\rho) + (v \mid \rho)), \qquad v \in X^*, \tag{13}$$

$$F(\rho) = \sup_{v \in X^*} (E(v) - (v \mid \rho)), \quad \rho \in X, \tag{14}$$

where $\overline{\mathbb{R}} = [-\infty, \infty]$ denotes the extended real numbers. The function F is the *Lieb universal density functional*. The Lieb functional F and the ground-state energy E are said to be *conjugate functions* or *Fenchel conjugates*. Since F can be calculated from E and *vice versa*, the two functions contain the same information, only encoded in different ways – each property of one function is exactly reflected in some property of its conjugate function. *Fenchel conjugation* is a generalization of the Legendre transformation – the relationship between the ground-state energy and the universal density functional in DFT is thus similar to the relationship between the Hamiltonian and the Lagrangian in classical mechanics.

Comparing the Hohenberg–Kohn variation principles in equation (9) and equation (13), we note that the variation is over all X in equation (13) but only over $\mathcal{I}_N$ in equation (9). If F_{DM} is extended from $\mathcal{I}_N$ to X by setting it equal to $+\infty$ on $X \setminus \mathcal{I}_N$, then the extended F_{DM} remains closed convex and satisfies the same Hohenberg–Kohn variation principle as does the Lieb functional F in equation (9). However, by the Fenchel–Moreau theorem, there exists only one closed convex function F on X that is conjugate to the closed concave function E on X^*, meaning that the Lieb functional in equation (14) and the density-matrix constrained-search functional in equation (9) are the same functional [21]:

$$F = F_{\mathrm{DM}}.$$

Furthermore, this functional is the lower bound to all admissible density functionals and the most well behaved such functional, being closed convex.

In short, DFT is essentially an exercise in convex analysis predicated on the concavity and continuity of the ground-state electronic energy $v \mapsto E(v)$. The concepts and tools of convex analysis are therefore well suited to DFT, from a computational as well as theoretical point of view.

1.11 Convex conjugation: A one-dimensional differentiable model

To understand the conjugate relationship between the ground-state energy and the universal density functional better, we will consider two one-dimensional models of the energy $\mathcal{E}\colon \mathbb{R} \to \mathbb{R}$ and the density functional $\mathcal{F}\colon \mathbb{R} \to \overline{\mathbb{R}}$ that satisfy the Fenchel conjugations

$$\mathcal{E}(v) = \inf_{\rho\in\mathbb{R}} (\mathcal{F}(\rho) + (v \mid \rho)), \quad \mathcal{F}(\rho) = \sup_{v\in\mathbb{R}} (\mathcal{E}(v) - (v \mid \rho)), \tag{15}$$

where the interaction between the potential and density is now the simple scalar product $(v|\rho) = v\rho$. We begin by considering in this section the differentiable functions plotted in Figure 3. In agreement with the true ground-state energy, we assume that $\mathcal{E} \leq 0$ with $\mathcal{E}(0) = 0$. In addition, we assume that $\mathcal{E}$ is twice differentiable with $\mathcal{E}'' < 0$ and $\lim_{v\to\pm\infty} \mathcal{E}'(v) = \mp\infty$. Strict concavity is unphysical since it implies that $\mathcal{E}(v) < 0$ for repulsive potentials $v > 0$, but we accept this unphysical behaviour for the time being in order to explore the consequences of strict concavity.

Consider now the condition for a maximizing potential in the Lieb variation principle in equation (15). By the derivative assumptions on $\mathcal{E}$, a maximizing potential $v \in \mathbb{R}$ exists for every $\rho \in \mathbb{R}$ as the solution to the stationarity condition $\rho = \mathcal{E}'(v)$ – in this case, even for the nonphysical negative densities. Since $\mathcal{E}'$ is strictly decreasing, this mapping from potentials to densities can be inverted to give the Hohenberg–Kohn mapping from densities to potentials, which we recognize as the stationary condition in the Hohenberg–Kohn variation principle: $v = -\mathcal{F}'(\rho)$. We thus have two equivalent optimality conditions:

$$-\mathcal{F}'(\rho) = v \iff \mathcal{E}'(v) = \rho. \tag{16}$$

The ground-state energy and the universal density functional are therefore *functions whose derivatives (to within a negative sign) are each other's inverse functions*. These inverse relationships are illustrated in Figure 3, which shows how we may generate

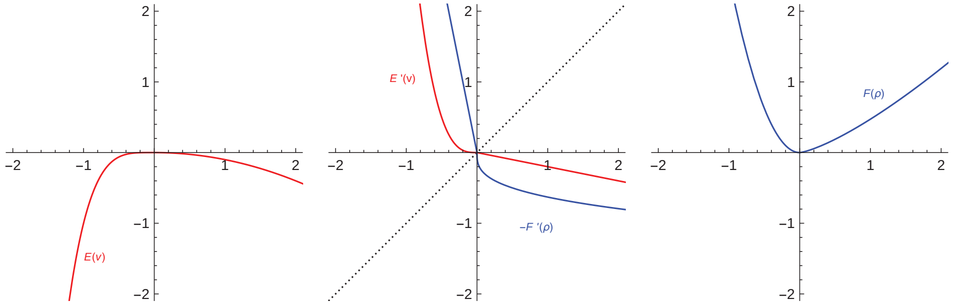

Figure 3. A strictly concave function $\mathcal{E}\colon \mathbb{R} \to (-\infty, 0]$ (left) with a strictly convex conjugate function $\mathcal{F} \to [0, +\infty)$ (right). Their derivatives are plotted in the middle.

$\mathcal{F}$ by differentiation of $\mathcal{E}$ (left plot), inversion of $\mathcal{E}'$ to obtain $-\mathcal{F}'$ (middle plot), followed by integration of $\mathcal{F}'$ to yield $\mathcal{F}$ (right plot). Although simple, this example illustrates the essence of DFT.

1.12 Convex conjugation: A one-dimensional nondifferentiable model

We have seen how, in a simple model, the Hohenberg–Kohn mapping follows from strict concavity of the ground-state energy. The true energy is not strictly concave, however, only concave. For example, when a scalar μ is added to the potential, the energy changes linearly, in the manner $\mu \mapsto E(v+\mu) = E(v) + \mu N$, where N is the number of electrons. Also, we have $E(v) = 0$ for all repulsive potentials $v \geq 0$, thereby violating strict concavity. There may be more ways in which strict concavity is violated, but these two violations can easily be explored by a one-dimensional model. In Figure 4, we have plotted a model energy $\mathcal{E}\colon \mathbb{R} \to (-\infty, 0]$ that is linear on $[-1, -\frac{1}{2}]$ with slope $\frac{3}{2}$ and linear on $[0, +\infty)$ with slope 0. Also, $\mathcal{E}$ is nondifferentiable at -1 and $-\frac{1}{2}$.

The effect on the density functional $\mathcal{F}\colon \mathbb{R} \to \overline{\mathbb{R}}$ by this modification is dramatic. First, $\mathcal{F}(\rho) = +\infty$ for nonphysical densities $\rho < 0$, which is a consequence of the fact that $\mathcal{E}(v) = 0$ for all repulsive potentials $v > 0$:

$$\mathcal{F}(\rho) = \sup_{v \in \mathbb{R}} (\mathcal{E}(v) - (v \mid \rho)) \geq \sup_{v>0} (\mathcal{E}(v) - (v \mid \rho)) = \sup_{v>0} (-(v \mid \rho)) = +\infty.$$

Second, $\mathcal{F}$ becomes nondifferentiable at $\rho = 0$ and $\rho = \frac{3}{2}$ – that is, at the slopes of the linear segments of $\mathcal{E}$. From the middle plots in Figure 4, we note that each horizontal segment of $v \mapsto (v, \mathcal{E}'(v))$ corresponds to a vertical segment of $\rho \mapsto (\rho, \mathcal{F}'(\rho))$ and *vice versa*, meaning that each linear segment of $v \mapsto (v, \mathcal{E}(v))$ corresponds to a kink of $\rho \mapsto (\rho, \mathcal{F}(\rho))$ and *vice versa*. Roughly speaking, therefore, we may associate nondifferentiability of $\mathcal{F}$ with nonstrict concavity of $\mathcal{E}$.

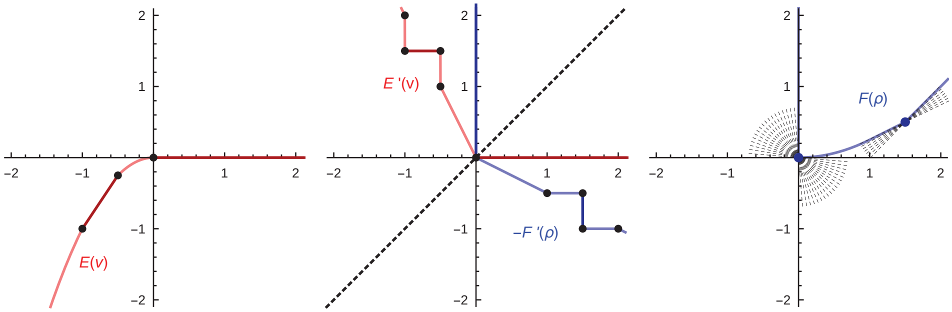

Figure 4. A concave function $\mathcal{E}\colon \mathbb{R} \to (-\infty, 0]$ (left) and its convex conjugate function $\mathcal{F}\colon \mathbb{R} \to [0, +\infty]$ (right) with their subgradients plotted in the middle.

Since E and F are nondifferentiable, we cannot express the optimality conditions in terms of derivatives, as done in equation (16). Instead, we introduce for a convex function $f : \mathbb{R} \to \overline{\mathbb{R}}$ the *subdifferential* of f at x by

$$\partial f(x) = [f'_{-}(x), f'_{+}(x)], \tag{17}$$

where $f'_{-}(x)$ and $f'_{+}(x)$ are the left and right derivatives, respectively, of f at x. Each $m \in \partial f(x)$ is then the slope of a supporting line to f and x, known as a *subgradient* of f at x. Returning to Figure 4, we see that $\partial \mathcal{F}(\rho) = \{\mathcal{F}'(\rho)\}$ everywhere except at the nondifferentiable points, where $\partial \mathcal{F}(0) = (-\infty, 0]$ and $\partial \mathcal{F}(\frac{3}{2}) = [\frac{1}{2}, 1]$. At these points, we have plotted supporting lines to $\mathcal{F}$ for illustration. For the ground-state energy, we have the equality $\partial \mathcal{E}(v) = \{\mathcal{E}'(v)\}$ everywhere except $\partial \mathcal{E}(-1) = [\frac{3}{2}, 2]$ and $\partial \mathcal{E}(-\frac{1}{2}) = [1, \frac{3}{2}]$.

Subgradients allow global minima (maxima) of convex (concave) functions to be identified in a simple manner: for a convex (concave) function $f : \mathbb{R} \to \overline{\mathbb{R}}$, a point $x \in \mathbb{R}$ is a global minimizer (maximizer) if and only if $0 \in \partial f(x)$. The condition for a minimum in the Hohenberg–Kohn variation principle in equation (15) is therefore that the subdifferential of the convex objective function $\rho \mapsto \mathcal{F}(\rho) + (v \mid \rho)$ contains zero. Since $\partial(\mathcal{F}(\rho) + (v \mid \rho)) = \partial \mathcal{F}(\rho) + v$, this condition is equivalent to $-v \in \partial \mathcal{F}(\rho)$; likewise, the condition for a global maximum in the Lieb variation principle becomes $\rho \in \partial \mathcal{E}(v)$. In terms of subgradients, the optimality conditions of the Hohenberg–Kohn and Lieb variation principles are therefore

$$-v \in \partial \mathcal{F}(\rho) \iff \rho \in \partial \mathcal{E}(v),$$

which reduce to the stationary conditions in equation (16) when both functions are differentiable. The ground-state energy $\mathcal{E}$ and the density functional $\mathcal{F}$ are therefore *functions whose subdifferential mappings $\partial \mathcal{E} : \mathbb{R} \rightrightarrows \mathbb{R}$ and $\partial \mathcal{F} : \mathbb{R} \rightrightarrows \mathbb{R}$ are each other's inverse multifunctions*, as illustrated in the middle plot of Figure 4.

1.13 Hohenberg–Kohn and Lieb optimality conditions

To set up the optimality conditions of the Hohenberg–Kohn and Lieb variation principles for the exact ground-state energy E and universal density functional F, we must generalize subdifferentials and subgradients to functions on vector spaces. For the concave ground-state energy $E : X^* \to \mathbb{R}$ and the convex density functional $F : X \to \overline{\mathbb{R}}$, the subdifferentials at $v \in X^*$ and $\rho \in X$ are, respectively, given by

$$\partial E(v) = \{\rho \in X \mid E(\tilde{v}) \leq E(v) + (\tilde{v} - v \mid \rho) \text{ for all } \tilde{v} \in X^*,\ E(v) \in \mathbb{R}\}, \tag{18}$$

$$\partial F(\rho) = \{v \in X^* \mid F(\tilde{\rho}) \geq F(\rho) + (v|\tilde{\rho} - \rho) \text{ for all } \tilde{\rho} \in X,\ F(\rho) \in \mathbb{R}\}. \tag{19}$$

The subdifferentials $\partial E(v) \subset X$ and $\partial F(\rho) \subset X^*$ are convex sets, which may or may not be empty. In equation (18), we recognize the subgradient inequality from equation (2), where it was defined on $\mathcal{V}_N \subset X$. Also, it is a simple exercise to check that $\partial F(\rho)$ in equation (19) reduces to that of equation (17) for convex functions on the real axis.

Proceeding as in the one-dimensional case, we find that the optimality conditions of the Hohenberg–Kohn and Lieb variation principles are given by

$$E(v) = F(\rho) + (v \mid \rho) \iff -v \in \partial F(\rho) \iff \rho \in \partial E(v). \tag{20}$$

The subdifferential mappings $\partial F(\rho)\colon X \rightrightarrows X^*$ and $\partial E(v)\colon X^* \rightrightarrows X$ clearly play an important role in DFT, setting up the mappings between densities and potentials.

It is instructive to consider in detail the subgradient condition on the density functional. Using the constrained-search expression for $F_{\mathrm{DM}} = F$ given in equation (9), we obtain-1mm

$$\begin{aligned} -v \in \partial F(\rho) &\iff (\forall \tilde{\rho} \in X)\colon F(\tilde{\rho}) \geq F(\rho) - (v|\tilde{\rho} - \rho) \\ &\iff (\forall \tilde{\rho} \in X)\colon F(\tilde{\rho}) + (v|\tilde{\rho}) \geq F(\rho) + (v \mid \rho) \\ &\iff (\forall \tilde{\rho} \in X)\colon \min_{\gamma \mapsto \tilde{\rho}} \operatorname{tr}(\gamma H(v)) \geq \min_{\gamma \mapsto \rho} \operatorname{tr}(\gamma H(v)) \\ &\iff \inf_{\gamma} \operatorname{tr}(\gamma H(v)) \geq \operatorname{tr}(\gamma_\rho H(v)), \end{aligned} \tag{21}$$

where γ_ρ is a ground-state density matrix of $H(v)$, thereby verifying that the conditions $-v \in \partial F(\rho) \iff \rho \in \partial E(v)$ hold if and only if $\rho \in \mathcal{B}_N$ is an ensemble ground-state density of $v \in \mathcal{V}_N$. In particular, $\rho \notin \mathcal{B}_N$ if and only if $\partial F(\rho) = \emptyset$. By a general result of convex analysis, F is subdifferentiable on a dense subset of $\mathcal{I}_N$. The set of ensemble v-representable densities is therefore dense in the set of N-representable densities.

Next, we verify that the subgradients in $\partial F(\rho)$ are identical up to a scalar. Let ρ be a ground-state density associated with the external potential $v_\rho \in \mathcal{V}_N$ so that $-v_\rho \in \partial F(\rho)$. Using equation (20) and the expression for the subgradient inequality given in (2), we obtain

$$\begin{aligned} -v_\rho \in \partial F(\rho) &\iff \rho \in \partial E(v_\rho) \\ &\iff \begin{cases} E(u) < E(v_\rho) + (u - v_\rho|\rho) & \text{for all } u \notin v_\rho + \mathbb{R}, \\ E(u) = E(v_\rho) + (u - v_\rho|\rho) & \text{for all } u \in v_\rho + \mathbb{R}, \end{cases} \\ &\iff \begin{cases} E(v_\rho) > E(u) + (v_\rho - u|\rho) & \text{for all } u \notin v_\rho + \mathbb{R}, \\ E(v_\rho) = E(u) + (v_\rho - u|\rho) & \text{for all } u \in v_\rho + \mathbb{R}. \end{cases} \end{aligned}$$

After this rearrangement of the subgradient inequality, we use equation (18) to give

$$\begin{aligned} -v_\rho \in \partial F(\rho) &\iff \begin{cases} \rho \notin \partial E(u) & \text{for all } u \notin v_\rho + \mathbb{R}, \\ \rho \in \partial E(u) & \text{for all } u \in v_\rho + \mathbb{R}, \end{cases} \\ &\iff \begin{cases} -u \notin \partial F(\rho) & \text{for all } u \notin v_\rho + \mathbb{R}, \\ -u \in \partial F(\rho) & \text{for all } u \in v_\rho + \mathbb{R}, \end{cases} \end{aligned} \tag{22}$$

in accordance with the Hohenberg–Kohn theorem. Our results from (21) and (22) are summarized as follows:

$$\partial F(\rho) = \begin{cases} \emptyset, & \rho \notin \mathcal{B}_N, \\ -v_\rho + \mathbb{R}, & \rho \in \mathcal{B}_N, \end{cases}$$

$$\partial E(v) = \begin{cases} \emptyset, & v \notin \mathcal{V}_N, \\ \left\{\sum_{i=1}^M \lambda_i \rho_{vi} \;\middle|\; \sum_{i=1}^M \lambda_i = 1,\, \lambda_i \geq 0\right\}, & v \in \mathcal{V}_N, \end{cases}$$

where v_ρ is an external potential such that $H(v_\rho)$ supports a ground state with ground-state density $\rho \in \mathcal{B}_N$, whereas $\{\rho_{vi}\}_{i=1}^M$ are the M degenerate pure ground-state densities of $v \in \mathcal{V}_N$. Note that $\partial E(v)$ is empty when v does not support a ground state.

1.14 Moreau–Yosida regularization

Realizing that nondifferentiability is an unavoidable feature of the density functional conjugate to the exact ground-state energy, we may instead work with the modified energy functional

$$E_\gamma(v) = E(v) - \frac{1}{2}\gamma\|v\|^2 \tag{23}$$

with $\gamma > 0$. Modified in this manner, the ground-state energy becomes strictly convex with a differentiable conjugate density functional

$$F_\gamma(\rho) = \min_{\tilde{\rho}}\left(F(\tilde{\rho}) + \frac{1}{2\gamma}\|\tilde{\rho} - \rho\|^2\right),$$

which is the *Moreau–Yosida regularization* of the Lieb functional [17]. The *smoothing parameter* $\gamma > 0$ can be adjusted, noting that $F_\gamma(\rho)$ approaches $F(\rho)$ pointwise from below as $\gamma \to 0$ from above. It is an attractive regularization in that, for each v, the true energy $E(v)$ can be recovered exactly from $E_\gamma(v)$ according to equation (23).

A slight complication of the Moreau–Yosida regularization is that Coulomb potentials are not square-integrable, as required to construct E_γ. This problem may be avoided by putting the system in a large box or by truncating the Coulomb potential at large separations. In any case, Moreau–Yosida regularization shows that, even though the exact density functional is not differentiable, it is “almost” differentiable, justifying the construction of differentiable density-functional approximations.

1.15 Four-way correspondence of density-functional theory

An important aspect of Lieb’s approach to DFT is its flexibility in describing extended density functional theories. In particular, we consider two cases, orbital-free DFT and DFT in the presence of a magnetic field. In both of these cases, the energy becomes a bifunctional with dependence on additional variables. In *orbital-free DFT*, the density functional in the context of grand canonical ensemble DFT depends on both the charge density ρ and the chemical potential μ. For DFT in the presence of an external magnetic field, two approaches have been suggested: *magnetic-field density-functional theory (BDFT)* [6, 28], in which the density functional depends on the charge density ρ and the external vector potential associated with the magnetic field $\mathbf{A}$; and *current-density functional theory (CDFT)* [39, 40], in which the density functional depends on the charge density ρ and the paramagnetic component of the current density $\mathbf{j}_{\mathrm{p}}$. A useful concept for these extended theories is the *four-way correspondence of DFT*, which helps to identify the relationships between the bifunctionals in each of these approaches.

Figure 5 shows the four-way correspondence for orbital-free DFT in the left panel and for DFT in a magnetic field in the right panel. In orbital-free DFT, the energy $\mathcal{E}(v, N)$ is a functional of the external potential v and the particle number N. The corresponding density functional depends on the charge density ρ and the chemical potential μ. This highlights the nature of the optimization problem in orbital-free DFT as a convex–concave saddle function. In [26], an optimization scheme was developed explicitly accounting for the saddle nature of the optimization, resulting in more rapid convergence for finite systems with all electrons included.

For DFT in the presence of a magnetic field, the four-way correspondence in Figure 5 clarifies the connections between the functionals involved in BDFT and CDFT. In particular, the variation principle used in BDFT corresponds to the partial conjugation $\rho \to u$ from $\mathcal{F}$ to $\mathcal{E}$, whilst in CDFT the variation principle corresponds to the full conjugation $(\rho, \mathbf{j}_{\mathrm{p}}) \to (u, \mathbf{A})$ from $\mathcal{G}$ to $\mathcal{E}$ – see [25] for a more detailed dis-

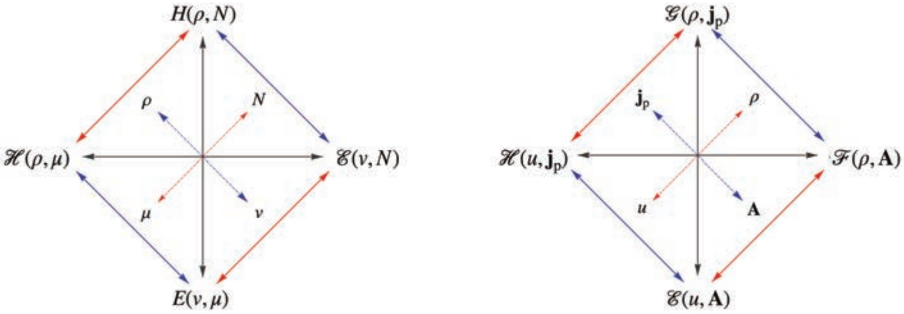
$H(\rho,N)$
ρ
N
$\mathcal{H}(\rho,\mu)$
$\mathcal{E}(v,N)$
μ
v
$E(v,\mu)$
$\mathcal{G}(\rho,\mathbf{j}_{\mathrm{p}})$
$\mathbf{j}_{\mathrm{p}}$
ρ
$\mathcal{H}(u,\mathbf{j}_{\mathrm{p}})$
$\mathcal{F}(\rho,\mathbf{A})$
u
$\mathbf{A}$
$\mathcal{E}(u,\mathbf{A})$

Figure 5. The four-way correspondence for orbital-free DFT (left) and DFT in a magnetic field (right). The black horizontal and vertical arrows represent bi-conjugations of both variables simultaneously. The solid diagonal arrows indicate skew conjugations of one of the variables independently. The dashed blue and red arrows indicate the dual relationships between the relevant variables.

cussion of these functionals. The full mathematical characterization of CDFT using convex analysis similar to Lieb's original formulation for standard DFT has recently been completed [18, 37].

In both of these cases, the extended density functional theories can be described using the techniques of convex analysis as in Lieb's pioneering 1983 paper [21]. In doing so, key features of the mathematical structure of the approaches and their interrelations are clearly revealed.

2 Applications of the Lieb variation principle

The work of Lieb to establish the variation principle of (14) has led not only to a clear and rigorous framework for understanding the concepts underpinning DFT, but also to a wide range of practical applications. We review various applications of the Lieb variation principle, demonstrating how it may be utilized to calculate the Kohn–Sham potential of atoms and molecules, to study the adiabatic connection and calculate the exchange–correlation functional from high-precision many-body methods, and to study the exchange–correlation hole and energy densities of atoms and molecules.

We here pragmatically treat the density functional as differentiable, both as a function of the density and of the interaction strength. We justify this practice by noting that all approximate density functionals are taken to be differentiable and that the exact density functional is "almost differentiable" in the sense that it may be approximated to any accuracy by a differentiable Moreau–Yosida regularized functional [17].

2.1 Studying the Kohn–Sham noninteracting system

Consider a simple generalization of the Hamiltonian in (1) to include an interaction-strength dependent two-electron interaction (in atomic units),

$$H_\lambda(v) = T + W_\lambda + \sum_{i=1}^{N} v(\mathbf{r}_i), \tag{24}$$

where the *interaction strength* between two electrons i, j is modulated by a parameter $\lambda \in \mathbb{R}$,

$$W_\lambda = \sum_{i<j} w_\lambda(r_{ij}), \quad w_0(r_{ij}) = 0, \quad w_1(r_{ij}) = \frac{1}{r_{ij}}, \tag{25}$$

here chosen such that $\lambda = 0$ corresponds to a noninteracting system of electrons and $\lambda = 1$ to the Hamiltonian for the usual physical system. The ground-state energy associated with this Hamiltonian is E_λ and the corresponding interaction-strength dependent Lieb variation principle [21] is

$$F_\lambda(\rho) = \sup_{v \in X^*} \left(E_\lambda(v) - (v \mid \rho)\right). \tag{26}$$

This interaction-strength dependent functional provides a powerful tool to study electronic systems at any interaction strength. In particular, in Kohn–Sham theory, we are often concerned with a noninteracting system of electrons ($\lambda = 0$) having the same density as that of the physical system ($\lambda = 1$). The properties of this system can be readily studied using equation (26) simply by setting $\lambda = 0$ and using an accurate density for the physical system as input to $F_0(\rho)$.

Savin and Colonna [3, 30] presented the first practical calculations of the Lieb variation principle in finite basis sets for atomic systems. In their approach, the optimization of the external potential v was carried out directly using the derivative-free Nelder–Mead method. To facilitate this optimization, the external potential $v_{\mathrm{ext}}(\mathbf{r})$ was expressed in terms of a basis-set expansion with an additional $\frac{C}{r}$ term, whose constant C was determined to account for the asymptotic form of the potential with increasing distance r from the nucleus. Following this work, Wu and Yang [41] employed a similar expansion for the external potential,

$$v(\mathbf{r}) = v_{\mathrm{ext}}(\mathbf{r}) + v_{\mathrm{ref}}(\mathbf{r}) + \sum_t b_t g_t(\mathbf{r}), \tag{27}$$

where $v_{\mathrm{ext}}(\mathbf{r})$ is the external potential due to the nuclei as usually evaluated, $v_{\mathrm{ref}}(\mathbf{r})$ is a choice of reference potential selected to capture the long-range asymptotic decay of the potential, and the final term is a linear combination of Gaussian basis functions $g_t(\mathbf{r})$ with coefficients b_t that are to be determined using the Lieb variation principle of equation (26). This expansion is convenient for implementation in finite-basis molecular codes and allows for direct optimization with respect to b_t using derivative-based (quasi-)Newton approaches. Calculations at the $\lambda = 0$ limit were presented in [41], with significantly accelerated convergence relative to the Nelder–Mead approach of [3]. Many choices of reference potential have been considered for use in equation (27). Commonly, the Fermi–Amaldi potential [4] has been used. However, since this potential is not size-consistent, alternatives such as the Slater [31], Krieger–Li–Iafrate [16], and localized-Hartree–Fock [27] exchange potentials have also been utilized as size-consistent alternatives.

The noninteracting case of equation (26) is somewhat computationally simpler than its (partially) interacting counterparts. To see this, consider the $\lambda = 0$ limit of $H_\lambda(v)$ in equation (24); in this limit, the electronic Schrödinger equation becomes separable into a set of one-electron equations

$$\left[-\frac{1}{2}\nabla_{\mathbf{r}}^2 + v(\mathbf{r})\right]\varphi_i(\mathbf{r}) = \varepsilon_i \varphi_i(\mathbf{r}). \tag{28}$$

which may be solved by diagonalization to give the canonical molecular orbitals $\varphi_i(\mathbf{r})$ and orbital energies $\varepsilon_i = \langle \varphi_i \mid -\frac{1}{2}\nabla_i^2 + v(\mathbf{r}) \mid \varphi_i \rangle$. The total noninteracting energy is then simply $E_0(v) = \sum_i^{n_{\mathrm{occ}}} \varepsilon_i$ with the associated noninteracting wave function $\Phi_0 = \det|\varphi_1 \cdots \varphi_{n_{\mathrm{occ}}}|$. Substitution of this expression into equation (26) for a v-representable density ρ leads to

$$F_0(\rho) = \max_{v \in X^*} \left(\sum_{i=1}^{n_{\mathrm{occ}}} \langle \varphi_i \mid -\tfrac{1}{2}\nabla_{\mathbf{r}}^2 \mid \varphi_i \rangle + (v|\rho_{\Phi_0} - \rho) \right), \tag{29}$$

where ρ_{Φ_0} is the density associated with Φ_0. The computational advantage of this form for $F_0(\rho)$ is clear since no two-electron integrals are required and no choice of (partially) interacting wave-function ansatz is required. Instead, all operations amount to evaluating matrix elements of one-electron operators and solving equation (28) with the potential of equation (27) at each step of the optimization procedure used to evaluate equation (29). At convergence, $\rho_{\Phi_0} = \rho$ and the Lieb functional becomes the noninteracting Kohn–Sham kinetic energy,

$$F_0(\rho) = T_{\mathrm{s}}(\rho).$$

Also, for $\lambda = 0$, the derivative $\frac{\partial^2 F_0(\rho)}{\partial b_t \partial b_u}$ takes a simple form and so second-order optimization schemes such as the trust-region Newton method can be readily employed, further accelerating convergence.

Given a sufficiently accurate input density ρ, which may be obtained from wave-function-based quantum-chemical methods, equation (29) yields accurate estimates of the Kohn–Sham noninteracting kinetic energy $T_{\mathrm{s}}(\rho)$, the Kohn–Sham orbitals φ_i and eigenvalues ε_i, and the Kohn–Sham effective potential $v_{\mathrm{s}}(\mathbf{r})$, which may be identified with the potential of equation (27) at convergence. However, two complications arise for finite (orbital) basis-set calculations.

Firstly, as pointed out by Harriman [9, 10], in a finite orbital basis, there is no Hohenberg–Kohn theorem. As a result, the optimizing potential in equation (26) may not be unique. In a complete orbital basis, the potential is determined up to a constant, leading to a well-defined shape of the Kohn–Sham potential from equation (29), for example, and the constant can then be chosen arbitrarily, usually so that the potentials vanish asymptotically. In finite basis sets, the extra nonuniqueness of the potential manifests itself in a near singular behavior of $\frac{\partial^2 F_0(\rho)}{\partial b_t \partial b_u}$ and convergence to oscillatory potentials.

Secondly, in finite basis sets, an input density arising from a correlated wave function may not be exactly represented by the density of a single Slater determinant in the same basis set; see [9, 10]. In practical calculations, this may manifest itself as $|\rho_{\Phi_0}(\mathbf{r}) - \rho(\mathbf{r})| > 0$ for some points $\mathbf{r}$. It is therefore necessary to carefully check the results of calculations to ensure that any residual density differences are sufficiently small. In practice, we find that, for small atomic and molecular systems, Gaussian basis sets of augmented triple-zeta quality are adequate to obtain close approximations to correlated input densities and that this problem is essentially removed for basis sets of quadruple-zeta quality and beyond.

To avoid oscillatory potentials in calculations in finite basis sets, it is necessary to perform calculations in a regularized manner. Strategies for regularization include using approaches such as singular-value decomposition or Tikhonov regularization in second-order optimizations to avoid issues with the singular components of $\frac{\partial^2 F_0(\rho)}{\partial b_t \partial b_u}$. An alternative approach, which emphasizes smoothness of the calculated potentials is the smoothing-norm approach of [11]. In this approach, a quadratic penalty function is applied to the energy that penalizes oscillatory solutions. This approach has the advantage that the objective function, gradient, and Hessian are then consistently modified leading to simple implementation with essentially any optimization procedure. Interestingly, this approach is closely related to the Moreau–Yosida regularization of convex analysis; see the discussion in [17]. Finally, we note that a similar regularization can be achieved by tailoring the basis set chosen for use in the expansion of equation (27) – see, for example, [13]. Whilst all of these approaches are effective in avoiding oscillatory solutions, the results of the calculations must be carefully assessed to ensure that the density $\rho_{\Phi_0}(\mathbf{r})$ remains sufficiently close to the input density ρ and that the associated potentials are not over-regularized.

Figure 6 shows the exchange–correlation potentials obtained from the Lieb variation principle at $\lambda = 0$. The main features of the potential are clearly captured, including the $-\frac{1}{r}$ asymptotic decay and inter-shell structure. The inset shows the valence region, where the potentials reproducing the electron density calculated from Hartree–Fock theory and coupled-cluster single–doubles–perturbative-triples (CCSD(T)) theory are most different.

2.2 Lieb calculations of the adiabatic connection

The adiabatic connection is a central concept in DFT, connecting the noninteracting Kohn–Sham system with the fully interacting physical system. The *density-fixed adiabatic connection*, in particular, can be determined by considering how the λ-interacting Lieb functional may be expressed in terms of its noninteracting limit and a correction term,

$$\begin{aligned} F_\lambda(\rho) &= F_0(\rho) + \int_0^\lambda F'_\nu(\rho)\,\mathrm{d}\nu \\ &= T_\mathrm{s}(\rho) + \int_0^\lambda \mathcal{W}_\nu(\rho)\,\mathrm{d}\nu; \end{aligned}$$

here the prime indicates differentiation with respect to interaction strength ν. Expressing $F_\lambda(\rho)$ in its constrained-search form

$$F_\lambda(\rho) = \min_{\gamma \mapsto \rho} \mathrm{tr}(\gamma H_\lambda(0)) = \mathrm{tr}(\gamma_\lambda^\rho H_\lambda(0)),$$

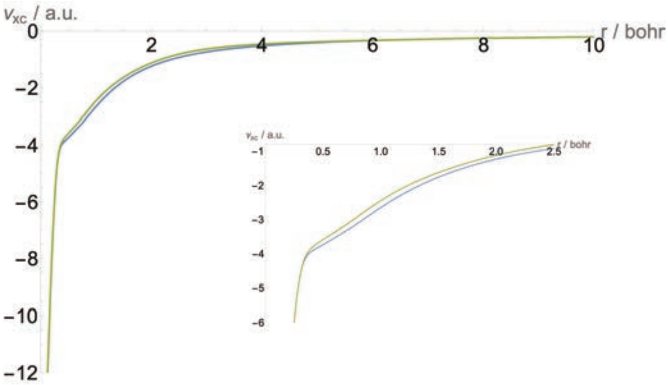

Figure 6. The exchange–correlation potentials obtained by performing the Lieb optimization at $\lambda = 0$ for the neon atom using the Hartree–Fock density (blue) and CCSD(T) density (green) in the uncontracted aug-cc-pCVQZ basis set. The inset highlights the differences in the valence region arising from the treatment of electron correlation. The potentials are determined using the smoothing-norm approach with a regularization parameter of 10^{-5} a.u.

differentiating with respect to the interaction strength, and using the well-known Hellmann–Feynman theorem, we obtain for the linear adiabatic-connection path with $w_\lambda(r_{ij}) = \frac{\lambda}{r_{ij}}$ in equation (25) the *adiabatic-connection integrand*

$$\mathcal{W}_\lambda(\rho) = \mathrm{tr}(\gamma_\lambda^\rho W_\lambda').$$

The Lieb variation principle therefore allows us to calculate the density-fixed adiabatic connection directly, by performing calculations for $0 \le \lambda \le 1$ using (26)–(27) as discussed in [3, 30, 32–36, 41].

The integrand $\mathcal{W}_\nu(\rho)$ accounts for all electron–electron interactions. It is customary to decompose this quantity further into the *classical Coulomb energy*,

$$J_\lambda(\rho) = \frac{1}{2} \iint w_\lambda(r_{12}) \rho(\mathbf{r}_1) \rho(\mathbf{r}_2) \, \mathrm{d}\mathbf{r}_1 \, \mathrm{d}\mathbf{r}_2,$$

the *exchange energy*

$$E_{\mathrm{x},\lambda}(\rho) = \mathrm{tr}\big(\gamma_0^\rho W_\lambda\big) - J_\lambda(\rho),$$

and the *correlation energy*

$$E_{\mathrm{c},\lambda}(\rho) = \int_0^\lambda \mathcal{W}_{\mathrm{c},\nu}(\rho) \, \mathrm{d}\nu, \quad \mathcal{W}_{\mathrm{c},\nu}(\rho) = \mathrm{tr}\big((\gamma_\nu^\rho - \gamma_0^\rho) W_\nu'\big), \tag{30}$$

so that

$$F_\lambda(\rho) = F_0(\rho) + J_\lambda(\rho) + E_{\mathrm{x},\lambda}(\rho) + E_{\mathrm{c},\lambda}(\rho).$$

Since the integrand is the derivative of each component with respect to the interaction strength, we see that, for the linear adiabatic connection path with $w_\lambda(r_{ij}) = \frac{\lambda}{r_{ij}}$, the contributions from the classical Coulomb energy and the exchange energy are constants. The shape of the adiabatic connection is therefore entirely determined by the correlation contribution of equation (30).

In fact, the shape of the linear-path adiabatic connection can be well understood from a perturbative analysis at $\lambda = 0$ using *Görling–Levy (GL) perturbation theory* [5]. The correlation energy can be expanded as

$$E_{c,\lambda}(\rho) = \sum_{n=2}^{\infty} \lambda^n E_{\mathrm{GL}}^{(n)}(\rho),$$

where $E_{\mathrm{GL}}^{(n)}$ is the nth-order GL correlation energy, for which explicit expressions may be readily derived and implemented. In particular, we see that the expansion of $\mathcal{W}_{c,\lambda}(\rho) = \mathcal{W}_\lambda(\rho) - J_\lambda(\rho) - E_{x,\lambda}(\rho)$ is given by

$$\begin{aligned} \mathcal{W}_{c,\lambda}(\rho) &= \sum_{n=1}^{\infty} \lambda^n \mathcal{W}_{\mathrm{GL}}^{(n)}(\rho) \\ &= \sum_{n=1}^{\infty} \lambda^n (n+1) E_{\mathrm{GL}}^{(n+1)}(\rho), \end{aligned} \tag{31}$$

from which is it clear that the slope of the linear-path adiabatic connection is $2E_{\mathrm{GL}}^{(2)}(\rho)$. For dynamically correlated systems, where low-order perturbative expansions provide a good approximation to $E_{c,\lambda}(\rho)$, plotting $\lambda \mapsto \mathcal{W}_\lambda(\rho)$ gives smooth, almost linear curves. For systems with more significant static correlation (i.e., where a single Kohn–Sham determinant is not close to the physical multi-determinantal wave function), the corresponding plots display significantly more curvature.

The Lieb variation principle is therefore a powerful tool to study not only the Kohn–Sham noninteracting system, but also the entire adiabatic connection, giving access to all of the associated λ-interacting energies, wave functions and associated density matrices. To perform practical studies of the adiabatic connection for many-electron systems in finite basis sets, a key step is to choose an appropriate, sufficiently accurate, ansatz to determine the input density ρ in equation (26) and to calculate the energy E_λ at each interaction strength. A natural choice for few-electron systems is full configuration-interaction (FCI) theory, as employed in [3, 33]. However, the entire repertoire of quantum-chemical methods can be utilized to study larger systems. Some care is required, however, when setting up optimization procedures to perform the Lieb maximization in equation (26) when the ansatz for E_λ is nonvariational, such as for Møller–Plesset and coupled-cluster theories. In particular, the required derivatives are most conveniently evaluated using the Lagrangian densities for these approaches [12]; see [33, 34] for further details.

2.3 The adiabatic connection for H_2

A simple illustrative example, for which FCI calculations can be readily performed, is the H_2 molecule. At short internuclear separations R, this system has a physical wave function that is well represented by a single Slater determinant. Correlation effects, whilst important for chemical accuracy, are relatively subtle and the CI expansion has one large coefficient on the Hartree–Fock determinant, with minor contributions from excited determinants – characteristic of dynamic correlation. At large R, the system has a FCI wave function with large coefficients on more than one determinant – characteristic of significant static correlation. At the dissociation limit, the physical wave function would be well represented by just two determinants, one for each hydrogen atom. Intermediate bond lengths capture the regime where the both dynamic and static correlation effects play a role.

The adiabatic-connection curves for $\lambda \mapsto \mathcal{W}_{c,\lambda}(\rho)$ calculated using the Lieb variation principle of equation (26) at the FCI level in the uncontracted aug-cc-pCVQZ basis (for both the orbital and potential expansions) are presented in Figure 7 for $R = 1.4$, 3.0, 5.0 and 7.0 bohr. These plots capture the transition from near the equilibrium geometry to almost dissociated hydrogen atoms and the corresponding transition from almost purely dynamic to static correlation.

As expected from the perturbative analysis of equation (31), the plot at equilibrium geometry displays only subtle curvature – characteristic of low-order dynamical correlation effects. As R increases, the plots exhibit more rapid curvature – indicative of more significant higher-order correlation contributions and the relevance of static correlation. Finally, at large R, close to dissociation, we see a characteristic “L” curve, a feature of (almost) pure static correlation. This shape arises because the restricted Kohn–Sham single determinant cannot well represent the λ-interacting wave function, which consists of essentially one determinant for each hydrogen atom at this geometry. Since the hydrogen atoms are widely separated, there is no dynamical correlation between the electrons. As a result, the curve abruptly changes for small λ values and then is essentially flat, reflecting the fact that, at the dissociation limit for a restricted Kohn–Sham reference wave function, $E_x(\rho) = -\frac{1}{2}J(\rho)$. Therefore, to cancel the electron–electron interactions at the dissociation limit, we must have $E_c(\rho) = -\frac{1}{2}J(\rho)$. The curve presented at $R = 7.0$ bohr shows that the behavior with respect to λ is already approximately constant, reflecting the linear behavior of $J(\rho)$ with respect to λ in equation (2.2).

2.4 An alternative path: Generalized adiabatic connections

As indicated in equation (25), the noninteracting and physical systems can be connected by any number of alternative paths – such adiabatic connections have been

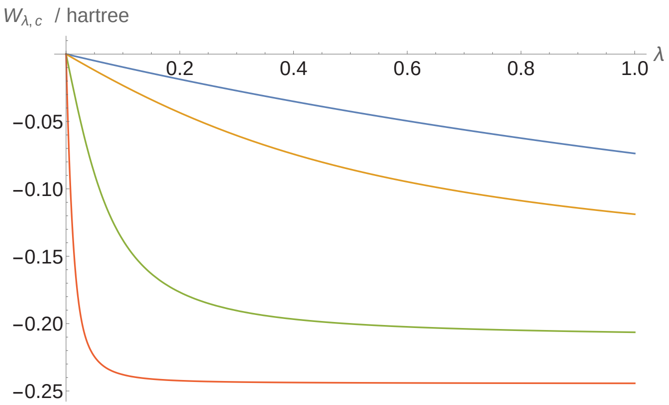

Figure 7. The linear adiabatic connection for the H_2 molecule at internuclear separations $R = 1.4, 3.0, 5.0$ and 7.0 bohr (top to bottom).

called generalized adiabatic connections [42]. The Lieb variation principle of equation (26) can for such connections be applied in the same manner as described in Section 2.2 but with $E_\lambda(\rho)$ calculated using $W_\lambda(\rho)$ defined with a different modification of the electron–electron interaction.

In [35], some common adiabatic paths of relevance to range-separated methods were examined. In range-separated methods [29], a distinction is made between long-ranged and short-ranged interactions of electrons. These interactions are typically partitioned by modifying the electron–electron interaction operator. A common approach is to use the error function to modulate the electron–electron interactions,

$$w_\lambda(r_{ij}) = \frac{\mathrm{erf}(\mu(\lambda) r_{ij})}{r_{ij}}, \qquad \mu(\lambda) = \frac{\lambda}{1-\lambda}.$$

For small values of λ, the long-range interactions are captured, with the shorter-range interactions being emphasized as λ increases towards 1. The value of the Lieb functional varies between $F_0(\rho) = T_s(\rho)$ and $F_1(\rho)$ in a manner that reflects the relative importance of different ranges of electron–electron interaction. In Figure 8, the contributions become progressively localized to small λ values as the electron–electron interactions become predominantly long-ranged at large internuclear separation.

The Lieb variation principle provides insight into the extent to which dynamic and static correlation can be divided according to the range of interaction. By exploiting the generalized adiabatic-connection paths in conjunction with knowledge of how the generalized adiabatic connection behaves for accurate calculations, it may be possible to devise new approximations to the exchange–correlation energy that more effectively exploit range-separation techniques.

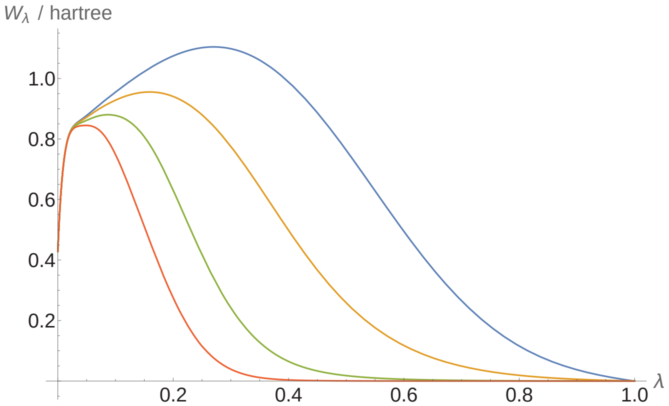

Figure 8. The error-function generalized adiabatic connection for the H_2 molecule at internuclear separations $R = 1.4, 3.0, 5.0$ and 7.0 bohr (top to bottom).

2.5 The exchange–correlation hole and energy densities

The exchange–correlation hole is a central concept in DFT and its analysis has given insight into why commonly used semi-local density-functional approximations can be so successful – see, for example, [1]. The exchange–correlation energy can be expressed in terms of the *exchange–correlation energy density* $w_{\mathrm{xc},\lambda}$ as

$$E_{\mathrm{xc}}(\rho) = \int_0^1 \int \rho(\mathbf{r}) w_{\mathrm{xc},\lambda}(\mathbf{r})\, \mathrm{d}\mathbf{r}\, \mathrm{d}\lambda, \quad w_{\mathrm{xc},\lambda}(\mathbf{r}) = \frac{1}{2} \int \frac{h_{\mathrm{xc},\lambda}(\mathbf{r},\mathbf{r}')}{|\mathbf{r}-\mathbf{r}'|}\, \mathrm{d}\mathbf{r}'. \tag{32}$$

The *exchange–correlation hole* at interaction strength λ is given by

$$h_{\mathrm{xc},\lambda}(\mathbf{r},\mathbf{r}') = \frac{P_{2,\lambda}(\mathbf{r},\mathbf{r}')}{\rho(\mathbf{r})} - \rho(\mathbf{r}'),$$

where the pair density $P_{2,\lambda}(\mathbf{r},\mathbf{r}')$ is obtained using the Lieb variation principle. At $\lambda = 0$, the exchange–correlation hole reduces to the Fermi (exchange) hole and so the Coulomb (correlation) hole can be identified as

$$h_{\mathrm{c},1}(\mathbf{r},\mathbf{r}') = h_{\mathrm{xc},1}(\mathbf{r},\mathbf{r}') - h_{\mathrm{xc},0}(\mathbf{r},\mathbf{r}').$$

In Figure 9, the exchange–correlation holes for the H_2 molecule at internuclear separations $R = 1.4$, 3.0, 5.0 and 7.0 bohr are presented, plotted along the bond axis with the position of the reference electron 0.3 bohr to the left of the second hydrogen atom, as done in [1]. All required quantities were calculated using the Lieb variation principle at the FCI level of theory.

At $R = 1.4$ bohr, it is clear that the exchange hole is delocalized with a significant amplitude on both atoms. The Coulomb hole is negative close to the reference

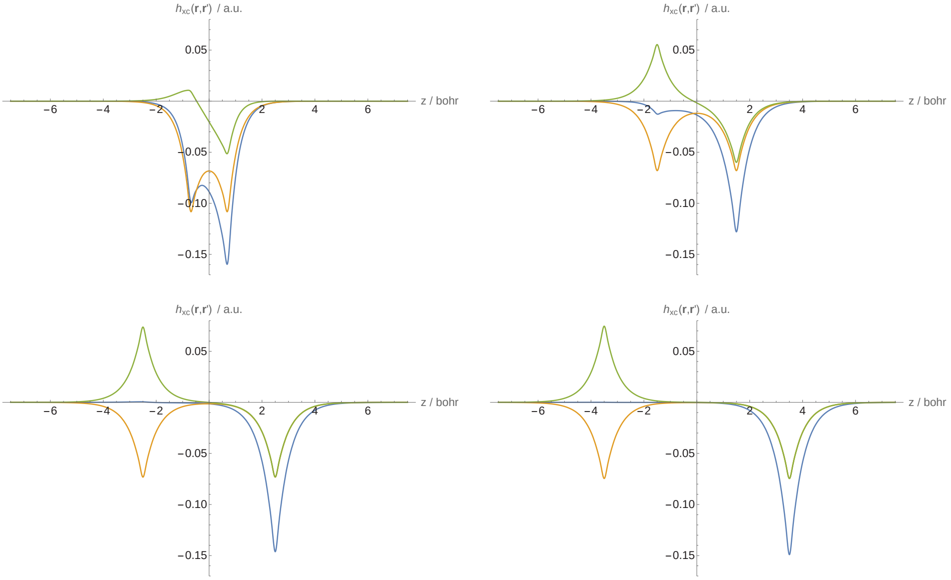

Figure 9. The Fermi (orange, $\lambda = 0$), Coulomb (green, $\lambda = 1$) and exchange–correlation (blue, $\lambda = 1$) holes for H_2 at internuclear separations $R = 1.4$ (top left), 3.0 (top right), 5.0 (lower left) and 7.0 (lower right) bohr plotted along the internuclear axis with the position of the reference electron chosen to be 0.3 bohr left of the right-hand hydrogen nucleus.

point and positive close to the other hydrogen nucleus, leading to an exchange–correlation hole that is more localized around the reference point than either of its constituent components. As the bond length increases, this effect becomes more and more pronounced, with the total exchange–correlation hole being relatively strongly localized around the reference point already at $R = 3.0$ bohr, with only relatively modest amplitude on the left hydrogen atom. For the longer bond lengths, the localization of the exchange–correlation hole is essentially complete, with cancellation of the Fermi and Coulomb holes far from the reference point.

This behavior of the exchange–correlation hole rationalizes to some extent the success of popular semi-local exchange–correlation functionals. Semi-local approximations such as generalized-gradient approximations cannot be expected to capture the strong nonlocality of the Fermi and Coulomb holes. However, when exchange and correlation are treated together, the overall exchange–correlation hole is more localized and can be effectively modelled by these simple density functionals. For more detailed discussion of these ideas, see [1].

The exchange–correlation hole is a rather complex quantity, depending explicitly on the coordinates of two electrons. It is therefore desirable to consider quantities that reduce the complexity, whilst retaining enough information to determine the exchange–correlation energy. One possibility is to consider the energy density $w_{\mathrm{xc},\lambda}$ directly as a target for functional development, since one electronic coordinate has already been removed. Furthermore, it is possible to interchange the order of integration in equation (32) to generate a *coupling-constant averaged energy density*, thereby also removing the λ dependence. This simpler function can then be analyzed in a similar manner to the exchange–correlation potentials generated by approximate functionals – see, for example, [15]. Another alternative is to consider the angle- and system-averaged exchange–correlation hole, which takes advantage of the isotropic nature of the Coulomb interaction,

$$\bar{h}_{\mathrm{xc}}(u) = \frac{1}{N}\int \rho(\mathbf{r}) \int h_{\mathrm{xc}}(\mathbf{r}, \mathbf{r}+\mathbf{u}) \frac{\mathrm{d}\Omega_{\mathbf{u}}}{4\pi}\, \mathrm{d}\mathbf{r},$$

where $\mathbf{u} = \mathbf{r}' - \mathbf{r}$ and the integration over the solid angle $\Omega_{\mathbf{u}}$ averages over all directions for $\mathbf{u}$. The exchange–correlation energy can then be expressed as

$$E_{\mathrm{xc}} = N \int_0^\infty 4\pi u^2 \frac{\bar{h}_{\mathrm{xc}}(u)}{2u}\, \mathrm{d}u.$$

Both of these alternatives provide simpler functions that can be targeted in functional development. Again, the Lieb variation principle provides access to all of the quantities required to compute these functions accurately using *ab-initio* methods. The resulting data can then serve as a benchmark for approximate models.

2.6 Beyond the ground state

The *Gross–Olivera–Kohn (GOK) approach* to excitation energies [7,8,23] in ensemble DFT can also be studied using the Lieb variation principle [22]. In the GOK approach, a density matrix of a statistical ensemble is formed as a convex combination of the M lowest states of the Hamiltonian,

$$\gamma_{\lambda,w} = \sum_{i=1}^{M} w_i |\Psi_{\lambda,i}\rangle\langle\Psi_{\lambda,i}|,$$

where the ensemble weights w_i satisfy $0 \leq w_{i+1} \leq w_i$ and $\sum_i w_i = 1$. The ground state is $|\Psi_{\lambda,0}\rangle$ and $|\Psi_{\lambda,i\geq 1}\rangle$ are excited states. For an implementation of the Lieb variation principle for such ensembles, see [2]. Excitation energies can be evaluated as

$$E_1 - E_0 = (\varepsilon_1 - \varepsilon_0) + \left.\frac{\partial E_{\mathrm{xc},w}}{\partial w}\right|_{\rho=\rho_w}.$$

Here $\varepsilon_1 - \varepsilon_0$ is a Kohn–Sham eigenvalue difference, which is corrected by a partial-derivative term evaluated as

$$\left.\frac{\partial E_{\mathrm{xc},w}}{\partial w}\right|_{\rho=\rho^{w}} = \frac{\partial}{\partial w}\left(F_{1,w}(\rho) - F_{0,w}(\rho)\right)\bigg|_{\rho=\rho_{w}} - J(\rho_{w}).$$

In this way, the Lieb variation principle may also offer insights into the behavior of the correction term in the GOK approach to excitation energies; see [2] for further details.

Other approaches to excited states have been developed using perturbation theory along the linear and generalized adiabatic connections [24], which were discussed in Sections 2.2 and 2.4. In these cases, the Lieb variation principle was used to determine the ground-state adiabatic connections and then excitation energies estimated by perturbation at each value of the interaction strength. The Lieb variation principle may provide a useful tool for further study of time-independent approaches to excitation energies in the future.

3 Conclusions

We have seen how DFT is predicated on the concavity and continuity of the ground-state energy in the external potential – from this fact, DFT follows by application of convex analysis. In particular, the universal density functional and the ground-state energy are conjugate functions, containing the same information but encoded in different ways. While the ground-state energy can be obtained from the universal density functional by the Hohenberg–Kohn variation principle, the density functional can be calculated from the ground-state energy by the Lieb variation principle.

The Lieb variation principle is not only theoretically important, giving insight into the structure of DFT, but it is also a practical computational tool. It allows us to calculate, for any N-representable density, the universal density functional and the external potential that supports this density (if such a potential exists). The utility of this approach has been briefly demonstrated here for determining the Kohn–Sham potential, the adiabatic connection and detailed information on the exchange and correlation holes, to high accuracy using *ab-initio* many-body wave function approaches. This tool can give valuable insight into the numerical behavior of the exact universal density functional, which may serve as a useful benchmark to guide the development of approximate models.

It is our view that the beautiful convex formulation of DFT first put forward by Lieb [21] has not yet received the attention it deserves, as a framework for teaching DFT and as a tool for further development of DFT – we hope the present overview goes some way towards rectifying this situation. We also highlighted how Lieb's approach can be utilized for extended DFTs such those required for systems in a magnetic field and for orbital-free DFT. This further illustrates how Lieb's convex formulation of DFT can be used to unify the presentation of different variants and extensions of DFT and clarify the relationships between them. Lieb's convex formulation of DFT continues to illuminate the development of state-of-the-art approaches in DFT almost forty years after its publication.

Acknowledgments. We thank Dr. Tom J. P. Irons and Dr. Bang Huynh for useful discussions in the preparation of this manuscript.

Funding. This work was partially supported by the Research Council of Norway through its Centres of Excellence scheme, project number 262695. We acknowledge financial support from the European Research Council under H2020/ERC Consolidator Grant top DFT (Grant No. 772259).

References

[1] E. J. Baerends and O. V. Gritsenko, A quantum chemical view on density functional theory. *J. Phys. Chem. A* **101** (1997), 5383–5403

[2] A. Borgoo, A. M. Teale, and T. Helgaker, Excitation energies from ensemble dft. *AIP Conf. Proc.* **1702** (2015), art. 090049

[3] F. Colonna and A. Savin, Correlation energies for some two- and four-electron systems along the adiabatic connection in density functional theory. *J. Chem.Phys.* **110** (1999), 2828–2835

[4] E. Fermi and E. Amaldi, Le orbite ∞ s degli elementi. *Mem. Accad. Italia* **117** (1934), 119–149

[5] A. Görling and M. Levy, Exact Kohn–Sham scheme based on perturbation theory. *Phys.Rev. A* **50** (1994), 196–204

[6] C. J. Grayce and R. A. Harris, Magnetic-field density-functional theory. *Phys. Rev. A* **50** (1994), 3089–3095

[7] E. K. U. Gross, L. N. Oliveira, and W. Kohn, Rayleigh–Ritz variational principle for ensembles of fractionally occupied states. *Phys. Rev. A* **37** (1988), 2805–2808

[8] E. K. U. Gross, L. N. Oliveira, and W. Kohn, Density-functional theory for ensembles of fractionally occupied states. I. Basic formalism. *Phys. Rev. A* **37** (1988), 2809–2820

[9] J. E. Harriman, Geometry of density-matrices. 4. The relationship between density-matrices and densities. *Phys. Rev. A* **27** (1983), 632–645

[10] J. E. Harriman, Densities, operators, and basis-sets. *Phys. Rev. A* **34** (1986), 29–39

[11] T. Heaton-Burgess, F. A. Bulat, and W. Yang, Optimized effective potentials in finite basis sets. *Phys. Rev. Lett.* **98** (2007), art. 256401

[12] T. Helgaker and P. Jørgensen, Configuration-interaction energy derivatives in a fully variational formulation. *Theor. Chim. Acta* **75** (1989), 111–127

[13] A. Heßelmann, A. W. Götz, F. Della Sala, and A. Görling, Numerically stable optimized effective potential method with balanced gaussian basis sets. *J. Chem. Phys.* **127** (2007), art. 054102

[14] P. Hohenberg and W. Kohn, Inhomogeneous electron gas. *Phys. Rev.* **136** (1964), B864–B871

[15] T. J. P. Irons and A. M. Teale, The coupling constant averaged exchange-correlation energy density. *Mol. Phys.* **114** (2016), 484–497

[16] J. B. Krieger, Y. Li, and G. J. Iafrate, Derivation and application of an accurate Kohn–Sham potential with integer discontinuity. *Phys. Lett. A* **146** (1990), 256–260

[17] S. Kvaal, U. Ekstrom, A. M. Teale, and T. Helgaker, Differentiable but exact formulation of density-functional theory, *J. Chem.Phys.* **140** (2014), art. 18A518

[18] S. Kvaal, A. Laestadius, E. Tellgren, and T. Helgaker, Lower semicontinuity of the universal functional in paramagnetic current-density functional theory. *J. Phys. Chem. Lett.* **12** (2021), 1421–1425

[19] P. E. Lammert, Differentiability of lieb functional in electronic density functional theory. *Int. J. Quantum Chem.* **107** (2007), 1943–1953

[20] M. Levy, Universal variational functionals of electron densities, first-order density matrices, and natural spin-orbitals and solution of the v-representability problem. *Proc. Natl. Acad. Sci. USA* **76** (1979), 6062–6065

[21] E. H. Lieb, Density functionals for Coulomb systems. *Int. J. Quantum Chem.* **24** (1983), 243–277

[22] E. H. Lieb, Density functionals for Coulomb systems. In *Inequalities*, edited by R. M. Dreizler, pp. 31–80, Nato Asi Ser. 123, Plenum, New York, 1985

[23] L. N. Oliveira, E. K. U. Gross, and W. Kohn, Density-functional theory for ensembles of fractionally occupied states. 2. Application to the He atom. *Phys. Rev. A* **37** (1988), 2821–2833

[24] E. Rebolini, J. Toulouse, A. M. Teale, T. Helgaker, and A. Savin, Excited states from range-separated density-functional perturbation theory. *Mol. Phys.* **113** (2015), 1740–1749

[25] S. Reimann, A. Borgoo, E. I. Tellgren, A. M. Teale, and T. Helgaker, Magnetic-field density-functional theory (BDFT): Lessons from the adiabatic connection. *J. Chem. Theory Comput.* **13** (2017), 4089–4100

[26] M. S. Ryley, M. Withnall, T. J. P. Irons, T. Helgaker, and A. M. Teale, Robust all-electron optimization in orbital-free density-functional theory using the trust-region image method. *J. Phys. Chem. A* **125** (2021), 459–475

[27] F. D. Sala and A. Görling, Efficient localized Hartree–Fock methods as effective exact-exchange Kohn–Sham methods for molecules. *J. Chem. Phys.* **115** (2001), 5718–5732

[28] F. R. Salsbury and R. A. Harris, The current in magnetic field density functional theory and its application to the chemical shielding and magnetic susceptibility. *J. Chem. Phys.* **107** (1997), 7350–7359

[29] A. Savin, On degeneracy, near-degeneracy and density functional theory. In *Recent developments and applications of modern density functional theory*, p. 327, Elsevier, Amsterdam, 1996

[30] A. Savin, F. Colonna, and R. Pollet, Adiabatic connection approach to density functional theory of electronic systems. *Int. J. Quantum Chem.* **93** (2003), 166–190

[31] J. C. Slater, A simplification of the Hartree–Fock method. *Phys. Rev.* **81** (1951), 385–390

[32] M. D. Strømsheim, N. Kumar, S. Coriani, E. Sagvolden, A. M. Teale, and T. Helgaker, Dispersion interactions in density-functional theory: An adiabatic-connection analysis. *J. Chem. Phys.* **135** (2011), art. 194109

[33] A. M. Teale, S. Coriani, and T. Helgaker, The calculation of adiabatic-connection curves from full configuration-interaction densities: Two-electron systems. *J. Chem. Phys.* **130** (2009), art. 104111

[34] A. M. Teale, S. Coriani, and T. Helgaker, Accurate calculation and modeling of the adiabatic connection in density functional theory. *J. Chem. Phys.* **132** (2010), art. 164115

[35] A. M. Teale, S. Coriani, and T. Helgaker, Range-dependent adiabatic connections. *J. Chem. Phys.* **133** (2010), art. 164112

[36] A. M. Teale, S. Coriani, and T. Helgaker, Range-dependent adiabatic connections. *AIP Conf. Proc.* **1504** (2012), 92–99

[37] E. I. Tellgren, S. Kvaal, E. Sagvolden, U. Ekstrom, A. M. Teale, and T. Helgaker, Choice of basic variables in current-density-functional theory. *Phys. Rev. A* **86** (2012), art. 062506

[38] J. van Tiel, *Convex analysis*. John Wiley & Sons, New York, 1984

[39] G. Vignale and M. Rasolt, Density-functional theory in strong magnetic fields. *Phys. Rev. Lett.* **59** (1987), 2360–2363

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

[41] Q. Wu and W. Yang, A direct optimization method for calculating density functionals and exchange-correlation potentials from electron densities. *J. Chem. Phys.* **118** (2003), 2498–2509

[42] W. Yang, Generalized adiabatic connection in density functional theory. *J. Chem. Phys.* **109** (1998), 10107–10110