

PERSPECTIVE • OPEN ACCESS

Perspective on Moreau–Yosida regularization in density-functional theory

To cite this article: Markus Penz *et al* 2026 *Electron. Struct.* **8** 022001

View the [article online](#) for updates and enhancements.

You may also like

- [Autler–Townes splitting in Rydberg atoms: transition dipole matrix element extraction and field efficiency analysis](#)
Brian C Holloway, Gavin M Chase, Lee E Harrell et al.
- [Long-term carbon dioxide removal potential from the application of wood biochar and basanite rock powder in sandy soil using the LiDELv2 process-based modeling approach](#)
Mikita Maslouski, Maria Ansari, Susanne E Hamburger et al.
- [2D material-based sensing technology innovation roadmap for biomedical applications](#)
Apostolos C Tsolakis, Thomas Zadrozny, Aristides Bakandritsos et al.

Electronic Structure



PERSPECTIVE

OPEN ACCESS

RECEIVED

9 November 2025

REVISED

11 March 2026

ACCEPTED FOR PUBLICATION

23 March 2026

PUBLISHED

8 April 2026

Original content from this work may be used under the terms of the [Creative Commons Attribution 4.0 licence](#).

Any further distribution of this work must maintain attribution to the author(s) and the title of the work, journal citation and DOI.



Perspective on Moreau–Yosida regularization in density-functional theory

Markus Penz^{1,2,3,4} , Michael F Herbst^{5,6} , Trygve Helgaker⁷ and Andre Laestadius^{3,7,*}

¹ Arnold Sommerfeld Center for Theoretical Physics, Ludwig-Maximilians-Universität München, Munich, Germany

² Munich Center for Quantum Science and Technology, Munich, Germany

³ Department of Computer Science, Oslo Metropolitan University, Oslo, Norway

⁴ Max Planck Institute for the Structure and Dynamics of Matter, Hamburg, Germany

⁵ Mathematics for Materials Modelling (MatMat), Institute of Mathematics & Institute of Materials, École Polytechnique Fédérale de Lausanne, Lausanne, Switzerland

⁶ National Centre for Computational Design and Discovery of Novel Materials (MARVEL), École Polytechnique Fédérale de Lausanne, Lausanne, Switzerland

⁷ Hylleraas Centre for Quantum Molecular Sciences, Department of Chemistry, University of Oslo, Oslo, Norway

* Author to whom any correspondence should be addressed.

E-mail: andre.laestadius@oslomet.no

Keywords: density-functional theory, Moreau–Yosida regularization, Kohn–Sham theory, density-potential inversion

Abstract

Within density-functional theory (DFT), Moreau–Yosida regularization enables both a reformulation of the theory and a mathematically well-defined definition of the Kohn–Sham approach. It is further employed in density–potential inversion schemes and, through the choice of topology for the density and potential space, can be directly linked to classical field theories. This perspective collects various appearances of the regularization technique within DFT alongside possibilities for their future development.

Contents

1. Introduction	2
2. Preliminaries	3
2.1. Functional analysis preliminaries	3
2.2. Convex analysis preliminaries	5
2.3. MY regularization	6
2.4. Formal DFT	8
2.5. KS method	12
3. MY regularization in DFT	14
3.1. MY based on the universal density functional	14
3.2. MY based on the total energy functional	15
3.3. MY based on density–potential mixing	16
4. Density–potential inversion	17
4.1. Proximal-point iteration and relation to ZMP	17
4.2. MY in inverse KS	18
4.3. Resolvent form and Dyson-like equation	19
5. Regularized KS theory	19
5.1. Iteration scheme	20
5.2. Transforming exchange–correlation potentials	21
5.3. Convergence proof	21
6. Auto-regularization in mean-field theories	23
6.1. Hartree auto-regularization	23
6.2. Maxwell–Schrödinger DFT	25
7. Application of MY-based KS inversion (MYiKS) to periodic settings	27

8. Outlook	31
Acknowledgments	32
Data availability statement	33
References	33

1. Introduction

It has been more than 10 years since Moreau–Yosida (MY) regularization was introduced into density-functional theory (DFT) [1], and it has arguably turned out to be an unexpectedly fruitful extension. First introduced as a conceptual tool to fix differentiability issues of the exact functional and to make the exchange–correlation potential mathematically well defined, it was later realized that it can be seen as the theoretical basis of density–potential inversion techniques. In order to connect it to existing methods, like the Zhao–Morrison–Parr inversion scheme [2], the topology of the density and potential spaces needed to be adjusted in a way that is critically different from the standard setting of DFT introduced by Lieb [3] based on Lebesgue spaces L^p . This recent realization [4] marked an important shift of perspective, since it became apparent that the topology of those spaces can entail physical meaning. More concretely, the duality map that canonically connects the density space and the potential space as its dual, is just the Poisson equation of electrostatics if the spaces are chosen as special Sobolev spaces. In this way, new physical meaning and connections to classical fields enter the theory and need to be interpreted accordingly. At this stage of development, many exciting routes can be taken, be it about the precise choice of topology, the extension of the spaces to include vector potentials, or the utilization of the regularized theory for computational purposes.

This perspective article aims at showing the current state of the field as well as mapping out several directions for continued research that we deem promising. We believe that, as the theory matures, it can establish itself as a mathematically rigorous but also computationally useful extension of DFT. Ideally, this means offering new and enhanced approximations for electronic-structure calculations, while also covering a wider range of physical phenomena. A main result of this work (discussed in section 2.4) is that the necessary function-space framework for MY regularization can be combined with the usual potential space of quantum mechanics only if the setting is such that it always guarantees existence of a ground state. Another interesting finding, laid out in section 3.3, is the possible connection of the topology to physics, in that the norm of the density and potential spaces exactly gives their energy content and that they are connected via the Poisson equation. These and many more insights into MY regularization applied to DFT are discussed in this perspective article, where we also make the relevant connections to the literature.

Since the original intention of introducing MY regularization into DFT was to make the theory mathematically sound, it is necessary to introduce a certain amount of mathematical concepts. We try to present this in the preliminary section 2 with minimal requirements, but the interested reader might want to follow up on certain details by consulting the vast mathematical literature on the fields that we mainly rely on. These are functional and convex analysis formulated on Banach and Hilbert spaces, as well as the more specialized theory of Sobolev spaces, and the reference list includes our favorite textbooks [5–12]. After covering the most important mathematical topics in sections 2.1–2.3, we continue in section 2.4 with a formalized presentation of DFT that introduces the universal density functional for a fully interacting system and for an auxiliary, non-interacting system. This is complemented by section 2.5 on the Kohn–Sham (KS) method, which we introduce with an additional offset potential to capture *a priori* information such as external binding potentials. Section 3 is concerned with introducing MY regularization into DFT in three different ways—namely, based on the universal density functional, based on the total energy functional, and based on density–potential mixing. The next two sections describe the main use cases for MY regularization in DFT to date: a technique for density–potential inversion in section 4 and a rigorous formulation of the KS method that takes a central role in the application of DFT in section 5. Section 6 shows how certain formulations of DFT, such as Hartree theory and Maxwell–Schrödinger DFT, get automatically regularized through the inclusion of a mean-field energy contribution. Section 7 deals with the numerical realization of density–potential inversion in the KS setting for periodic systems. Finally, we conclude with an extended outlook in section 8. We provide a brief notational aid for the most frequently used symbols in table 1.

Table 1. Notational aid for frequently used symbols.

X	Generic Banach space for densities	} Section 2.1 &
X^*	Generic Banach space for potentials	
D	Set of N -representable densities	Equation (29)
J	Duality map from X to X^*	Equation (5)
$f^*(u)$	Legendre–Fenchel transform (convex conjugate)	Equation (9)
$\partial f(x)$	Sub/superdifferential of a convex/concave functional	Equation (11)
$f_\varepsilon(x)$	MY regularization	Equation (15)
$\nabla f_\varepsilon(x)$	Functional derivative	Section 2.3
$\text{prox}_{f_\varepsilon}(x)$	Proximal map	Equation (16)
$F^\lambda(\rho)$	Universal functional	Equation (31)
$E^\lambda(v)$	Ground-state energy functional	Equation (24)
$F_\varepsilon^\lambda(x)$	MY-regularized universal functional	Equation (49)
$E_\varepsilon^\lambda(v)$	Convex conjugate of $F_\varepsilon^\lambda(x)$	Equation (50)

2. Preliminaries

The following five subsections introduce the mathematical and theoretical prerequisites for performing MY regularization in DFT and for studying related topics. Section 2.1 deals with some basic notions of functional analysis, most importantly certain properties of Banach spaces that are later required for MY regularization as well as defining the duality map. Section 2.2 gives some central results needed for the convex analysis of DFT. Section 2.3 is the last mathematical part and introduces MY regularization with its most important properties. Section 2.4 continues with a presentation of formal DFT and contains an important result that connects the properties of the function spaces with the existence of ground states. Finally, section 2.5 briefly introduces the KS method of DFT.

2.1. Functional analysis preliminaries

Our development of MY regularization applied to DFT relies heavily on topological properties of the involved density and potential spaces, the Banach space X and its topological dual X^* . We keep them as abstract spaces for most of our presentation, since this allows the application of the results to a huge variety of different settings. It will be indispensable to repeat important results from functional analysis, especially from the theory of Banach spaces, since this will be the first pillar on which we base DFT mathematically. The properties of the spaces X and X^* are collected in three different assumption sets of increasing restrictiveness, to which later results will refer.

The basic setting for the whole presentation is a real Banach space X with the norm denoted as $\|\cdot\|_X$. Banach spaces have been developed from the beginning as spaces of functions [13] and as such typically are infinite dimensional, which complicates their mathematical study. The feature that makes a normed vector space a Banach space is that it is *complete*, meaning that every Cauchy sequence in X has a well-defined limit within the space X . Convergence to a limit value usually means *strong convergence*—that is, convergence in the norm $\|\cdot\|_X$, where $x_n \rightarrow x$ in X if and only if

$$\lim_{n \rightarrow \infty} \|x_n - x\|_X = 0. \quad (1)$$

Since the norm of the space naturally induces a topology (a notion of ‘closeness’ and ‘continuity’), we are able to define the *topological dual* of X as the space of linear and continuous maps from X to $\mathbb{R}$ (called ‘functionals’),

$$X^* = \{u: X \rightarrow \mathbb{R} \mid u \text{ linear and continuous}\}. \quad (2)$$

Now, X^* gets equipped with a norm itself, the operator norm on X , which we will accordingly denote with $\|\cdot\|_{X^*}$,

$$\|u\|_{X^*} = \sup_{\|x\|_X \leq 1} |u(x)|. \quad (3)$$

The completeness of X then carries over to X^* , making it a Banach space as well. With the help of X^* , we can define *weak convergence* $x_n \rightarrow x$ in X as the property that, for all $u \in X^*$, it holds that

$$\lim_{n \rightarrow \infty} u(x_n - x) = \lim_{n \rightarrow \infty} u(x_n) - u(x) = 0. \quad (4)$$

As suggested by the terminology, strong convergence of a sequence implies weak convergence. Further, let

$$J(x) = \{u \in X^* \mid u(x) = \|x\|_X^2 = \|u\|_{X^*}^2\} \quad (5)$$

be a set-valued map that maps each element in X to a subset of X^* , called the *duality map*. This generalizes the Riesz representation for Hilbert spaces, that is a one-to-one map, to the case of Banach spaces. The Hahn–Banach theorem guarantees that the image of every $x \in X$ under J is non-empty and it is easy to see that J is homogeneous (i.e. $J(\alpha x) = \alpha J(x)$ for every $\alpha \in \mathbb{R}$), but we would like to impose further topological properties on X so as to make the duality mapping nicer. A Banach space for which $J(x)$ is always single-valued (and thus has *not* to be considered as a set-valued map) is called *smooth*. In a smooth Banach space, the duality map is unique and given by the Gâteaux (functional) derivative of the convex function $\phi(x) = \frac{1}{2}\|x\|_X^2$,

$$J(x) = \nabla \phi(x). \quad (6)$$

Interestingly, the smoothness of a space X implies that its dual X^* is *strictly convex* (meaning that the line segment between any two points $u \neq v$, $\|u\|_{X^*} = \|v\|_{X^*} = 1$, lies strictly within the unit ball). This is an expression of the general connection of smoothness properties of a Banach space to convexity features on the dual side.

The next important property for our space X is that of *reflexivity*, which means that the bidual $X^{**} = (X^*)^*$ is isometrically isomorphic to the space X itself and can be identified with it. Hence, if X is reflexive, then X can be recovered from X^* as $X = X^{**} = (X^*)^*$.

If X is reflexive, then X^* is also reflexive. We may therefore also set up a duality map J^* from X^* to $X = X^{**}$. For every $u \in X^*$, we can consider an $x \in J^*(u) = \{x \in X \mid u(x) = \|x\|_X^2 = \|u\|_{X^*}^2\}$ and with this recover $u \in J(x)$, so J is also surjective if X is reflexive. For a surjective function, we can consider its set-valued inverse and we thus get $J^{-1} = J^*$. Now note that, when X^* is also smooth (X strictly convex), we have J^{-1} single-valued as well, so J is indeed bijective [5, proposition 1.117]. We collect all these properties as the minimal assumptions for our Banach spaces X and X^* :

Assumption 1. The real Banach space X is reflexive, with X and X^* strictly convex. The corresponding duality map $J: X \rightarrow X^*$ is then single-valued with a well-defined inverse $J^{-1}: X^* \rightarrow X$.

The prototypical example for such Banach spaces, which is also relevant in the context of DFT, are the Lebesgue spaces L^p , $1 < p < \infty$. It is important to note that $p = 1$ and $p = \infty$ are excluded by assumption 1, since those spaces are neither reflexive nor strictly convex. At this point, it already becomes obvious that we need to adapt the typical choice of density space in DFT [3], $L^1(\mathbb{R}^3) \cap L^3(\mathbb{R}^3)$, since by including L^1 the space becomes non-reflexive and thus cannot fulfill assumption 1. It is precisely for this reason that, for example, Eschrig [14] in his discussion of thermal DFT switches to the reflexive space $L^3(\Omega)$ on a bounded three-dimensional periodic domain, which anyway entails $L^3(\Omega) \subset L^1(\Omega)$.

Assumption 1 has been our standard assumption in several previous publications on MY regularization in DFT [4, 15, 16] and is sufficient for many use cases. Yet, in some instances, additional regularity is required and we add the condition that X^* is also uniformly convex. Note that uniform convexity implies reflexivity. A space X^* is called uniformly convex if, for each $\varepsilon \in (0, 2]$, there is a $\delta(\varepsilon) > 0$ such that, for all $u, v \in X^*$ with $\|u\|_{X^*} = \|v\|_{X^*} = 1$ and $\|u - v\|_{X^*} = \varepsilon$, we have $1 - \frac{1}{2}\|u + v\|_{X^*} \geq \delta(\varepsilon)$. Adding the requirement of uniform convexity, we obtain:

Assumption 2. X and X^* fulfill all requirements of assumption 1 and additionally X^* is assumed uniformly convex. The duality map $J: X \rightarrow X^*$ is then additionally uniformly continuous.

Finally, we can tighten those assumptions even more and move to a Hilbert space X , equipped with an inner product denoted as $\langle \cdot, \cdot \rangle_X$. Hilbert spaces are automatically uniformly convex and thus also reflexive. Since then also X^* is a Hilbert space, we have all the properties in assumption 2 fulfilled. For

Hilbert spaces, the duality map $J: X \rightarrow X^*$ is exactly the isometric isomorphism from the Riesz representation theorem and linear as such. Nevertheless, we do not want to identify a Hilbert space with its dual in the given context. The original work introducing MY regularization to DFT of Kvaal *et al* [1] considered a Hilbert-space setting and it is also employed elsewhere [17–21]. We also highlight a comprehensive chapter by Kvaal [22] in a recent textbook on DFT.

Assumption 3. The Banach space X and consequently also X^* are Hilbert spaces. The duality map $J: X \rightarrow X^*$ is then linear and Lipschitz continuous with constant 1, with a well-defined inverse $J^{-1}: X^* \rightarrow X$ that has the same properties.

For each $u \in X^*$, the element $x = J^{-1}(u) \in X$ is then just the unique element given from the Riesz representation theorem and we have for every $y \in X$ that

$$u(y) = \langle u, y \rangle = \langle x, y \rangle_X. \quad (7)$$

Note carefully that the notation $\langle u, y \rangle$, which we will also employ in the Banach-space case from here on, describes the dual pairing of $u \in X^*$ with $y \in X$ (i.e. the application of the linear functional u to y), and should not be confused with $\langle x, y \rangle_X$, which is the inner product of elements in the Hilbert space X . This notation fits well with the previously mentioned case $X = L^p(\Omega)$ on some domain $\Omega \subseteq \mathbb{R}^n$, $1 < p < \infty$, where $X^* = L^{p^*}(\Omega)$ with $1/p + 1/p^* = 1$, and

$$u(x) = \langle u, x \rangle = \int_{\Omega} u(\mathbf{r}) x(\mathbf{r}) d\mathbf{r}, \quad (8)$$

which equals $\langle u, x \rangle_{L^2}$ for $p = 2$. So, in this case, we arrive back at the inner product on the Hilbert space $L^2(\Omega)$, which is called a *pivot space* between $L^p(\Omega)$ and $L^{p^*}(\Omega)$. In other words, the inner product of the pivot space serves as the duality pairing.

This construction can be generalized to arbitrary X and X^* within the formalism of a *rigged Hilbert space* (Gel'fand triple). Applied to the Hilbert space of wavefunctions, this approach provides a rigorous foundation for Dirac's bra–ket formalism and can be employed to describe continuous spectra and scattering states in quantum mechanics [23, 24]. Those are different objects than the densities and potentials considered here, but we find the similarities still striking.

Note that the three assumption sets are increasingly strong—that is, assumption 3 implies assumption 2, which in turn implies our minimal requirement assumption 1.

2.2. Convex analysis preliminaries

The second mathematical pillar of DFT is without any doubt convex analysis, since DFT admits a beautiful formulation in terms of convex and concave functionals on Banach spaces. On a Banach space X that fulfills assumption 1, we now consider convex functions $f: X \rightarrow \mathbb{R} \cup \{+\infty\}$ that are not $+\infty$ everywhere. We also usually assume lower semicontinuity, which means that, for each $x_0 \in X$, it holds that $\liminf_{x \rightarrow x_0} f(x) \geq f(x_0)$. Since the argument $x \in X$ of $f(x)$ is the element of a Banach space of functions that gets mapped to the scalar field, f is sometimes called a *functional* (especially in the context of DFT). The *Legendre–Fenchel transform* (convex conjugate) of any $f: X \rightarrow \mathbb{R} \cup \{-\infty, +\infty\}$ is a function $f^*: X^* \rightarrow \mathbb{R} \cup \{-\infty, +\infty\}$ given by

$$f^*(u) = \sup_{x \in X} \{ \langle u, x \rangle - f(x) \}. \quad (9)$$

The convex conjugate of a convex and lower semicontinuous $f: X \rightarrow \mathbb{R} \cup \{+\infty\}$ is again convex and lower semicontinuous. By applying this transformation twice and restricting the result again to $X \subseteq X^{**}$ (biconjugate),

$$f^{**}(x) = \sup_{u \in X^*} \{ \langle u, x \rangle - f^*(u) \}, \quad (10)$$

one arrives back at the original function, $f^{**} = f$, if and only if f is convex and lower semicontinuous.

The *subdifferential* of f at some $x \in X$ is the set-valued map

$$\partial f(x) = \{u \in X^* \mid \forall y \in X: f(y) \geq f(x) + \langle u, y - x \rangle\}. \quad (11)$$

On the real line, the set $\partial f(x)$ can be interpreted as the collection of all slopes for which the corresponding tangent of f at x never exceeds the function f itself (globally, since the function is assumed convex). It can be defined for concave functions as the *superdifferential* by just reverting the inequality, and we use the same notation in this case. Note that $\partial f(x)$ is always a convex set, and that it includes infinitely many elements where f has a kink and is thus non-differentiable. It can also be the empty set. To illustrate this, take $X = \mathbb{R}$ and $f(x) = -\sqrt{x}$ for $x \geq 0$ and $f(x) = +\infty$ for $x < 0$ (which actually connects to reduced-density-matrix functional theory [25]) and then note that $\partial f(0) = \emptyset$. Another one-dimensional example can be taken from a DFT-treatment of the quantum Rabi model [26], where $\partial f(\pm 1) = \emptyset$ for $f(x) = -\sqrt{1-x^2}$ for $x \in [-1, 1]$ and $f(x) = +\infty$ outside. If, on the other hand, f is differentiable at x , then the subdifferential $\partial f(x)$ is single-valued and consists only of the functional derivative $\nabla f(x)$. An graphical illustration with a function on $\mathbb{R}$ is provided in figure 1.

The subdifferential fulfills the usual role of a derivative, by qualifying the minimum of a convex (not necessarily lower semicontinuous) functional,

$$0 \in \partial f(x_*) \iff x_* \in \underset{x \in X}{\operatorname{argmin}} f(x). \quad (12)$$

Yet, unlike the derivative, the subdifferential does not require the function to be differentiable. For a function g on the dual space X^* , we define the subdifferential at $u \in X^*$ as a convex subset of the predual X rather than the bidual X^{**} ,

$$\partial g(u) = \{x \in X \mid \forall v \in X^*: g(v) \geq g(u) + \langle v - u, x \rangle\}. \quad (13)$$

With these definitions in place, and if f is lower semicontinuous, the subdifferentials of a pair of conjugate functions (f, f^*) are connected in the following way (inverse subdifferential relation),

$$(\partial f)^{-1} = \partial f^*, \quad (14)$$

where the inverse is meant as a set-valued mapping.

2.3. MY regularization

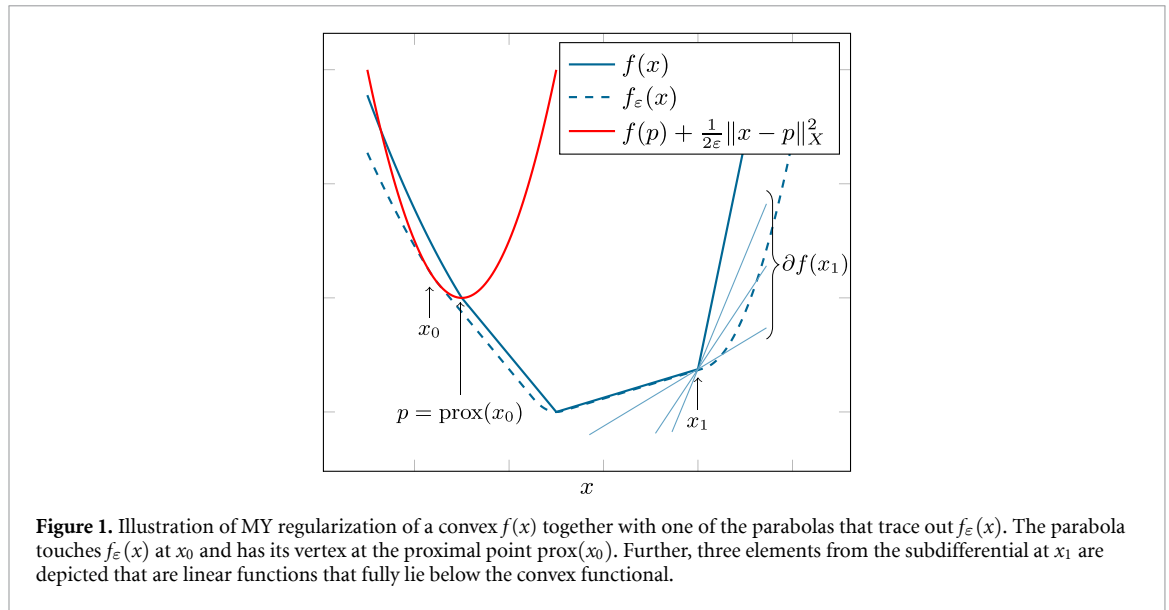
In order to deal with problems arising from a possible non-differentiability of $f: X \rightarrow \mathbb{R} \cup \{+\infty\}$, we introduce the MY regularization with smoothing parameter $\varepsilon > 0$ of a convex and lower semicontinuous functional f as

$$f_\varepsilon(x) = \inf_{y \in X} \left\{ f(y) + \frac{1}{2\varepsilon} \|x - y\|_X^2 \right\}. \quad (15)$$

MY regularization was originally introduced by Moreau [27] and can be linked to the previously discovered Yosida approximation of maximal monotone operators [28]. How such a MY regularized functional looks like is shown with one-dimensional examples in figures 1 and 2 in section 3.1. This construction has a long list of favorable properties that we reproduce here, while the proofs can be found in standard textbooks [5, 7]. It is important to remember that we always demand at least assumption 1, ensuring that X and X^* are both reflexive and strictly convex:

- (i) f_ε is convex, finite, and Gâteaux differentiable. If in addition X^* is uniformly convex (assumption 2), then f_ε is also Fréchet differentiable.
- (ii) For each $x \in X$ and all $\varepsilon, \varepsilon'$ such that $0 < \varepsilon' < \varepsilon$, we have $\inf f(X) \leq f_\varepsilon(x) \leq f_{\varepsilon'}(x) \leq f(x)$. In particular, $\inf f(X) = \inf f_\varepsilon(X) = \inf f_{\varepsilon'}(X)$. The minimal value and the minimizers, if they exist, are preserved by the regularization.
- (iii) $\lim_{\varepsilon \rightarrow 0} f_\varepsilon(x) = f(x)$ for every $x \in X$.
- (iv) The infimum in equation (15) is attained at a unique point (*proximal point*) that defines the *proximal map*

$$\begin{aligned} \operatorname{prox}_{\varepsilon f}: X &\rightarrow X, \\ \operatorname{prox}_{\varepsilon f}(x) &= \underset{y \in X}{\operatorname{argmin}} \left\{ f(y) + \frac{1}{2\varepsilon} \|x - y\|_X^2 \right\}, \end{aligned} \quad (16)$$



and $\lim_{\varepsilon \rightarrow 0} \text{prox}_{\varepsilon f}(x) = x$ if $f(x) < \infty$.

- (v) $\nabla f_\varepsilon(x) = \varepsilon^{-1} J(x - \text{prox}_{\varepsilon f}(x)) \in \partial f(\text{prox}_{\varepsilon f}(x))$.
- (vi) $f_\varepsilon(x) = f(\text{prox}_{\varepsilon f}(x)) + \frac{\varepsilon}{2} \|\nabla f_\varepsilon(x)\|_{X^*}^2$.
- (vii) For each $x \in X$ with $\partial f(x)$ non-empty, it holds that $\lim_{\varepsilon \rightarrow 0} \nabla f_\varepsilon(x) = w$ (weakly) where $w \in X^*$ is the (unique) element in $\partial f(x)$ with minimal norm. If X^* is uniformly convex (assumption 2), then the convergence is strong.
- (viii) For the Legendre–Fenchel transform of f_ε , it holds that

$$(f_\varepsilon)^*(u) = f^*(u) + \frac{\varepsilon}{2} \|u\|_{X^*}^2 \quad \text{and} \quad (17)$$

$$\partial (f_\varepsilon)^*(u) = \partial f^*(u) + \varepsilon J^{-1}(u) \quad (18)$$

for each $u \in X^*$. By adding $\frac{\varepsilon}{2} \|u\|_{X^*}^2$ to $f^*(u)$, the $(f_\varepsilon)^*$ automatically becomes a strictly convex function.

Note that X strictly convex was already needed for the definition of the proximal map since it means that $x \mapsto \|x\|_X^2$ is strictly convex and thus the minimum in equation (16) is reached at a unique point. The reason that we can write $\text{prox}_{\varepsilon f}$ instead of $\text{prox}_f^\varepsilon$ (or similar) in order to include the regularization parameter ε in the notation, is that the definition of the proximal map could have also been written as

$$\text{prox}_{\varepsilon f}(x) = \arg \min_{y \in X} \left\{ \varepsilon f(y) + \frac{1}{2} \|x - y\|_X^2 \right\}. \quad (19)$$

In a Hilbert-space setting (assumption 3), a few additional useful properties of the MY regularization and the proximal map can be added [6]:

- (ix) f_ε is Fréchet differentiable and its functional derivative ∇f_ε is Lipschitz continuous with constant ε^{-1} .
- (x) The proximal map is given by the resolvent of the subdifferential, $\text{prox}_{\varepsilon f}(x) = (\text{id} + \varepsilon \partial f)^{-1}(x)$.
- (xi) The proximal map is firmly nonexpansive, meaning that, for all $x, y \in X$, it holds

$$\begin{aligned} \langle \text{prox}_{\varepsilon f}(x) - \text{prox}_{\varepsilon f}(y), x - y \rangle_X \\ \geq \|\text{prox}_{\varepsilon f}(x) - \text{prox}_{\varepsilon f}(y)\|_X^2. \end{aligned} \quad (20)$$

The same property holds for $\text{id} - \text{prox}_{\varepsilon f}$.

(xii) $\partial(f_\varepsilon)^*$ is strongly monotonous, meaning that, for all $u, v \in X^*$ and $x \in \partial(f_\varepsilon)^*(u)$, $y \in \partial(f_\varepsilon)^*(v)$, it holds

$$\langle u - v, x - y \rangle \geq \varepsilon \|u - v\|_{X^*}^2. \quad (21)$$

While the notion of firm nonexpansiveness itself can be generalized to Banach spaces [29], it does not necessarily hold for the proximal map anymore. Nevertheless, a result of uniform continuity can be achieved for the proximal map also in a certain Banach-space setting [30]. Combining (v) from the previous properties and the resolvent form for the proximal map, we obtain

$$\nabla f_\varepsilon(x) = \varepsilon^{-1} \left(x - (\text{id} + \varepsilon \partial f)^{-1}(x) \right), \quad (22)$$

which exactly corresponds to the previously mentioned Yosida approximation [6, definition 23.1] of the multivalued mapping ∂f . Together with Moreau's contributions to convex analysis this justifies the naming of the regularization method.

2.4. Formal DFT

The standard setting of DFT is the N -particle Hamiltonian with an external, scalar potential v on $\mathbb{R}^3$ (atomic units):

$$\begin{aligned} \hat{H}_v^\lambda &= \hat{T} + \lambda \hat{W} + \sum_{j=1}^N v(\mathbf{r}_j), \\ \hat{T} &= -\frac{1}{2} \sum_{j=1}^N \nabla_j^2, \quad \hat{W} = \sum_{j < k} |\mathbf{r}_j - \mathbf{r}_k|^{-1}. \end{aligned} \quad (23)$$

Here, we allow different choices of the interaction strength λ , which can range from the non-interacting case ($\lambda = 0$) to the fully interacting case with a repulsive force ($\lambda = 1$). For all choices of $\lambda \in \mathbb{R}$ and $v \in L^{3/2}(\mathbb{R}^3) + L^\infty(\mathbb{R}^3)$, this Hamiltonian is self-adjoint and bounded below on its usual form domain by the Kato–Lions–Lax–Milgram–Nelson theorem [31, theorems X.17 and X.19]. It was already mentioned that the methods and results presented here readily apply to various settings outside of standard DFT, although subtleties around the involved spaces and definitions of the basic functionals have to be taken into account. A setting that also involves vector potentials is briefly discussed in section 6.2.

The ground-state problem of quantum mechanics is given by the Rayleigh–Ritz variational principle,

$$E^\lambda(v) = \inf_{\Psi} \langle \Psi, \hat{H}_v^\lambda \Psi \rangle = \inf_{\hat{\Gamma}} \text{Tr} \left(\hat{H}_v^\lambda \hat{\Gamma} \right), \quad (24)$$

where the variation is over all normalized wave functions or density matrices with finite kinetic energy and the required antisymmetry. The result, $E^\lambda(v)$, is then the ground-state energy of the system for a given external potential v and chosen interaction strength λ . The principal idea of DFT is now to separate off the only density-dependent part inside this variation, while the remaining part is *universal* in the sense that it does not depend on v [3, 32–34]. An elegant way to achieve this is by Levy's constrained search—that is, one varies over all wave functions Ψ or density matrices that yield a given one-particle density that we define for wave functions and density matrices $\hat{\Gamma} = \sum_j p_j |\Psi_j\rangle \langle \Psi_j|$ as

$$\rho_\Psi(\mathbf{r}) = N \int |\Psi(\mathbf{r}, \mathbf{r}_2, \dots, \mathbf{r}_N)|^2 d\mathbf{r}_2 \dots d\mathbf{r}_N, \quad (25)$$

$$\rho_{\hat{\Gamma}}(\mathbf{r}) = \sum_j p_j \rho_{\Psi_j}(\mathbf{r}), \quad p_j \in [0, 1], \quad \sum_j p_j = 1. \quad (26)$$

We define the universal density functional by a constrained search over pure states and ensembles, respectively,

$$\tilde{F}^\lambda(\rho) = \inf_{\Psi \mapsto \rho} \langle \Psi, \hat{H}_0^\lambda \Psi \rangle, \quad (27)$$

$$F_{\text{DM}}^\lambda(\rho) = \inf_{\hat{\Gamma} \mapsto \rho} \text{Tr} \left(\hat{H}_0^\lambda \hat{\Gamma} \right), \quad (28)$$

where $\hat{H}_0^\lambda = \hat{T} + \lambda \hat{W}$ and where we write $\Psi \mapsto \rho$ for an N -particle wave function with $\rho_\Psi = \rho$ and analogously for density matrices. Since the constrained search over ensembles includes the search over pure

states, we have $F_{\text{DM}}^\lambda(\rho) \leq \tilde{F}^\lambda(\rho)$. The functional $F_{\text{DM}}^\lambda(\rho)$ is convex since it is given as the infimum of a linear functional over a convex set, unlike $\tilde{F}^\lambda(\rho)$, which is *not* necessarily convex. At non-pure-state ν -representable ρ it holds $\tilde{F}^\lambda(\rho) \neq F_{\text{DM}}^\lambda(\rho)$ [3, 35], and explicit examples are known [36, section VI.E].

The effective domain of a convex functional is the set of points where $f(x) < +\infty$. For the constrained-search functionals $\tilde{F}^\lambda(x)$ and $F_{\text{DM}}^\lambda(x)$, the effective domains agree and are also independent of λ and we set

$$\begin{aligned} D &= \{x \in X \mid \tilde{F}^\lambda(x) < \infty\} \\ &= \{x \in X \mid F_{\text{DM}}^\lambda(x) < \infty\}. \end{aligned} \quad (29)$$

From the definition of the density ρ_Ψ and the definition of the functionals, it follows that any $\rho \in D$ must be non-negative and normalized to the number of particles. It is a fact that D corresponds to the N -representable densities [3, theorem 1.2]. Note that we generally use ρ for elements in D that are real densities, while we use x for more general elements in X that will later rise to prominence as ‘mixed densities’. By definition, we have $\tilde{F}^\lambda(x) = F_{\text{DM}}^\lambda(x) = +\infty$ for $x \in X \setminus D$. For this reason, the usual convention will be to use the symbol ρ for elements of D that are ‘real’ densities, and to use $x, y \in X$ otherwise.

The pure- and ensemble-state density functionals given in equations (27) and (28), respectively, yield the same ground-state energy in the Hohenberg–Kohn variational principle,

$$\begin{aligned} E^\lambda(\nu) &= \inf_{\rho \in X} \{\tilde{F}^\lambda(\rho) + \langle \nu, \rho \rangle\} \\ &= \inf_{\rho \in X} \{F_{\text{DM}}^\lambda(\rho) + \langle \nu, \rho \rangle\}, \end{aligned} \quad (30)$$

where the dual pairing of potential and density yields the potential energy, $\langle \nu, \rho \rangle = \int \nu(\mathbf{r})\rho(\mathbf{r})\mathbf{d}\mathbf{r}$. Recognizing that equation (30) has the structure of the Legendre–Fenchel transform introduced in section 2.2, we define the back transform

$$F^\lambda(\rho) = (-E^\lambda)^*(-\rho) \quad (31)$$

as a new functional that is convex and lower semicontinuous by construction. It needs to be stressed here that although $F^\lambda = (F_{\text{DM}}^\lambda)^{**}$ by the two previous equations, this does *not* necessarily mean that the two functionals agree, since $f^{**} = f$ only holds if the functional is convex *and* lower semicontinuous. It nevertheless generally holds that $F^\lambda(\rho) \leq F_{\text{DM}}^\lambda(\rho)$. But in any case we can get $E^\lambda(\nu)$ back from $F^\lambda(\rho)$ by

$$E^\lambda(\nu) = -(F^\lambda)^*(-\nu). \quad (32)$$

Now, if the convex $F_{\text{DM}}^\lambda(\rho)$ is also lower semicontinuous (assumption 4 below), this implies

$$F^\lambda(\rho) = F_{\text{DM}}^\lambda(\rho). \quad (33)$$

Unlike convexity, lower semicontinuity of $F_{\text{DM}}^\lambda(\rho)$ must be checked for every choice of topology.

As said before, in settings where $F_{\text{DM}}^\lambda(\rho)$ is not lower-semicontinuous, the convex conjugate of $E^\lambda(\nu)$ does notably *not* retrieve the original constrained-search functional $F_{\text{DM}}^\lambda(\rho)$ in equation (31). Thus it loses connection to the variational principle in equation (24) and consequently also to the equivalent ground-state problem involving the Schrödinger equation. This means that minimizers $\rho \in X$ in equation (30) can occur where the convex conjugate $F^\lambda(\rho)$ of $E^\lambda(\nu)$ is finite, yet the constrained-search functional has $F_{\text{DM}}^\lambda(\rho) = +\infty$. If, on the other hand, ν in $E^\lambda(\nu)$ supports a ground state Ψ , then $F_{\text{DM}}^\lambda(\rho_\Psi)$ is also finite. For this reason, it needs to be stressed that, albeit the requirement of lower semicontinuity for the ensemble constrained-search functional seems like a mathematical sophistry, it is a critical feature of the theory, and we will elaborate further on this on page 13. The convex $F^\lambda(\rho) = F_{\text{DM}}^\lambda(\rho)$ on X and the concave $E^\lambda(\nu)$ on X^* , as a conjugate pair, then serve as the backbone of a convex-analysis formulation of DFT.

Based on the above discussion, we formulate a separate assumption set for a consistent DFT framework.

Assumption 4. X and X^* are such that the Hamiltonian $\hat{H}_\nu^\lambda$ is self-adjoint with a ν -independent domain and bounded below, and the ensemble constrained-search functional $F_{\text{DM}}^\lambda(\rho)$ is lower semicontinuous. This implies $F_{\text{DM}}^\lambda(\rho) = F^\lambda(\rho)$.

To fulfill this assumption, we can think about Lieb's choice $L^1(\mathbb{R}^3) \cap L^3(\mathbb{R}^3)$ for the space of densities X , which indeed makes $F_{\text{DM}}^\lambda(\rho)$ lower semicontinuous [3]. Yet, this space does not fulfill assumption 1, which is not necessary for a consistent formulation of DFT but needed for our later constructions involving MY regularization. Instead, we suggest to switch to a more general yet reflexive space $X \supsetneq L^1(\mathbb{R}^3) \cap L^3(\mathbb{R}^3)$. This, on the other hand, can break lower semicontinuity of the ensemble constrained-search functional if we switch to a coarser topology.

The dual pairing between potentials and densities in X suggests a matching space for potentials that consists of the topological dual X^* . For $L^1(\mathbb{R}^3) \cap L^3(\mathbb{R}^3)$, the dual space is $L^{3/2}(\mathbb{R}^3) + L^\infty(\mathbb{R}^3)$. If instead $X \supsetneq L^1(\mathbb{R}^3) \cap L^3(\mathbb{R}^3)$, then $X^* \subsetneq L^{3/2}(\mathbb{R}^3) + L^\infty(\mathbb{R}^3)$. Interestingly, $L^{3/2}(\mathbb{R}^3)$ is exactly the boundary case in terms of L^p spaces and the Hamiltonian fails to be self-adjoint on its usual domain for $p < 3/2$ when the potentials can become too singular. This is the case for a potential $r^{-\alpha}$, $\alpha \geq 2$, where *different* self-adjoint realizations on the punctured domain $\mathbb{R}^3 \setminus \{0\}$ become possible. To ensure a self-adjoint and bounded-below Hamiltonian on the usual domain, we thus need to limit the potential space to $X^* \subseteq L^{3/2}(\mathbb{R}^3) + L^\infty(\mathbb{R}^3)$.

At this point, another problem appears due to this limitation. In certain cases, this space will not accommodate a given physical setting—for example, it might exclude singular Coulomb potentials and thus molecular systems in a clamped-nuclei or Born–Oppenheimer approximation. We can solve this problem by just including a fixed *offset potential* from our previous class

$$v_{\text{off}} \in L^{3/2}(\mathbb{R}^3) + L^\infty(\mathbb{R}^3). \quad (34)$$

It can also be taken from any other potential class that still admits a self-adjoint and bounded-below Hamiltonian according to assumption 4, including non-local potentials. In particular, v_{off} can fulfill the task of a binding potential for N electrons. But one has to take careful note that such a potential can change the domain of the self-adjoint Hamiltonian if it is not from the class of equation (34) [31, section X.2]. Instead of writing $E^\lambda(v_{\text{off}} + v)$, we redefine the energy functional as

$$E^\lambda(v) = \inf_{\Psi} \langle \Psi, \hat{H}_{v_{\text{off}}+v}^\lambda \Psi \rangle, \quad (35)$$

including the fixed offset automatically. Note that the offset potential then must also enter the definition of the universal density functional so that it remains compatible to $E^\lambda(v)$,

$$F_{\text{DM}}^\lambda(\rho) = \inf_{\hat{\Gamma} \mapsto \rho} \text{Tr} \left(\hat{H}_{v_{\text{off}}}^\lambda \hat{\Gamma} \right). \quad (36)$$

While the inclusion of this offset potential does not change anything within the basic structure of DFT, which is the reason for not including it explicitly into the notation of the functionals, it cannot be ignored when solving the underlying equations, and we will highlight its occurrence at several places in the following sections.

Whenever a minimizer can be found in the Hohenberg–Kohn variation principle of equation (30) with the functional $F_{\text{DM}}^\lambda(\rho)$, we call this $\rho \in D$ a *v -representable* density. This means, in other words, that there is a $v \in X^*$ such that the (ensemble) ground-state solution $\hat{\Gamma}$ to the Schrödinger equation with Hamiltonian $\hat{H}_{v_{\text{off}}+v}^\lambda$ yields $\hat{\Gamma} \mapsto \rho$ and ground-state energy $E^\lambda(v) = F_{\text{DM}}^\lambda(\rho) + \langle v, \rho \rangle$. We next write out the relation in equation (31) as the Lieb variational principle,

$$F^\lambda(\rho) = \sup_{v \in X^*} \{ E^\lambda(v) - \langle v, \rho \rangle \}. \quad (37)$$

Since in general $F^\lambda(\rho) \leq F_{\text{DM}}^\lambda(\rho)$ and since a v -representable ρ is a minimizer in equation (30) and must fulfill $F_{\text{DM}}^\lambda(\rho) = E^\lambda(v) - \langle v, \rho \rangle$, we see that $F^\lambda(\rho) = F_{\text{DM}}^\lambda(\rho)$ always holds for v -representable densities, even when $F_{\text{DM}}^\lambda(\rho)$ is not lower semicontinuous. In what follows, we want to relate this property to the functional $F^\lambda(\rho)$ and its subdifferential.

The notions of sub- and superdifferentials serve an important purpose in the formulation of DFT. If there exists an optimizer $v \in X^*$ for a given $\rho \in X$ in equation (37), then the following statements are all equivalent for the lower-semicontinuous $F^\lambda(\rho)$,

$$\begin{aligned} v \in -\partial F^\lambda(\rho) &\iff \rho \in \partial E^\lambda(v) \\ &\iff E^\lambda(v) = F^\lambda(\rho) + \langle v, \rho \rangle. \end{aligned} \quad (38)$$

We can combine this into the inverse subdifferential relation,

$$(-\partial F^\lambda)^{-1}(v) = \partial E^\lambda(v). \quad (39)$$

Yet, it is by no means guaranteed that the set $-\partial F^\lambda(\rho)$ is non-empty for all $\rho \in D$. But since $F^\lambda(\rho)$ is convex and lower semicontinuous by definition, it is known that its subdifferential is non-empty on a dense subset of $D \subset X$, the effective domain of the functional [5, corollary 2.44].

If $F_{\text{DM}}^\lambda(\rho)$ is lower semicontinuous, then $F_{\text{DM}}^\lambda(\rho) = F^\lambda(\rho)$, and what was said above also applies to $F_{\text{DM}}^\lambda(\rho)$. In addition, in this case, each $\rho \in \partial E^\lambda(v)$ is an optimizer in the second line of equation (30) and therefore $\rho \in D$. The convex set $\partial E^\lambda(v) \subseteq D$ then yields all possible ground-state densities for the potential $v \in X^*$. This $\partial E^\lambda(v)$ will be empty if v does not support a ground state for interaction strength λ , includes a single element if v supports a nondegenerate ground state, and contains multiple elements in case of degeneracy. We may now identify the points at which $F^\lambda(\rho)$ has a non-empty subdifferential with the (v -representable) ground-state densities of a Schrödinger equation. Hence, assuming a lower semicontinuous $F_{\text{DM}}^\lambda(\rho)$, it also holds that the set of v -representable densities is dense in the set of N -representable densities. In this way, it can be highlighted in a different manner why the lower semicontinuity of the ensemble constrained-search functional is important. Else, with a $F_{\text{DM}}^\lambda(\rho)$ that is not lower semicontinuous, it is possible to find $v \in -\partial F^\lambda(\rho)$ when at the same time $F_{\text{DM}}^\lambda(\rho) = +\infty$, so this density has $\rho \in X \setminus D$ and would clearly be not v -representable. This is why assumption 4 is important when dealing with the functionals $F_{\text{DM}}^\lambda(\rho)$ and $F^\lambda(\rho)$. See also the example given after the proposition below.

But it is also known since a long time that v -representability, even if it is secured for a dense set, is a problematic notion since most densities $\rho \in D$ will not have this property [37, 38]. In many cases (including standard DFT), the set of v -representable densities is known only implicitly (exceptions are lattice systems [36, 39, 40] and particles on a one-dimensional torus [18, 19]). Even worse, in the standard setting of DFT with density space $L^1 \cap L^3$, arbitrarily close to each v -representable density there lies a non- v -representable one, which makes $F^\lambda(\rho)$ discontinuous and non-differentiable [41]. Such an irregular behavior of the universal functional $F^\lambda(\rho)$ will create complications in the optimization procedures that DFT relies on. Further, full (interacting and non-interacting) v -representability is imperative for the construction of the KS iteration scheme (see section 2.5). If, on the other hand, $F^\lambda(\rho)$ would be differentiable, the derivative $v = -\nabla F^\lambda(\rho)$ would yield an immediate solution to the v -representability question, because this potential v then exactly gives ρ as a ground-state density by equation (38). We will see in section 3 how MY regularization offers a way out of this dilemma.

The Hohenberg–Kohn theorem [32] states that, for a v -representable density $\rho \in D$, the corresponding potential is unique up to the addition of a constant. In standard DFT, a rigorous proof of the Hohenberg–Kohn theorem [42] has been achieved for $v \in L^p(\mathbb{R}^3) + L^\infty(\mathbb{R}^3)$, with $p > 2$, and it extends to closely related settings [43]. This leaves $3/2 \leq p \leq 2$ as undecided (the lower bound is given by the argument for self-adjoint Hamiltonians from before), although results on unique continuation [44, 45] suggest that the Hohenberg–Kohn theorem holds also in this range. In non-standard DFT settings, like finite lattices or when currents are included, counterexamples are known [36, 46].

Combining the assumptions given in this section and in section 2.1 (which will later serve as the basis for MY regularized DFT), we arrive at the following important result:

Proposition. *Assumption 1 together with assumption 4 with a continuously embedded potential space $X^* \subset L^{3/2}(\mathbb{R}^3) + L^\infty(\mathbb{R}^3)$ imply that every $v \in X^*$ supports a ground state—that is, for each $v \in X^*$, the superdifferential $\partial E^\lambda(v)$ is non-empty.*

This claim can be seen as follows. The choice $L^{3/2}(\mathbb{R}^3) + L^\infty(\mathbb{R}^3)$ as a potential space makes $E^\lambda(v)$ continuous [3, theorem 3.1(iii)], implying that each continuously embedded $X^* \subset L^{3/2}(\mathbb{R}^3) + L^\infty(\mathbb{R}^3)$ gives a continuous energy functional $E^\lambda(v)$ as well. But a continuous and concave functional on X^* admits a non-empty subdifferential belonging to X^{**} everywhere in the interior of its domain (here all X^*) [5, proposition 2.36]. In general, $X^{**} \supseteq X$, but with assumption 1 we have reflexivity and thus $X = X^{**}$. By the additional requirement that $F_{\text{DM}}^\lambda(\rho)$ is lower semicontinuous, we have $F^\lambda(\rho) = F_{\text{DM}}^\lambda(\rho)$. Each $\rho \in \partial E^\lambda(v)$ is now in $D \subset X$ and through $v \in -\partial F^\lambda(\rho) = -\partial F_{\text{DM}}^\lambda(\rho)$ it is connected to an actual ground-state solution of the Schrödinger equation as argued above. This demonstrates the claim of the proposition.

We illustrate this subtle and important point with an example. Take the Hamiltonian from equation (23) with just a single particle and set $v_{\text{off}} = v = 0$. Then there is no bound state and thus $\partial E^\lambda(v)$ should be empty (note that the λ parameter would not be needed for single particles since there are no interactions). Indeed, this system is incompatible with reflexivity and lower semicontinuity assumed together. If we give up on reflexivity and in contrast to equation (13) allow a superdifferential

in X^{**} , then we can find an $x \in \partial E^\lambda(0)$ where $x \in X^{**} \setminus X$, which does not represent a ground state but instead corresponds to a scattering state. Next, we keep reflexivity but give up on lower semicontinuity of $F_{\text{DM}}(\rho)$. We pick a minimizing sequence $\rho_i \rightarrow x$ in equation (30) consisting, for example, of normalized Gaussian-shaped densities $\rho_i \in D$ that get broader with increasing i , while $x \notin D$. Note that this means $x = 0$, which cannot hold if X includes the L^1 norm (and which would make it non-reflexive), since all ρ_i were assumed normalized to the same number and clearly $N = \|\rho_i\|_{L^1} \rightarrow 0$ is a contradiction. But it is not necessarily a contradiction in a reflexive X without the L^1 norm. Now the chosen sequence fulfills $\lim F_{\text{DM}}(\rho_i) = 0$ but at the same time $F_{\text{DM}}(x) = +\infty$, since $x \notin D$, so lower semicontinuity indeed cannot hold.

If, on the other hand, we add a fixed offset potential to the Hamiltonian that always guarantees existence of a ground state, for example, a harmonic oscillator potential $v_{\text{off}}(\mathbf{r}) = |\mathbf{r}|^2$, then the sequence ρ_i of Gaussian-shaped densities from before cannot minimize $F_{\text{DM}}(\rho)$ any more. This is because the definition of $F_{\text{DM}}(\rho)$ includes the offset potential by equation (36) and, as the densities get broader, the potential energy $\langle v_{\text{off}}, \rho_i \rangle$ diverges. Every potential that has $\lim_{|\mathbf{r}| \rightarrow \infty} v_{\text{off}}(\mathbf{r}) = \infty$ guarantees existence of a ground state [47, XIII.47] and we call it a *confining potential*.

This section summarized the most important features of DFT in the language of functional and convex analysis. The reader may wonder if assumption 1 (or one of the two stronger ones) and assumption 4 can be satisfied simultaneously. After all, the classical setting of DFT, with $X = L^1(\mathbb{R}^3) \cap L^3(\mathbb{R}^3)$ for densities and its dual $X^* = L^{3/2}(\mathbb{R}^3) + L^\infty(\mathbb{R}^3)$ for potentials, allows for assumption 4 but breaks reflexivity, a property not needed for a consistent formulation of DFT but necessary for our later results related to MY regularization. One solution is to choose a bounded domain Ω , which implies a continuous embedding $L^p(\Omega) \subseteq L^q(\Omega)$, $q < p$. Given that the important result of lower semicontinuity due to Lieb [3, theorem 4.4 and corollary 4.5] holds with respect to the L^1 topology, it extends to each $L^p(\Omega)$, $p > 1$. We can then choose any $X = L^p(\Omega)$ with $1 < p < \infty$ with dual $X^* = L^{p^*}(\Omega)$, $1/p + 1/p^* = 1$, and thus have assumptions 2 and 4 simultaneously satisfied. By selecting $p = 2$, we satisfy both assumptions 3 and 4. The previous works of Eschrig [14] and Kvaal et al [1] used this route, while many other works on the topic had to give up on at least one of the assumptions. Rather than switching to a bounded domain, we can use the offset potential v_{off} to introduce a confining potential for the whole system and thus ensure that every $v \in X^*$ yields a ground-state density. This, at the same time, serves to break the full translational and rotational symmetry of the setting, which is necessary to allow for any localized ground state in the first place [48].

Summary. To allow for any of assumptions 1–3 together with 4 simultaneously, possible strategies are to switch to a bounded domain $\Omega \subset \mathbb{R}^3$ or to add a confining offset potential v_{off} .

2.5. KS method

DFT reformulates the ground-state problem of quantum mechanics by replacing the many-particle wave function or density matrix as the central variable with the one-particle density, thereby significantly reducing the complexity of the description. Yet, the usefulness of such a reduced description entirely depends on whether $F^\lambda(\rho)$ can be evaluated accurately and efficiently. The $F^0(\rho)$ of a non-interacting system amounts to just the kinetic energy and its conjugate functional $E^0(v)$ can be computed effectively by means of orbitals. For the interacting case, however, $F^1(\rho)$ is in principle unavailable since its evaluation corresponds to a quantum version of an NP-hard problem [49]. The difference $F^1(\rho) - F^0(\rho)$, which relates to the exchange–correlation functional, has moreover turned out to be difficult to model in terms of ρ [50]. Thus, although the exact DFT formulation is very appealing by avoiding the full description by a wave function, further steps are needed to attain a useful reformulation. In the KS method [51], an auxiliary system of non-interacting electrons is introduced and its external potential v_s is chosen in such a way that it reproduces the same density as the interacting system. Yet, we will argue below that it is by no means clear that this approach is always possible.

The *exchange–correlation potential* is defined as the potential that must be added to the external potential and the Hartree potential of the auxiliary, non-interacting system in order to get the same ground-state density as the interacting system. This means that the method requires $\rho \in D$ to be simultaneously interacting and non-interacting v -representable, which is equivalent to $\partial F^1(\rho)$ and $\partial F^0(\rho)$ both non-empty. We already noted on the general issues regarding v -representability in section 2.4, and this additional requirement of simultaneous v -representability comes as a further restriction [52].

Assuming for now that this is ensured, we can take $v \in -\partial F^1(\rho)$ and $v_s \in -\partial F^0(\rho)$ and define the exchange–correlation potential

$$v_{xc}(\rho) = v_s - v - v_H(\rho). \quad (40)$$

Here we already subtracted the Hartree potential

$$v_H(\rho) = \nabla E_H(\rho) = \int \frac{\rho(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|} d\mathbf{r}', \quad (41)$$

which itself is the functional derivative of the Hartree mean-field energy

$$E_H(\rho) = \frac{1}{2} \iint \frac{\rho(\mathbf{r})\rho(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|} d\mathbf{r}d\mathbf{r}'. \quad (42)$$

The total external potential that needs to be applied to the non-interacting system in order to get ρ is then, not forgetting the previously introduced offset potential,

$$v_{\text{off}} + v_s = v_{\text{off}} + v + v_H(\rho) + v_{xc}(\rho). \quad (43)$$

As was already noted, the assumption of (simultaneous interacting and non-interacting) v -representability comes as a severe limitation. But putting this issue aside for the moment and assuming the existence of $v_{xc}(\rho)$ for all ρ in question, equation (43) can be translated into the KS self-consistent iteration scheme,

$$\begin{array}{ccc} v_s = v + v_H(\rho) + v_{xc}(\rho) \\ \updownarrow & \updownarrow & \updownarrow \\ v_{i+1} = v + v_H(\rho_i) + v_{xc}(\rho_i) \end{array} \quad (44)$$

and

$$\rho_{i+1} \in \partial E^0(v_{i+1}). \quad (45)$$

The second step corresponds to solving the non-interacting Schrödinger equation with an external potential v_{i+1} and determining a ground-state density ρ_{i+1} . If the Hamiltonian is defined on a bounded domain $\Omega \subset \mathbb{R}^3$ or contains a confining offset potential v_{off} , this approach is always possible. A different convex-analytical argument for the non-emptiness of the superdifferential is given by our proposition building on assumption 1 together with assumption 4 and this was similarly used before [15, proof of theorem 12]. The two steps of equations (44) and (45) can be combined into a single step, either for the potentials or, as we will do here, for the densities,

$$\rho_{i+1} \in \partial E^0(v + v_H(\rho_i) + v_{xc}(\rho_i)). \quad (46)$$

For initialization, we choose an initial density ρ_1 or, equivalently, an initial potential v_1 for the auxiliary system that roughly leads to the desired ground-state density ρ . Since the auxiliary system does not include any interactions between the particles, we choose (in the simplest case) its ground state as a Slater determinant and solve equation (45) by filling up the orbitals of a one-particle Schrödinger equation with external potential $v_{\text{off}} + v_{i+1}$. If the procedure converges, we get

$$\rho_i \rightarrow \rho, \quad v_i \rightarrow v_s, \quad (47)$$

so indeed both systems share the same ground-state density if, in the auxiliary system, an external potential $v_{\text{off}} + v_s$ is applied. Finally, in order to conveniently retrieve the interacting energy $E^1(v)$ from a converged density and the expressions defined above, we can just use $E^0(v_s) = F^0(\rho) + \langle v_s, \rho \rangle$ and calculate

$$\begin{aligned} E^1(v) &= F^1(\rho) + \langle v, \rho \rangle \\ &= F^0(\rho) + F^1(\rho) - F^0(\rho) + \langle v - v_s + v_s, \rho \rangle \\ &= E^0(v_s) + E_{xc}(\rho) + E_H(\rho) - \langle v_{xc}(\rho) + v_H(\rho), \rho \rangle \\ &= E^0(v_s) + E_{xc}(\rho) - E_H(\rho) - \langle v_{xc}(\rho), \rho \rangle, \end{aligned} \quad (48)$$

where we have introduced the *exchange–correlation functional* $E_{\text{xc}}(\rho) = F^1(\rho) - F^0(\rho) - E_{\text{H}}(\rho)$. We remark that practical KS-DFT methods are usually based on approximating $E_{\text{xc}}(\rho)$ and assuming differentiability, we then have $v_{\text{xc}} = \nabla E_{\text{xc}}(\rho)$.

Note that this formulation of the KS method does not come with any guarantee of convergence. Indeed, in practice, an additional damping step [53–55], like taking $\rho_i + \alpha(\rho_{i+1} - \rho_i)$ with $\alpha < 1$ as the next density, is introduced to facilitate convergence. A mathematical proof of convergence has only been achieved in the MY-regularized setting, which is in any case needed to avoid the v -representability problem, as will be discussed in section 5.3.

3. MY regularization in DFT

Having the stage set, we now present three different ways to introduce MY regularization into DFT, which are equivalent and thus have exactly the same outcome—namely, replacing the universal density functional $F^\lambda(\rho)$ with a differentiable $F_\varepsilon^\lambda(x)$ for all $x \in X$. Which way is to be preferred, is up to the reader, but we think it is illustrative to put them all side to side. Some readers might find one or the other easier to understand or more or less conceptually appealing. In section 3.2, we show that $F_\varepsilon^\lambda(x)$ is the convex conjugate of an energy functional $E_\varepsilon^\lambda(v)$ that includes the field energy of v (if appropriately defined), while section 3.3 identifies every $x \in X$ as a mix of internal and external densities. In this section, we only require assumption 1, but it is useful to also have assumption 4 with $F^\lambda = F_{\text{DM}}^\lambda$ fulfilled, so that no ambiguity about the respectively used universal functional arises.

3.1. MY based on the universal density functional

The most straightforward way to introduce MY regularization into DFT follows from its definition in equation (15): By directly applying the method to the non-differentiable (but convex and lower semicontinuous) $F^\lambda(\rho)$, yielding a regularized universal density functional $F_\varepsilon^\lambda(x)$ that is well-defined and finite for all $x \in X$,

$$F_\varepsilon^\lambda(x) = \inf_{\rho \in X} \left\{ F^\lambda(\rho) + \frac{1}{2\varepsilon} \|x - \rho\|_X^2 \right\}, \quad (49)$$

where $\varepsilon > 0$ is the smoothing parameter. An illustration of a regularized functional in one dimension together with the potentially non-convex $\tilde{F}^\lambda(x)$ and convex $F^\lambda(x)$ is shown in figure 2. The regularized functional has a well-defined functional derivative (see (v) from the properties of MY regularization in section 2.3) and the regularization is lossless in the sense that it conserves the overall infimum of the functional (see (ii)).

We next introduce the corresponding energy functional $E_\varepsilon^\lambda(v)$ as the convex conjugate of $F_\varepsilon^\lambda(x)$,

$$E_\varepsilon^\lambda(v) = - (F_\varepsilon^\lambda)^* (-v) = \inf_{x \in X} \{ F_\varepsilon^\lambda(x) + \langle v, x \rangle \}. \quad (50)$$

Note that the resulting $E_\varepsilon^\lambda(v)$ is *not* the MY regularization of $E^\lambda(v)$. Instead, the following relation holds (from equation (17), where the different sign is due to equation (32)),

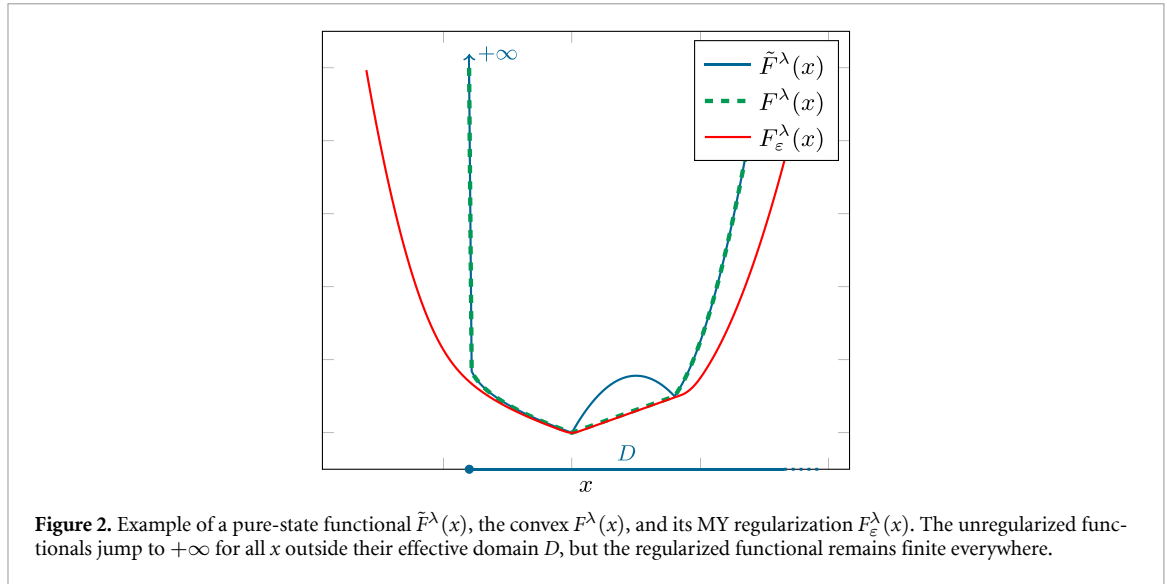
$$E_\varepsilon^\lambda(v) = E^\lambda(v) - \frac{\varepsilon}{2} \|v\|_{X^*}^2. \quad (51)$$

Since we subtract the norm-square of v from the concave $E^\lambda(v)$, the MY energy $E_\varepsilon^\lambda(v)$ is *strongly concave*, which implies strict concavity. Conversely, it holds that the Legendre–Fenchel transform of a strictly concave energy functional is always differentiable.

With $v = -\nabla F_\varepsilon^\lambda(x)$ we now have a *unique* representing potential $v \in X^*$ for every $x \in X$. We here write $\nabla F_\varepsilon^\lambda$ instead of $\partial F_\varepsilon^\lambda$ because we know that the regularized functional is differentiable. This then automatically yields a well-defined and unique mapping from ‘densities’ (elements of X) to potentials (elements of X^*) without any v -representability issue. We are not even limited to the N -representability set $D \subset X$ any more, and section 3.3 will help to make sense of the generalized densities $x \in X$. Uniqueness of the potential also means we automatically have a Hohenberg–Kohn-type result.

In the Hilbert-space setting (assumption 3) we further obtain a *quantitative* bound that shows how the difference between two potentials is bound by the difference between the corresponding densities. This comes in form of the Lipschitz estimate from property (ix):

$$\|\nabla F_\varepsilon^\lambda(x) - \nabla F_\varepsilon^\lambda(y)\|_{X^*} \leq \varepsilon^{-1} \|x - y\|_X. \quad (52)$$



This bound is lost in the limit $\epsilon \rightarrow 0$, in agreement with the fact that, in *standard* DFT, two densities that lie very close to each other can come from vastly different potentials [56], making density–potential inversion ill-posed and thus notoriously difficult. A similar bound is also possible in more general Banach spaces if the MY regularization is adapted to the ‘modulus of convexity’ of the space [57]. Furthermore, MY regularization allows for a well-defined formulation of the KS iteration procedure (section 5).

3.2. MY based on the total energy functional

A different viewpoint starts from the concave ground-state energy functional $E^\lambda(v)$ and introduces a *strictly concave* energy $E_\epsilon^\lambda(v)$ by just subtracting $\epsilon\phi(v) = \frac{\epsilon}{2}\|v\|_{X^*}^2$:

$$E_\epsilon^\lambda(v) = E^\lambda(v) - \frac{\epsilon}{2}\|v\|_{X^*}^2. \quad (53)$$

Convex conjugation of this strictly concave functional then yields the regularized density functional of equation (49) as shown by equation (17),

$$F_\epsilon^\lambda(x) = (-E_\epsilon)^*(-x) = \sup_{v \in X^*} \{E_\epsilon^\lambda(v) - \langle v, x \rangle\}. \quad (54)$$

We have thus defined MY regularization starting from the energy functional instead of the universal density functional. But talking about the total energy allows to give this manipulation a physical interpretation. For a given external potential $v \in X^*$, we introduce a *field energy functional*

$$E^f(v) = -\frac{\epsilon}{2}\|v\|_{X^*}^2, \quad (55)$$

where the minus sign that makes the functional concave, just like the concavity of the matter energy functional $E^\lambda(v)$. Of course, to retrieve the correct physical electrostatic *field energy*, $\frac{\epsilon_0}{2}\|\nabla v\|_{L^2}^2$, an appropriate choice of Banach space X^* must be made; see section 6.1. We then introduce the new energy functional (that could be interpreted as the total energy of matter *and* field)

$$E_\epsilon^\lambda(v) = E^\lambda(v) + E^f(v), \quad (56)$$

which corresponds to equation (53). In this way, MY regularization arises *naturally* by attaching more physical relevance to the external potential and by assigning to it an energy that has to be taken into account in the total energy functional.

Regarding the negative sign in equation (55), remember that in equation (50) it is $F_\epsilon^\lambda(x) + \langle v, x \rangle$ that is varied in order to find the ground state. There, the additional term is $\frac{1}{2\epsilon}\|x - \rho\|_X^2$ as seen in equation (49) and using the relation motivated in equation (57) below this is equal to $\frac{\epsilon}{2}\|v\|_{X^*}^2$, precisely the expression for the (positive) field energy. Viewed in this manner, it is the energy content of the external field that regularizes $F^\lambda(\rho)$. Conceptually, this is an extremely attractive viewpoint, since it explains the regularization as a physical effect.

3.3. MY based on density–potential mixing

Remember that we defined the N -representability set $D \subset X$ in equation (29) as the set of all $\rho \in X$ where $F_{\text{DM}}^\lambda(\rho) < \infty$, known as the effective domain of $F_{\text{DM}}^\lambda(\rho)$. For the sake of argument, we now suppose assumption 4 under which $F_{\text{DM}}^\lambda(\rho) = F^\lambda(\rho)$. Now even for $\rho \in D$, we can have issues with $\partial F^\lambda(\rho)$. First, $\partial F^\lambda(\rho)$ may be empty. In such cases, no potential $v \in X^*$ exists such that $v \in -\partial F^\lambda(\rho) \iff \rho \in \partial E^\lambda(v)$, implying that ρ is not v -representable. Also, at v -representable densities a small variation of ρ can lead outside the set of v -representable densities [41, 56]. There are further examples on small lattice systems where even large variations of v lead to almost (or even exactly) the same ground-state density ρ (and thus give counterexamples to the Hohenberg–Kohn theorem) [36, 58].

It would be better altogether if we could ignore this difficult notion of ‘ v -representability’ and instead are able to give a physical interpretation to all elements $x \in X$. This is achieved with the inverse duality map $J^{-1}: X^* \rightarrow X$ as the canonical way of mapping potentials to densities. For any v -representable $\rho \in \partial E^\lambda(v)$ coming from $v \in X^*$ define the *mixed density*

$$x = \rho - \varepsilon J^{-1}(v). \quad (57)$$

This definition is unique only if a Hohenberg–Kohn-type result is available so that the corresponding $v \in -\partial F^\lambda(\rho)$ is unique, but even without uniqueness we can make this assignment and will simply allow multiple such combinations. Choosing $\varepsilon > 0$ small means that in order to get the generalized density, we only mix in a small quantity of the corresponding potential v . We note that the resulting x does not need to be positive, normalizable, or have a finite kinetic energy, only the v -representable density ρ has all those properties. This is immediately visible from the definition in equation (57), where x is seen to be the sum of the electronic density ρ and $-\varepsilon J^{-1}(v)$. With $v \in X^*$ as an *external* potential with assigned density $-\varepsilon J^{-1}(v)$ that gets added to the *internal* electronic density of a quantum system, one could call x a ‘bidensity’. It has been given different names before, such as ‘depot’ (for density–potential), ‘pseudodensity’ [1], or ‘quasidensity’ [59]. We will not commit to a singular nomenclature here, and instead refer to it as a general element $x \in X$ or a ‘mixed density’ when its decomposition like in equation (57) matters.

One physically motivated possibility is to choose the spaces such that J corresponds to solving the Poisson equation of electrostatics and $\varepsilon J^{-1}(v) = -\varepsilon \Delta v$. Then $\rho_{\text{ext}} = \varepsilon J^{-1}(v)$ can be physically interpreted as just the external density that enters the Poisson equation to yield the electrostatic potential v . The mixed density is then the total observable density of the system, $x = \rho - \rho_{\text{ext}}$ (the minus sign is needed since we have used the convention that the electronic density is positive, $\rho \geq 0$). Interestingly, even the MY parameter ε then receives a physical interpretation as the vacuum permittivity. This choice was previously suggested in the context of density–potential inversion [4]; it will be covered in more detail in section 6.1 and has already found multiple use in ameliorating the KS method [60–63], albeit not with the connection to the topology of density and potential space.

Since $\rho \in \partial E^\lambda(v) = \partial(F^\lambda)^*(-v)$ from equation (32) (not forgetting a minus from the chain rule) we note that remarkably, equation (57) reproduces equation (18) from the properties of the MY regularization. We just have to use $E_\varepsilon^\lambda(v) = -(F_\varepsilon^\lambda)^*(-v)$ from equation (50) and the homogeneity of the duality map to get

$$x \in \partial E_\varepsilon^\lambda(v) = \partial E^\lambda(v) - \varepsilon J^{-1}(v), \quad (58)$$

which is just the superdifferential of equation (53). Since $\partial E_\varepsilon^\lambda$ has a well-defined inverse, $-\nabla F_\varepsilon^\lambda$, it follows that for any $x \in X$ we can find a $v \in X^*$ such that the representation in equation (57) holds. In other words, *every* $x \in X$ can be represented as a mixed density. The regularizing effect of the mixing for $F^\lambda(\rho)$ is now easily explained: If $v \neq v'$ lead to (almost) the same ground-state densities $\rho \approx \rho'$, then this will not be true any more for $x = \rho - \varepsilon J^{-1}(v)$ and $x' = \rho' - \varepsilon J^{-1}(v')$ due to the additional external term.

Also the proximal map receives a simple interpretation. If we set $v = -\nabla F_\varepsilon^\lambda(x)$ in equation (57) and solve for ρ , it follows from property (v) in section 2.3 that

$$\rho = x - \varepsilon J^{-1}(\nabla F_\varepsilon^\lambda(x)) = \underset{\varepsilon F^\lambda}{\text{prox}}(x). \quad (59)$$

Hence the internal density is just the proximal point of the mixed density. The consequence for DFT is that, to be able to work with the MY-regularized functionals, we also have to switch from internal densities $\rho \in D$ to general, mixed densities $x \in X$. The usual internal densities can then always be recovered as the proximal point or by performing the limit $\varepsilon \rightarrow 0$, where the limit procedure is well-defined through the properties of the MY regularization.

4. Density–potential inversion

The limit procedure in property (vii) of the MY regularization in section 2.3 that says $\lim_{\varepsilon \rightarrow 0} \nabla F_\varepsilon^\lambda(x) \in \partial F^\lambda(x)$ will be the basis for a method to calculate the corresponding external potential from a given ground-state density. The method itself is already known since a long time as the ‘ZMP method’ [2], but our derivation as well as the accompanying interpretation is only feasible by resorting to MY regularization. We also add a more speculative subsection that makes an interesting and previously unexplored connection to the Dyson equation. In this section we will require assumption 1 but sometimes, when it is mentioned, demand higher regularity of the involved spaces for additional results. Assumption 4 ($F^\lambda = F_{\text{DM}}^\lambda$) is further necessary in order to stay connected to solutions of the Schrödinger equation as explained in section 2.4.

As a motivation, we start with the variational problem for the universal density functional from equation (37),

$$F^\lambda(\rho) = \sup_{v \in X^*} \{E^\lambda(v) - \langle v, \rho \rangle\}, \quad (60)$$

whose stationarity condition $\partial E^\lambda(v) + \rho \ni 0$ gives $v \in -\partial F^\lambda(\rho)$ as an optimizer if ρ is assumed v -representable. This approach to density–potential inversion was first applied to atoms by Colonna and Savin [64] in 1999 and by Teale, Coriani and Helgaker to molecules [65–67] in 2009. It is sometimes called ‘Lieb optimization’, ‘Lieb maximization’ [66–68] or the ‘dual problem’ [69] in the literature. The method of Wu and Yang [70] turns this principle into a practical constrained optimization problem for the KS system. Convergence of Lieb optimization can be facilitated by making the functional strictly concave by subtracting $\frac{\varepsilon}{2} \|v\|_{X^*}^2$, which already corresponds to MY regularization,

$$\begin{aligned} F_\varepsilon^\lambda(\rho) &= \sup_{v \in X^*} \{E_\varepsilon^\lambda(v) - \langle v, \rho \rangle\} \\ &= \sup_{v \in X^*} \left\{ E^\lambda(v) - \frac{\varepsilon}{2} \|v\|_{X^*}^2 - \langle v, \rho \rangle \right\}. \end{aligned} \quad (61)$$

The stationarity condition for an optimizer v_ε is then

$$0 \in \partial E^\lambda(v_\varepsilon) - \varepsilon J^{-1}(v_\varepsilon) - \rho. \quad (62)$$

We use equation (58) to translate this to $\rho \in \partial E_\varepsilon^\lambda(v_\varepsilon)$ which is equivalent to $v_\varepsilon = -\nabla F_\varepsilon^\lambda(\rho)$ by using the analogous relation to equation (39) for the regularized functionals. By property (v) from section 2.3, we can connect v_ε to the proximal map,

$$v_\varepsilon = \varepsilon^{-1} J \left(\text{prox}_{\varepsilon F^\lambda}(\rho) - \rho \right). \quad (63)$$

This equation then serves as the basis for MY density–potential inversion as an iterative procedure involving the proximal point of ρ , to find v_ε , which in the limit $\varepsilon \rightarrow 0$ gives the representing potential of ρ . The method will be more properly derived from MY regularization in the following section.

4.1. Proximal-point iteration and relation to ZMP

The inverse problem of finding the corresponding potential v for a given ground-state density ρ is central to DFT and to achieve this, we ‘only’ have to evaluate the subdifferential of $F^\lambda(\rho)$ (with a minus sign, see equation (38)),

$$v \in -\partial F^\lambda(\rho). \quad (64)$$

As mentioned before, it is not guaranteed that this subdifferential is non-empty even for $\rho \in D \subset X$ and, in non-standard settings for DFT (like DFT on finite lattices [36]), it may even be multi-valued. For the regularized functional, no such problems occur, and we can evaluate

$$v_\varepsilon = -\nabla F_\varepsilon^\lambda(\rho) \quad (65)$$

for any $\rho \in D$ and even for $\rho \in X \setminus D$. Of course, this does in general not yield the same potential as equation (64), but it does if $\rho = \text{prox}_{\varepsilon F^\lambda}(\rho)$ by property (v) in section 2.3 or as $\varepsilon \rightarrow 0$ by property (vii). We therefore know how the functional derivative in equation (65) relates back to the non-regularized form, but we may wonder how it can be computed.

Noting relation (v) for the proximal map and introducing $p = \text{prox}_{\varepsilon F^\lambda}(\rho)$, we obtain

$$\nabla F_\varepsilon^\lambda(\rho) = \varepsilon^{-1}J(\rho - p) \in \partial F(p). \quad (66)$$

which by equation (38) gives

$$p \in \partial E^\lambda(\varepsilon^{-1}J(p - \rho)). \quad (67)$$

Now we have the unknown $p \in D$ on both sides, so this form lends itself to an iteration scheme. Beginning with a guess p_1 , we iterate

$$p_{i+1} \in \partial E^\lambda(\varepsilon^{-1}J(p_i - \rho)). \quad (68)$$

Due to the possibility of degeneracy in the ground state, the next step p_{i+1} is possible non-unique, which creates a potential issue that has not been explored up to date. If converged, this process yields the proximal point $p = \text{prox}_{\varepsilon F^\lambda}(\rho)$ and we retrieve equation (63),

$$v_\varepsilon = -\nabla F_\varepsilon^\lambda(\rho) = \varepsilon^{-1}J(p - \rho), \quad (69)$$

where the resulting potential v_ε depends on the choice of the smoothing parameter ε . Repeating this iterative scheme for a sequence of decreasing $\varepsilon > 0$, we can extrapolate to the limit $\varepsilon = 0$ and obtain $v_\varepsilon \rightharpoonup v$. Under the minimal assumption 1, convergence $v_\varepsilon \rightharpoonup v$ is weak; under the stronger assumption 2, convergence is strong. Also by virtue of (vii), the resulting potential v is the element of $\partial F^\lambda(\rho)$ with the smallest norm.

The benefit of this method is that we do not need to evaluate the (usually unknown) $F^\lambda(\rho)$, relying instead on $\partial E^\lambda(v)$ by solving the Schrödinger equation for the ground state. This density-potential inversion procedure using MY regularization and its relation to the closely related ZMP method [2] was previously detailed in Penz *et al* [4].

4.2. MY in inverse KS

While the procedure from the preceding section remains a hard task for interacting particles, density-potential inversion is often considered for the non-interacting KS system ($\lambda = 0$) and then typically specified as the ‘inverse KS method’. Starting with a given ground-state density ρ for some external potential v in an interacting system that is also non-interacting v -representable, the KS potential must satisfy

$$v_s \in -\partial F^0(\rho). \quad (70)$$

We can approach this potential according to equation (69) with

$$v_s = \lim_{\varepsilon \rightarrow 0} \frac{1}{\varepsilon} J \left(\text{prox}_{\varepsilon F^0}(\rho) - \rho \right), \quad (71)$$

where the proximal map can be evaluated by using the iteration in equation (68). To obtain $\rho \in \partial E^0(v)$, we solve the non-interacting Schrödinger equation, introducing orbitals.

A more intuitive viewpoint for this reasoning is to compare the corresponding minimization of the MY regularization and the very KS idea. Since ρ is by assumption a minimizer of $F^0(\cdot) + \langle v_s, \cdot \rangle$, and $\text{prox}_{\varepsilon F^0}(\rho)$ is the minimizer of $F^0(\cdot) + \frac{1}{\varepsilon} \|\cdot - \rho\|^2$, we must have simultaneously

$$0 \in \partial F^0(\rho) + v_s, \quad (72)$$

$$0 \in \partial F_\varepsilon^0 \left(\text{prox}_{\varepsilon F^0}(\rho) \right) + \frac{1}{\varepsilon} J \left(\text{prox}_{\varepsilon F^0}(\rho) - \rho \right). \quad (73)$$

Motivated by the fact that $\text{prox}_{\varepsilon F^0}(\rho) \rightarrow \rho$ as $\varepsilon \rightarrow 0$, one can then under technical assumptions prove that indeed v_s corresponds to the limit of $\frac{1}{\varepsilon} J(\text{prox}_{\varepsilon F^0}(\rho) - \rho)$ as the equations above suggest.

Now exactly this method was proposed and implemented as a density-potential inversion scheme by Zhao *et al* [2] (and was later termed ‘ZMP method’), where the choice of spaces corresponds to homogeneous Sobolev spaces for potentials and their duals for densities (also see section 6.1 and an upcoming publication for periodic domains like in section 7). But in the original publication the method was developed ad hoc and only proven to work empirically. Only quite recently it was shown [4] that the proximal-point iteration with $\varepsilon \rightarrow 0$ relates closely to the ZMP method for the usual Sobolev spaces and that it also reproduces a method by van Leeuwen and Baerends [71] if employed on a Hilbert space with $J = J^{-1} = \text{id}$. It is further equivalent to yet another density-potential inversion method by Kumar and

Harbola [72], where different choices for their penalty function $S[\rho]$ would correspond to different topologies for X . The method has been analyzed numerically [20, 21] (see also section 7 here) and adds itself to a long list of KS inversion techniques [73]. More than that, it shows that MY regularization offers a mathematical framework that allows for an elegant unification of numerous density–potential inversion methods.

4.3. Resolvent form and Dyson-like equation

Instead of establishing a practical scheme for density–potential inversion, this section will be devoted to an interesting reformulation of the basic relation $v = -\nabla F_\varepsilon^\lambda(x)$ that becomes possible for a regularized functional. Starting from equation (18) for the universal density functional F^λ and using equation (14) we get

$$(\nabla F_\varepsilon^\lambda)^{-1}(v) = \partial(F^\lambda)^*(v) + \varepsilon J^{-1}(v). \quad (74)$$

Even though $\nabla F_\varepsilon^\lambda(x)$ is single-valued, its inverse is in general a set-valued map, so the equation above is actually between sets. From equation (32) we get (not forgetting the minus from the chain rule) $\partial(F^\lambda)^*(-v) = \partial E^\lambda(v)$ and we apply this to the function above after switching $v \mapsto -v$, where we also use that J^{-1} is homogeneous. Then for $v = -\nabla F_\varepsilon^\lambda(x)$ (or, more directly, from equation (58)) we have

$$(\nabla F_\varepsilon^\lambda)^{-1}(-v) = \partial E^\lambda(v) - \varepsilon J^{-1}(v) = x. \quad (75)$$

Finally, we invert the whole expression, arriving at a resolvent-type formula that can be applied to all elements from X , the density space, to yield the corresponding potential,

$$-\nabla F_\varepsilon^\lambda = (\partial E^\lambda - \varepsilon J^{-1})^{-1}. \quad (76)$$

Note that in the Hilbert-space setting we already saw a resolvent form for the proximal map in property (x) in section 2.3. By transforming equation (75) into an iterative procedure to solve for v , we arrive at

$$-\nabla F_\varepsilon^\lambda = -\varepsilon^{-1}J + \varepsilon^{-1}J \circ \partial E^\lambda \circ (-\nabla F_\varepsilon^\lambda). \quad (77)$$

Here we used ‘ $\circ$ ’ to denote the concatenation of operations and, unsurprisingly, this result is equivalent to the proximal-point iteration in equation (68) if we combine it with equation (69). But writing it in the form above exposes a striking parallel to the Dyson equation [74], where $-\nabla F_\varepsilon^\lambda$ takes the role of the interacting Green’s function (dressed propagator), $-\varepsilon^{-1}J$ that of the free Green’s function (bare propagator), and $-\partial E^\lambda$ (with an additional minus to get the sign right) would be the self-energy. Applying $-\varepsilon J^{-1}$ from the left and $(-\nabla F_\varepsilon^\lambda)^{-1} = \partial E_\varepsilon^\lambda$ from the right yields

$$\partial E_\varepsilon^\lambda = \partial E^\lambda - \varepsilon J^{-1}, \quad (78)$$

which is just equation (58) again and shows another path to arrive at equation (77). In our setting of MY applied to DFT, the mapping $-\varepsilon^{-1}J$, that gets defined by the topology of the space X itself, gives the ‘bare’ connection between densities and potential. Compared to this, $-\nabla F_\varepsilon^\lambda$ is a ‘dressed’ way how to connect the quantities, where the dressing occurs through the presence of particles the potential can interact with. Lastly, $-\partial E^\lambda$ derives from the functional $E^\lambda(v)$, the energy of the matter system. This last element introduces a possible ambiguity into the formulation that needs to be considered if degeneracy occurs and $\partial E^\lambda(v)$ is a multi-valued set. Nevertheless, this reformulation allows for an interesting alternative viewpoint that links it to Green’s function methods and diagrammatic expansions. We hope that this can serve as a starting point for future investigations, either for concrete approximations to the potential $-\nabla F_\varepsilon^\lambda$ or into how to effectively link this back to the corresponding unregularized form $-\partial F^\lambda$.

5. Regularized KS theory

We already noted in section 2.5 that the requirement of having all involved densities simultaneously interacting and non-interacting v -representable puts a severe limitation on the possibility of a rigorous formulation of the KS method. This limitation can also influence numerical applications of the method, since it means that basis-set limits are ill-defined. The MY regularized functional $F_\varepsilon^\lambda(x)$ offers a way how to avoid representability issues since the functional is differentiable and every $x \in X$ becomes representable through $-\nabla F_\varepsilon^\lambda(x)$. The corresponding regularized KS method was first formulated in Kvaal [1] for a Hilbert-space setting and later extended to Banach spaces [15]. In this section, we not only give its

basic formulation, but also show how usual exchange–correlation can be translated to the new method. Finally, we repeat the main steps of a convergence proof for the method with an additional damping step. As usual, we require assumption 1 but have to move to the Hilbert-space setting (assumption 3) and even finite-dimensional spaces for the convergence proof. Since in the KS method we ought to stay connected to the Schrödinger equation and related KS equation, also assumption 4 ($F^\lambda = F_{\text{DM}}^\lambda$) is necessary.

5.1. Iteration scheme

We will build directly on the iteration scheme introduced in section 2.5 but, with the help of MY regularization, can completely avoid the critical issues within the usual formulation of the KS method. As a reminder, we were able to introduce a well-defined exchange–correlation potential $v_{\text{xc}}(\rho)$ only if ρ is simultaneously interacting and non-interacting v -representable, i.e. at every density ρ that occurs within the KS iteration both subdifferentials $\partial F^1(\rho)$ and $\partial F^0(\rho)$ must be non-empty. From what we learned about MY regularization, this can be directly guaranteed if instead we switch to the regularized functionals $F_\varepsilon^1(x)$ and $F_\varepsilon^0(x)$. However, we then need to make sense of the resulting regularized KS iteration procedure. So let us define in exact analogy to equation (40) for any $y \in X$ with $v(y) = -\nabla F_\varepsilon^1(y)$ and $v_s(y) = -\nabla F_\varepsilon^0(y)$,

$$\begin{aligned} v_{\text{xc},\varepsilon}(y) &= v_s(y) - v(y) - v_H(y) \\ &= \nabla F_\varepsilon^1(y) - \nabla F_\varepsilon^0(y) - v_H(y) \\ &= \nabla \left(\underbrace{F_\varepsilon^1(y) - F_\varepsilon^0(y) - E_H(y)}_{E_{\text{xc},\varepsilon}(y)} \right). \end{aligned} \quad (79)$$

We now write y instead of ρ to highlight that this is an arbitrary element in X that does not need to have the properties of an actual density (see also section 3.3). We can now define the same, self-consistent iteration procedure as in equations (44) and (45), where we can use equation (58) (which is equivalent to equation (18)) for the superdifferential of the E_ε^0 functional. Let the external potential that acts on the fully interacting system now be fixed as $v_{\text{off}} + v$ with $v = -\nabla F_\varepsilon^1(x)$ (as before, the offset v_{off} will be included into the functionals), then iterate starting from an initial guess,

$$v_{i+1} = v + v_H(x_i) + v_{\text{xc},\varepsilon}(x_i), \quad (80)$$

$$x_{i+1} \in \partial E_\varepsilon^0(v_{i+1}) = \partial E^0(v_{i+1}) - \varepsilon J^{-1}(v_{i+1}). \quad (81)$$

In solving for the next x_{i+1} in equation (81), we rely on exactly the same non-interacting Schrödinger equation that allows an orbital construction as in the usual KS procedure, we just have to shift the resulting density by $-\varepsilon J^{-1}(v_{i+1})$ as the density–potential mixing dictates. If the procedure converges (see section 5.3 for a convergence proof that applies to finite-dimensional X if an additional damping step is added), we get

$$x_i \rightarrow x, \quad v_i \rightarrow v_s, \quad (82)$$

and the resulting x can then be mapped to the ground-state density with the proximal map according to equation (59). But now an important difference to the usual KS procedure appears that was already apparent in the formulation of the regularized KS iteration: The method aims at a shared $x \in X$ for the full and auxiliary system, and not at the internal ground-state density. This means that when $v_{\text{off}} + v_s$ is applied to the auxiliary system we get the same $x = \rho_s - \varepsilon J^{-1}(v_s)$ as when $v_{\text{off}} + v$ is applied to the full system. Therefore, in general $\rho \neq \rho_s$ and instead we have the ground-state densities

$$\rho = x + \varepsilon J^{-1}(v) = \text{prox}_{\varepsilon F^1}(x), \quad (83)$$

$$\rho_s = x + \varepsilon J^{-1}(v_s) = \text{prox}_{\varepsilon F^0}(x). \quad (84)$$

That it must hold $\rho, \rho_s \in D$ serves as a consistency check for the method or can be used as an *a posteriori* correction. To get an expression for the ground-state energy, we can rely on the analogous expression to equation (48),

$$E_\varepsilon^1(v) = E_\varepsilon^0(v_s) + E_{\text{xc},\varepsilon}(x) - E_H(x) - \langle v_{\text{xc},\varepsilon}(x), x \rangle, \quad (85)$$

that we just need to shift as in equation (17),

$$E^1(v) = E_\varepsilon^1(v) + \frac{\varepsilon}{2} \|v\|_{X^*}^2. \quad (86)$$

Note, however, that while we recover the usual ground-state energy $E^1(v)$ this way, it does not mean that the regularized KS method will yield the same orbitals as the usual method. This is clear since the orbitals in regularized KS will produce a density ρ_s that is different to the interacting density ρ , and instead the generalized densities $x = \rho - \varepsilon J^{-1}(v) = \rho_s - \varepsilon J^{-1}(v_s)$ agree.

5.2. Transforming exchange–correlation potentials

While this reformulation of the KS iteration scheme with the help of MY regularization is conceptually already very appealing, the question remains how it relates back to the unregularized setting. Having again $v = -\nabla F_\varepsilon^1(x)$, in the first KS step that is equation (80) we rely on an exchange–correlation potential for the regularized scheme at every step x_i for which so far only the defining relation equation (79) was given. More concretely, we want to know how $v_{xc,\varepsilon}$ relates to the usual v_{xc} or any approximation to it. For this we take equations (83) and (84) and combine them into

$$v_{xc,\varepsilon}(x) = \varepsilon^{-1}J(\rho_s - x) - \varepsilon^{-1}J(\rho - x) - v_H(x). \quad (87)$$

Note that if X is a Hilbert space then J is a linear mapping and the expression above can be simplified. Within the regularized KS iteration we already have access to ρ_s in every step as an element of $\partial E^0(v)$, so we just need the corresponding $\rho \in \partial E^1(v)$ in order to evaluate the above equation. Since we of course want to avoid solving the full problem, the v_{xc} from standard KS-DFT (or any approximation to it) enters the picture, since it allows getting the same density from the auxiliary system by adding just the usual exchange–correlation potential, $\rho \in \partial E^0(v + v_H(\rho) + v_{xc}(\rho))$. Together with $v = \varepsilon^{-1}J(\rho - x)$, we can derive a self-consistent iteration that works similar to KS itself,

$$\begin{array}{c} \rho \in \partial E^0(\varepsilon^{-1}J(\rho - x) + v_H(\rho) + v_{xc}(\rho)) \\ \updownarrow \qquad \qquad \updownarrow \qquad \qquad \updownarrow \qquad \qquad \updownarrow \\ \rho_{i+1} \in \partial E^0(\varepsilon^{-1}J(\rho_i - x) + v_H(\rho_i) + v_{xc}(\rho_i)). \end{array} \quad (88)$$

If this iteration is integrated into the regularized KS scheme from section 5.1 together with determining $v_{xc,\varepsilon}(x)$ from equation (87), we can recycle any given exchange–correlation approximation for the regularized setting. It must be added that this transformation of exchange–correlation potentials for the regularized KS method was not implemented to date.

5.3. Convergence proof

Now that a rigorous formulation of the KS method has been achieved, the question of convergence of the method to the desired ground-state density, when using the exact exchange–correlation functional, remains to be answered. First results on convergence of the usual KS method (section 2.5) required a v -representability assumption and only showed convergence in energy [75, 76]. With the help of MY regularization the problem of v -representability could be avoided and by the additional inclusion of a damping step a full convergence proof was achieved at least for a setting where X is a (finite-dimensional) Euclidean space [22, 77, 78]. This setting means that assumption 3 holds, and we slightly generalize the result to any finite-dimensional Hilbert space X in the following presentation.

Implementations of the KS method regularly come with additional steps like damping that facilitate convergence. Such a damping step must also be included here, which means that in equation (81) not the full step from x_i to x_{i+1} , but one with a reduced step length is applied. This step length will be such that one reaches exactly the apex of the regularization parabola that is tangentially aligned to $F_\varepsilon^1 + v$ at x_i , as illustrated in figure 3. Yet, it must be stressed that such a choice of step-length does not yield a practical guideline, since it requires knowledge about the fully interacting functional. Its sole purpose here is to be able to establish a mathematical convergence proof. An actual implementation of the KS method must decide about the step length from the available information from the previous steps. Different schemes with an *a priori* or adaptive choice of step length can be constructed as well [15, 16, 55] and their convergence rates have been compared [16, 77]. We also note on a strong parallel to the proximal-point algorithm [79], where the minimum of $f(x)$ is found by iterating $x_{i+1} = \text{prox}_{\varepsilon f}(x_i)$. By the definition of the proximal map, equation (16), this corresponds to jumping to the apex of the full regularization function $\varepsilon\phi(x - x_i)$ attached to x_i in every step. The important difference to the regularized KS method is that the method fixes the step direction in the previous step (by solving the non-interacting system with the exchange–correlation potential) and then jumps to the apex of the regularization parabola restricted to this direction.

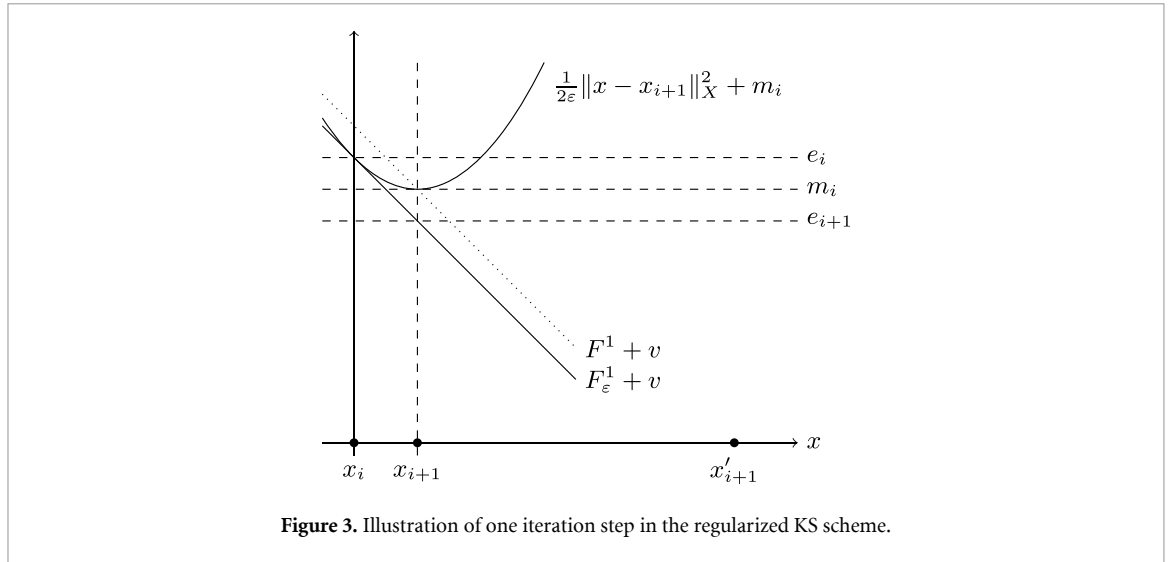


Figure 3. Illustration of one iteration step in the regularized KS scheme.

The regularized KS scheme with additional damping step is then given by the iteration

$$v_{i+1} = v + v_H(x_i) + v_{xc,\varepsilon}(x_i), \quad (89)$$

$$x'_{i+1} \in \partial E_\varepsilon^0(v_{i+1}) = \partial E^0(v_{i+1}) - \varepsilon J^{-1}(v_{i+1}) \quad (90)$$

and set

$$y_i = (x'_{i+1} - x_i) / \|x'_{i+1} - x_i\|_X, \\ x_{i+1} = x_i + \tau_i y_i,$$

where

$$\tau_i = -\varepsilon \langle \nabla F^1_\varepsilon(x_i) + v, y_i \rangle. \quad (91)$$

In the following, we provide most steps and explanations from the original proof [77] that was subsequently revised due to an error in the first version [78]. For additional details we refer to the original publications.

From the first step and using the definition of $v_{xc,\varepsilon}$ in equation (79) we have that

$$\nabla F^1_\varepsilon(x_i) + v = v_{i+1} + \nabla F^0_\varepsilon(x_i). \quad (92)$$

The iteration is converged if either side of this equation is zero, which by the inverse subdifferential relation, equation (14), means that $x_i \in \partial E_\varepsilon^1(v)$ and also $x_i \in \partial E_\varepsilon^0(v_{i+1})$ (the interacting and the non-interacting system share the same x as in equations (83) and (84)). In order to track progression of the method, we define

$$e_i = F^1_\varepsilon(x_i) + \langle v, x_i \rangle. \quad (93)$$

In case of correct convergence $x_i \rightarrow x \in \partial E_\varepsilon^1(v)$ we then have $\lim e_i = E_\varepsilon^1(v)$, which equals the ground-state energy of the interacting system minus $\frac{\varepsilon}{2} \|v\|_{X^*}^2$. We take the expression of equation (93) that is differentiable and convex in x_i and rewrite its directional derivative in the chosen direction $x'_{i+1} - x_i$ with the help of equation (92),

$$\langle \nabla F^1_\varepsilon(x_i) + v, x'_{i+1} - x_i \rangle = \langle v_{i+1} + \nabla F^0_\varepsilon(x_i), x'_{i+1} - x_i \rangle. \quad (94)$$

Notice that $x'_{i+1} \in \partial E_\varepsilon^0(v_{i+1})$ and $x_i \in \partial E_\varepsilon^0(-\nabla F^0_\varepsilon(x_i))$ by definition, so from section 2.3 we can use the strong monotonicity property (xii) of $\partial E_\varepsilon^0(v) = \partial(F^0_\varepsilon)^*(-v)$ (note the change of sign) to get

$$\begin{aligned} \langle \nabla F^1_\varepsilon(x_i) + v, x'_{i+1} - x_i \rangle &\leq -\varepsilon \|v_{i+1} + \nabla F^0_\varepsilon(x_i)\|_{X^*}^2 \\ &= -\varepsilon \|\nabla F^1_\varepsilon(x_i) + v\|_{X^*}^2. \end{aligned} \quad (95)$$

This means that by a step in the direction $x'_{i+1} - x_i$, which is parallel to y_i , we can always decrease the energy, if the step length is not taken too large. For this reason we limit the step length to τ_i such that

we just reach the apex of the regularization parabola (at value m_i). We get this value from equating the directional derivatives of $F^1_\varepsilon(x_i) + \nu$ and the regularization parabola at x_i , which yields

$$\langle \nabla F^1_\varepsilon(x_i) + \nu, y_i \rangle = -\frac{1}{\varepsilon} \|x_{i+1} - x_i\|_X = -\frac{\tau_i}{\varepsilon} \quad (96)$$

and guarantees $e_{i+1} \leq m_i < e_i$. Since the energy is bounded below and thus strictly decreasing, we already showed convergence in energy. By determining $e_i - m_i$ from the regularization parabola and using equation (96) we get

$$\frac{\tau_i^2}{2\varepsilon} = \frac{1}{2\varepsilon} \|x_{i+1} - x_i\|_X^2 = e_i - m_i \leq e_i - e_{i+1} \rightarrow 0 \quad (97)$$

that shows $\tau_i \rightarrow 0$. Taking equations (95) and (96) together gives the estimate

$$\frac{\tau_i}{\varepsilon} \|x'_{i+1} - x_i\|_X \geq \varepsilon \|\nabla F^1_\varepsilon(x_i) + \nu\|_{X^*}^2. \quad (98)$$

If we can be sure that $\|x'_{i+1} - x_i\|_X$ is bounded this means that $-\nabla F^1_\varepsilon(x_i) \rightarrow \nu$.

In order to show convergence in densities and potentials, a compactness argument must be used. The revised proof [78] uses the fact that $F^1_\varepsilon(x)$ increases like $\|x\|_X^2$ asymptotically outside of its effective domain D as a result of MY regularization. For X finite-dimensional, which critically enters here, all norms are equivalent and thus D is bounded since it only includes normalized densities. Since $e_i = F^1_\varepsilon(x_i) + \langle \nu, x_i \rangle < e_1$ from the result that the e_i are strictly decreasing, we arrive at $(x_i)_i$ bounded. Again, the finite dimensionality of X is critical in the next step, because it guarantees that a bounded sequence always has convergent subsequences. This means $(x_i)_i$ has (possibly infinitely many) convergent subsequences of which we exemplarily take $x_{\alpha(i)} \rightarrow z$. Since we are in a Hilbert-space setting (assumption 3), property (ix) holds and ∇F^1_ε is Lipschitz continuous. Now the first step equation (89) asserts

$$\begin{aligned} \nu_{\alpha(i)+1} &= \nu + \nu_H(x_{\alpha(i)}) + \nu_{\text{xc},\varepsilon}(x_{\alpha(i)}) \\ &= \nu + \nabla F^1_\varepsilon(x_{\alpha(i)}) - \nabla F^0_\varepsilon(x_{\alpha(i)}), \end{aligned} \quad (99)$$

so by continuity $\nu_{\alpha(i)+1} \rightarrow \nu + \nabla F^1_\varepsilon(z) - \nabla F^0_\varepsilon(z)$ and the subsequence is bounded. By the same argument as before also the sequence given by $x'_{\alpha(i)+1} \in \partial E^0_\varepsilon(\nu_{\alpha(i)+1})$ is bounded which means $\|x'_{\alpha(i)+1} - x_{\alpha(i)}\|_X$ is bounded and so we really have $\lim_{i \rightarrow \infty} (-\nabla F^1_\varepsilon(x_{\alpha(i)})) = -\nabla F^1_\varepsilon(z) = \nu$ from equation (98). This already implies $\lim_{i \rightarrow \infty} \nu_{\alpha(i)+1} = -\nabla F^0_\varepsilon(z)$, so we have found a potential for the non-interacting KS system reproducing $z \in \partial E^1_\varepsilon(\nu)$. Taking a different subsequence $x_{\beta(i)}$ can then yield another element from $\partial E^1_\varepsilon(\nu)$ as well as its corresponding KS potential.

6. Auto-regularization in mean-field theories

In certain, specialized settings, like the Hartree-approximation to DFT or DFT with induced magnetic fields (Maxwell–Schrödinger DFT), the special form of the additional energy terms leads to automatic MY regularization. Interestingly, in the first case this is not on the side of the universal density functional, as in equation (49), but affects the ground-state energy $E^\lambda(\nu)$. In the case of Maxwell–Schrödinger DFT we can apply Legendre–Fenchel transformations to two different arguments and thus have the possibility to define four different functionals from convex conjugation, in some analogy to the potentials of thermodynamics that are also linked by Legendre–Fenchel transformations. There, it turns out that two of those functionals can even be linked to a certain form of double MY regularization called Lasry–Lions regularization.

6.1. Hartree auto-regularization

Through a special choice for the potential space previously mentioned in sections 3.2 and 3.3, we can achieve automatic regularization of the energy functional in the Hartree approximation of DFT. This choice is the homogeneous Sobolev space [80–82],

$$X^* = \{u \in L^1_{\text{loc}}(\mathbb{R}^3) \mid \nabla u \in L^2(\mathbb{R}^3; \mathbb{R}^3)\}, \quad (100)$$

that must be taken modulo constants in order to make $\|u\|_{X^*} = \sqrt{\varepsilon_0} \|\nabla u\|_{L^2}$ a valid norm. This includes the customary ‘up to an additive constant’ already into the definition of the potential space. Note that

we also built the vacuum permittivity $\varepsilon_0 (= 1/(4\pi)$ in atomic units) into the definition of the norm for reasons that will soon become evident. This space is actually a Hilbert space with inner product $\langle u, v \rangle_{X^*} = \varepsilon_0 \langle \nabla u, \nabla v \rangle_{L^2}$, so we have assumption 3 fulfilled. Applying partial integration in the L^2 inner product gives

$$\langle u, v \rangle_{X^*} = -\varepsilon_0 \int u(\mathbf{r}) \Delta v(\mathbf{r}) d\mathbf{r} = \langle u, J^{-1}(v) \rangle, \quad (101)$$

where the last term stems from the Riesz representation equation (7) applied to X^* and means that $u \in X^*$ is applied to $J^{-1}(v) \in X$. But this allows us to write the inverse duality map as

$$J^{-1}(v) = -\varepsilon_0 \Delta v, \quad (102)$$

so evaluating $v = J(x)$ is equivalent to solving the Poisson equation $-\varepsilon_0 \Delta v = x \in X$. Strictly speaking, the integral in equation (101) makes only sense if Δv is still an L^2 function, but we can still take equation (102) in the sense of distributions. This implies that a general element $x \in X$ can be a distribution. Since J is the canonical mapping between densities and potentials, and is itself fully defined by the topology (norm) of the space, we needed to include ε_0 in the norm's definition already to get the units right (although in atomic units one would not notice). Now for regular distributions $x \in X$ (that can be represented by an $f \in L^1_{\text{loc}}(\mathbb{R}^3)$) as $\langle u, x \rangle = \int u(\mathbf{r}) f(\mathbf{r}) d\mathbf{r}$ we know the solution of $-\varepsilon_0 \Delta v = x$ to be exactly the Hartree potential (in this context also called the Riesz potential; note again that in atomic units $4\pi\varepsilon_0 = 1$),

$$v(\mathbf{r}) = \int \frac{f(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|} d\mathbf{r}' = v_{\text{H}}(f)(\mathbf{r}). \quad (103)$$

Thus we have found an explicit expression for the duality map, $J(x) = v_{\text{H}}(f)$, when $x \in X$ is a regular distribution represented by $f \in L^1_{\text{loc}}(\mathbb{R}^3)$. Note that we usually use the symbols for distributional densities ' x ' and their representing functions ' f ' for elements of X interchangeably, so we could have also written $J(x) = v_{\text{H}}(x)$. Further note that since the potentials are defined modulo a constant it must also hold $\langle u, x \rangle = \langle u + c, x \rangle$ which implies $\int f(\mathbf{r}) d\mathbf{r} = 0$, i.e. all integrable f representing an $x \in X$ must integrate to zero. This may seem problematic if we think of densities f as normalized to a certain particle number, but it can be compensated by the introduction of a fixed reference density and shifting X to an affine space with the same dual X^* . Now take the norm in X^* with one-half in front,

$$\frac{1}{2} \|v\|_{X^*}^2 = \frac{\varepsilon_0}{2} \|\nabla v\|_{L^2}^2 = \frac{\varepsilon_0}{2} \int |\nabla v(\mathbf{r})|^2 d\mathbf{r}, \quad (104)$$

and we can recognize this as the electrostatic field energy. We can equivalently rewrite the norm in X with the dual norm as in equation (5),

$$\begin{aligned} \frac{1}{2} \|v\|_{X^*}^2 &= \frac{1}{2} \|x\|_X^2 = \frac{1}{2} \langle J(x), x \rangle \\ &= \frac{1}{2} \iint \frac{f(\mathbf{r}) f(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|} d\mathbf{r} d\mathbf{r}' = E_{\text{H}}(f), \end{aligned} \quad (105)$$

which retrieves the Hartree mean-field energy from equation (42) that is also the energy of the field induced by the charge distribution $f \in X$. Through these identifications, it becomes apparent that this choice of density and potential space is a very natural one for DFT and we believe that it can have a future importance for the formulation of DFT. Indeed, the use of similar spaces in DFT was already suggested around 2012, based on an appearance in regularization of optimized-effective-potential methods [61], but never published. Still, several important mathematical questions are left to be studied in this setting, especially the self-adjointness and boundedness below of the Hamiltonian and lower semicontinuity of the ensemble constrained-search functional (assumption 4). Like the summary of section 2.4 states, to make the the last point even possible to hold together with assumption 1, one could switch to a bounded domain or include a confining offset potential v_{off} .

A very basic approximation for an interacting system is the Hartree functional

$$F^{\text{Hartree}}(\rho) = F^0(\rho) + E_{\text{H}}(\rho) = F^0(\rho) + \frac{1}{2} \|\rho\|_X^2, \quad (106)$$

which becomes strongly (and thus strictly) convex through the addition of the Hartree mean-field energy. We are now prepared to show the auto-regularization of the corresponding energy functional

from the Legendre–Fenchel transform (the attentive reader might already have noticed that a strictly convex functional $F^{\text{Hartree}}(\rho)$ implies a MY-regularized convex conjugate functional). With $v = J(x)$ we get

$$\begin{aligned}
 E^{\text{Hartree}}(v) &= -\left(F^{\text{Hartree}}\right)^*(-v) \\
 &= \inf_{\rho} \left\{ F^{\text{Hartree}}(\rho) + \langle v, \rho \rangle \right\} \\
 &= \inf_{\rho} \left\{ F^0(\rho) + \frac{1}{2} \|\rho\|_X^2 + \langle v, \rho \rangle \right\} \\
 &= \inf_{\rho} \left\{ F^0(\rho) + \frac{1}{2} \|\rho\|_X^2 + \langle v, \rho \rangle + \frac{1}{2} \|v\|_{X^*}^2 \right\} - \frac{1}{2} \|v\|_{X^*}^2 \\
 &= \inf_{\rho} \left\{ F^0(\rho) + \frac{1}{2} \|\rho + x\|_X^2 \right\} - \frac{1}{2} \|v\|_{X^*}^2 \\
 &= F^0_{\varepsilon=1}(-x) - \frac{1}{2} \|v\|_{X^*}^2 = F^0_{\varepsilon=1}(-J^{-1}(v)) - \frac{1}{2} \|v\|_{X^*}^2.
 \end{aligned} \tag{107}$$

We can note several things here. In the third line we recognize that the considered energy content of the system is the (non-interacting) kinetic energy $F^0(\rho)$ of the system, the potential energy from coupling to the external v , but also the energy of the induced field $E_H(\rho)$. This sums up to give the self-interaction error of the Hartree approximation. The regularization parameter can be read off directly from the fifth line of equation (107) as $\varepsilon = 1$. Without including the vacuum permittivity into the norm we can have $\varepsilon = \varepsilon_0$, but then the duality map does not directly map densities to potentials any more. The new argument of the regularized functional is $-x = -J^{-1}(v) = \varepsilon_0 \Delta v$ and can of course be negative as a general element in X . Note that now the energy functional $E^{\text{Hartree}}(v)$ is the result of MY regularization, not the universal density functional as before. Even though we subtract $\frac{1}{2} \|v\|_{X^*}^2$, which reminds of the way MY regularization was introduced in section 3.2, the energy functional does not necessarily become strictly concave (yet it must be concave by construction) since $F^0_{\varepsilon}(-x)$ is convex. Since $\rho = \nabla E^{\text{Hartree}}(v)$ is now unique, there arises no ambiguity in the ground-state density.

6.2. Maxwell–Schrödinger DFT

Maxwell–Schrödinger DFT [83] is an extension of standard DFT with electromagnetic fields and internal currents that also considers induced magnetic fields. In this respect it is similar to the Hartree method discussed before in section 6.1, adding a current–current integral instead of the density–density integral of equation (105). Since the currents couple to an external vector potential, we include those into the description, just like in current DFT [46]. The vector potentials live in a special homogeneous and divergence-free (Coulomb gauged) Sobolev space that is similar to the previously mentioned homogeneous Sobolev space,

$$Y^* = \{ \mathbf{a} \in L^2(\mathbb{R}^3, \mathbb{R}^3) \mid \nabla \times \mathbf{a} \in L^2(\mathbb{R}^3, \mathbb{R}^3), \nabla \cdot \mathbf{a} = 0 \}. \tag{108}$$

This space must further be taken modulo constant vector fields in order to make $\|\mathbf{a}\|_{Y^*} = \mu_0^{-1/2} \|\nabla \times \mathbf{a}\|_{L^2}$ a valid norm. Here, we built the vacuum permeability μ_0 ($= 4\pi/c^2$ in atomic units) into the norm of the space for vector potentials and it is again a Hilbert space. Since $\mathbf{b} = \nabla \times \mathbf{a}$ describes a magnetic field, $\frac{1}{2} \|\mathbf{a}\|_{Y^*}^2 = \frac{1}{2\mu_0} \|\mathbf{b}\|_{L^2}^2$ exactly corresponds to the magnetostatic field energy. The corresponding space for current densities is then its dual space Y , where the duality map $J: Y \rightarrow Y^*$ then means solving the Maxwell equation

$$-\Delta \mathbf{a} = \nabla \times \mathbf{b} = \mu_0 \mathbf{j}. \tag{109}$$

It was possible to rewrite this as a set of component-wise Poisson equations since in the application of the vector Laplacian $\Delta \mathbf{a} = \nabla(\nabla \cdot \mathbf{a}) - \nabla \times (\nabla \times \mathbf{a})$ the divergence term vanishes for $\mathbf{a} \in Y^*$ due to the chosen Coulomb gauge. The current space Y includes distributional vector fields and the norm for regular distributions $\mathbf{j} \in Y$ features a current–current integral,

$$\frac{1}{2} \|\mathbf{a}\|_{Y^*}^2 = \frac{1}{2} \|\mathbf{j}\|_Y^2 = \frac{\mu_0}{2} \iint \frac{\mathbf{j}(\mathbf{r}) \cdot \mathbf{j}(\mathbf{r}')}{4\pi |\mathbf{r} - \mathbf{r}'|} d\mathbf{r} d\mathbf{r}'. \tag{110}$$

After setting the functional-analytic stage, let us introduce the coupled Maxwell–Schrödinger system. We redefine the kinetic-energy operator so that it includes vector potentials $\mathbf{a} \in Y^*$ through minimal coupling,

$$\hat{T}_{\mathbf{a}} = -\frac{1}{2} \sum_{j=1}^N (-i\nabla_j + \mathbf{a})^2. \quad (111)$$

In contrast to the original publication [83], we do not include spin degrees-of-freedom, since this keeps the notation more concise but does not change the main argument about auto-regularization. We now define a Hamiltonian $\hat{H}_{v,\mathbf{a}}^{\lambda=1}$ like in equation (23) just replacing $\hat{T}$ with $\hat{T}_{\mathbf{a}}$. We further always choose the fully interacting case with $\lambda=1$ here and will thus suppress this parameter in the notation. The ground-state energy for a potential pair $(v, \mathbf{a}) \in X^* \times Y^*$ is then

$$E(v, \mathbf{a}) = \inf_{\Psi} \langle \Psi, \hat{H}_{v,\mathbf{a}} \Psi \rangle, \quad (112)$$

which can equivalently be defined with variation over density matrices. This shows that this functional is again concave in v , but has no special convexity property in $\mathbf{a}$, a fact related to diamagnetism [83]. We define the corresponding constrained-search functional over ensembles [84–86],

$$F(\rho, \mathbf{j}^p) = \inf_{\hat{\Gamma} \mapsto (\rho, \mathbf{j}^p)} \text{Tr} \left(\hat{H}_{0,0} \hat{\Gamma} \right), \quad (113)$$

that is now defined on $X \times Y$, the space of pairs of densities and paramagnetic currents. The notation $\hat{\Gamma} \mapsto (\rho, \mathbf{j}^p)$ means that in the variation all density matrices that map to a given one-particle density ρ and paramagnetic current $\mathbf{j}^p$ are considered. We then have

$$E(v, \mathbf{a}) = \inf_{\substack{\rho \in X \\ \mathbf{j}^p \in Y}} \left\{ F(\rho, \mathbf{j}^p) + \langle v + \frac{1}{2} \mathbf{a}^2, \rho \rangle + \langle \mathbf{a}, \mathbf{j}^p \rangle \right\}. \quad (114)$$

One immediately sees that the density now additionally couples to $\frac{1}{2} \mathbf{a}^2$, which means that the resulting energy functional is also not concave in $\mathbf{a}$. This can easily be repaired by substituting $u = v + \frac{1}{2} \mathbf{a}^2$, which makes the energy functional jointly concave in u and $\mathbf{a}$. But note that this construction requires $\mathbf{a}^2 \in X^*$, the space of potentials, which only holds when the spaces X and Y have a special compatibility property [16].

Without requiring joint concavity, the energy functional for Maxwell–Schrödinger DFT is defined as

$$E_M(v, \mathbf{A}) = \inf_{\alpha \in Y^*} \left\{ E(v, \alpha + \mathbf{A}) + \frac{1}{2} \|\alpha\|_{Y^*}^2 \right\}. \quad (115)$$

Here and in the following, we exactly follow the notation of Tellgren [83] in order to allow for direct comparison. The stationarity condition for this energy functional then gives the Schrödinger equation with Hamiltonian $\hat{H}_{v,\alpha+\mathbf{A}}$ coupled to the Poisson equations for all components of the induced field $\alpha \in Y^*$,

$$-\Delta \alpha = \mu_0 J^{-1}(\alpha) = \mu_0 \mathbf{j}_{\Psi}. \quad (116)$$

Here, $\mathbf{j}_{\Psi}$ is the physical current from the solution of the Schrödinger equation. Next, we define $\bar{E}_M$ by subtracting the field energy $\frac{1}{2} \|\mathbf{A}\|_{Y^*}^2$. Also substituting $\mathbf{a} = \alpha + \mathbf{A}$, we see that we have arrived at a formula in close analogy to the Hartree auto-regularization in equation (107) (line 5), just for the vector potential,

$$\begin{aligned} \bar{E}_M(v, \mathbf{A}) &= E_M(v, \mathbf{A}) - \frac{1}{2} \|\mathbf{A}\|_{Y^*}^2 \\ &= \inf_{\mathbf{a} \in Y^*} \left\{ E(v, \mathbf{a}) + \frac{1}{2} \|\mathbf{a} - \mathbf{A}\|_{Y^*}^2 \right\} - \frac{1}{2} \|\mathbf{A}\|_{Y^*}^2. \end{aligned} \quad (117)$$

This functional is now jointly concave in v and $\mathbf{A}$, yet we do *not* have the same effect of auto-regularization as for E^{Hartree} since $E(v, \mathbf{a})$ is not convex (nor concave) in $\mathbf{a}$. For a regularized functional a further step is necessary.

By Legendre–Fenchel transform in just one or both arguments, one can construct three further functionals,

$$\bar{F}_M(\rho, \mathbf{A}) = \sup_{v \in X^*} \{ \bar{E}_M(v, \mathbf{A}) - \langle v, \rho \rangle \}, \quad (118)$$

$$\bar{e}_M(v, \mathbf{j}^p) = \sup_{\mathbf{A} \in Y^*} \{ \bar{E}_M(v, \mathbf{A}) - \chi \langle \mathbf{A}, \mathbf{j}^p \rangle \}, \quad (119)$$

$$\bar{f}_M(\rho, \mathbf{j}^p) = \sup_{\mathbf{A} \in Y^*} \{ \bar{F}_M(\rho, \mathbf{A}) - \chi \langle \mathbf{A}, \mathbf{j}^p \rangle \}. \quad (120)$$

In this step we introduced a new magnetic coupling constant χ . Now it turns out that the last two functionals, $\bar{e}_M(v, \mathbf{j}^p)$ and $\bar{f}_M(\rho, \mathbf{j}^p)$ can be given in the form of Lasry–Lions regularizations [87], a form of double MY regularization, that has the formula

$$f_{\varepsilon, \delta}(x) = \sup_{z \in X} \inf_{y \in X} \left\{ f(y) + \frac{1}{2\varepsilon} \|x - y\|_X^2 - \frac{1}{2\delta} \|x - z\|_X^2 \right\}. \quad (121)$$

By a theorem due to Attouch and Aze [88] they both become Fréchet differentiable if $\chi > 1$. This differentiability allows for a unique mapping from paramagnetic currents to vector potentials $\mathbf{A}$, either for given v or ρ , and thus gives a form of Hohenberg–Kohn theorem in the context of Maxwell–Schrödinger DFT. The details of this procedure can be found in the original publication by Tellgren [83].

If we remember that the Coulomb potential arises from a QED treatment of the longitudinal component of the electric field in Coulomb gauge [89], then the relation between the Hartree approximation (with only mean-field interaction) and standard DFT is the same as between Maxwell–Schrödinger DFT and a *full* (non-relativistic) QED treatment in terms of a functional theory. The development of such a theory, called QEDFT, is an ongoing program [26, 90–93] with important implications for fields like polaritonic chemistry [94], and it might greatly benefit from MY regularization on its theoretical side.

7. Application of MY-based KS inversion (MYiKS) to periodic settings

Crucial for the density–potential inversion described in section 4 is the computation of the proximal point $\text{prox}_{\varepsilon F^\lambda}(x)$. For the full universal density functional, including interactions, this is a daunting task. Considering the inversion for the non-interacting functional $F^0(\rho)$ is considerably easier and the resulting method is called ‘inverse KS’ since it relates to the KS system. For a periodic setting a KS inversion algorithm based on MY regularization has recently been implemented and analyzed leading to first error bounds [20].

A key advantage of the periodic setting for exploration is that natural choices for the density and potential spaces, X and X^* , are the periodic Sobolev spaces $H_{\text{per}}^s(\Omega)$ with $\Omega \subset \mathbb{R}^3$ denoting the periodically repeated unit cell [17, section 2]. These spaces are Hilbert spaces (thus fulfill assumption 3) and are easily characterized mathematically using Fourier modes $e_{\mathbf{G}} = e^{i\mathbf{G} \cdot \mathbf{r}} / \sqrt{|\Omega|}$, namely by

$$H_{\text{per}}^s(\Omega) = \left\{ u = \sum_{\mathbf{G}} \hat{u}_{\mathbf{G}} e_{\mathbf{G}} \mid \hat{u}_{\mathbf{G}}^* = \hat{u}_{-\mathbf{G}}, \|u\|_{H_{\text{per}}^s} < \infty \right\} \quad (122)$$

with

$$\|u\|_{H_{\text{per}}^s}^2 = \sum_{\mathbf{G}} (1 + |\mathbf{G}|^2)^s |\hat{u}_{\mathbf{G}}|^2, \quad (123)$$

where the sum is over the usual reciprocal lattice vectors $\mathbf{G}$. This characterization is very much in line with the usual discretization schemes based on plane waves, where essentially one just selects a subspace of $H_{\text{per}}^s(\Omega)$ by limiting the largest $|\mathbf{G}|$. Note that the space with $s = +1$ is different to the homogeneous Sobolev space of equation (100), apart from consisting of periodic functions, since the $1 + |\mathbf{G}|^2$ (instead of just $|\mathbf{G}|^2$) in the definition of the norm in equation (123) means the uniform $\mathbf{G} = 0$ component is included as well.

As a result of the above definition of the norm in terms of a Fourier expansion, key quantities of the MY program, such as the duality map J , can be obtained in computable, closed-form expressions. For example, with the particular choice $X = H_{\text{per}}^{-1}(\Omega)$ for densities and $X^* = H_{\text{per}}^1(\Omega)$ for potentials one obtains the duality map in the Fourier representation as [20]

$$J(\rho) = \sum_{\mathbf{G}} \frac{\hat{\rho}_{\mathbf{G}} e_{\mathbf{G}}}{1 + |\mathbf{G}|^2}. \quad (124)$$

Noticing that $1/(\mu + |\mathbf{G}|^2)$ is the Fourier representation of the Yukawa potential $\Phi^\mu(\mathbf{r}) = \exp(-\mu|\mathbf{r}|)/(4\pi|\mathbf{r}|)$ then admits the closed-form formulation as a convolution,

$$J(\rho)(\mathbf{r}) = (\Phi^{\mu=1} * \rho)(\mathbf{r}) = \int_{\mathbb{R}^3} \frac{e^{-|\mathbf{r}-\mathbf{r}'|} \rho(\mathbf{r}')}{4\pi|\mathbf{r}-\mathbf{r}'|} d\mathbf{r}', \quad (125)$$

which is efficiently computable in a plane-wave discretization using equation (124) and fast-Fourier transforms. We remark that the validity of assumption 4 must still explicitly be demonstrated with this choice of spaces, although the periodic setting naturally guarantees the existence of a ground state for every choice of potential $v \in H_{\text{per}}^1(\Omega)$.

Targeting KS inversion in this setting we (as usual) assume that we are given a simultaneously interacting and non-interacting v -representable ground-state density ρ_{gs} , meaning that there exists a KS potential $v_s \in -\partial F^0(\rho_{\text{gs}})$ and a potential $v \in -\partial F^1(\rho_{\text{gs}})$. While $F^0(\rho)$ could be directly subjected to the MY program, results seem to generally improve if a maximal amount of known information about the problem is explicitly employed [2, 73]. We therefore choose a ‘guiding functional’

$$\mathcal{F}(\rho) = F^0(\rho) + E_{\text{H}}(\rho) + \langle v, \rho \rangle, \quad (126)$$

which additionally includes the Hartree mean-field energy and the external potential. Noting that $E_{\text{H}}(\rho)$ in the periodic setting is modified to include a compensating background, that is $E_{\text{H}}(\rho) = \sum_{\mathbf{G} \neq 0} |\hat{\rho}_{\mathbf{G}}|^2 / |\mathbf{G}|^2$, one shows $\mathcal{F}(\rho)$ to remain convex [20]. Lower semicontinuity of $F^0(\rho)$ and thus the entire functional $\mathcal{F}(\rho)$ in the H_{per}^{-1} topology remains to be shown. Similarly, we remark that the ideal choice of guiding functional is currently also an open research question.

Using property (v) of MY regularization in section 2.3 then yields

$$\varepsilon^{-1} J(\rho_{\text{gs}, \varepsilon} - \rho_{\text{gs}}) \in -\partial \mathcal{F}(\rho_{\text{gs}, \varepsilon}), \quad (127)$$

where we introduced the shorthand $\rho_{\text{gs}, \varepsilon} = \text{prox}_{\varepsilon \mathcal{F}}(\rho_{\text{gs}})$. We remark that our assumptions on ρ_{gs} allow the definition of an exchange–correlation potential $v_{\text{xc}} = v_s - v - v_{\text{H}}(\rho_{\text{gs}}) \in -\partial \mathcal{F}(\rho_{\text{gs}})$ that we obtain in the limit $\varepsilon \rightarrow 0$ (just like in section 4.1)

$$v_{\text{xc}} = \lim_{\varepsilon \rightarrow 0} v_{\text{xc}, \varepsilon} \quad \text{with} \quad v_{\text{xc}, \varepsilon} = \varepsilon^{-1} J(\rho_{\text{gs}, \varepsilon} - \rho_{\text{gs}}). \quad (128)$$

To compute this numerically we require two ingredients:

- (i) an approach to compute the proximal points $\rho_{\text{gs}, \varepsilon}$, which we will discuss next, and,
- (ii) a suitable sequence of decreasing values for ε enabling an extrapolation of the $\varepsilon \rightarrow 0$ limit. Despite progress in the development of analytical error bounds for MYiKS, the choice of such an ε -sequence remains empirical to date. An explicit convergence study is thus still required as we will outline below.

One approach to achieve (i) is to follow the fixed-point iteration suggested in equation (68). However, in our previous work [20], we directly solve the minimization problem encoded in the proximal-point computation, i.e. we minimize

$$\mathcal{F}(\rho) + \frac{1}{2\varepsilon} \|\rho - \rho_{\text{gs}}\|_{H_{\text{per}}^{-1}}^2 \quad (129)$$

over all $\rho \in D$, which automatically excludes non-physical densities (e.g. with infinite kinetic energies) from the minimization. Notably, for the numerical approximation of the supremum in equation (37) for $F^0(\rho)$ we employ the usual KS parametrization of the density in terms of orthonormal orbitals. This parametrization reads $\Phi = (\phi_1, \dots, \phi_N)$ with $\langle \phi_i, \phi_j \rangle = \delta_{ij}$ and occupations $\mathbf{f} = (f_1, \dots, f_N)$ and gives

$$\rho_{\Phi}(\mathbf{r}) = \sum_i f_i |\phi_i(\mathbf{r})|^2. \quad (130)$$

We can then write the proximal-point computation as the minimization of

$$\mathcal{E}(\Phi, \mathbf{f}) = \sum_i f_i \|\nabla \phi_i\|_{L^2}^2 + E_{\text{H}}(\rho_{\Phi}) + \langle v, \rho_{\Phi} \rangle + \frac{1}{2\varepsilon} \|\rho_{\Phi} - \rho_{\text{gs}}\|_{H_{\text{per}}^{-1}}^2 \quad (131)$$

with respect to the orbitals ϕ_i and occupations f_i . In these expressions we do not make explicit any details about spin or Brillouin zone sampling, i.e. the choice of the $\mathbf{k}$ -point grid to model periodic supercells [95, chapter 2.8]. In that sense one could think of the index i to run over all occupied orbitals of all spins and all reducible $\mathbf{k}$ -points.

We remark that this energy expression equation (131) is closely related to the usual KS-DFT energy expression, where essentially the exchange–correlation energy term is just replaced by the penalty term of the regularization. Standard direct minimization techniques from KS-DFT are therefore immediately applicable [96–98]. A first implementation of this approach in DFTK [20, 99] is capable of performing proximal-point computations for non-trivial 3D solid-state systems, such as silicon, gallium arsenide or potassium chloride, thus providing a practical MYiKS procedure. Right now the treatment is still restricted to insulating solids, since the algorithm currently makes the assumption that all $f_i = 1$ in equation (130) and respectively only minimizes over the N (spin) orbitals in equation (131). A generalization to $f_i \neq 1$, required for the treatment of metals, is possible taking inspiration from the KS-DFT literature [97, 100–102].

While the current practical MYiKS scheme is still far from optimized in terms of employed numerical algorithms and implementation, we want to provide a high-level idea of their computational cost. Similar to standard KS-DFT, the main computational bottlenecks of MYiKS are the matrix-vector products involving the gradient of $\mathcal{E}$, which is an operator similar to the usual KS Hamiltonian. For plane-wave basis sets a gradient-vector product scales as $O(N_b \log N_b)$, where N_b is the number of plane waves. One iteration step in the minimization of equation (131) typically requires one gradient-vector product. With current algorithms a few tens (for large ε) to a few hundred (for small ε) of minimization steps are required to compute one proximal density. The advantage of the MYiKS approach is therefore that no explicit diagonalization of a KS-like Hamiltonian is required—in contrast to for example the ZMP [2] or the Wu–Yang [70] methods. However, the disadvantage of current MYiKS techniques are (i) the sometimes large number of minimization steps per proximal point computation and (ii) the need to estimate the limit $\varepsilon \rightarrow 0$, i.e. the need to perform multiple proximal point computations for different ε values. We now describe some ideas to overcome these challenges.

Crucial to reduce the number of minimization steps is an appropriate preconditioning of the minimization of $\mathcal{E}$. For the usual KS problems it is primarily the kinetic-energy term with its $|\mathbf{G}|^2$ -dependence that leads to ill-conditioned problems in large plane-wave basis sets. Appropriate preconditioners have been developed to effectively tame this source of ill-conditioning [97, 103] for KS DFT problems. For the MY proximal-point computation we currently rely on the same kinetic-energy preconditioners. For large ε values this is indeed sufficient. However, for small ε the penalty term $\frac{1}{2\varepsilon} \|\rho_\Phi - \rho_{\text{gs}}\|_{H_{\text{per}}^{-1}}^2$ can dominate, such that kinetic-energy preconditioning alone is no longer appropriate. This explains why sometimes our minimization procedure requires hundreds of steps. Refined preconditioning techniques for the MY setting therefore need to be developed. These should be ε -dependent and take both the kinetic-energy and the penalty term into account.

An additional complication in practical inverse KS calculations is that the input density ρ_{gs} itself is usually obtained from an inexact procedure, such as an experiment or a computation from a higher-level theory. As a result, we typically do not have access to the exact ρ_{gs} , but only to a perturbed version $\tilde{\rho}_{\text{gs}} = \rho_{\text{gs}} + \delta\rho$. Computing the proximal point for a given ε with this perturbed ground-state density thus only provides us with an estimate $\tilde{v}_{\text{xc},\varepsilon}$ instead of the correct $v_{\text{xc},\varepsilon}$. Previous analysis [20] lets us deduce the bound

$$\begin{aligned} \|v_{\text{xc}} - \tilde{v}_{\text{xc},\varepsilon}\|_{H_{\text{per}}^1} &\leq \|v_{\text{xc}} - v_{\text{xc},\varepsilon}\|_{H_{\text{per}}^1} + \|v_{\text{xc},\varepsilon} - \tilde{v}_{\text{xc},\varepsilon}\|_{H_{\text{per}}^1} \\ &\leq \|v_{\text{xc}} - v_{\text{xc},\varepsilon}\|_{H_{\text{per}}^1} + \frac{C}{\varepsilon} \|\delta\rho\|_{H_{\text{per}}^{-1}} \end{aligned} \quad (132)$$

for the error between the numerically computed $\tilde{v}_{\text{xc},\varepsilon}$ and the desired true v_{xc} . We notice two error contributions. The last term is due to the inherent error $\delta\rho$ in the available ground-state density data and increases with decreasing ε . We observe (see figure 4) that the potential $\tilde{v}_{\text{xc},\varepsilon}$ initially converges towards the true v_{xc} , but then the distance increases again: there is an optimal ε value which yields the most accurate estimate of the inverted potential. We remark that on top of the effects of a noisy input density outlined in equation (132) and which we studied in [20], there is an additional contribution to the total error that we have not made explicit here. This is the error due to the proximal density being only converged up to a certain finite tolerance in practical iterative schemes for minimizing $\mathcal{E}$. Including this term consistently in the analysis is work in progress.

The bound in equation (132) provides a qualitative explanation of the observed behavior as $\varepsilon \rightarrow 0$. However, it is neither sharp nor does it fully match the practical setting of periodic calculations, such

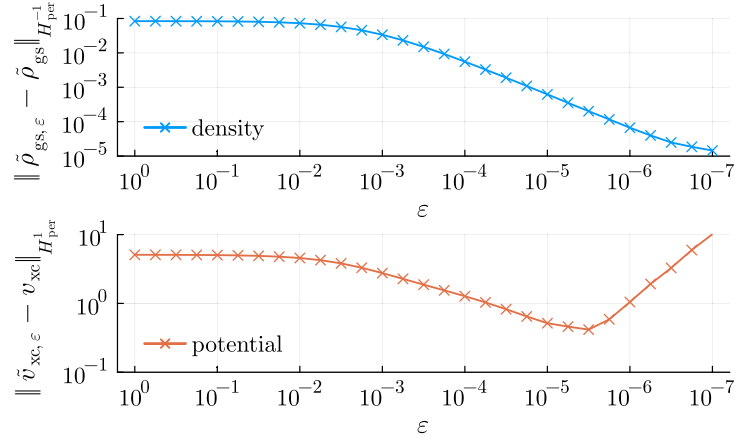


Figure 4. Convergence of $\tilde{\rho}_{\text{gs},\varepsilon}$ and $\tilde{v}_{\text{xc},\varepsilon}$ as a function of ε for the bulk silicon test case and a basis truncation of $E_{\text{cut}} = 30$, leading to an error $\delta\rho = 2.7 \times 10^{-5}$. We refer the reader to the original [20] for details of the computational setup and numerical parameters.

that it is not yet sufficient to suggest an optimal ε value *a priori*, let alone an entire ε sequence for potential extrapolation. For example, while equation (132) lets us observe that ε should not be taken smaller than the error $\delta\rho$, this error is usually not known in practice. Furthermore, the use of a pseudo-potential approximation is required for efficient plane-wave calculations. This implies the addition of a non-local (Kleinman–Bylander) term [104, chapter 7] to the energy functional equation (131), which is outside of the scope of the current mathematical framework [20].

While the current analysis thus already provides first insights into the optimal choice of an ε sequence (ingredient (ii)), it still requires considerable refinement until an *a priori* mathematically guided selection of such a sequence is feasible. To date, an explicit numerical study of the $\varepsilon \rightarrow 0$ limit, e.g. based on a logarithmically spaced ε grid, is thus generally required to empirically identify the optimal ε . However, the progress towards first analytical error bounds such as equation (132) as well as the rich literature on *a priori* and *a posteriori* error analysis for the closely related problem of periodic KS-DFT [17, 105–111] make it hopeful that a more refined and comprehensive analysis is feasible. A tantalizing outlook would be sharper error bounds and thus a better estimation of an optimal ε , or the combination of convergence properties of $\tilde{v}_{\text{xc},\varepsilon}$ with extrapolation techniques to accelerate convergence.

Beyond refinements of the inversion algorithm itself, a number of possible alternative choices in the mathematical setup have neither been fully explored. This includes a different choice than $H_{\text{per}}^{-1}(\Omega)$ for the density space X and $H_{\text{per}}^1(\Omega)$ for the potential space X^* . For example, the inverse choice ($H_{\text{per}}^{-1}(\Omega)$ for the potential and $H_{\text{per}}^1(\Omega)$ for the density) would in principle work as well and exactly fits to recent results about v -representability on periodic one-dimensional domains [18, 19]. Yet, it has to be noted that this will break the self-adjointness requirement from assumption 4 in more than one dimension [112]. We believe that systematically exploring the question of function spaces is a promising endeavor to precisely target the inversion procedure towards reproducing the physically relevant features of the potential: as visible from equation (131), the function space choice is directly reflected in the norms employed during the proximal point computation. From a physical point of view, the regions close to the nucleus are generally less relevant, such that an ideal function space X should feature a norm that puts less weight on the core and larger weight on the valence regions.

In closing, we remark that the exceptionally close agreement between the plane-wave discretization basis and the Fourier characterization of the periodic Sobolev spaces make the periodic setting ideally suited for further joint numerical and analytical study of MY-regularized DFT. In combination with a flexible tool such as DFTK [99], mathematical ideas can be seamlessly implemented and tested both on toy problems as well as realistic materials systems. Considering specifically the similarity of MYIKS to other inversion approaches (such as the ZMP method), see section 4, as well as the critical lack of mathematical analysis for practical inverse KS, this provides an accessible framework to close this important gap.

8. Outlook

This work aimed at showing how, since its introduction into the field of DFT, the technique of MY regularization grew ever more useful and versatile. At this point, it is interesting to ask, which are the most promising routes to follow and what more can be expected. Here, we collect a number of ideas and possible research directions that we deem especially interesting and worthwhile to investigate. This will at the same time serve as a summary of the presented concepts relating to MY regularization applied to DFT.

- In section 6 it was demonstrated how in certain mean-field approximations a choice of spaces for densities and potentials that is adapted such that $\|\cdot\|_{X^*}^2$ yields the field energy automatically leads to MY regularization. A very similar choice of spaces was made in the ZMP method for density–potential inversion (section 4) and in the numerical discussion of the periodic setting (section 7). In general, we believe that fitting the mathematical setting to physical principles, while taking all the functional-analytic considerations of section 2.4 into account, is extremely useful as a guiding principle. The original Lieb setting [3] with $X = L^1 \cap L^3$ in some sense also had this in mind via the estimate of the L^3 -norm of the density by the kinetic energy, but this is just comparative, while an exact relation to the field energy is possible. Equation (102) demonstrated how this leads to an equivalence between the duality map and the electrostatic Poisson equation. It is thus possible to *encode important physical relations fully into the geometry* of the density and potential spaces. (Note that such an encoding of physical laws into geometry is also present quite generally in gauge theories through the introduction of fibre bundles.) This basic idea will be pursued further for a new formulation of QEDFT in a forthcoming publication, where the adaption of the norm to the field energy means that all of Maxwell’s equations will be intimately linked to the theory. This then gives rise to the exact macroscopic Maxwell formulation of non-relativistic QED and has important consequences for renormalization.
- The spaces of section 6 just mentioned also happen to be Hilbert spaces, thus fulfilling our assumption 3. This means that additionally to the Banach-space structure several extremely beneficial properties hold true, like the linearity of the duality map and firm nonexpansiveness of the proximal map. Up to this point, this was not used to full capability. It is for example possible that the Hilbert-space setting can help to generalize the proof of convergence of the regularized KS scheme from section 5.3 to infinite dimensions.
- A generalization in the opposite direction is also possible. This means considering Banach spaces and adapting the MY regularization itself to what is called the ‘modulus of convexity’ of the respective space. This is possible by replacing the $(\cdot)^2$ in equation (15) by $(\cdot)^p$, $p > 1$, or right away by a Young function $\phi(\cdot)$. One then also needs to switch to a corresponding generalized duality map according to equation (6). While this direction was explored mathematically [30, 57, 113], it was not really implemented into the study of DFT yet.
- The original motivation for devising the regularized KS scheme was to be able to prove its convergence. Yet, clearly, a guaranteed convergence is also beneficial for practical applications, so one could try and develop functional approximations $v_{xc,\varepsilon}(\rho)$ that are already adapted to the regularized KS iteration of section 5.1. A small step in this direction was presented in section 5.2 that demonstrated a way how to transform available exchange–correlation potentials such that they can be applied in the regularized KS scheme. But this must be seen just as a first approach to the question. Instead, a construction of such approximations from first principles could also be tried. On the other hand, the practical implications of switching to the regularized form of the KS method have not really been explored yet.
- Speaking of convergence, to date we do not have a convergence proof for the proximal-point iteration (section 4.1) that is also the basis for the inverse KS method applied to periodic settings (section 7). We can maybe expect it in the form of a fixed-point proof that is typical for Banach-space structures. We also note at this point that the presented method can easily be adapted to other DFT settings, like spin DFT, current DFT, or QEDFT.
- As mentioned, the homogeneous Sobolev spaces introduced in section 6 and their duals might have a central role in future reformulations of (regularized) DFT, but they are not that well studied in

functional analysis, especially not on the dual side. What is known is that they are Hilbert spaces and thus assumption 3 is secured. We noted in section 3.3 that they are already used in some practical schemes to enhance the KS method. Further functional theoretic insights on these spaces would thus be a valuable contribution to ongoing work. In the course of such study, the important questions of self-adjointness and boundedness below of the Hamiltonian and, in particular, the lower semicontinuity of the ensemble constrained-search density functional must also be clarified in order to have the correct setting for MY-regularized DFT according to assumption 4. In section 2.4 it was noted that possible strategies to have a lower semicontinuous $F_{\text{DM}}^\lambda(\rho)$ consist in choosing a bounded domain Ω (where it is possible to add periodic boundary condition) or to include an offset potential like $v_{\text{off}}(\mathbf{r}) = |\mathbf{r}|^2$. Both options are currently investigated.

- The whole regularized KS framework in section 5.1 used the difference between the interacting functional F_ε^1 and the non-interacting F_ε^0 , as typically done in KS-DFT. Yet, it was noted in section 7 that switching to a different guiding functional that already includes *a priori* information might be beneficial at least numerically. In a way we did this already in section 2.4 through the inclusion of the offset potential v_{off} that would include information about the external potential. But other, more complicated terms, like the Hartree mean-field energy $E_{\text{H}}(\rho)$ could be included as well before regularization, thus choosing $(F^0 + E_{\text{H}})_\varepsilon$ as a regularized guiding functional. One only has to make sure that the functional is still convex and lower semicontinuous before regularization, in order to stay within assumption 4. Further studies about these questions are already underway.
- In the periodic setting, the computation of proximal points, respectively the inversion of KS densities based on MY techniques has recently become feasible for practical systems, see section 7. However, the currently considered algorithm still suffers from a number of restrictions: (i) it is only applicable to insulating systems, (ii) proper preconditioners have not yet been developed, and (iii) the choice of regularization parameter ε requires careful manual tuning or an explicit convergence study. Overcoming these restrictions is crucial to broaden the scope of MY-based numerical experiments for mathematical research on DFT as well as for practical KS inversion computations.
- While recasting KS inversion via MY regularization yields an *a priori* orbital-free density-functional formulation, see section 4, practical implementations of MYiKS still rely on an explicit orbital parametrization to solve the proximal-point optimization problem equation (131). The same is true for many standard KS inversion algorithms [2, 70, 73]. Consequently, it would be advantageous to develop a mathematical framework in which the regularization is expressed directly in terms of the orbitals to provide a closer match between MY-based theory and practical inversion schemes. We remark that such an orbital-based formulation is also more naturally suited to provide a fully consistent treatment of the nonlocal term present in most pseudopotential approximations (see section 7).
- As previously discussed in section 7, the close match between the mathematical framework and the numerical implementation in the periodic setting is remarkable. This enables joint analytical and numerical investigations, where mathematical developments can be guided and tested using numerical experiments, and suggested theoretical relationships can be directly exploited in the development of inversion algorithms with improved efficiency and reliability. Considering the many possibilities for formulating MY-regularized DFT (e.g. the considerable flexibility in selecting appropriate function spaces), we further expect numerical investigations to be decisive for suggesting which theoretical extensions are worthwhile to pursue. Moreover, we remark that since the MYiKS framework can be viewed as a unified framework encompassing a wide range of inversion schemes, see section 4, we expect the resulting mathematical and algorithmic advances from such studies to be broadly applicable and to directly contribute to the development of reliable and computationally efficient inversion schemes.

Acknowledgments

The authors thank Erik I Tellgren, Mihály A Csirik, Thiago Carvalho Corso, and Michael Ruggenthaler for valuable discussions. MP acknowledges support from the German Research Foundation (Grant SCHI 1476/1-1). MFH acknowledges support by the Swiss National Science Foundation (SNSF, Grant Nos. 221186 and 10002757) as well as the NCCR MARVEL, a National Centre of Competence in

Research, funded by the SNSF (Grant No. 205602). AL and MP have received funding from the ERC-2021-STG under grant agreement No. 101041487 REGAL. AL and TH were funded by the Research Council of Norway through CoE Hylleraas Centre for Quantum Molecular Sciences Grant No. 262695.

Data availability statement

No new data were created or analysed in this study.

Author contributions

Markus Penz  [0000-0001-8363-1980](#)

Conceptualization (equal), Formal analysis (equal), Investigation (equal), Methodology (equal), Validation (equal), Writing – original draft (equal), Writing – review & editing (equal)

Michael F Herbst  [0000-0003-0378-7921](#)

Formal analysis (equal), Investigation (equal), Methodology (equal), Validation (equal), Writing – original draft (equal), Writing – review & editing (equal)

Trygve Helgaker  [0000-0002-5032-8392](#)

Formal analysis (equal), Investigation (equal), Methodology (equal), Validation (equal), Writing – original draft (equal), Writing – review & editing (equal)

Andre Laestadius  [0000-0001-7391-0396](#)

Conceptualization (equal), Formal analysis (equal), Funding acquisition (equal), Investigation (equal), Methodology (equal), Validation (equal), Writing – original draft (equal), Writing – review & editing (equal)

References

- [1] Kvaal S, Ekström U, Teale A M and Helgaker T 2014 Differentiable but exact formulation of density-functional theory *J. Chem. Phys.* **140** 18A518
- [2] Zhao Q, Morrison R C and Parr R G 1994 From electron densities to Kohn–Sham kinetic energies, orbital energies, exchange–correlation potentials and exchange–correlation energies *Phys. Rev. A* **50** 2138–42
- [3] Lieb E H 1983 Density functionals for Coulomb-systems *Int. J. Quantum Chem.* **24** 243–77
- [4] Penz M, Csirik M A and Laestadius A 2023 Density-potential inversion from Moreau–Yosida regularization *Electron. Struct.* **5** 014009
- [5] Barbu V and Precupanu T 2012 *Convexity and Optimization in Banach Spaces* 4th (Springer)
- [6] Bauschke H H and Combettes P L 2011 *Convex Analysis and Monotone Operator Theory in Hilbert Spaces* (Springer)
- [7] Zalinescu C 2002 *Convex Analysis in General Vector Spaces* (World Scientific)
- [8] Rockafellar R T and Wets R J-B 2009 *Variational Analysis* vol 317 (Springer)
- [9] Chidume C 2009 *Geometric Properties of Banach Spaces and Nonlinear Iterations* vol 1965 (Springer)
- [10] Blanchard P and Brüning E 2015 *Mathematical Methods in Physics: Distributions, Hilbert Space Operators, Variational Methods and Applications in Quantum Physics* 2nd edn (Birkhäuser)
- [11] Adams R A and Fournier J J 2003 *Sobolev Spaces* (Elsevier)
- [12] van Tiel J 1984 *Convex Analysis – An Introductory Text* (Wiley)
- [13] Banach S 1922 Sur les opérations dans les ensembles abstraits et leur application aux équations intégrales *Fundam. Math.* **3** 133–81
- [14] Eschrig H 2010 $T > 0$ ensemble-state density functional theory via Legendre transform *Phys. Rev. B* **82** 205120
- [15] Laestadius A, Penz M, Tellgren E I, Ruggenthaler M, Kvaal S and Helgaker T 2018 Generalized Kohn–Sham iteration on Banach spaces *J. Chem. Phys.* **149** 164103
- [16] Laestadius A, Tellgren E I, Penz M, Ruggenthaler M, Kvaal S and Helgaker T 2019 Kohn–Sham theory with paramagnetic currents: compatibility and functional differentiability *J. Chem. Theory Comput.* **15** 4003–20
- [17] Cancès E, Chakir R and Maday Y 2012 Numerical analysis of the planewave discretization of some orbital-free and Kohn–Sham models *ESAIM: Math. Model. Numer. Anal.* **46** 341–88
- [18] Sutter S M, Penz M, Ruggenthaler M, Leeuwen R v and Giesbertz K J H 2024 Solution of the v-representability problem on a one-dimensional torus *J. Phys. A: Math. Theor.* **57** 475202
- [19] Carvalho Corso T 2025 v-representability and Hohenberg–Kohn theorem for non-interacting Schrödinger operators with distributional potentials on the one-dimensional torus *J. Phys. A: Math. Theor.* **58** 125203
- [20] Herbst M F, Bakkestuen V H and Laestadius A 2025 Kohn–Sham inversion with mathematical guarantees *Phys. Rev. B* **111** 205143
- [21] Bohle O M, Lotfigolian M, Laestadius A and Tellgren E I 2026 Regularized density-potential inversion for periodic systems: application to exact exchange in one dimension *J. Chem. Phys.* **164** 064104
- [22] Kvaal S 2022 Moreau–Yosida regularization in DFT *Density Functional Theory: Modeling, Mathematical Analysis, Computational Methods and Applications* (Springer) pp 267–306

- [23] Böhm A and Dollard J D 1978 *The Rigged Hilbert Space and Quantum Mechanics: Lectures in Mathematical Physics at the University of Texas at Austin* (Springer)
- [24] Madrid R d l 2005 The role of the rigged Hilbert space in quantum mechanics *Euro. J. Phys.* **26** 287–312
- [25] Schilling C and Schilling R 2019 Diverging exchange force and form of the exact density matrix functional *Phys. Rev. Lett.* **122** 013001
- [26] Bakkestuen V H, Falmår V, Lotfigolian M, Penz M, Ruggenthaler M and Laestadius A 2025 Quantum-electrodynamical density-functional theory exemplified by the quantum Rabi model *J. Phys. Chem. A* **129** 2337–60
- [27] Moreau J-J 1965 Proximité et dualité dans un espace hilbertien *Bull. Soc. Math. France* **93** 273–99
- [28] Yosida K 1948 On the differentiability and the representation of one-parameter semi-group of linear operators *J. Math. Soc. Japan* **1** 15–21
- [29] Kohsaka F and Takahashi W 2008 Existence and approximation of fixed points of firmly nonexpansive-type mappings in Banach spaces *SIAM J. Optim.* **19** 824–35
- [30] Bacák M and Kohlenbach U 2018 On proximal mappings with Young functions in uniformly convex Banach spaces *J. Convex Anal.* **25** 1291–318
- [31] Reed M and Simon B 1975 *Methods of Modern Mathematical Physics. II: Fourier Analysis, Self-Adjointness* (Academic)
- [32] Hohenberg P and Kohn W 1964 Inhomogeneous electron gas *Phys. Rev.* **136** B864–71
- [33] Percus J 1978 The role of model systems in the few-body reduction of the N-fermion problem *Int. J. Quantum Chem.* **13** 89–124
- [34] Levy M 1979 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** 6062–5
- [35] Levy M 1982 Electron densities in search of hamiltonians *Phys. Rev. A* **26** 1200
- [36] Penz M and van Leeuwen R 2021 Density-functional theory on graphs *J. Chem. Phys.* **155** 244111
- [37] Englisch H and Englisch R 1983 Hohenberg–Kohn theorem and non- V -representable densities *Physica A* **121** 253–68
- [38] van Leeuwen R 2003 Density functional approach to the many-body problem: Key concepts and exact functionals *Advances in Quantum Chemistry* vol 43 (Elsevier) pp 25–94
- [39] Chayes J, Chayes L and Ruskai M B 1985 Density functional approach to quantum lattice systems *J. Stat. Phys.* **38** 497–518
- [40] Penz M and van Leeuwen R 2023 Geometry of degeneracy in potential and density space *Quantum* **7** 918
- [41] Lammert P E 2007 Differentiability of Lieb functional in electronic density functional theory *Int. J. Quantum Chem.* **107** 1943–53
- [42] Garrigue L 2018 Unique continuation for many-body Schrödinger operators and the Hohenberg–Kohn theorem *Math. Phys. Anal. Geom.* **21** 27
- [43] Garrigue L 2019 Hohenberg–Kohn theorems for interactions, spin and temperature *J. Stat. Phys.* **177** 415–37
- [44] Jerison D and Kenig C E 1985 Unique continuation and absence of positive eigenvalues for Schrödinger operators *Ann. Math.* **121** 463
- [45] Sogge C D 1990 Strong uniqueness theorems for second order elliptic differential equations *Amer. J. Math.* **112** 943
- [46] Penz M, Tellgren E I, Csirik M A, Ruggenthaler M and Laestadius A 2023 The structure of density-potential mapping. part II: including magnetic fields *ACS Phys. Chem. Au* **3** 492–511
- [47] Reed M and Simon B 1978 *Methods of Modern Mathematical Physics. IV: Analysis of Operators* (Academic)
- [48] Sutcliffe B T and Woolley R G 2012 On the quantum theory of molecules *J. Chem. Phys.* **137** 22A544
- [49] Schuch N and Verstraete F 2009 Computational complexity of interacting electrons and fundamental limitations of density functional theory *Nat. Phys.* **5** 732–5
- [50] Toulouse J 2022 Review of approximations for the Exchange–Correlation Energy in DFT *Density Functional Theory: Modeling, Mathematical Analysis, Computational Methods and Applications* (Springer) pp 1–90
- [51] Kohn W and Sham L J 1965 Self-consistent equations including exchange and correlation effects *Phys. Rev.* **140** A1133
- [52] Trushin E, Erhard J and Görling A 2024 Violations of the v -representability condition underlying Kohn–Sham density-functional theory *Phys. Rev. A* **110** L020802
- [53] Karlström G 1979 Dynamical damping based on energy minimization for use in ab initio molecular orbital SCF calculations *Chem. Phys. Lett.* **67** 348–50
- [54] Cancès E and Le Bris C 2000 Can we outperform the diis approach for electronic structure calculations? *Int. J. Quantum Chem.* **79** 82–90
- [55] Cancès E 2001 Self-consistent field algorithms for Kohn–Sham models with fractional occupation numbers *J. Chem. Phys.* **114** 10616–22
- [56] Garrigue L 2021 Some properties of the potential-to-ground state map in quantum mechanics *Commun. Math. Phys.* **386** 1803–44
- [57] Penz M and Laestadius A 2025 Adaptation of Moreau–Yosida regularization to the modulus of convexity *J. Math. Anal. Appl.* **554** 129956
- [58] Penz M and van Leeuwen R 2024 Geometrical perspective on spin–lattice density-functional theory *J. Chem. Phys.* **161** 150901
- [59] Penz M, Tellgren E I, Csirik M A, Ruggenthaler M and Laestadius A 2023 The structure of density-potential mapping. part I: standard density-functional theory *ACS Phys. Chem. Au* **3** 334–47
- [60] Görling A 1999 New KS method for molecules based on an exchange charge density generating the exact local KS exchange potential *Phys. Rev. Lett.* **83** 5459–62
- [61] Heaton-Burgess T, Bulat F A and Yang W 2007 Optimized effective potentials in finite basis sets *Phys. Rev. Lett.* **98** 256401
- [62] Callow T J, Pearce B J, Pitts T, Lathiotakis N N, Hodgson M J P and Gidopoulos N I 2020 Improving the exchange and correlation potential in density-functional approximations through constraints *Faraday Discuss.* **224** 126–44
- [63] Callow T J, Lathiotakis N N and Gidopoulos N I 2020 Density-inversion method for the Kohn–Sham potential: role of the screening density *J. Chem. Phys.* **152** 164114
- [64] Colonna F and Savin A 1999 Correlation energies for some two- and four-electron systems along the adiabatic connection in density functional theory *J. Chem. Phys.* **110** 2828–28235
- [65] Teale A M, Coriani S and Helgaker T 2009 The calculation of adiabatic-connection curves from full configuration-interaction densities: two-electron systems *J. Chem. Phys.* **130** 104111
- [66] Teale A M, Coriani S and Helgaker T 2010 Accurate calculation and modeling of the adiabatic connection in density functional theory *J. Chem. Phys.* **132** 164115
- [67] Teale A M, Coriani S and Helgaker T 2010 Range-dependent adiabatic connections *J. Chem. Phys.* **133** 164112

- [68] Helgaker T and Teale A 2022 Lieb variation principle in density-functional theory *The Physics and Mathematics of Elliott Lieb*, ed R L Frank, A Laptev, M Lewin and R Seiringer (EMS Press) pp 527–59
- [69] Garrigue L 2022 Building Kohn–Sham potentials for ground and excited states *Arch. Ration. Mech. Anal.* **245** 949–1003
- [70] Wu Q and Yang W 2003 A direct optimization method for calculating density functionals and exchange–correlation potentials from electron densities *J. Chem. Phys.* **118** 2498–509
- [71] van Leeuwen R and Baerends E J 1994 Exchange–correlation potential with correct asymptotic behavior *Phys. Rev. A* **49** 2421
- [72] Kumar A and Harbola M K 2020 A general penalty method for density to potential inversion *Int. J. Quantum Chem.* **120** e26400
- [73] Shi Y and Wasserman A 2021 Inverse Kohn–Sham density functional theory: progress and challenges *J. Phys. Chem. Lett.* **12** 5308–18
- [74] Fetter A L and Walecka J D 2012 *Quantum Theory of Many-Particle Systems* (Courier Corporation)
- [75] Wagner L O, Stoudenmire E M, Burke K and White S R 2013 Guaranteed convergence of the Kohn–Sham equations *Phys. Rev. Lett.* **111** 093003
- [76] Wagner L O, Baker T E, Stoudenmire E M, Burke K and White S R 2014 Kohn–Sham calculations with the exact functional *Phys. Rev. B* **90** 045109
- [77] Penz M, Laestadius A, Tellgren E I and Ruggenthaler M 2019 Guaranteed convergence of a regularized Kohn–Sham iteration in finite dimensions *Phys. Rev. Lett.* **123** 037401
- [78] Penz M, Laestadius A, Tellgren E I, Ruggenthaler M and Lammert P E 2020 Erratum: guaranteed convergence of a regularized Kohn–Sham iteration in finite dimensions *Phys. Rev. Lett.* **125** 249902
- [79] Parikh N and Boyd S 2014 Proximal algorithms *Found. Trends® Optim.* **1** 127–239
- [80] Maz'ya V 2011 *Sobolev Spaces With Applications to Elliptic Partial Differential Equations* 2nd edn (Springer)
- [81] Galdi G 2011 *An Introduction to the Mathematical Theory of the Navier-Stokes Equations: Steady-State Problems* 2nd (Springer)
- [82] Ortner C and Süli E 2012 A note on linear elliptic systems on $\mathbb{R}^d$ (arXiv:1202.3970[math.AP])
- [83] Tellgren E I 2018 Density-functional theory for internal magnetic fields *Phys. Rev. A* **97** 012504
- [84] Vignale G and Rasolt M 1987 Density-functional theory in strong magnetic fields *Phys. Rev. Lett.* **59** 2360–3
- [85] Vignale G and Rasolt M 1988 Current- and spin-density-functional theory for inhomogeneous electronic systems in strong magnetic fields *Phys. Rev. B* **37** 10685–96
- [86] Laestadius A 2014 Density functionals in the presence of magnetic field *Int. J. Quantum Chem.* **114** 1445–56
- [87] Lasry J M and Lions P L 1986 A remark on regularization in Hilbert spaces *Isr. J. Math.* **55** 257–66
- [88] Attouch H and Aze D 1993 Approximation and regularization of arbitrary functions in Hilbert spaces by the Lasry–Lions method *Ann. Inst. Henri Poincaré C*, **10** 289–312
- [89] Greiner W and Reinhardt J 2013 *Field Quantization* (Springer)
- [90] Ruggenthaler M, Flick J, Pellegrini C, Appel H, Tokatly I V and Rubio A 2014 Quantum-electrodynamical density-functional theory: Bridging quantum optics and electronic-structure theory *Phys. Rev. A* **90** 012508
- [91] Ruggenthaler M 2015 Ground-state quantum-electrodynamical density-functional theory (arXiv:1509.01417 [quant-ph])
- [92] Lu I-T, Ruggenthaler M, Tancogne-Dejean N, Latini S, Penz M and Rubio A 2024 Electron-photon exchange-correlation approximation for quantum-electrodynamical density-functional theory *Phys. Rev. A* **109** 052823
- [93] Bakkestuen V H, Csirik M A, Laestadius A and Penz M 2025 Density-functional theory for the Dicke Hamiltonian *J. Stat. Phys.* **192** 61
- [94] Ruggenthaler M, Sidler D and Rubio A 2023 Understanding polaritonic chemistry from ab initio quantum electrodynamics *Chem. Rev.* **123** 11191–229
- [95] Lin L and Lu J 2019 *A Mathematical Introduction to Electronic Structure Theory* (Society for Industrial and Applied Mathematics) (<https://doi.org/10.1137/1.9781611975802>)
- [96] Payne M C, Teter M P, Allan D C, Arias T A and Joannopoulos J D 1992 Iterative minimization techniques for ab initio total-energy calculations: molecular dynamics and conjugate gradients *Rev. Mod. Phys.* **64** 1045–97
- [97] Kresse G and Furthmüller J 1996 Efficient iterative schemes for ab initio total-energy calculations using a plane-wave basis set *Phys. Rev. B* **54** 11169–86
- [98] Dai X, Liu Z, Zhang L and Zhou A 2017 A conjugate gradient method for electronic structure calculations *SIAM J. Sci. Comput* **39** A2702–40
- [99] Herbst M F, Levitt A and Cancès E 2021 DFTK: a Julia approach for simulating electrons in solids *Proc. JuliaCon Conf.* **3** 69
- [100] Marzari N, Vanderbilt D and Payne M C 1997 Ensemble density-functional theory for ab initio molecular dynamics of metals and finite-temperature insulators *Phys. Rev. Lett.* **79** 1337–40
- [101] Freysoldt C, Boeck S and Neugebauer J 2009 Direct minimization technique for metals in density functional theory *Phys. Rev. B* **79** 241103
- [102] Gonze X, Rostami S and Tantardini C 2024 Variational density functional perturbation theory for metals *Phys. Rev. B* **109** 014317
- [103] Teter M P, Payne M C and Allan D C 1989 Solution of Schrödinger's equation for large systems *Phys. Rev. B* **40** 12255–63
- [104] Cancès E and Friesecke G 2023 *Density Functional Theory* (Springer) (<https://doi.org/10.1007/978-3-031-22340-2>)
- [105] Cancès E, Chakir R and Maday Y 2010 Numerical analysis of nonlinear eigenvalue problems *J. Sci. Comput.* **45** 90–117
- [106] Cancès E, Chakir R, He L and Maday Y 2018 Two-grid methods for a class of nonlinear elliptic eigenvalue problems *IMA J. Numer. Anal.* **38** 605–45
- [107] Herbst M F, Levitt A and Cancès E 2020 A posteriori error estimation for the non-self-consistent Kohn–Sham equations *Faraday Discuss.* **224** 227
- [108] Cancès E, Dusson G, Maday Y, Stamm B and Vohralík M 2021 Post-processing of the planewave approximation of Schrödinger equations. Part I: linear operators *IMA J. Numer. Anal.* **41** 2423–55
- [109] Dusson G and Maday Y 2023 An overview of a posteriori error estimation and post-processing methods for nonlinear eigenvalue problems *J. Comput. Phys.* **491** 112352
- [110] Bordignon A, Dusson G, Cancès E, Kemlin G, Reyes R A L and Stamm B 2024 Fully guaranteed and computable error bounds on the energy for periodic Kohn–Sham equations with convex density functionals (arXiv:2409.11769 [math.NA])
- [111] Cancès E, Kemlin G and Levitt A 2024 A priori error analysis of linear and nonlinear periodic schrödinger equations with analytic potentials *J. Sci. Comput.* **98** 25
- [112] Herczyński J 1989 On Schrödinger operators with distributional potentials *J. Operator Theory* **21** 273–95
- [113] Bacho A 2023 A generalization of the Moreau–Yosida regularization *J. Math. Anal. Appl.* **524** 127139