Fundamentals of Density-Functional Theory

Trygve Helgaker

Centre for Theoretical and Computational Chemistry
Department of Chemistry, University of Oslo, Norway

GdR CORREL Meeting 2015

Amphi Sciences Naturelles, Site Saint Charles, Aix–Marseille Université,
Marseille, France, 8–10 July 2015
Introduction

1. Hohenberg–Kohn theory
   - the Hohenberg–Kohn theorem and concavity
   - the Hohenberg–Kohn and Lieb variation principles
   - $\nu$-representable densities and ground-state potentials

2. Levy–Lieb constrained-search theory
   - $\mathcal{N}$-representable densities
   - the Levy–Lieb constrained-search functional

3. Lieb convex-conjugate theory
   - convex functions and convex conjugation
   - the Lieb convex-conjugate functional
   - Hohenberg–Kohn and Lieb variation principles

4. Discontinuity and nondifferentiability of the universal functional
   - subgradients and subdifferentials
   - Hohenberg–Kohn and Lieb optimality conditions

Literature:
Consider the electronic Hamiltonian of an $N$-electron atom or molecule

$$H(\nu) = T + W + \sum_i \nu(r_i)$$

- $T$ and $W$ are system-independent operators

$$T = -\frac{1}{2} \sum_i \nabla_i^2 \quad W = \sum_{i>j} r_{ij}^{-1}$$

- $\nu(r)$ is a system-dependent multiplicative external potential

Typically, we are interested in potentials that can bind $N$ electrons

$$\nu_N = \{ \nu \mid H(\nu) \text{ has an } N\text{-electron ground state} \}$$

- If $\nu \in \nu_N$, then an $N$-electron ground state exists

$$H(\nu)\Psi = E(\nu)\Psi, \quad \nu \in \nu_N$$

- We shall later also consider potentials that cannot bind $N$ electrons

The energy may be calculated as an expectation value:

$$E(\nu) = \langle \Psi | H(\nu) | \Psi \rangle = \langle \Psi | T + W | \Psi \rangle + (\nu | \rho)$$

- The system interacts with the external potential $\nu$ through the density $\rho$:

$$(\nu | \rho) = \int \nu(r)\rho(r) \, dr \quad \leftarrow \text{interaction or pairing}$$

- The relationship between $\nu$ and $\rho$ is at the heart of DFT
Different potentials have different wave functions

- We denote by $\Psi_v$ a normalised ground-state wave function associated with $v$:

$$H(v)\Psi_v = E(v)\Psi_v$$

- Assume that $v_1, v_2 \in \mathcal{V}_N$ have a common ground-state wave function $\Psi$:

$$\begin{align*}
H(v_1)\Psi_{v_1} &= E(v_1)\Psi_{v_1} \\
H(v_2)\Psi_{v_2} &= E(v_2)\Psi_{v_2}
\end{align*}$$

$$\Psi_{v_1} = \Psi_{v_2} = \Psi$$

- Subtracting the two Schrödinger equations, we obtain

$$[H(v_1) - H(v_2)]\Psi = \sum_i [v_1(r_i) - v_2(r_i)]\Psi = [E(v_1) - E(v_2)]\Psi$$

- Eliminating $\Psi$ from both sides of the last equation, we find

$$\sum_i [v_1(r_i) - v_2(r_i)] = E(v_1) - E(v_2)$$

- Identical ground-state wave functions have potentials that are identical up to a constant:

$$\Psi_{v_1} = \Psi_{v_2} \implies v_1(r) = v_2(r) + c$$

Different external potentials have different wave functions

$$v_1(r) \neq v_2(r) + c \implies \Psi_{v_1} \neq \Psi_{v_2}$$
The Hohenberg–Kohn theorem

- Consider two different potentials $v_1$ and $v_2$ with ground states $\Psi_1$ and $\Psi_2$:
  \[ v_1(r) \neq v_2(r) + c \implies \Psi_1 \neq \Psi_2, \quad v_i \in \mathcal{V}_N \]

- are the corresponding densities $\rho_1$ and $\rho_2$ also different?

- Invoking the Rayleigh–Ritz variation principle for the two ground states, we obtain
  \[
  E(v_1) < \langle \Psi_2 | H(v_1) | \Psi_2 \rangle = \langle \Psi_2 | H(v_2) | \Psi_2 \rangle + (v_1 - v_2 | \rho_2) = E(v_2) + (v_1 - v_2 | \rho_2) \\
  E(v_2) < \langle \Psi_1 | H(v_2) | \Psi_1 \rangle = \langle \Psi_1 | H(v_1) | \Psi_1 \rangle - (v_1 - v_2 | \rho_1) = E(v_1) - (v_1 - v_2 | \rho_1)
  \]

- Adding the two inequalities, we arrive at the strict inequality
  \[
  E(v_1) + E(v_2) < E(v_1) + E(v_2) + (v_1 - v_2 | \rho_2 - \rho_1)
  \]

- This result gives a contradiction unless the two densities are different

The Hohenberg–Kohn (HK) theorem: different potentials have different densities

- A key point in the HK proof are the following relations of the ground-state energy
  \[
  u \neq v + c \in \mathcal{V}_N \implies E(u) < E(v) + (u - v | \rho) \\
  u = v + c \in \mathcal{V}_N \implies E(u) = E(v) + (u - v | \rho)
  \]
HK theorem from concavity of ground-state energy

- The ground-state energy is **concave** in the external potential

$$E(v) = \min_{\Psi} \langle \Psi | H(v) | \Psi \rangle$$

- The concavity of $E(v)$ may be understood in the following two-step manner:
  1. From A to B, the energy increases linearly since $H(v)$ is linear in $v$ and $\Psi_0$ is fixed.
  2. From B to C, the energy decreases as the wave function relaxes to the ground state $\Psi_1$

$$\langle \Psi_1 | H(v_1) | \Psi_1 \rangle \leq \langle \Psi_0 | H(v_1) | \Psi_0 \rangle = \langle \Psi_0 | H(v_0) | \Psi_0 \rangle + (v_1 - v_0 | \rho_0)$$

- Note: the density $\rho_0$ is the slope of the ground-state energy at $v_0$:

$$E(v_1) \leq E(v_0) + (v_1 - v_0 | \rho_0)$$
HK theorem: strict and non-strict concavity

- The concavity of the energy $E(\nu)$ follows from two circumstances:
  - the linearity of $H(\nu)$ changes the energy linearly from $\nu_0$ to $\nu_1$ for fixed $\Psi_0$
  - the variation principle lowers the energy from $\Psi_0$ to $\Psi_1$ for fixed $\nu_1$

- There are two cases to consider: strict concavity (left) and nonstrict concavity (right)
  
  \[
  [H(\nu_0), H(\nu_1)] \neq 0 \Rightarrow \Psi_1 \neq \Psi_0 \Rightarrow \rho_1 \neq \rho_0 \quad \text{strict concavity}
  \]
  \[
  [H(\nu_0), H(\nu_1)] = 0 \Rightarrow \Psi_1 = \Psi_0 \Rightarrow \rho_1 = \rho_0 \quad \text{non-strict concavity}
  \]

- We have strict concavity and different densities except if $\nu_1 - \nu_0 = c$ is a scalar
  - the Hohenberg–Kohn theorem: the density determines the potential up to a constant
  - with vector potentials, non-strict concavity occurs more generally
HK mapping between potentials and densities

**v**-representable densities

Each density $\rho$ is the ground-state density of at most one external potential $v + c$.

- those that arise from some potential $v_\rho$ are said to be **v**-representable
- those that do not are said to be non-**v**-representable

HK theory thus sets up a mapping between the following two sets:

\[

v \in \mathcal{V}_N = \{ v \mid H(v) \text{ has an } N\text{-electron ground state} \} \\
\rho \in \mathcal{A}_N = \{ \rho \mid \rho \text{ comes from an } N\text{-electron ground state} \}

\]

- the density determines the potential up to a constant
- conversely, the potential determines the density up to a degenerate set
Hohenberg–Kohn and Lieb variation principles

- In HK theory, the potential and wave function are determined by the density

\[ \rho \in \mathcal{A}_N \rightarrow v_{\rho} + c \rightarrow \gamma \Psi_{\rho} \]

- the additive constant \( c \) and the phase factor \( \gamma \) with \( \gamma^* \gamma = 1 \) are undetermined

- We now introduce the Hohenberg–Kohn functional:

\[ F_{HK}(\rho) = \langle \Psi_{\rho} | T + W | \Psi_{\rho} \rangle = E(v_{\rho}) - (v_{\rho} | \rho) , \quad \rho \in \mathcal{A}_N \]

- it is independent of \( c \) and \( \gamma \) and unique also for degenerate systems

- From the Rayleigh–Ritz variation principle, we obtain for arbitrary \( v \in \mathcal{V}_N \) the inequality

\[ F_{HK}(\rho) + (v | \rho) = \langle \Psi_{\rho} | T + W | \Psi_{\rho} \rangle + (v | \rho) = \langle \Psi_{\rho} | H(v) | \Psi_{\rho} \rangle \geq E(v) \]

which may be written in two equivalent ways:

\[ E(v) \leq F_{HK}(\rho) + (v | \rho) , \quad F_{HK}(\rho) \geq E(v) - (v | \rho) \]

- These inequalities may be sharpened into equalities, yielding two variation principles:

<table>
<thead>
<tr>
<th>Hohenberg–Kohn and Lieb variation principles</th>
</tr>
</thead>
<tbody>
<tr>
<td>[ E(v) = \min_{\rho \in \mathcal{A}<em>N} \left( F</em>{HK}(\rho) + (v</td>
</tr>
<tr>
<td>[ F_{HK}(\rho) = \max_{v \in \mathcal{V}_N} \left( E(v) - (v</td>
</tr>
</tbody>
</table>
Hohenberg–Kohn and Lieb variation principles

\[ F(\rho) = \max_v \left( E(v) - (v|\rho) \right) \]

\[ E(v) = \min_\rho \left( F(\rho) + (v|\rho) \right) \]
Hohenberg–Kohn theory summarised

The Hohenberg–Kohn functional and the Hohenberg–Kohn variation principle

\[
F_{HK}(\rho) = \langle \Psi_{\rho} | T + W | \Psi_{\rho} \rangle, \quad \rho \in \mathcal{A}_N
\]

\[
E(\nu) = \min_{\rho \in \mathcal{A}_N} (F_{HK}(\rho) + (\rho | \nu)), \quad \nu \in \mathcal{V}_N
\]

- We may obtain the ground-state energy by a variation over densities
  - there is no need to involve the wave function!

- However, certain difficulties remain
  - the explicit form of \( F_{HK} \) is unknown: this does not concern us here
  - neither of the sets \( \mathcal{A}_N \) and \( \mathcal{V}_N \) are explicitly known

- shall first see how \( \mathcal{A}_N \) and \( \mathcal{V}_N \) may be extended to explicitly known sets
  - the Levy–Lieb constrained-search functional

- Next we shall explore Lieb’s variation principle, extending it to these explicitly known sets

\[
F_{HK}(\rho) = \max_{\nu \in \mathcal{V}_N} (E(\nu) - (\nu | \rho)), \quad \rho \in \mathcal{A}_N
\]
Bound and unbound systems

- In HK theory, we have restricted ourselves to electronic eigenstates
  \[ H(\nu)\Psi_\nu = E(\nu)\Psi_\nu, \quad \nu \in \mathcal{V}_N \]

- We now broaden our scope and consider the Rayleigh–Ritz variation principle
  \[ E(\nu) = \inf_{\Psi} \langle \Psi | T + W + \sum_i \nu(r_i) | \Psi \rangle, \quad \nu \in \mathcal{U} \supset \mathcal{V}_N \]

  - here \( \mathcal{U} \) is the set of all potentials such that the energy is finite
  - note: this set is explicitly known: \( \mathcal{U} = L^{3/2} + L^\infty \)

- Since \( \nu \in \mathcal{V}_N \) is not assumed, there may not be a minimising wave function (an eigenstate)
  - we therefore determine an infimum (greatest lower bound) rather than a minimum

- Example: the oxygen atom has an electronic ground state only for \( N \leq 9 \)

  - for \( N > 9 \), no ground state exists and the excess electrons are not bound
  - the infimum is equal to the energy of \( O^- \), with the excess electrons at rest infinitely far away
Levy–Lieb constrained-search theory

We now perform the Rayleigh–Ritz variation principle for $N$ electrons in two nested steps:

$$
E(v) = \inf_{\Psi} \langle \Psi | T + W + \sum_i v(r_i) | \Psi \rangle
$$

$$
= \inf_{\rho} \inf_{\Psi \mapsto \rho} \langle \Psi | T + W + \sum_i v(r_i) | \Psi \rangle
$$

$$
= \inf_{\rho} \left[ \inf_{\Psi \mapsto \rho} \langle \Psi | T + W | \Psi \rangle + (\rho | v) \right], \quad v \in \mathcal{U}
$$

- an outer minimisation over $\rho$ and an inner minimisation over $\Psi \mapsto \rho$

Introducing the Levy–Lieb constrained-search functional, we obtain

$$
F_{LL}(\rho) = \inf_{\Psi \mapsto \rho} \langle \Psi | T + W | \Psi \rangle, \quad \rho \in \mathcal{I}_N \subset X = L^3 \cap L^1
$$

$$
E(v) = \inf_{\rho} (F_{LL}(\rho) + (\rho | v)), \quad v \in \mathcal{U} = X^* = L^{3/2} + L^{\infty}
$$

where we search over all $N$-representable densities:

$$
\mathcal{I}_N = \{ \rho(r) | \rho \text{ can be obtained from some } N\text{-electron wave function } \Psi \}
$$

The set of $N$-representable densities is explicitly known:

$$
\mathcal{I}_N = \{ \rho(r) | \rho(r) \geq 0, \int \rho(r) \, dr = N, \int |\nabla \rho^{1/2}(r)|^2 \, dr < \infty \}
$$

- among all $\Psi \mapsto \rho$, there is always a determinantal wave function $\Psi_{\text{det}} \mapsto \rho$
- the ‘unknown set’ $\mathcal{A}_N$ is dense in the ‘known’ set $\mathcal{I}_N$
Hohenberg–Kohn and Levy–Lieb theories compared

**Hohenberg–Kohn theory:**

\[
F_{HK}(\rho) = \langle \Psi_\rho | T + W | \Psi_\rho \rangle, \quad \rho \in \mathcal{A}_N
\]

\[
E(\nu) = \min_{\rho \in \mathcal{A}_N} \left( F_{HK}(\rho) + (\rho|\nu) \right), \quad \nu \in \mathcal{V}_N
\]

**The Levy–Lieb theory:**

\[
F_{LL}(\rho) = \inf_{\Psi \mapsto \rho} \langle \Psi | T + W | \Psi \rangle, \quad \rho \in \mathcal{I}_N
\]

\[
E(\nu) = \inf_{\rho \in \mathcal{I}_N} \left( F_{LL}(\rho) + (\rho|\nu) \right), \quad \nu \in \mathcal{U}
\]

- We have avoided the unknown domain of Hohenberg–Kohn theory
  - search is over all “reasonable” densities for all “reasonable” potentials
  - however, this theory is still not fully satisfactory

- We obtain the energy by a variational minimisation of \( F_{LL}(\rho) + (\rho|\nu) \)
  - the functional to be minimised should then be as simple as possible

- In particular, we would like it to at most one minimizer (except by degeneracy)
  - this cannot be guaranteed for the Levy–Lieb functional

- We now turn our attention to a density functional with a unique solution: Lieb’s functional
  - we first gives some background in convex analysis and convex conjugation
A function is said to be **convex** if it satisfies the inequality

$$
\lambda f(x_1) + (1 - \lambda) f(x_2) \geq f(\lambda x_1 + (1 - \lambda) x_2), \quad 0 < \lambda < 1
$$

a linear interpolation always overestimates a convex function

For a **strictly convex function**, we may replace $\geq$ by $>$ above

A function $f(x)$ is **concave** if $-f(x)$ is convex
Concavity of the ground-state energy

- Concavity of the ground-state energy is simple to prove
  - it follows from the Rayleigh–Ritz variation principle and the linearity of $H(v)$
- Consider the variationally optimised ground-state energy
  \[ E(v) = \inf_\Psi \langle \Psi | H(v) | \Psi \rangle \]
  - a minimising ground state is not assumed
- Insert $v = \lambda v_1 + (1 - \lambda) v_2$ with $0 < \lambda < 1$ and use linearity of Hamiltonian
  \[ H(\lambda v_1 + (1 - \lambda) v_2) = \lambda H(v_1) + (1 - \lambda) H(v_2) \]
- Insert this Hamiltonian into the Rayleigh–Ritz variation principle:
  \[ E(\lambda v_1 + (1 - \lambda) v_2) = \inf_\Psi \langle \Psi | \lambda H(v_1) + (1 - \lambda) H(v_2) | \Psi \rangle \]
  \[ = \inf_\Psi (\lambda \langle \Psi | H(v_1) | \Psi \rangle + (1 - \lambda) \langle \Psi | H(v_2) | \Psi \rangle) \]
  \[ \geq \lambda \inf_\Psi_1 \langle \Psi_1 | H(v_1) | \Psi_1 \rangle + (1 - \lambda) \inf_\Psi_2 \langle \Psi_2 | H(v_2) | \Psi_2 \rangle \]
  \[ = \lambda E(v_1) + (1 - \lambda) E(v_2) \]
  - separate minimisation of the two terms lowers the energy
- We have now proved concavity:
  \[ E(\lambda v_1 + (1 - \lambda) v_2) \geq \lambda E(v_1) + (1 - \lambda) E(v_2) \]
  - a linear interpolation always underestimates the true ground-state energy
Amusing consequence of concavity: united atom

Consider now the following diatomic potential:

\[ v_{\text{mol}}(r) = -\frac{Z_A}{r_A} - \frac{Z_B}{r_B} = \lambda \left( -\frac{Z_A + Z_B}{r_A} \right) + (1 - \lambda) \left( -\frac{Z_A + Z_B}{r_B} \right) = \lambda v_A(r) + (1 - \lambda) v_B(r) \]

where \( \lambda = \frac{Z_A}{Z_A + Z_B} \) and \( 1 - \lambda = 1 - \frac{Z_A}{Z_A + Z_B} = \frac{Z_B}{Z_A + Z_B} \).

Since \( 0 < \lambda < 1 \), we obtain from the concavity of the energy

\[ E(v_{\text{mol}}) \geq \lambda E(v_A) + (1 - \lambda) E(v_B) = \lambda E(v_A) + (1 - \lambda) E(v_A) = E(v_A) \]

the energy of the molecule is an upper bound to the energy of the united atom.

Conclusion: without nuclear–nuclear repulsion, all molecules would collapse into atoms.
A convex function is **continuous except possibly at the boundary points of its domain**
- in many dimensions all points may be boundary points

A convex function is **not necessarily everywhere differentiable**
- in fact, it may be nowhere differentiable

The universal density functional is nowhere continuous and nowhere differentiable...
Convex functions: supporting lines and stationary points

- Consider a function $f : \mathbb{R} \mapsto \mathbb{R}$ on the real axis
  - a supporting line $h$ to $f$ touches the graph of $f$ and is nowhere above it

- A convex function $f$ has a set of supporting lines everywhere in the interior of its domain
  - the slope of a supporting line at $(x, f(x))$ is called a subgradient of $f$ at $x$

- The condition for a (global) minimum at $x$ is the existence of zero subgradient at $x$
  - note: all local minima of a convex function are global minima
Convex functions constructed from supporting lines I

- A graph of a convex function \( f : \mathbb{R} \mapsto \mathbb{R} \) has supporting lines for all \( x \)
  \( h_y(x) = xy - g(y) \) ← line of slope \( y \) and intercept \(-g(y)\) with the ordinate axis

- Consider the piecewise linear function \( f(x) \) plotted below

- At each \( x \), the function \( f(x) \) is equal to its largest supporting linear line \( h_y(x) \)

Characterisation of convex functions

A function \( f : \mathbb{R} \mapsto \mathbb{R} \) is convex if and only if it can be written in the form

\[
f(x) = \sup_y [xy - g(y)] \quad \leftarrow \text{pointwise supremum of all supporting lines}
\]
A function $f : \mathbb{R} \mapsto \mathbb{R}$ is convex if and only if it can be written in the form

$$f(x) = \sup_y [xy - g(y)] \leftarrow \text{pointwise supremum of all supporting lines}$$

The plots below illustrate this construction for $x^2$, $x + x^4$, $|x| + x^2$ and $\exp(x)$.

- The supporting lines (not the functions) have been plotted at intervals of 0.1.
For a convex function $f : \mathbb{R} \mapsto \mathbb{R}$, we obtain

$$f(x) = \sup_y [xy - g(y)] \implies f(x) \geq xy - g(y)$$

$$\iff g(y) \geq xy - f(x) \iff g(y) = \sup_x [xy - f(x)]$$

Convex conjugate functions

For each convex $f : \mathbb{R} \mapsto \mathbb{R}$, there is an associated convex function $f^* = g$ such that

$$f(x) = \sup_y [xy - f^*(y)] \iff f^*(y) = \sup_x [xy - f(x)]$$

- $f^*$ is called the convex conjugate or Legendre–Fenchel transform of $f$
- conjugate pairs: $f$ and $f^*$ are each other’s conjugate functions
- each convex function contains all information about its conjugate partner

Convex conjugation of $f$ on a general vector space $X$

- we must then require both convexity and lower semi-continuity
- its conjugate function $f^*$ is convex and lower semi-continuous on the dual space $X^*$

Lower semi-continuity is a weak form of continuity

- a continuous functions cannot jump as a limit is reached
- a lower semi-continuous function can jump down but not up as a limit is reached
DFT by convex conjugation $E(v) \leftrightarrow F(\rho)$

- The negative ground-state energy $E$ is convex and continuous
  - it therefore has a convex conjugate: Lieb’s universal density functional $F$
    \[
    E(v) = \inf_{\rho \in X} \left( F(\rho) + (v|\rho) \right) \iff F(\rho) = \sup_{v \in X^*} \left( E(v) - (v|\rho) \right)
    \]
  - these transformations are the Hohenberg-Kohn and Lieb variation principles, respectively
  - the vector spaces are $X = L^3 \cap L^1$ and $X^* = L^{3/2} + L^\infty$

Trygve Helgaker (CTCC, University of Oslo)
The concave envelope $E(v) \rightarrow F(\rho) \leftrightarrow \text{co } E(v)$

- Assume now that $E(v)$ is not concave (not variationally minimised)
  - it still generates a convex $F(\rho)$, conjugate to the concave envelope $\text{co } E(v) \geq E(v)$

The concave envelope $\text{co } E(v)$ is the least concave upper bound to $E(v)$

- excited-state energies are in general not concave
- approximate electronic ground-state energies are in general not concave
Biconjugation and the convex envelope

- The conjugate function $f^*$ is well defined also when $f$ is not convex:
  \[ f^*(y) = \sup_x (xy - f(x)) \]

- We can therefore always form the biconjugate function:
  \[ f^{**}(x) = \sup_y (xy - f^*(y)) \]

  - note: $f = f^{**}$ holds only when $f$ is convex

- We have the following conjugation relationships
  \[ f(x) \rightarrow f^*(y) \leftrightarrow f^{**}(x) \]

- $f^{**}$ is the largest convex lower bound to $f$, known as its convex envelope:
  \[ f^{**} \leq f \quad \text{(arbitrary } f), \quad f^{**} = f \quad \text{(convex } f) \]

- Lieb’s functional is the convex envelope to the Levy-Lieb constrained-search functional
  \[ F = F_{LL}^{**} \leq F_{LL} \]
Conjugate functions $L(\dot{x}) \leftrightarrow H(p)$

- If $f$ is strictly convex and differentiable, $f^*$ is called a Legendre transform
  - Legendre transforms are ubiquitous in physics
- The Lagrangian of classical mechanics is convex in the velocity $\dot{x}$
  \[ L(\dot{x}) = \frac{1}{2} m \dot{x}^2 - V_{\text{pot}} \]
- Its Legendre transform (convex conjugate) is the Hamiltonian:
  \[ H(p) = L^*(p) = \max_{\dot{x}} \left( p\dot{x} - \frac{1}{2} m \dot{x}^2 + V_{\text{pot}} \right) \]
  - the stationary condition identifies the momentum
  \[ p = m\dot{x} \iff \dot{x} = \frac{p}{m} \]
  - substituting $\dot{x} = p/m$ into $H(p)$, we obtain the Hamiltonian
    \[ H(p) = \frac{p^2}{2m} + V_{\text{pot}} \]
- The reciprocal relation (see next slide) is satisfied:
  \[ L'(\dot{x}) = m\dot{x} = p \quad \& \quad H'(p) = \frac{p}{m} = \dot{x} \]
- Legendre transforms are also used in thermodynamics
Reciprocal relations of conjugate functions

- Conjugate functions are related by the conjugate variation principles:
  \[ f(x) = \sup_y [xy - f^*(y)] \iff f^*(y) = \sup_x [xy - f(x)] \]

- From these relations, Fenchel's inequality follows directly:
  \[ f(x) + f^*(y) \geq xy \iff \text{valid for all pairs } (x, y) \]

- Assuming that the maxima can be achieved above, we obtain
  \[ f(xy) + f^*(yx) = xy \iff \text{valid for conjugate pairs } (xy, yx) \]
  - this may not be possible for all \( x \) or all \( y \)

- Differentiation with respect to \( xy \) and \( yx \) yields the reciprocal relations:
  \[ f'(xy) = yx \iff (f^*)'(yx) = xy \]
  - we have here ASSUMED differentiability

- We conclude that the first derivatives of conjugate functions are inverse functions:
  \[ (f')^{-1} = (f^*)' \]
  - this relationship holds in a wider sense in the more general case
Examples of convex conjugate functions

\[ f(x) = 1 + |x|, \quad f^*(y) = \begin{cases} -1, & |y| \leq 0 \\ +\infty, & |y| > 1 \end{cases} \]

\[ g(x) = \sqrt{1 + x^2}, \quad g^*(y) = \begin{cases} -\sqrt{1 - y^2}, & |y| \leq 0 \\ +\infty, & |y| > 1 \end{cases} \]
Discontinuity of the universal density functional

- The ground-state energy $E$ is continuous but what about $F$?
- For a one-electron system, the universal density functional has a simple explicit form:

$$F(\rho) = \frac{1}{2} \int |\nabla \rho^{1/2}(r)|^2 \, dr \quad \leftarrow \text{one-electron kinetic energy}$$

- A one-electron Gaussian density of unit exponent has a finite kinetic energy:

$$\rho(r) = \pi^{-3/2} \exp(-r^2), \quad F(\rho) = 3/4$$

- Let $\{\rho_n\}$ be a sequence that approaches $\rho$ in the norm,

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

while developing increasingly rapid oscillations of increasingly small amplitude:

- The kinetic energy $F(\rho_n)$ is driven arbitrarily high in the sequence and $F$ is not continuous:

$$\lim_{n} F(\rho_n) = +\infty \neq F \left( \lim_{n} \rho_n \right) = 3/4$$

- The universal density functional is everywhere discontinuous and hence nondifferentiable

Nondifferentiability of the universal density functional

- Concavity of $E$ and convexity of $F$ imply great simplicity
  - the Hohenberg–Kohn and Lieb variation principles have only global extrema
    
    $$E(v) = \inf_{\rho \in \mathcal{X}} (F(\rho) + (v|\rho))$$
    
    $$F(\rho) = \sup_{v \in \mathcal{X}^*} (E(v) - (v|\rho))$$

  - advanced methods of convex optimisation theory can be used

- We would like to set up optimality conditions for a given potential $v$
  - typically, the Euler–Lagrange equation is set up
    
    $$\frac{\delta F(\rho)}{\delta \rho(r)} = -v(r) - \mu$$  Euler–Lagrange equation with chemical potential $\mu$

- However, convex functions are not necessarily differentiable
  - in particular, $F$ is discontinuous and therefore not differentiable
  - the solutions therefore cannot be characterised by derivatives

- Two approaches are possible, taking advantage of the convexity of $F$
  - express optimality conditions in terms of subgradients
  - generate a differentiable density functional by Moreau–Yosida regularisation
Differentiability and subdifferentiability

- Convex functions are not necessarily differentiable
  - minima are not characterised by derivatives but instead by subgradients
- The subgradients $y_0$ of $f$ at $x_0$ are the slopes of the supporting lines to $f$ at $x_0$: 
  \[ f(x) \geq f(x_0) + y_0(x - x_0), \quad \forall x \]
- The set off all subgradients of $f$ at $x_0$ is the subdifferential $\partial f(x_0)$ of $f$ at $x_0$
  \[ \partial f_1(0) = [-1, 1], \quad \partial f_2(0) = \{0\}, \quad \partial f_3(0) = \{0\} \]
- A minimum occurs $x_0$ if and only if the subdifferential contains zero
  \[ 0 \in \partial f(x_0) \quad \text{(horizontal supporting line)} \]
- Differentiability follows when $\partial f$ is a singleton and $f$ is continuous
Hohenberg–Kohn optimality conditions

The Hohenberg–Kohn variation principle is a convex minimisation problem:

\[ E(v) = \inf_{\rho \in \mathcal{X}} (F(\rho) + (v|\rho)) \]

- the (global) minimum is attained when the subdifferential of right-hand side contains zero:
  \[ E(v) = F(\rho) + (v|\rho) \iff 0 \in \partial_\rho (F(\rho) + (v|\rho)) \]

- evaluation of the subdifferential:
  \[ \partial_\rho (F(\rho) + (v|\rho)) = \partial_\rho F(\rho) + \partial_\rho (v|\rho) = \partial F(\rho) + \{v\} \]

- subgradient optimality conditions for the global minimum:
  \[ 0 \in \partial F(\rho) + \{v\} \iff -v \in \partial F(\rho) \]

\( \partial F(\rho) \) contains all potentials associated with density \( \rho \): it may be empty or nonempty

- if \( \partial F(\rho) = \emptyset \), then \( \rho \) is not a ground-state density
- if \( \partial F(\rho) \neq \emptyset \), then \( \rho \) is ground-state density and the subdifferential is unique up to scalar:
  \[ \partial F(\rho) = \{-v - \mu \mid \mu \in \mathbb{R}\} \quad \leftarrow \text{Hohenberg–Kohn theorem} \]

- Important result: \( \partial F(\rho) \neq \emptyset \) on a dense subset of \( \mathcal{X} \)
Lieb optimality conditions and reciprocal relations

- The Lieb variation principle is a concave maximisation problem:

  \[ F(\rho) = \sup_\nu (E(\nu) - (\nu|\rho)) \]

- Exactly the same considerations yield the conditions for optimality

### Lieb optimality conditions

\[ F(\rho) = E(\nu) - (\nu|\rho) \iff \rho \in \partial E(\nu) \]

- \( \partial E(\nu) \) contains precisely all ensemble ground-state densities associated with \( \nu \)

  \[ \partial E(\nu) = \text{co}\{\rho_1, \rho_2, \ldots \rho_n\} \]

  - if \( \partial E(\nu) = \{\rho\} \) (nondegenerate state), then it is equal to the derivative
  - if \( \partial E(\nu) = \emptyset \), then \( \nu \) does not support a ground state
  - important result: \( \partial E(\nu) \neq \emptyset \) on a dense subset of \( \mathcal{X}^* \)

- Comparison of the HK and Lieb optimality conditions yields:

### Reciprocal relations: \( -\partial F(\rho) \) and \( \partial E(\nu) \) are inverse multifunctions

\[ E(\nu) = F(\rho) - (\nu|\rho) \iff -\nu \in \partial F(\rho) \iff \rho \in \partial E(\nu) \]
DFT by convex conjugation summarised

The ground-state energy may be represented in two alternative forms:

\[ F(\rho) = \sup_{v \in X^*} \left\{ E(v) - (v|\rho) \right\} \quad \leftarrow \text{energy as a function of density} \]

\[ E(v) = \inf_{\rho \in X} \left\{ F(\rho) + (v|\rho) \right\} \quad \leftarrow \text{energy as a function of potential} \]

- here \( X = L^3 \cap L^1 \) and \( X^* = L^{3/2} + L^\infty \)
- analogous to the energy represented in terms of velocity \( L(\dot{x}) \) and momentum \( H(p) \)

The potential \( v \) and the density \( \rho \) are conjugate variables

- they belong to dual vector spaces such that \((v|\rho) = \int v(r)\rho(r) \, dr\) is finite
- they satisfy the reciprocal relations (assuming well-defined derivatives)

\[ \rho \in \partial E(v) \quad \leftarrow \text{determines } v \text{ when calculating } F(\rho) \text{ from } E(v) \]

\[ -v \in \partial F(\rho) \quad \leftarrow \text{determines } \rho \text{ when calculating } E(v) \text{ from } F(\rho) \]

- since the functionals are either convex or concave, their solutions (if they exist) are unique

The Hohenberg–Kohn theorem:

\[ \partial F(\rho) = \begin{cases} 
\{-v - \mu \mid \mu \in \mathbb{R}\} & (v\text{-representable density}) \\
\emptyset & (\text{non } v\text{-representable density}) 
\end{cases} \]

Convex conjugation highlights the duality of \( \rho \) and \( v \)

- sometimes it is best to work with \( F(\rho) \), other times with \( E(v) \)
- DFT parameterises \( F(\rho) \), molecular mechanics parameterises \( E(v) \)
Comparison of density functionals

We have introduced three universal density functionals

\[
F_{\text{HK}}(\rho) = \langle \Psi_\rho | T + W | \Psi_\rho \rangle, \quad \rho \in \mathcal{A}_N
\]

\[
F_{\text{LL}}(\rho) = \inf_{\Psi \to \rho} \langle \Psi | T + W | \Psi \rangle, \quad \rho \in \mathcal{I}_N
\]

\[
F(\rho) = \sup_{\nu \in \mathcal{X}} \{ E(\nu) - (\nu | \rho) \}, \quad \rho \in \mathcal{X} = L^3 \cap L^1
\]

- These functionals give the same results for ground-state densities:

\[
F(\rho) = F_{\text{LL}}(\rho) = F_{\text{HK}}(\rho), \quad \rho \in \mathcal{A}_N
\]

- Only the Levy–Lieb and Lieb functionals are defined for other densities

\[
F(\rho) = F_{\text{LL}}^{**}(\rho) \leq F_{\text{LL}}(\rho), \quad \rho \in \mathcal{I}_N
\]

- The Lieb functional is the constrained-search functional for ensembles

\[
F(\rho) = \inf_{\Gamma \to \rho} \text{tr} \Gamma (T + W)
\]

- The Levy–Lieb constrained search functional does not obey the reciprocal relations:

\[
E(\nu) = F(\rho) - (\nu | \rho) \iff -\nu \in \partial F(\rho) \iff \rho \in \partial E(\nu)
\]

\[
E(\nu) = F_{\text{LL}}(\rho) - (\nu | \rho) \iff -\nu \in \partial F_{\text{LL}}(\rho) \implies \rho \in \partial E(\nu)
\]
Grand-canonical density–functional theory

- We have studied DFT for a fixed (integral) particle number
  - it may be generalised to arbitrary (variable) particle numbers
  - system described by a grand-canonical ensemble density matrix
    \[
    \hat{\gamma} = \sum_{iN} p_{iN} |\Psi_{iN}\rangle\langle\Psi_{iN}|, \quad p_{iN} \geq 0, \quad \sum_{iN} p_{iN} = 1
    \]
- The energy \( E(\nu, N) \) is concave in \( \nu \) and convex in \( N \)
  - it may be transformed in a similar manner, yielding
    \[
    E(\nu, N) = \sup_{\mu} \inf_{\rho} (F(\rho) + (\nu - \mu|\rho) + \mu N)
    \]
  - the \( \mu \) conjugate to \( N \) is the chemical potential
  - the universal density function \( F \) is convex and defined for all particle numbers
- The optimality conditions are then
  \[
  -\nu - \mu \in \partial F(\rho), \quad N = \int \rho(r) \, dr
  \]
- The Hohenberg–Kohn theorem now becomes
  \[
  \partial F(\rho) = \begin{cases}
    \{-\nu - \mu\} & -I_N \leq \mu \leq -I_{N+1}, \quad \text{(integral } N, \nu\text{-representable } \rho) \\
    \{-\nu\} & \text{(nonintegral } N, \nu\text{-representable } \rho) \\
    \emptyset & \text{(non-\nu\text{-representable } \rho)}
  \end{cases}
  \]
- \( I_N = E(\nu, N - 1) - E(\nu, N) \) is the ionisation potential
- \( \nu \) is uniquely determined by \( \rho \) at nonintegral \( N \)
- \( \nu \) is determined to within a scalar in the interval \( [-I_N, -I_{N+1}] \) at integral \( N \)
Conclusions

- **Convex analysis** is the natural mathematical framework for DFT
  - convex conjugation
  - subgradient and subdifferentials

- **Density-functional theory** follows from concavity and continuity of the ground-state energy
  - Lieb’s functional is the conjugate to the ground-state energy
  - it is convex but neither differentiable nor continuous
  - optimality conditions are best given in terms of subdifferentials
  - subdifferentials give the mapping from density to potentials and vice versa

- **Acknowledgements:**
  - Alex Borgoo, Ulf Ekström, Simen Kvaal, Elisa Rebolini, Sarah Reimann, Espen Sagvolden, Andy Teale, Erik Tellgren
  - Poul Jørgensen, Jeppe Olsen

- **Support:**
  - ERC advanced grant ABACUS
  - Norwegian Research Council for Centre of Excellence CTCC