

Differentiability in density-functional theory

Ingvar Lindgren and Sten Salomonson

Department of Physics, Chalmers University of Technology
and the Göteborg University, Göteborg, Sweden

August 31, 2003

Abstract

The differentiability of different functionals used in density-functional theory (DFT) is investigated, and it is shown that the so-called Levy-Lieb functional $F_{LL}[\rho]$ and the Lieb functional $F_L[\rho]$ are Gâteaux differentiable at pure-state v -representable and ensemble v -representable densities, respectively. The conditions for the Fréchet differentiability of these functionals is also discussed. The Gâteaux differentiability of the Lieb functional has been demonstrated by Englisch and Englisch (Phys. Stat. Solidi **123**, 711 and **124**, 373 (1984)), but the differentiability of the Levy-Lieb functional has not been shown before.

Contents

1	Introduction	2
2	General concepts	3
2.1	The space of wavefunctions and densities	3
2.2	The Hohenberg-Kohn theorems	4
2.3	The minimization process	5
2.4	Comments on the treatment of Nesbet	6
3	Functional derivatives	8
3.1	Gâteaux derivative	8
3.2	Fréchet derivative	9
3.3	Comparison between the Fréchet and Gâteaux differentiabilitys	10

4	Application to density functionals	11
4.1	The Levy-Lieb functional	11
4.2	The Lieb functional	16
5	The Kohn-Sham model	18
6	Conclusions	21
A	Basic topological concepts	21
A.1	Notations	22
A.2	Vector spaces	22

1 Introduction

The differentiability of density functionals is of fundamental importance in *Density-Functional Theory* (DFT) and forms the basis for models of Kohn-Sham type [HK64,KS65][1, 2]. In standard DFT an energy functional, $E[\rho]$, is minimized under the constraint that the density $\rho(\mathbf{r})$ is normalized to the number of electrons, which requires that the functional is differentiable with respect to the density at the minimum. In the Kohn-Sham model an interacting system is simulated by a system of noninteracting electrons moving in a *local* potential, which requires that the derivative of the functionals involved can be represented by a local function.

The locality of density-functional derivatives has been a controversial issue for some time. It was demonstrated almost 20 years ago by Englisch and Englisch [EE84,EE84a][3, 4], based upon works of Lieb [L183][5], that a very large class of functionals is (Gâteaux) differentiable with the derivative in the form of a local function. This result has been challenged by Nesbet [Ne98,NC99,Ne01,Ne03][6, 7, 8, 9], who in a series of papers claims that the energy-functional derivative cannot be represented by a local potential even for a noninteracting system with more than two electrons. In a recent Comment to the Physical Review [LS03a][10] we claim that the results of Nesbet are incorrect. There we have demonstrated in a simple way that such a derivative does exist for noninteracting systems in general – in accordance with the more general results of Englisch and Englisch – and, in addition, pointed out where we believe the mistake of Nesbet has been made. In the present work we shall extend our treatment in the Comment to systems of *interacting electrons* and with a ground state that can also be *degenerate*.

Englich and Englisch [EE84,EE84a][3, 4] have demonstrated the Gâteaux differentiability of the so-called Lieb functional, $F_L[\rho]$, utilizing the convexity of the functional. The question of the differentiability of the so-called Levy-Lieb functional, $F_{LL}[\rho]$, which is not necessarily convex, is left open in their work.

The general situation concerning the differentiability of density functionals has been reviewed by van Leeuwen in this volume [11], confirming the results of Englisch and Englisch. In the present work we shall show that the $F_L[\rho]$ and $F_{LL}[\rho]$ functionals are both Gâteaux differentiable at pure-state- v - (PS- v) and ensemble- v - (E- v) representable densities, respectively. We shall also discuss the conditions for the Fréchet differentiability of these functionals and show that the possible difference between the Gâteaux and Fréchet differentiability is quite subtle.

2 General concepts

sec:General

2.1 The space of wavefunctions and densities

We consider a system of N interacting electrons with the Hamiltonian (in Hartree atomic units, i.e., $m = e = \hbar = 4\pi\epsilon_0 = 1$)

$$\hat{H}_v = \hat{T} + \hat{W} + \hat{V} = \sum_{i=1}^N -\frac{1}{2} \nabla_i^2 + \sum_{i<j=1}^N \frac{1}{r_{ij}} + \sum_{i=1}^N v(\mathbf{r}_i). \quad (1) \quad \text{Ham}$$

Here, $\hat{T} = \sum_i -\frac{1}{2} \nabla_i^2$ represents the kinetic energy, $\hat{W} = \sum_{i<j} 1/r_{ij}$ the interaction between the electrons and $\hat{V} = \sum_i v(\mathbf{r}_i)$ the external (usually nuclear) field. The wavefunctions are assumed to be normalizable (but not necessarily normalized)¹

$$\int d\mathbf{r}_1 \int d\mathbf{r}_2 \cdots \int d\mathbf{r}_N |\Psi(\mathbf{r}_1, \mathbf{r}_2 \cdots \mathbf{r}_N)|^2 < \infty. \quad (2) \quad \text{Norm}$$

The electron density is defined as the diagonal of the first-order density matrix,

$$\rho(\mathbf{r}) = N \int d\mathbf{r}_2 \int d\mathbf{r}_3 \cdots \int d\mathbf{r}_N |\Psi(\mathbf{r}, \mathbf{r}_2 \cdots \mathbf{r}_N)|^2, \quad (3) \quad \text{Density}$$

and we shall use this definition also for wavefunctions that are not normalized. Integration of the density then leads to

$$\int d\mathbf{r} \rho(\mathbf{r}) = N \langle \Psi | \Psi \rangle, \quad (4) \quad \text{Dens}$$

which shows that normalizing the density to N , automatically implies that the wavefunction is normalized to unity. Regardless of normalization, we have the relation

$$\langle \Psi | \hat{V} | \Psi \rangle = \int d\mathbf{r} \rho(\mathbf{r}) v(\mathbf{r}). \quad (5) \quad \text{Vint}$$

¹The integrals are generally assumed to be of Lebesgue type and normally include a sum over spin coordinates.

We shall also restrict the space of wavefunctions to those with *finite kinetic energy*, which implies that

$$\sum_i \int d\mathbf{r}_1 \int d\mathbf{r}_2 \cdots \int d\mathbf{r}_N |\nabla_i \Psi(\mathbf{r}_1, \mathbf{r}_2 \cdots \mathbf{r}_N)|^2 < \infty. \quad (6) \quad \boxed{\text{FinKE}}$$

The wavefunctions then belong to the *Sobolev space* $H^1(\mathbf{R}^{3N})$ ². The densities corresponding to these wavefunctions form the space

$$\mathcal{S} = \{\rho | \rho \geq 0; \sqrt{\rho} \in H^1(\mathbf{R}^3)\}, \quad (7) \quad \boxed{\text{S}}$$

which is a subset of the intersection of the L^1 and L^3 spaces, $\mathcal{S} \subset X = L^1 \cap L^3$ ^[183, Lee03]. All densities in \mathcal{S} can be generated by at least one function in $H^1(\mathbf{R}^{3N})$. The \mathcal{S} space has the property of being *convex*, which implies that if $\rho_1, \rho_2 \in \mathcal{S}$, then also $\lambda\rho_1 + (1 - \lambda)\rho_2 \in \mathcal{S}$ with $\lambda \in [0, 1]$.

For future reference we shall define the concept of *pure-state v -representable* (PS- v -representable) densities, being the densities corresponding to a single ground-state wavefunction of a Hamiltonian (I)^{Ham} with the potential v , which is in the dual space of X , i.e., $v \in X^* = L^\infty + L^{3/2}$ ^[183, DG90, Lee03] [5, 12, 11]. The ground state can be degenerate or nondegenerate.

2.2 The Hohenberg-Kohn theorems

According to the Hohenberg-Kohn (HK) theorems^[HK64] [1], the ground-state energy (E_0) of an electronic system (I)^{Ham} is a functional of the ground-state density (ρ_0),

$$E_0 = E_{\text{HK}}[\rho_0], \quad (8) \quad \boxed{\text{EO}}$$

and the ground-state energy is obtained by minimizing the energy functional

$$E_0 = \min E_{\text{HK}}[\rho] \quad (9) \quad \boxed{\text{HK2}}$$

over the set of v -representable normalized densities.

The energy functional can in the HK model be expressed

$$E_{\text{HK}}[\rho_0] = F_{\text{HK}}[\rho_0] + \int d\mathbf{r} \rho_0(\mathbf{r}) v(\mathbf{r}), \quad (10) \quad \boxed{\text{EHK}}$$

where

$$F_{\text{HK}}[\rho_0] = \langle \Psi_0 | \hat{T} + \hat{W} | \Psi_0 \rangle \quad (11) \quad \boxed{\text{FHK}}$$

²Some of the topological concepts used here are defined in the Appendix. See also the review article by van Leeuwen in this volume^[Lee03] [11], where some of these concepts are further discussed.

is the *universal Hohenberg-Kohn functional*, which is independent of the external potential $v(\mathbf{r})$.

Originally, the HK theorem was derived only for nondegenerate ground states, and the densities were restricted to v -representable densities of such states. Later it has been shown that the theorems hold also if the ground-state is degenerate. ρ_0 can then be any of the ground-state densities and Ψ_0 any ground-state wavefunction yielding this density ^{PG90} [12].

2.3 The minimization process

Often the variation of a functional $F[\rho]$ at a density ρ_0 due to a small density change $\delta\rho$ can be expressed in the form

$$\delta F(\rho_0, \delta\rho) = F[\rho_0 + \delta\rho] - F[\rho_0] = \int d\mathbf{r} \left(\frac{\delta F[\rho]}{\delta\rho(\mathbf{r})} \right)_{\rho=\rho_0} \delta\rho(\mathbf{r}) + \text{higher order terms.} \quad (12)$$

FuncDiff

Then we refer to $\left(\frac{\delta F[\rho]}{\delta\rho(\mathbf{r})} \right)_{\rho=\rho_0}$ as the *functional derivative* at the density ρ_0 . There are different definitions of this concept, as we shall discuss in the next section. If the functional has an *extremum* (maximum or minimum) at the density ρ_0 , then the functional derivative will vanish in that point.

In order to be able to perform the minimization, using standard variational principles, it is necessary that we can make arbitrarily small variations of the density. With the densities restricted by the v -representability condition, this is not necessarily the case, and standard procedures cannot be generally applied. Therefore, in order to find workable forms of the theory, the definition of the functionals has to be generalized to a larger group of densities. This will be discussed in section ^{sec:DensFunc} 4.

We have assumed here that the variations are performed within the domain of normalized densities. Alternatively, the minimization can be performed using the *Euler-Lagrange procedure*. Then the densities are allowed to vary also outside the normalization domain. This we shall do by relaxing the normalization constraint of the wavefunctions and by using the definition ^{Density} (3) of the density also in the extended domain. The normalization constraint is enforced by means of a *Lagrange multiplier* (μ),

$$\delta \left(F[\rho_0] - \mu \int d\mathbf{r} \rho(\mathbf{r}) \right) = 0, \quad (13)$$

Lagrange

which leads to the Euler equation

$$\left(\frac{\delta F[\rho]}{\delta\rho(\mathbf{r})} \right)_{\rho=\rho_0} - \mu = 0. \quad (14)$$

Euler

The extension of the functionals into the domain of unnormalized densities can be made in different ways, and the value of the Lagrange multiplier will depend on the way this is done. The process is not trivial, however, as we shall demonstrate below. Certain rules have to be followed in making such an extension (c.f., for instance, analytical continuation in function analysis). The expression must above all fulfill the conditions for a functional in the extended region, which, among other things, implies that it has to be uniquely defined – a certain density must always lead to a unique value of the functional.

2.4 Comments on the treatment of Nesbet

At this point we want to make a brief comment upon the approach of Nesbet, reviewed in another article of this volume [9]. Nesbet's main conclusion is that one of the main fundamentals of DFT, the so-called *locality hypothesis* – the assumption that the derivative of the density functionals can be expressed in the form of multiplicative local function – is not generally valid. As we have pointed out in a separate Comment to the Physical Review [10], we believe that the arguments of Nesbet are incorrect and that the mistake is connected to the above-mentioned question of extending the functionals into the domain of unnormalized densities. We summarize our main arguments here and refer to our Comment for further details.

Nesbet considers a system of N noninteracting electrons with a nondegenerate ground state and with the wavefunctions in the form of Slater determinants $\Phi = 1/\sqrt{N!} \text{Det}\{\phi_1, \phi_2, \dots\}$. Restricting ourselves here to a two-electron system, the kinetic energy is then given by the orbital functional

$$T_s[\phi_1, \phi_2] = \langle \Phi | \hat{T} | \Phi \rangle = \langle \phi_1 | \hat{t} | \phi_1 \rangle + \langle \phi_2 | \hat{t} | \phi_2 \rangle \quad (15) \quad \boxed{\text{TN}}$$

with $\hat{t} = -\frac{1}{2} \nabla^2$ and the density by

$$\rho(\mathbf{r}) = |\phi_1(\mathbf{r})|^2 + |\phi_2(\mathbf{r})|^2. \quad (16) \quad \boxed{\text{DensN}}$$

Differentiating the expression (15), yields the orbital derivative

$$\frac{\delta T_s[\phi_1, \phi_2]}{\delta \phi_i^*(\mathbf{r})} = \hat{t} \phi_i(\mathbf{r}) = (\varepsilon_i - v(\mathbf{r})) \phi_i(\mathbf{r}), \quad (17) \quad \boxed{\text{OrbDerN}}$$

using the orbital equation

$$(\hat{t} + v(\mathbf{r})) \phi_i(\mathbf{r}) = \varepsilon_i \phi_i(\mathbf{r}).$$

Considering the kinetic energy (15) also as a *density* functional,

$$T_s[\phi_1, \phi_2] = T_s[\rho[\phi_1, \phi_2]], \quad (18) \quad \boxed{\text{DensOrb}}$$

Nesbet applies the chain rule and obtains

$$\frac{\delta T_s[\rho[\phi_1, \phi_2]]}{\delta \phi_i^*(\mathbf{r})} = \frac{\delta T_s[\rho]}{\delta \rho(\mathbf{r})} \frac{\delta \rho(\mathbf{r})}{\delta \phi_i^*(\mathbf{r})} = \frac{\delta T_s[\rho]}{\delta \rho(\mathbf{r})} \phi_i(\mathbf{r}). \quad (19) \quad \boxed{\text{ChainN}}$$

Identification then leads to an *orbital-dependent derivative*

$$\frac{\delta T_s[\rho]}{\delta \rho(\mathbf{r})} \phi_i(\mathbf{r}) = (\varepsilon_i - v(\mathbf{r})) \phi_i(\mathbf{r}), \quad (20) \quad \boxed{\text{OrbDep}}$$

which Nesbet interprets so that the *locality hypothesis* is not valid if the system contains different orbital energy eigenvalues.

As we have pointed out in our Comment [10], the expression (15) is in combination with the density expression (16) *not a density functional* outside the normalization domain. Therefore, the identity (18) is not generally valid and the chain rule cannot be used in the way Nesbet does. If the variations are restricted to the normalization domain, then the constant term (orbital eigenvalue) disappears, and the locality hypothesis is restored.

Another approach is to modify the kinetic-energy expression so that it will be a density functional also in the unnormalized domain. This we have performed in our Comment by defining the density according to (3)

$$\rho(\mathbf{r}) = |\phi_1(\mathbf{r})|^2 \langle \phi_2 | \phi_2 \rangle + |\phi_2(\mathbf{r})|^2 \langle \phi_1 | \phi_1 \rangle, \quad (21) \quad \boxed{\text{DensDet}}$$

and the kinetic energy similarly by

$$T_s[\phi_1, \phi_2] = \langle \Phi | \hat{T}_s | \Phi \rangle = \langle \phi_1 | \hat{t} | \phi_1 \rangle \langle \phi_2 | \phi_2 \rangle + \langle \phi_2 | \hat{t} | \phi_2 \rangle \langle \phi_1 | \phi_1 \rangle. \quad (22) \quad \boxed{\text{TDet}}$$

(We here relaxed the normalization constraint of the orbitals but maintained the orthogonality constraint.) Then we have shown that the kinetic energy is in fact a density as well as an orbital functional in the close neighborhood of the ground-state density, and hence the relation (18) is valid in that neighborhood. Direct derivation now yields the orbital derivative

$$\begin{aligned} \frac{\delta T_s[\phi_1, \phi_2]}{\delta \phi_1^*} &= [\hat{t} \langle \phi_2 | \phi_2 \rangle + \langle \phi_2 | \hat{t} | \phi_2 \rangle] \phi_1(\mathbf{r}) \\ &= [(\varepsilon_1 - v(\mathbf{r})) \langle \phi_2 | \phi_2 \rangle + \langle \phi_2 | \varepsilon_2 - v(\mathbf{r}) | \phi_2 \rangle] \phi_1(\mathbf{r}) \end{aligned} \quad (23) \quad \boxed{\text{OrbDer}}$$

and application of the chain rule

$$\frac{\delta T_s[\rho[\phi_1, \phi_2]]}{\delta \phi_1^*(\mathbf{r})} = \int d\mathbf{r}' \frac{\delta T_s[\rho]}{\delta \rho(\mathbf{r}')} \frac{\delta \rho(\mathbf{r}')}{\delta \phi_1^*(\mathbf{r})} = \left[\frac{\delta T_s[\rho]}{\delta \rho(\mathbf{r})} \langle \phi_2 | \phi_2 \rangle + \left\langle \phi_2 \left| \frac{\delta T_s[\rho]}{\delta \rho} \right| \phi_2 \right\rangle \right] \phi_1(\mathbf{r}). \quad (24) \quad \boxed{\text{Chain}}$$

Identification then yields

$$\frac{\delta T_s[\rho]}{\delta \rho(\mathbf{r})} = \frac{\varepsilon_1 + \varepsilon_2}{2} - v(\mathbf{r}), \quad (25) \quad \text{Chain2}$$

which is *orbital independent* and in the form of a *multiplicative local potential*, and again the locality hypothesis is restored.

This demonstrates the importance of extending the functionals into the domain of unnormalized densities in a proper way in applying the Euler-Lagrange procedure.

3 Functional derivatives

sec:Deriv

The functional derivative is a particular important concept in DFT, and we shall here investigate that in some detail. Two forms of functional derivatives are here of interest, the *Gâteaux* derivative and the *Fréchet* derivative.

3.1 Gâteaux derivative

sec:Ga

We consider a functional $f : M \rightarrow \mathbf{R}$, which is a mapping of a normed space (M) on the space of real numbers (\mathbf{R}). If at some point $x_0 \in M$ there exists a mapping $df(x_0, \cdot) : M \rightarrow \mathbf{R}$ such that

$$df(x_0, h) = \lim_{\lambda \rightarrow 0} \frac{f(x_0 + \lambda h) - f(x_0)}{\lambda}, \quad (26) \quad \text{Ga}$$

then $df(x_0, h)$ is the *Gâteaux differential* at the point x_0 in the direction h [BB92, p. 46], [PM1999, p. 448]. An equivalent definition is

$$f(x_0 + h) = f(x_0) + df(x_0, h) + \omega(x_0, h), \quad (27) \quad \text{Gaa}$$

where

$$\lim_{\lambda \rightarrow 0} \frac{\omega(x_0, \lambda h)}{\lambda} = 0. \quad (28) \quad \text{omega0}$$

The Gâteaux differential is also termed the *weak differential* [LS61, p. 293]. In principle, this mapping need neither be linear nor continuous, but we shall follow the convention [See03, Eq. 60] that

a functional is *Gâteaux differentiable* at a point x_0 if and only if the mapping in (26) is linear and continuous in h for all h at this point.

In applying the formalism to DFT, we express the variation of a density functional $F[\rho]$ at a density ρ_0 due to a density change $\delta\rho$ in the form [FuncDiff, Eq. 12]

$$\delta F(\rho_0, \delta\rho) = F[\rho_0 + \delta\rho] - F[\rho_0] = \int d\mathbf{r} v([\rho_0]; \mathbf{r}) \delta\rho(\mathbf{r}) + \omega(\rho_0, \delta\rho), \quad (29) \quad \text{Diff}$$

and the differential ^(Ga)₍₂₆₎ then becomes

$$df(\rho_v, \delta\rho) = \int d\mathbf{r} v([\rho_0]; \mathbf{r}) \delta\rho(\mathbf{r}) + \lim_{\lambda \rightarrow 0} \frac{1}{\lambda} \omega(\rho_0, \lambda\delta\rho). \quad (30) \quad \text{GaDiff}$$

This mapping is linear and continuous if (1) $v([\rho_0]; \mathbf{r})$ is a single-valued *bounded* function of \mathbf{r} that may depend on the density ρ_0 but is independent of $\delta\rho$, and (2)

$$\lim_{\lambda \rightarrow 0} \frac{1}{\lambda} \omega(\rho_0, \lambda\delta\rho) = 0. \quad (31) \quad \text{Omega}$$

(Note that a linear operator is continuous if and only if it is bounded ^(Gr81)_[16, p. 197, 213], ^(DMik99)_[14, p. 22].) We shall refer to the function above as the *Gâteaux derivative* at the density ρ_0 ^(Lee03)_[11, Eq. 61]

$$\left(\frac{\delta F[\rho]}{\delta \rho(\mathbf{r})} \right)_{\rho=\rho_0} = v([\rho_0]; \mathbf{r}). \quad (32) \quad \text{GaDer}$$

A necessary condition for the Gâteaux differentiability is also that the functional is defined for $\rho_0 + \lambda\delta\rho$ – for all $\lambda \in [0, 1]$ – if it is defined for ρ_0 and $\rho_0 + \delta\rho$. This is the case if the functional is defined on a space that is *convex*, as the \mathcal{S} space ^(S)₍₇₎

The form ⁽³⁰⁾₍₃₀₎ represents a differential that is linear and continuous if the conditions specified are fulfilled. A question relevant for the discussion of the differentiability in section ^(sec:DensFunc)₄, is if the differential of a differentiable functional always can be written in this form. We believe that this can be answered positively.³

3.2 Fréchet derivative

Another form of functional differential is the *Fréchet differential*, also termed the *strong differential*, which can be defined as follows ^(BB92)_[13, p. 37], ^(DMik99)_[14, p. 451]. With M being a subset of a Banach space E with the norm $\|\cdot\|$, the function

³In order to show that this is at least plausible, we consider a *discrete* coordinate space, $\mathbf{r} = \mathbf{r}_i$, $i = 1 \dots n$ and $\delta\rho_i = \delta\rho(\mathbf{r}_i)$. A functional $l[\delta\rho]$ then becomes a *function* of n variables, $l[\delta\rho] \rightarrow l(\delta\rho_1, \dots, \delta\rho_n)$. If the functional is *linear*, then

$$l(\delta\rho_1, \dots, \delta\rho_n) = l(\delta\rho_1, 0, \dots) + l(0, \delta\rho_2, 0, \dots) + \dots = \sum_{i=1}^n f_i \delta\rho_i$$

with $l(0, \dots, 0, \delta\rho_i, 0, \dots) = f_i \delta\rho_i$. In the limit, this approaches an integral,

$$l[\delta\rho] \rightarrow l(\delta\rho_1, \dots, \delta\rho_n) = \sum_{i=1}^n f_i \delta\rho_i \rightarrow \int d\mathbf{r} f(\mathbf{r}) \delta\rho(\mathbf{r}).$$

If the functional is *continuous*, then all f_i must be finite and the function $f(\mathbf{r})$ *bounded*.

$f : M \rightarrow \mathbf{R}$ is *Fréchet differentiable* at a point $x_0 \in M$, if there exists a *continuous and linear operator*, $L(\cdot) : E \rightarrow \mathbf{R}$, such that (1)

$$f(x_0 + h) - f(x_0) = L(h) + \omega(x_0, h) \quad (33) \quad \boxed{\text{Fr}}$$

for all $h \in E$ and $x_0 + h \in M$ and (2) $\omega(x_0, 0) = 0$ and

$$\lim_{\|h\| \rightarrow 0} \frac{\omega(x_0, h)}{\|h\|} = 0. \quad (34) \quad \boxed{\text{omega}}$$

The function $L(h)$ is the *Fréchet differential* at the point x_0 in the direction h , and the operator $L(\cdot)$ is sometimes termed the *Fréchet derivative* at x_0 . Here, however, we shall reserve the term *Fréchet derivative* to a function equivalent to the Gâteaux derivative (32). ^{Gader} Fréchet differentiability is a stronger requirement than that of Gâteaux differentiability, but – with the interpretations we make – the difference is quite subtle.

3.3 Comparison between the Fréchet and Gâteaux differentiabilitys

It can be shown that if a Gâteaux derivative exists in a neighborhood of the point x_0 , i.e., at all points $\|x - x_0\| < \delta$, and is (uniformly) *continuous* in x in this neighborhood, then the functional is *Fréchet differentiable* at the point x_0 [BB92, p.47], [LS61, p.295]. (For a linear mapping the concepts of 'continuity' and 'uniform continuity' are equivalent [Mik99, p. 23].) In this case there exists a δ for each ε so that the difference between the Gâteaux derivatives at the points x_0 and x is less than ε for all x with $\|x - x_0\| < \delta$. Therefore, a functional that is Gâteaux but not Fréchet differentiable at a point x_0 must have some kind of discontinuity at this point.

In order to demonstrate the difference further, we shall assume that the functional is Gâteaux differentiable and then find out what the additional requirement is to make it Fréchet differentiable. If we make the replacement $h \rightarrow \lambda h$ and set $\|h\| = 1$, the expression ^{omega} (34) becomes

$$\lim_{\lambda \rightarrow 0} \frac{\omega(x_0, \lambda h)}{\lambda} = 0 \quad (35) \quad \boxed{\text{omega1}}$$

or formally identical to the corresponding condition ^{omega0} (28) in the Gâteaux case. The difference is that in the Gâteaux case it is sufficient that this relation is satisfied for each individual value of h separately, while in the Fréchet case the relation must be fulfilled for the entire neighborhood. In other words, the limit must in the Fréchet case be *uniform*, implying that for each value of $\delta > 0$

there exists a $\lambda > 0$, so that

$$\frac{|\omega(x_0, \lambda h)|}{\lambda} < \delta \quad \text{for all } h. \quad (36) \quad \boxed{\text{omegas}}$$

If we assume that $\omega(x_0, \lambda h) = C(h)\lambda^2$, where $C(h)$ is a number that depends on h , then the condition for Gâteaux differentiability is that $C(h)$ is *finite*, while for Fréchet differentiability it has to be *bounded*, implying that there exists a number $D < \infty$, such that $C(h) < D$ for all h . This implies that also $\omega(x_0, h)$ must be bounded within the neighborhood $\|h\| < \varepsilon$. Thus, if the functional is not Fréchet differentiable, this term is unbounded in any neighborhood of x_0 , which indicates some kind of singularity.

Fréchet differentiability is obviously a stronger condition than Gâteaux differentiability. Therefore, *if a functional is Fréchet differentiable, it is also Gâteaux differentiable, and the Fréchet and the Gâteaux derivatives are identical.*

4 Application to density functionals

`sec:DensFunc`

As mentioned, in order to be able to apply the variational principle in DFT, it is necessary to extend the definition of the functionals beyond the domain of v -representable densities, and the standard procedure is here to apply the *Levy constrained-search procedure* [Le79]. This has led to the functionals known as the Levy-Lieb ($F_{\text{LL}}[\rho]$) and Lieb ($F_{\text{L}}[\rho]$) functionals, respectively, and we shall now investigate the differentiability of these functionals. This will represent the main part of our paper.

4.1 The Levy-Lieb functional

`sec:FLL`

We consider a general electronic system with the Hamiltonian (Ham)

$$\hat{H}_v = \hat{T} + \hat{W} + \hat{V} \quad (37) \quad \boxed{\text{Ham1}}$$

with a ground-state energy E_v that can be degenerate. By applying the constrained-search procedure to the Hohenberg-Kohn functional (HK) , Levy [Le79] and Lieb [Li83] have defined the functional

$$F_{\text{LL}}[\rho] = \inf_{\Psi \rightarrow \rho} \langle \Psi | \hat{T} + \hat{W} | \Psi \rangle, \quad (38) \quad \boxed{\text{FLL}}$$

referred to as the *Levy-Lieb functional*. Here, Ψ is any normalized wavefunction in the Sobolev space $H^1(\mathbf{R}^{3N})$ [equation (6)] that generates the density $\rho(\mathbf{r})$.

This extends the definition of the HK functional to all densities of the space \mathcal{S} . The energy functional (10) is generalized accordingly,

$$E_{\text{LL}}[\rho] = F_{\text{LL}}[\rho] + \int d\mathbf{r} \rho(\mathbf{r}) v(\mathbf{r}). \quad (39) \quad \boxed{\text{ELL}}$$

Here, we shall extend these definitions further by relaxing the normalization restriction of the wavefunctions and densities by using the definition of the density (3) also outside the normalization domain.

Within the normalization domain the energy functional (39) has its minimum equal to the exact ground-state energy E_v of the system, when the density is equal to any of the exact ground-state densities ρ_v ,

$$E_v = \min_{\rho \in \mathcal{S}} E_{\text{LL}}[\rho] = E_{\text{LL}}[\rho_v]. \quad (40) \quad \boxed{\text{Emin}}$$

We shall now investigate the differentiability of the Levy-Lieb functional (38), and in doing so we shall largely follow the arguments of our recent Comment to the work of Nesbet [10], extended to the more general situation.

With the normalization constraint, any ground-state wavefunction Ψ_v of \hat{H}_v minimizes $\langle \Psi | \hat{H}_v | \Psi \rangle$ according to the variational principle of quantum mechanics. Hence, any ground-state wavefunction, corresponding to a particular ground-state density ρ_v , also minimizes this quantity when the variation is restricted to that density. In such a restricted variation the potential contribution (5) is constant, and it then follows that also $\langle \Psi | \hat{T} + \hat{W} | \Psi \rangle$ is minimized, or

$$F_{\text{LL}}[\rho_v] = \langle \Psi_v | \hat{T} + \hat{W} | \Psi_v \rangle. \quad (41) \quad \boxed{\text{FLLmin}}$$

We then have, using (4) and (5),

$$F_{\text{LL}}[\rho_v] = \langle \Psi_v | \hat{H}_v - \hat{V} | \Psi_v \rangle = \int d\mathbf{r} \left(\frac{E_v}{N} - v(\mathbf{r}) \right) \rho_v(\mathbf{r}). \quad (42) \quad \boxed{\text{FLL1}}$$

In order to find the differential of the functional (38) at a ground-state density, ρ_v , we make a small change in the density around that density, $\rho = \rho_v + \delta\rho$, and investigate the corresponding change of the functional

$$\begin{aligned} \delta F_{\text{LL}}(\rho_v, \delta\rho) &= F_{\text{LL}}[\rho_v + \delta\rho] - F_{\text{LL}}[\rho_v] \\ &= \inf_{\Psi \rightarrow \rho_v + \delta\rho} \langle \Psi | \hat{T} + \hat{W} | \Psi \rangle - \int d\mathbf{r} \left(\frac{E_v}{N} - v(\mathbf{r}) \right) \rho_v(\mathbf{r}). \end{aligned} \quad (43) \quad \boxed{\text{Fvar}}$$

We write the minimizing wavefunction as $\Psi = \Psi_v + \delta\Psi$ with Ψ_v being a normalized ground-state wave function, corresponding to the density ρ_v . We do *not* assume that $\delta\Psi$ has to be 'small', and we also leave the question of normalization of Ψ open.

Generally, we have

$$\langle \Psi | \hat{T} + \hat{W} | \Psi \rangle = \langle \Psi | \hat{H}_v - \hat{V} | \Psi \rangle = \langle \Psi | \hat{H}_v | \Psi \rangle - \int d\mathbf{r} v(\mathbf{r}) \rho(\mathbf{r}) \quad (44) \quad \boxed{\text{Tave2}}$$

and

$$\begin{aligned} \langle \Psi | \hat{H}_v | \Psi \rangle &= E_v \langle \Psi | \Psi \rangle + \langle \Psi_v + \delta\Psi | \hat{H}_v - E_v | \Psi_v + \delta\Psi \rangle \\ &= E_v \langle \Psi | \Psi \rangle + \langle \delta\Psi | \hat{H}_v - E_v | \delta\Psi \rangle. \end{aligned} \quad (45) \quad \boxed{\text{Have}}$$

This yields

$$\langle \Psi | \hat{T} + \hat{W} | \Psi \rangle = \langle \Psi | \hat{H}_v - \hat{V} | \Psi \rangle = \int d\mathbf{r} \left(\frac{E_v}{N} - v(\mathbf{r}) \right) \rho(\mathbf{r}) + \langle \delta\Psi | \hat{H}_v - E_v | \delta\Psi \rangle \quad (46)$$

and with the definition $\boxed{\text{FLL}}$ (38)

$$F_{\text{LL}}[\rho_v + \delta\rho] = \int d\mathbf{r} \left(\frac{E_v}{N} - v(\mathbf{r}) \right) (\rho_v(\mathbf{r}) + \delta\rho(\mathbf{r})) + \inf_{\Psi \rightarrow \rho_v + \delta\rho} \langle \delta\Psi | \hat{H}_v - E_v | \delta\Psi \rangle. \quad (47)$$

The variation $\boxed{\text{Fvar}}$ (43) now becomes, using $\boxed{\text{FLL1}}$ (42),

$$\begin{aligned} \delta F_{\text{LL}}(\rho_v, \delta\rho) &= F_{\text{LL}}[\rho_v + \delta\rho] - F_{\text{LL}}[\rho_v] = \int d\mathbf{r} \left(\frac{E_v}{N} - v(\mathbf{r}) \right) \delta\rho(\mathbf{r}) \\ &+ \inf_{\Psi_v + \delta\Psi \rightarrow \rho_v + \delta\rho} \langle \delta\Psi | \hat{H}_v - E_v | \delta\Psi \rangle. \end{aligned} \quad (48) \quad \boxed{\text{DeltaF}}$$

This is of the form $\boxed{\text{Diff}}$ (29), with the potential $v(\mathbf{r}) \in L^\infty + L^{3/2}$ being finite. Therefore, the functional is Gâteaux differentiable, if the last term

$$\omega(\rho_v, \delta\rho) = \inf_{\Psi_v + \delta\Psi \rightarrow \rho_v + \delta\rho} \langle \delta\Psi | \hat{H}_v - E_v | \delta\Psi \rangle \quad (49) \quad \boxed{\text{omega3}}$$

vanishes faster than $\delta\rho$ or

$$\lim_{\lambda \rightarrow 0} \frac{1}{\lambda} \inf_{\Psi_v + \delta\Psi \rightarrow \rho_v + \lambda\delta\rho} \langle \delta\Psi | \hat{H}_v - E_v | \delta\Psi \rangle = 0 \quad (50) \quad \boxed{\text{Limit}}$$

for all $\delta\rho$.

We know that the kinetic energy is finite with the wavefunctions of the space we consider [equation $\boxed{\text{FinKE}}$ (6)], and since the electrostatic energy is finite and the potential energy can be expected to be finite, it follows that the matrix element in (50) is finite. Then we can show the differentiability by scaling.

We consider one specific wavefunction $\Psi_v + \delta\Psi_{\delta\rho}$ that yields a particular density $\rho_v + \delta\rho$. We know that all densities in the \mathcal{S} space (7) can be generated by at least one wavefunction in our space. We then scale the wavefunction

increment by $\lambda \in [0, 1]$, i.e., $\Psi_v \rightarrow \Psi_v + \lambda \delta \Psi_{\delta \rho}$, which generates the density change

$$\lambda \delta \rho_1 + \lambda^2 \delta \rho_2 = \lambda \delta \rho'. \quad (51) \quad \boxed{\text{ScalDens}}$$

With the scaled wavefunction we have

$$\lim_{\lambda \rightarrow 0} \frac{1}{\lambda} \langle \lambda \delta \Psi_{\delta \rho} | \hat{H}_v - E_v | \lambda \delta \Psi_{\delta \rho} \rangle = 0,$$

which means that

$$\lim_{\lambda \rightarrow 0} \frac{1}{\lambda} \inf_{\Psi_v + \delta \Psi \rightarrow \rho_v + \lambda \delta \rho'} \langle \delta \Psi | \hat{H}_v - E_v | \delta \Psi \rangle = 0. \quad (52) \quad \boxed{\text{Limit2}}$$

The infimum in ^(Limit2)(52) is searched over the density $\rho_v + \lambda \delta \rho'$, which is not identical to the density $\rho_v + \lambda \delta \rho$ of ^(Limit)(50). For that reason we consider instead of ^(DeltaF)(48) the variation

$$\begin{aligned} \delta F_{\text{LL}}(\rho_v, \lambda \delta \rho') &= F_{\text{LL}}[\rho_v + \lambda \delta \rho'] - F_{\text{LL}}[\rho_v] \\ &= \int d\mathbf{r} \left(\frac{E_v}{N} - v(\mathbf{r}) \right) \lambda \delta \rho'(\mathbf{r}) + \omega(\rho_v, \lambda \delta \rho') \end{aligned} \quad (53) \quad \boxed{\text{DeltaF1}}$$

where

$$\omega(\rho_v, \lambda \delta \rho') = \inf_{\Psi_v + \delta \Psi \rightarrow \rho_v + \lambda \delta \rho'} \langle \delta \Psi | \hat{H}_v - E_v | \delta \Psi \rangle$$

vanishes faster than λ in view of ^(Limit2)(52). The expression ^(DeltaF1)(53) can be rewritten as

$$\frac{\delta F_{\text{LL}}(\rho_v, \lambda \delta \rho')}{\lambda} = \int d\mathbf{r} \left(\frac{E_v}{N} - v(\mathbf{r}) \right) \delta \rho'(\mathbf{r}) + \frac{\omega(\rho_v, \lambda \delta \rho')}{\lambda}, \quad (54) \quad \boxed{\text{GaM1}}$$

where the last term goes to zero as $\lambda \rightarrow 0$. This leads to

$$\lim_{\lambda \rightarrow 0} \frac{\delta F_{\text{LL}}(\rho_v, \lambda \delta \rho')}{\lambda} = \int d\mathbf{r} \left(\frac{E_v}{N} - v(\mathbf{r}) \right) \delta \rho_1(\mathbf{r}). \quad (55) \quad \boxed{\text{GaM2}}$$

The r.h.s of ^(GaM2)(55) is *continuous* with respect to changes in $\delta \rho'$, which we can show by replacing $\delta \rho'$ in ^(GaM1)(54) by $\delta \rho' + \delta^2 \rho'$ and let $\delta^2 \rho' \rightarrow 0$. Then the corresponding change of the r.h.s. vanishes in the limit $\lambda \rightarrow 0$. Furthermore, the various 'paths' $\delta \rho'$ cover all possible densities, so that we can make infinitesimal changes. By setting $\delta^2 \rho' = -\lambda \delta \rho_2$, we then obtain

$$\lim_{\lambda \rightarrow 0} \frac{\delta F_{\text{LL}}(\rho_v, \lambda \delta \rho')}{\lambda} = \lim_{\lambda \rightarrow 0} \frac{\delta F_{\text{LL}}(\rho_v, \lambda \delta \rho_1)}{\lambda} \quad (56) \quad \boxed{\text{GaM3}}$$

and

$$\lim_{\lambda \rightarrow 0} \frac{\delta F_{\text{LL}}(\rho_v, \lambda \delta \rho_1)}{\lambda} = \int d\mathbf{r} \left(\frac{E_v}{N} - v(\mathbf{r}) \right) \delta \rho_1(\mathbf{r}). \quad (57) \quad \boxed{\text{GaM4}}$$

This is exactly the definition of the Gâteaux differential (26) in the direction $\delta\rho_1$, and hence we have

$$\boxed{dF_{\text{LL}}(\rho_v, \delta\rho_1) = \int d\mathbf{r} \left(\frac{E_v}{N} - v(\mathbf{r}) \right) \delta\rho_1(\mathbf{r}).} \quad (58) \quad \boxed{\text{GaDiff2}}$$

This relation holds for any direction, which implies that *the functional F_{LL} is Gâteaux differentiable with the Gâteaux derivative*

$$\left(\frac{\delta F_{\text{LL}}[\rho]}{\delta\rho(\mathbf{r})} \right)_{\rho=\rho_v} = \frac{E_v}{N} - v(\mathbf{r}), \quad (59) \quad \boxed{\text{TDer}}$$

which is *equivalent to a local multiplicative function*.

Since E_v is the ground-state energy, it follows that the expression (50) is non-negative. This implies that the functional is *locally convex* in the neighborhood of the density ρ_v . In the standard method for proving the differentiability the convexity of the functional is used [5, 4, 11]. Since the Levy-Lieb functional is not necessarily convex, this procedure does not work. The reason that in spite of this fact it has been possible to demonstrate the differentiability here could be connected to the fact that the functional is locally convex in the neighborhood of the points of interest.

We have here made no assumption about the normalization of the wavefunction or density. If the density is normalized, the integral over $\delta\rho$ vanishes,

$$\int d\mathbf{r} \delta\rho(\mathbf{r}) = 0,$$

and the derivative will be undetermined up to a constant. The constant in (59) depends on the particular way the functional is extended into the unnormalized domain and has no physical significance. This constant corresponds to the Lagrange multiplier in the Euler equation (14), when the Euler-Lagrange procedure is used.

In the treatment here we have assumed that Ψ_v is a ground-state wavefunction of a Hamiltonian \hat{H}_v corresponding to the ground-state density ρ_v , but no further assumption has been made. The results hold for nondegenerate as well as degenerate ground states.

We can then conclude that *the Levy-Lieb functional F_{LL} is Gâteaux differentiable at any PS- v -representable density and that the derivative can be represented by a multiplicative local function*.

If the density is *not* that of an eigenstate of a Hamiltonian \hat{H}_v – ground or excited state – then the differential cannot be of the form (58), which we conclude will exclude differentiability. On the other hand, we cannot from our analysis exclude the possibility that the functional is differentiable at a density

due to an *excited* state of some \hat{H}_v . This question, though, requires further study.

Having considered the Gâteaux differentiability of $F_{LL}[\rho]$ at PS- v -representable densities, we may now turn to the question of possible *Fréchet differentiability* at these densities. The condition (34) is then

$$\lim_{\delta\rho\rightarrow 0} \frac{\omega(\rho_v, \delta\rho)}{\|\delta\rho\|} = 0 \quad \text{(60)} \quad \text{omega2}$$

with $\omega(\rho_v, \delta\rho)$ given by (49). As discussed above, this requires that the expression (49) is *bounded* in the neighborhood $\|\delta\rho\| < \varepsilon$ as $\varepsilon \rightarrow 0$. The norm could here be $\|\delta\rho\|_1$ or $\|\delta\rho\|_3$ of the spaces L^1 and L^3 , respectively.

Since the electrostatic energy is bounded and we can assume that the external potential is such that the potential energy is bounded, it follows that the residual term is bounded, if the kinetic energy is bounded. We have assumed that the kinetic energy is *finite* (6), but this does not necessarily mean that it is bounded. Generally, we cannot exclude the possibility that the kinetic energy is unbounded, but it should be observed that it is the *infimum* that appears in the expression (49). Therefore, a sufficient condition for the Fréchet differentiability is that for *each* density in any neighborhood of the ground-state density ρ_v there exists *one* wavefunction that generates bounded kinetic energy. It seems clear that all densities in the neighborhood of ρ_v can be generated by functions close to the ground-state function Ψ_v , generating the density ρ_v .⁴ On the other hand, it is possible to construct a function, generating a density arbitrary close to ρ_v with arbitrarily high kinetic energy by adding a small rapidly oscillating component to the wavefunction. In such a case the functional would *not* be Fréchet differentiable. From a DFT standpoint, though, it seems that the possible difference between Gâteaux and Fréchet differentiability is of no practical importance.

4.2 The Lieb functional

sec:Lieb

In the degenerate case we can have a situation, where a linear combination of ground-state densities is not necessarily itself a ground-state density. This has the consequence that the Hohenberg-Kohn (11) and the Levy-Lieb (38) functionals are not necessarily *convex*, which for many applications is a disadvantage. A convex functional can be constructed by considering *ensemble- v -representable* (E- v -representable) densities [3, 12, 11]

$$\rho_{Ev}(\mathbf{r}) = \sum_k \lambda_k \rho_{vk}(\mathbf{r}) \quad \lambda_k \geq 0 \quad \sum_k \lambda_k = 1, \quad \text{(61)} \quad \text{rhoEv}$$

⁴This problem is closely related to the 'Question 2', raised by Lieb [5, p. 247], which Lieb believes has an affirmative answer, although a rigorous proof has not been found.

which is a linear combination of ground-state densities ρ_{vk} of the Hamiltonian \hat{H}_v [equation (I)]. It can be shown that the Hohenberg-Kohn theorems are essentially valid also for such densities. The corresponding *extended Hohenberg-Kohn functional* (EHK) is defined

$$F_{\text{EHK}}[\rho_{Ev}] = E_v - \int d\mathbf{r} v(\mathbf{r}) \rho_{Ev}(\mathbf{r}). \quad (62) \quad \boxed{\text{FEHK}}$$

The ground-state energy then becomes [Lee03, Theorem 2]

$$E_v = \inf_{\rho \in \mathcal{B}} \left\{ F_{\text{EHK}}[\rho] + \int d\mathbf{r} v(\mathbf{r}) \rho(\mathbf{r}) \right\}, \quad (63) \quad \boxed{\text{HKE}}$$

where \mathcal{B} is the space of all E- v -representable densities.

Using the constrained-search procedure, Lieb [5] has in analogy with the Levy-Lieb functional (38) extended the EHK functional (10) to

$$F_{\text{L}}[\rho] = \inf_{\lambda_k, \Psi_k \rightarrow \rho} \sum_k \lambda_k \langle \Psi_k | \hat{T} + \hat{W} | \Psi_k \rangle, \quad (64) \quad \boxed{\text{FL}}$$

where $\{\Psi_k\}$ is any set of orthonormal eigenfunctions of some \hat{H}_v (not necessarily degenerate) and λ_k has the same restrictions as in (61). This is usually referred to as the *Lieb functional*. Both the functionals F_{EHK} (62) and F_{L} (64) are convex.

The energy functional corresponding to the functional F_{L} is

$$E_{\text{L}}[\rho] = E_{\text{L}}[\rho] + \int d\mathbf{r} v(\mathbf{r}) \rho(\mathbf{r}) = \inf_{\lambda_k, \Psi_k \rightarrow \rho} \sum_k \lambda_k \langle \Psi_k | \hat{T} + \hat{W} + \hat{V} | \Psi_k \rangle,$$

which has its minimum, equal to the ground-state energy, when all Ψ_k are ground-state eigenfunctions of \hat{H}_v . Therefore,

$$E_v = \min_{\rho \in \mathcal{S}} E_{\text{L}}[\rho] = E_{\text{L}}[\rho_{Ev}], \quad (65) \quad \boxed{\text{EL}}$$

where $\rho_{Ev}(\mathbf{r})$ is according to (61) composed of any combination of ground-state wavefunctions of the Hamiltonian H_v .

We can now demonstrate the Gâteaux differentiability of the Lieb functional (64) for all E- v densities using the same procedure as in the previous section. We know that the energy functional (65) has its lowest value when all functions belong to the ground state. It then follows that the Lieb functional for an E- v density becomes

$$F_{\text{L}}[\rho_{Ev}] = \sum_k \lambda_k \langle \Psi_{vk} | \hat{T} + \hat{W} | \Psi_{vk} \rangle \quad \lambda_k, \Psi_{vk} \rightarrow \rho_{Ev}, \quad (66) \quad \boxed{\text{FL1}}$$

which can be expressed

$$F_L[\rho_{Ev}] = \sum_k \lambda_k \langle \Psi_{vk} | \hat{H}_v - \hat{V} | \Psi_{vk} \rangle = \int d\mathbf{r} \left(\frac{E_v}{N} - v(\mathbf{r}) \right) \rho_{Ev}(\mathbf{r}). \quad (67) \quad \boxed{\text{FL2}}$$

In exactly the same way as in the previous case, it can then be shown that the differential becomes

$$\begin{aligned} dF_L(\rho_{Ev}, \delta\rho) &= \sum_k \lambda_k \langle \delta\Psi_{vk} | \hat{T} + \hat{W} | \Psi_{vk} \rangle + c.c. = \sum_k \lambda_k \langle \delta\Psi_{vk} | \hat{H}_v - \hat{V} | \Psi_{vk} \rangle + c.c. \\ &= E_v \langle \delta\Psi_{vk} | \Psi_{vk} \rangle + c.c. - \int d\mathbf{r} v(\mathbf{r}) \delta\rho(\mathbf{r}) = \int d\mathbf{r} \left(\frac{E_v}{N} - v(\mathbf{r}) \right) \delta\rho(\mathbf{r}). \end{aligned} \quad (68) \quad \boxed{\text{DeltaFL}}$$

This shows that the functional $F_L[\rho]$ is Gâteaux differentiable at any density generated by an ensemble of ground-state wave functions, i.e., at any E - v -representable density,

$$\left(\frac{\delta F_L[\rho]}{\delta \rho(\mathbf{r})} \right)_{\rho=\rho_{Ev}} = \frac{E_v}{N} - v(\mathbf{r}) \quad (69) \quad \boxed{\text{FLDer}}$$

Using the same arguments as in the previous case, it seems that functionals of physical interest are also Fréchet differentiable at the E - v -representable densities.

If the density in (64) is *not* E - v representable, i.e., cannot be expressed as a linear combination of degenerate ground-state densities of any Hamiltonian, H_v , then the minimum cannot be represented by ground-state wavefunctions only – at least one must belong to an excited state. If not all wavefunctions belong to the same energy eigenvalue, then the differential cannot be written in the form (68). Using the same arguments as in the previous case, this means that the functional is *not* differentiable. If, on the other hand, all wavefunctions belong to the same *excited* energy eigenvalue, our analysis does not at present exclude differentiability. As in the previous case, however, this question requires further study.

5 The Kohn-Sham model

sec:KS

We shall now show how the formalism examined above can be used to derive the standard Kohn-Sham scheme. We start by considering the Levy-Lieb energy functional (39), which is minimized under the normalization constraint

$$E_v = \min_{\rho \in \mathcal{S}} E_{LL}[\rho] = E_{LL}[\rho_v] \quad (70) \quad \boxed{\text{Emin2}}$$

by varying the density in the space \mathcal{S} [equation (7)].

The minimization of $E_{\text{LL}}[\rho]$, using (39), leads to

$$\delta E_{\text{LL}}[\rho] = \delta \left\{ F_{\text{LL}}[\rho] + \int d\mathbf{r} \rho(\mathbf{r}) v(\mathbf{r}) \right\} = 0,$$

and staying within the normalization domain to

$$\frac{\delta E_{\text{LL}}[\rho]}{\delta \rho} = \frac{\delta F_{\text{LL}}[\rho]}{\delta \rho} + v(\mathbf{r}) + \text{constant} = 0. \quad (71) \quad \boxed{\text{MinELL}}$$

The constant is here undetermined, due to the normalization constraint, $\int d\mathbf{r} \rho(\mathbf{r}) = 0$. If the functional $E_{\text{LL}}[\rho]$ is Gâteaux (Fréchet) differentiable, then its Gâteaux (Fréchet) derivative will vanish at the minimum [14, p. 460].

In the Kohn-Sham model [2] we consider a system of *noninteracting electrons*, moving in a *local* potential $v_{\text{KS}}(\mathbf{r})$,

$$\hat{H}_{\text{KS}} = \hat{T} + \hat{V}_{\text{KS}} = \sum_{i=1}^N -\frac{1}{2} \nabla_i^2 + \sum_{i=1}^N v_{\text{KS}}(\mathbf{r}_i), \quad (72) \quad \boxed{\text{HKS}}$$

and the basic idea of the model is to determine the potential $v_{\text{KS}}(\mathbf{r})$ so that the ground-state density of the noninteracting system becomes the same as that of the interacting system we are considering. This requires that the density is *noninteracting v -representable*, i.e., can be reproduced by the ground-state density of a noninteracting system with a local potential. van Leeuwen has shown that this is always possible with arbitrary accuracy [11].

Using the constrained search, the correspondence of the Levy-Lieb functional (38) is for the noninteracting system the minimum of the kinetic energy,

$$T_{\text{KS}}[\rho] = \min_{\Psi \rightarrow \rho} \langle \Psi | \hat{T} | \Psi \rangle, \quad (73) \quad \boxed{\text{TS}}$$

and the corresponding energy functional becomes

$$E_{\text{KS}}[\rho] = T_{\text{KS}}[\rho] + \int d\mathbf{r} \rho(\mathbf{r}) v_{\text{KS}}(\mathbf{r}). \quad (74) \quad \boxed{\text{EKS}}$$

The ground-state energy of the noninteracting system is obtained by minimizing this energy functional under the normalization constraint,

$$E_v = \min_{\rho \in \mathcal{S}} E_{\text{KS}}[\rho], \quad (75) \quad \boxed{\text{KS2}}$$

where the search is over the same set of densities as in the interacting case. This leads to the equation

$$\frac{\delta E_{\text{KS}}[\rho]}{\delta \rho} = \frac{\delta T_{\text{KS}}[\rho]}{\delta \rho} + v_{\text{KS}}(\mathbf{r}) + \text{constant} = 0. \quad (76) \quad \boxed{\text{MinEKS}}$$

The derivative exists here for densities that are PS- v representable for a non-interacting system.

We now require that the solution of the equation $(76)^{\text{MinEKS}}$ should lead to the same ground-state density as the equation $(71)^{\text{MinELL}}$, which yields the condition for the potential v_{KS} ,

$$v_{\text{KS}}(\mathbf{r}) = \frac{\delta E_{\text{LL}}[\rho]}{\delta \rho} - \frac{\delta T_{\text{KS}}[\rho]}{\delta \rho} = \frac{\delta F_{\text{LL}}[\rho]}{\delta \rho} - \frac{\delta T_{\text{KS}}[\rho]}{\delta \rho} + v(\mathbf{r}) + \text{constant}. \quad (77) \quad \boxed{\text{KS1}}$$

If this potential is inserted in $(76)^{\text{MinEKS}}$, we see that $(71)^{\text{MinELL}}$ is automatically satisfied.

We can express F_{LL} as

$$F_{\text{LL}}[\rho] = T_{\text{KS}}[\rho] + E_{\text{Coul}}[\rho] + E_{\text{xc}}[\rho], \quad (78)$$

where the second term on the r.h.s. represents the Coulomb (Hartree) interaction energy and the last term the exchange-correlation energy, including also the (hopefully small) difference between the T_{KS} and the kinetic-energy part of the F_{LL} functional. This leads to the Kohn-Sham potential

$$\boxed{v_{\text{KS}}(\mathbf{r}) = v(\mathbf{r}) + \frac{\delta E_{\text{Coul}}[\rho]}{\delta \rho} + \frac{\delta E_{\text{xc}}[\rho]}{\delta \rho}} \quad (79) \quad \boxed{\text{KS}}$$

leaving out the undetermined constant term.

With the potential $(79)^{\text{KS}}$ the Kohn-Sham model $(72)^{\text{HKS}}$ yields the same ground-state density as the original problem, and inserting this density into the HK functional $(10)^{\text{EHK}}$ leads, in principle, to the ground-state energy. This is the basic *Kohn-Sham procedure*.

Above we have assumed that the minimization is carried out within the domain of normalized densities. Alternatively, we can perform the minimization, using the Euler-Lagrange procedure. Then we use the extension of the functionals valid also outside the normalization domain and enforce the normalization constraint by a Lagrange multiplier.⁵ For the Levy-Lieb energy functional $(70)^{\text{Emin2}}$ this leads to

$$\delta \left\{ E_{\text{LL}}[\rho] - \mu \int d\mathbf{r} \rho(\mathbf{r}) \right\} = 0 \quad (80) \quad \boxed{\text{ELKS}}$$

and to the Euler equation

$$\frac{\delta F_{\text{LL}}[\rho]}{\delta \rho} + v(\mathbf{r}) = \mu \quad (81) \quad \boxed{\text{EulerLL}}$$

⁵As demonstrated in section [2](#),^{[sec:General](#)} the extension of the functionals into the unnormalized domain is nontrivial.

and, similarly, minimization of the KS energy functional to

$$\frac{\delta T_{\text{KS}}[\rho]}{\delta \rho} + v_{\text{KS}}(\mathbf{r}) = \mu'. \quad (82) \quad \boxed{\text{EulerKS}}$$

This yields

$$v_{\text{KS}}(\mathbf{r}) = \frac{\delta F_{\text{LL}}[\rho]}{\delta \rho} - \frac{\delta T_{\text{KS}}[\rho]}{\delta \rho} + v(\mathbf{r}) + \text{constant}, \quad (83) \quad \boxed{\text{KS3}}$$

which is identical to the equation [\(77\)](#) ^{[KS1](#)}.

Instead of basing the treatment on the Levy-Lieb functional [\(38\)](#) ^{[FLL](#)}, the corresponding result can be obtained by using the Lieb functional [\(64\)](#) ^{[FL](#)}.

6 Conclusions

`sec:Concl`

We have shown that the Lieb functional [\(69\)](#) ^{[FLDer](#)} is *Gâteaux* differentiable at all E- v -representable densities, which is consistent with the result of Englisch and Englisch [\[3, 4\]](#) ^{[EE84, EE84a](#)}, who demonstrated the differentiability by using the convexity of the functional. The same procedure is used by van Leeuwen [\[11\]](#) ^{[Lee03](#)}. This procedure cannot be used for the Levy-Lieb functional, $F_{\text{LL}}[\rho]$, which is not manifestly convex. According to Englisch and Englisch, the differentiability of this functional is an open question.

The procedure we have applied does not depend on the (global) convexity of the functional, and we have been able to demonstrate the *Gâteaux* differentiability of the Levy-Lieb functional at all PS- v -representable densities, where this functional is locally convex. It seems plausible that both these functionals are also *Fréchet* differentiable at the same densities, although we have not been able to find a rigorous proof.

Acknowledgments

Communications with Evert Jan Baerends, Erkki Brändas, and Robert van Leeuwen are gratefully acknowledged. This work has been supported in part by a grant from the Swedish Research Council.

A Basic topological concepts

`sec:GeneralA`

In this Appendix we shall define some notations and basic topological concepts, which are needed in treating density functionals in a formal way. (We also refer to the article by van Leeuwen in this volume [\[11\]](#) ^{[Lee03](#)}, where some of these concepts are discussed in more detail.)

A.1 Notations

$X, Y, ..$ are sets with elements $x, y, ..$

$x \in X$ means that x is an element in the set X .

\mathbf{N} is the set of nonnegative integers. \mathbf{R} is the set of real numbers. \mathbf{C} is the set of complex numbers.

\mathbf{R}^n is the set of real n -dimensional vectors. \mathbf{C}^n is the set of complex n -dimensional vectors.

$A \subset X$ means that A is a subset of X .

$A \cup B$ is the *union* of A and B . $A \cap B$ is the *intersection* of A and B .

$A = \{x \in X : P\}$ means that A is the set of all elements x in X that satisfy the condition P .

$f : X \rightarrow Y$ represents a *function* or *operator*, which means that f maps *uniquely* the elements of X onto elements of Y .

A *functional* is a unique mapping $f : X \rightarrow \mathbf{R}(\mathbf{C})$ of a function space on the space of real (complex) numbers.

The set of arguments $x \in A$ for which the function $f : A \rightarrow B$ is defined is the *domain*, and the set of results $y \in B$ which can be produced is the *range*.

(a, b) is the *open interval* $\{x \in \mathbf{R} : a < x < b\}$. $[a, b]$ is the *closed interval* $\{x \in \mathbf{R} : a \leq x \leq b\}$.

sup represents the *supremum*, the least upper bound of a set

inf represents the *infimum*, the largest lower bound of a set.

A.2 Vector spaces

A real (complex) *vector space* or function space X is an infinite set of elements, x , referred to as *points* or *vectors*, which is *closed* under addition, $x + y = z \in X$, and under multiplication by a real (complex) number c , $cx = y \in X$. The continuous functions $f(x)$ on the interval $x \in [a, b]$ form a vector space, also with some boundary conditions, like $f(a) = f(b) = 0$.

A subset of X is a *subspace* of X if it fulfills the criteria for a vector space.

A *norm* of a vector space X is a function $p : X \rightarrow [0, \infty]$ with the properties (1) $p(\lambda x) = |\lambda|p(x)$, (2) $p(x + y) \leq p(x) + p(y)$ for all real λ ($\lambda \in \mathbf{R}$) and all $x, y \in X$, and (3) that $p(x) = 0$ always implies $x = 0$. The norm is written $p(x) = \|x\|$. We then have $\|\lambda x\| = |\lambda|\|x\|$ and $\|x + y\| \leq \|x\| + \|y\|$ and $\|x\| = 0 \Rightarrow x = 0$. If the last condition is not fulfilled, it is a *seminorm*.

A vector space with a norm for all its elements is a *normed space*, denoted $(X, \|\cdot\|)$. The continuous functions, $f(x)$, on the interval $[a, b]$ form a normed space by defining a norm, for instance, $\|f\| = \left[\int_a^b dt |f(t)|^2 \right]^{1/2}$. By means of the Cauchy-Schwartz inequality, it can be shown that this satisfies the criteria

for a norm ^{Gr81} [16, p. 93].

If f is a function $f : A \rightarrow Y$ and $A \subset X$, then f is defined in the neighborhood of $x_0 \in X$, if there is an $\epsilon > 0$ such that the entire sphere $\{x \in X : \|x - x_0\| < \epsilon\}$ belongs to A ^{Gr81} [16, p. 309].

A function/operator $f : X \rightarrow Y$ is *bounded*, if there exists a number C such that

$$\sup_{0 \neq x \in X} \left[\frac{\|fx\|}{\|x\|} \right] = C < \infty.$$

Then $C = \|f\|$ is the *norm* of f . Thus, $\|fx\| \leq \|f\| \|x\|$.

A function f is *continuous* at the point $x_0 \in X$, if for every $\delta > 0$ there exists an $\epsilon > 0$ such that for every member of the set $x : \|x - x_0\| < \epsilon$ we have $\|fx - fx_0\| \leq \delta$ ^{Gr81} [16, p. 139]. This can also be expressed so that f is continuous at the point x_0 , if and only if $fx \rightarrow fx_0$ whenever $x_n \rightarrow x_0$, $\{x_n\}$ being a sequence in X , meaning that fx_n converges to fx_0 , if x converges to x_0 ^{Ta57} [18, p. 70].

A linear function/operator is continuous if and only if it is bounded ^{Gr81} [16, p. 197, 213], ^{DMik99} [14, p. 22].

A functional $f : X \rightarrow \mathbf{R}$ is *convex* if

$$f(tx + (t - 1)y) \leq tf(x) + (t - 1)f(y)$$

for all $x, y \in X$ and $t \in (0, 1)$.

A subset $A \subset X$ is *open*, if for every $x \in A$ there exists an $\epsilon > 0$ such that the entire ball $B_r(x) = \{y \in X : \|y - x\| < \epsilon\}$ belongs to A , i.e., $B_r(x) \subset A$ ^{BB92} [13, p. 363], ^{Gr81} [16, p. 98], ^{Ta57} [18, p. 57].

A *sequence* $\{x_n\}$, where n is an integer ($n \in \mathbf{N}$), is an infinite numbered list of elements in a set or a space. A *subsequence* is a sequence, which is a part of a sequence.

A sequence $\{x_n \in A\}$ is (strongly) *convergent* towards $x \in A$, if and only if for every $\epsilon > 0$ there exists an N such that $\|x_n - x\| < \epsilon$ for all $n > N$ ^{Gr81} [16, p. 95, 348].

A sequence is called a *Cauchy sequence* if and only if for every $\epsilon > 0$ there exists an N such that $\|x_n - x_m\| < \epsilon$ for all $m, n > N$. If a sequence $\{x_n\}$ is convergent, then it follows that for $n, m > N$

$$\|x_m - x_n\| = \|(x_n - x) + (x - x_m)\| \leq \|x_n - x\| + \|x_m - x\| < 2\epsilon$$

which means that a convergent sequence is always a Cauchy sequence. The opposite is not necessarily true, since the point of convergence need not be an element of X ^{DW74} [19, p. 44].

A *normed space in which every Cauchy sequence converges (to a point in the space) is said to be complete and termed a Banach space.*

A subset A of a normed space is termed *compact*, if every infinite sequence of elements in A has a subsequence, which converges to an element in A . The closed interval $[0,1]$ is an example of a compact set, while the open interval $(0,1)$ is noncompact, since the sequence $1, 1/2, 1/3, \dots$ and all of its subsequences converge to 0, which lies outside the set [18, p. 149]. This sequence satisfies the Cauchy convergence criteria but not the (strong) convergence criteria.

L^p represents the Banach space with the norm $\|f(x)\| = [\int dx |f(x)|^p]^{1/p} < \infty$ [13].

A *dual space* or *adjoint space* of a vector space X , denoted X^* , is the space of all functions on X .

An *inner or scalar product* in a vector space X is a function $\langle \cdot, \cdot \rangle : X \times X \rightarrow \mathbf{R}$ with the properties (1)

$$\langle x, \lambda_1 y_1 + \lambda_2 y_2 \rangle = \lambda_1 \langle x, y_1 \rangle + \lambda_2 \langle x, y_2 \rangle, \quad \langle x, y \rangle = \langle y, x \rangle$$

for all $x, y, y_1, y_2 \in X$ and all $\lambda_j \in \mathbf{R}$, and (2) $\langle x, x \rangle = 0$ only if $x = 0$. A Banach space with the norm $x \rightarrow \|x\| = +\sqrt{\langle x, x \rangle}$ is called a *Hilbert space* [13, p. 364].

A *Sobolev space* $H^{(m)}(\mathbf{R}^{3N})$ is the space of all functions which together with their partial derivatives up to order m are square integrable in the space \mathbf{R}^{3N} , i.e., belong to the space L^2 [14, p. 97].

References

- [HK64] [1] P. Hohenberg, W. Kohn, *Inhomogeneous Electron Gas*, Phys. Rev. 136 (1964) B864–71.
- [KS65] [2] W. Kohn, L. J. Sham, *Self-Consistent Equations Including Exchange and Correlation Effects*, Phys. Rev. 140 (1965) A1133–38.
- [EE84] [3] H. Englisch, R. Englisch, *Exact Density Functionals for Ground-State Energies. I. General Results*, Phys. Stat. Sol. 123 (1984) 711–21.
- [EE84a] [4] H. Englisch, R. Englisch, *Exact Density Functionals for Ground-State Energies. II. Details and Remarks*, Phys. Stat. Sol. 124 (1984) 373–79.
- [Li83] [5] E. H. Lieb, *Density Functionals for Coulomb Systems*, Int. J. Quantum Chem. 24 (1983) 243–77.
- [Ne98] [6] R. K. Nesbet, *Kinetic energy in density-functional theory*, Phys. Rev. A 58 (1998) R12–15.

- NC99** [7] R. K. Nesbet, R. Colle, *Tests of the locality of exact Kohn-Sham exchange potentials*, Phys. Rev. A 61 (1999) 012503.1–7.
- Ne01** [8] R. K. Nesbet, *Kinetic energy as a density functional*, Phys. Rev. A 65 (2001) 010502(R).1–2.
- Ne03** [9] R. K. Nesbet, *Local and nonlocal potential functions in density functional theory*, Adv. Quantum Chem. 43 (2003) 1–23.
- LS03a** [10] I. Lindgren, S. Salomonson, *Comments on the Locality in Density-Functional Theories*, Phys. Rev. A 67 (2003) 056501.1–6.
- Lee03** [11] R. van Leeuwen, *Density functional approach to the many-body problem: key concepts and exact functionals*, Adv. Quantum Chem. 43 (2003) 25–94.
- DG90** [12] R. M. Dreizler, E. K. U. Gross, *Density Functional Theory*, Springer-Verlag, Berlin, 1990.
- BB92** [13] P. Blanchard, E. Brüning, *Variational Methods in Mathematical Physics. A Unified Approach*, Springer-Verlag, Berlin, 1992.
- DMik99** [14] L. Debnath, P. Mikusiński, *Introduction to Hilbert Spaces with Applications.*, 2nd Edition, Academic Press, New York, 1999.
- LS61** [15] L. A. Lusternik, V. J. Sobolev, *Elements of Functional Analysis*, Wiley, N.Y., 1961.
- Gr81** [16] D. H. Griffel, *Applied Functional Analysis*, Wiley, N.Y., 1981.
- Le79** [17] M. Levy, *Universal variational functionals of electron densities, first-order density matrices, and natural spin-orbitals and solutions of the v -representability problem*, Proc. Natl. Acad. Sci. USA 76 (1979) 6062–65.
- Ta57** [18] A. E. Taylor, *Introduction to Functional Analysis*, Wiley and Sons, 1957.
- DW74** [19] L. M. Delves, J. Welsh, *Numerical Solution of Integral Equations*, Clarendon Press, Oxford, 1974.