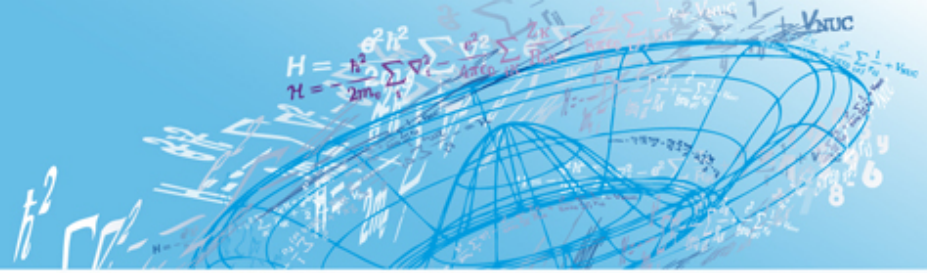


Differentiable but Exact Formulation of Density-Functional Theory

Simen Kvaal[†], Ulf Ekström[†], Andy Teale[†‡], Trygve Helgaker[†]

[†] Centre for Theoretical and Computational Chemistry (CTCC), Department of Chemistry, University of Oslo, Norway
[‡] School of Chemistry, University of Nottingham, Nottingham, UK

Current Topics in Theoretical Chemistry (CTTC),
Sunrise Nha Trang Hotel, Nha Trang, Vietnam, 25–29 August 2014

Density-functional theory (DFT)

- In DFT, the ground-state energy for a given potential v and electron number N is given by

$$E(v) = \min_{\rho \to N} \big(F(\rho) + (v|\rho) \big) \quad \text{Hohenberg–Kohn (HK) variation principle}$$

  where ρ is the density and

$$F(\rho) = \inf_{\Psi \mapsto \rho} \langle \Psi \,|\, T + W \,|\, \Psi \rangle \quad \text{universal density functional}$$

$$(v|\rho) = \int v(\mathbf{r}) \rho(\mathbf{r}) \, \mathrm{d}\mathbf{r} \quad \text{interaction of density with potential}$$

- The solution of the HK minimisation problem is given by

$$\frac{\delta F(\rho)}{\delta \rho(\mathbf{r})} = -v(\mathbf{r}) - \mu \quad \text{Euler–Lagrange equation with chemical potential } \mu$$

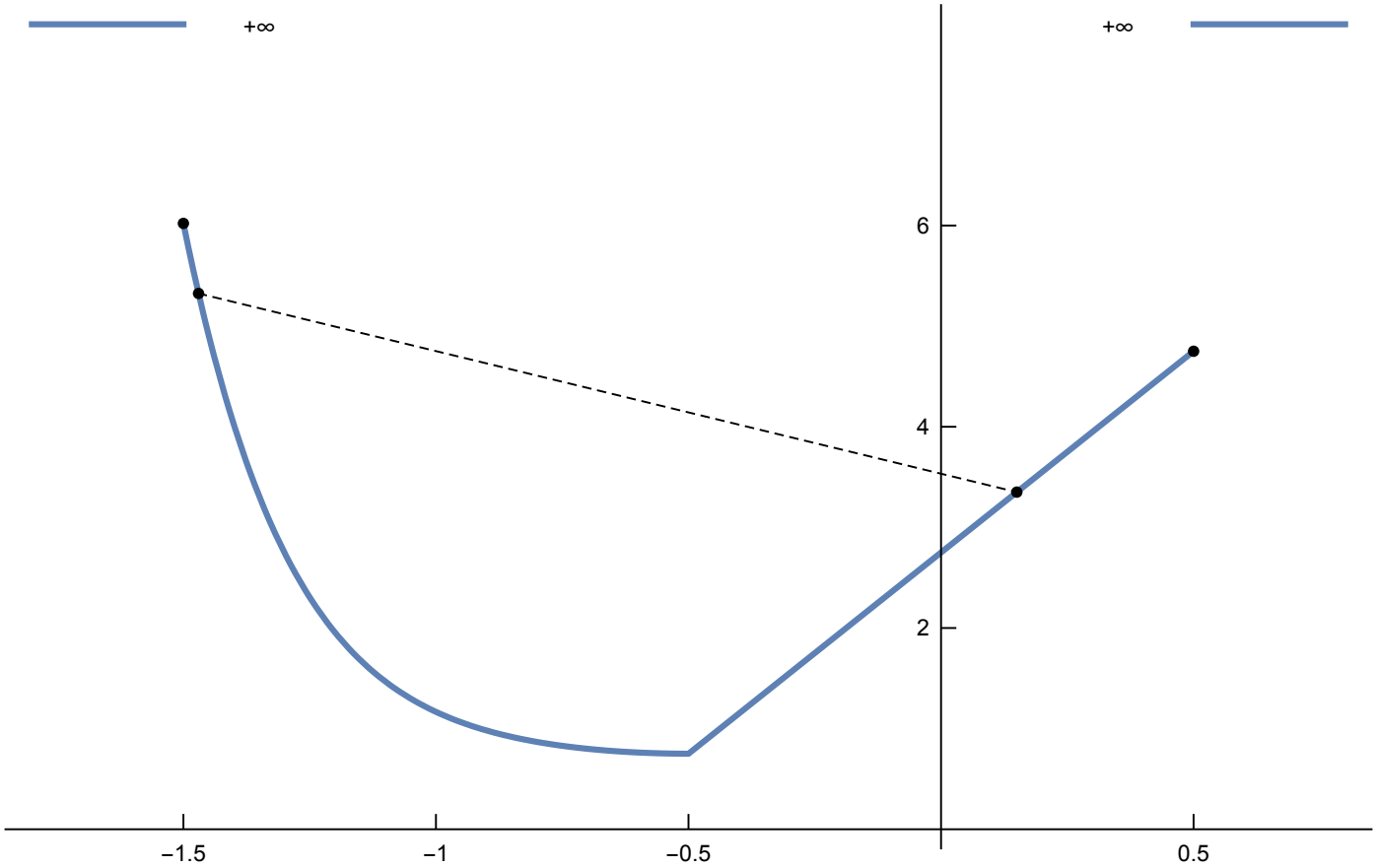
  - clearly, an underlying assumption is that F is differentiable, at least for ground-state densities
  - however, F is everywhere discontinuous and therefore nondifferentiable

- Outline:
  - nondifferentiability of the universal density functional in standard DFT
  - Moreau–Yosida regularised (but exact) DFT with a differentiable density functional
  - noninteracting representability problem of Kohn–Sham theory
  - relationship to regularisation in OEP calculations

- S. Kvaal, U. Ekström, A. Teale, and T. Helgaker, JCP **140**, 18A518 (2014)

Properties of universal density functional

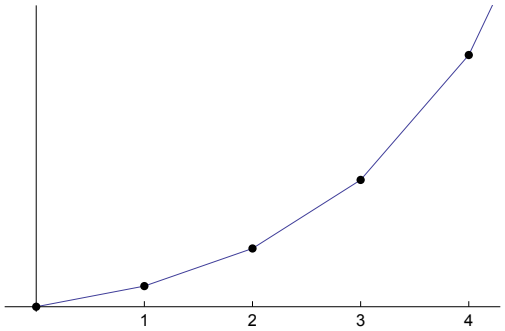
- The universal density function F is less regular than we could hope for
- In its most regular form, F is convex and lower semi-continuous
  - convexity ensures that all local minima are global in the HK variation principle
  - lower semi-continuity allows functions to jump down but not up as limits are taken



- But F is neither continuous nor differentiable

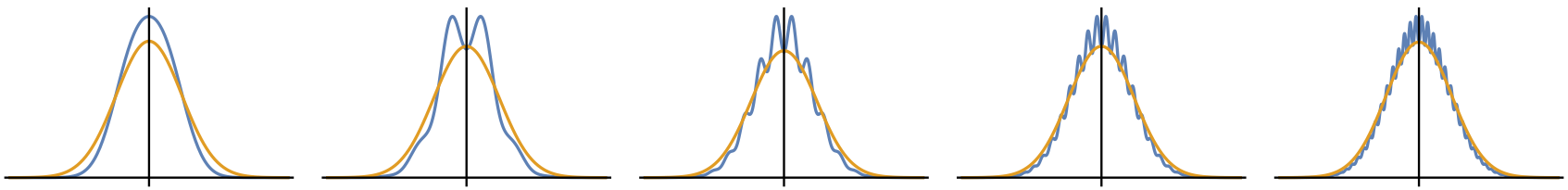
Nondifferentiability of universal density functional

- Derivative discontinuities at integral particle numbers
  - kinks in F at integral particle numbers (gives jump discontinuities in the KS potential)
  - J. P. Perdew, R. G. Parr, M. Levy, and J. L. Balduz, *Phys. Rev. Lett.* **49**, *1691 (1982)*



- Nearby densities of infinite kinetic energy
  - F is unbounded above in the neighbourhood of any density
  - arbitrarily close to any 'reasonable density' are densities of arbitrarily high kinetic energy
  - kinetic energy can be driven arbitrarily high by adding rapid oscillations of small amplitude
  - P. E. Lammert, *Int. J. Quantum Chem.* **107**, *1943 (2007)*

$$\int \rho(\mathbf{r})\,\mathrm{d}\mathbf{r} = 1, \quad T_\mathrm{W}(\rho) = 3/4, \quad \lim_{n\to+\infty} \|\rho - \rho_n\|_p = 0, \quad \lim_{n\to+\infty} T_\mathrm{W}(\rho_n) = +\infty$$



Lieb variation principle

- The nondifferentiability of F makes it awkward to work with
  - we shall develop a more regular DFT with a differentiable density functional
  - we must first examine the relationship between E and F more carefully
- Lieb showed that the energy and density functional are symmetrically related:

$$E(v) = \inf_{\rho} \big(F(\rho) + (v|\rho)\big) \quad \text{the Hohenberg–Kohn variation principle (1964)}$$

$$F(\rho) = \sup_{v} \big(E(v) - (v|\rho)\big) \quad \text{the Lieb variation principle (1983)}$$

- These are alternative attempts at sharpening the same inequality into an equality

$$E(v) \leq F(\rho) + (v|\rho) \iff F(\rho) \geq E(v) - (v|\rho)$$

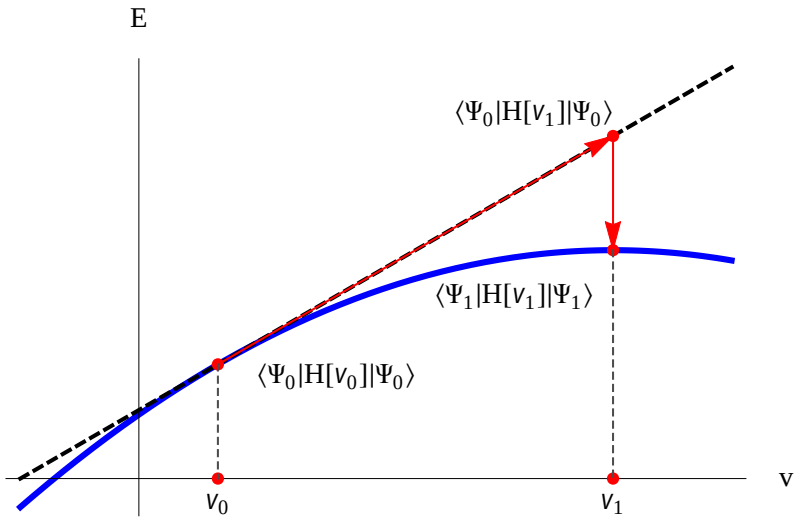
- In the language of convex analysis, E and F are said to be conjugate functions
  - Fenchel conjugates, Legendre–Fenchel transforms...

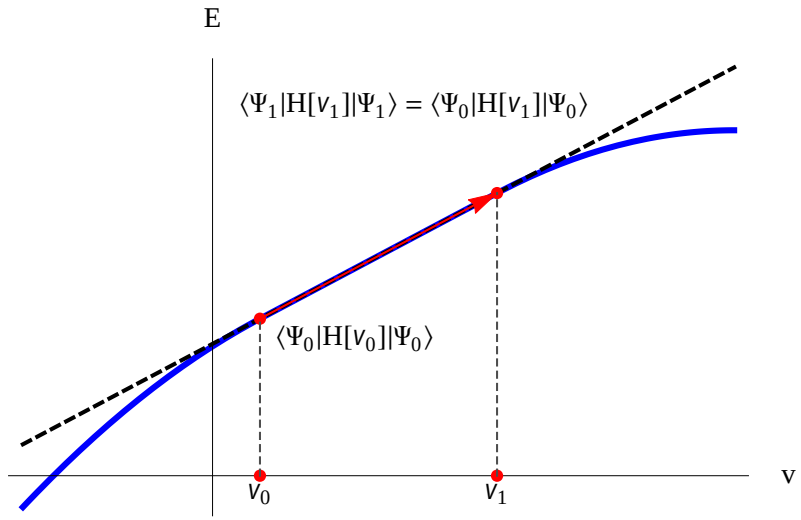
$$E(\text{concave}) \leftrightarrow F(\text{convex})$$

  - each upper semi-continuous concave function has a lower semi-continuous convex partner function
  - conjugate functions contain the same information, represented in different manners
  - each property of one function is exactly reflected in some property of its conjugate function
- Being continuous and concave, E has a lower semi-continuous partner function F
  - the nondifferentiability of F follows from the properties of E
  - we shall modify E in a trivial manner to make F differentiable

Concavity of the ground-state energy

- The concavity of the energy $E(v)$ follows from two circumstances:
  - the linearity of $H(v)$ changes the energy linearly from v_0 to v_1 for fixed Ψ_0
  - the variation principle lowers the energy concavely from Ψ_0 to Ψ_1 for fixed v_1





- There are two cases to consider: strict concavity (left) and nonstrict concavity (right)

$$[H(v_0), H(v_1)] \neq 0 \;\Rightarrow\; \Psi_1 \neq \Psi_0 \;\Rightarrow \rho_1 \neq \rho_0 \quad \text{strict concavity}$$
$$[H(v_0), H(v_1)] = 0 \;\Rightarrow\; \Psi_1 = \Psi_0 \;\Rightarrow \rho_1 = \rho_0 \quad \text{non-strict concavity}$$

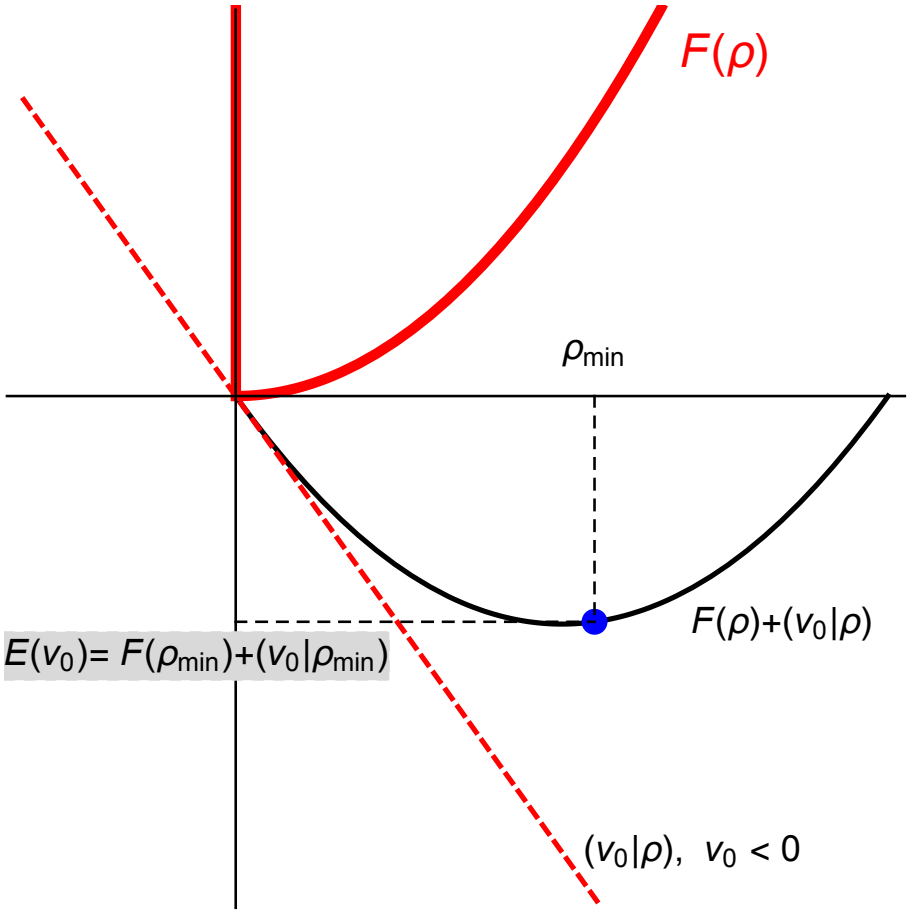
  - note: the density is given by the derivative of the blue curve

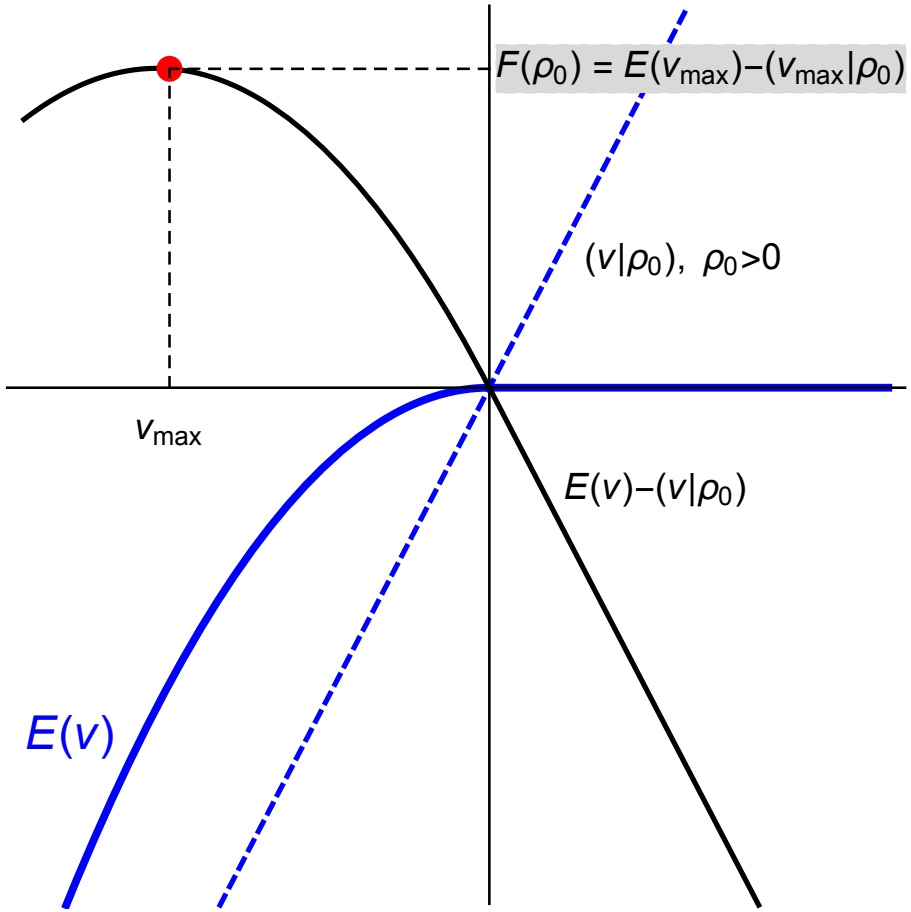
- We have strict concavity and different densities except if $v_1 - v_0 = c$ is a scalar
  - the Hohenberg–Kohn theorem: the density determines the potential up to a constant
  - with vector potentials, non-strict concavity occurs more generally

Hohenberg–Kohn and Lieb variation principles

- Hohenberg–Kohn (left) and Lieb (right) variation principles

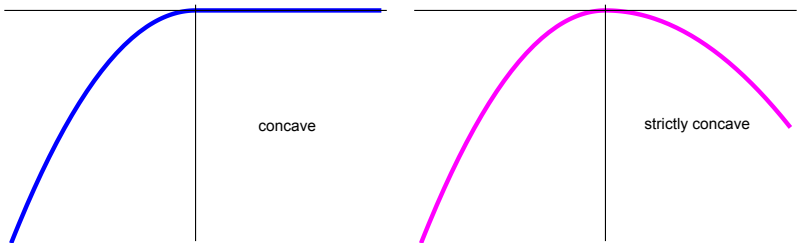
$$E(v) = \inf_\rho \left(F(\rho) + (v|\rho)\right) \iff F(\rho) = \sup_v \left(E(v) - (v|\rho)\right)$$





DFT with a differentiable density functional

- The ground-state energy $E(v)$ is concave but not strictly concave



- Consider now the following modified ground-state energy E_γ, easily obtained from E:

$$E_\gamma(v) = E(v) - \tfrac{1}{2}\gamma\|v\|^2, \quad \gamma > 0$$

  - E_γ is strictly concave and supercoercive: $E_\gamma(v)/\|v\| \to -\infty$ as $\|v\| \to +\infty$

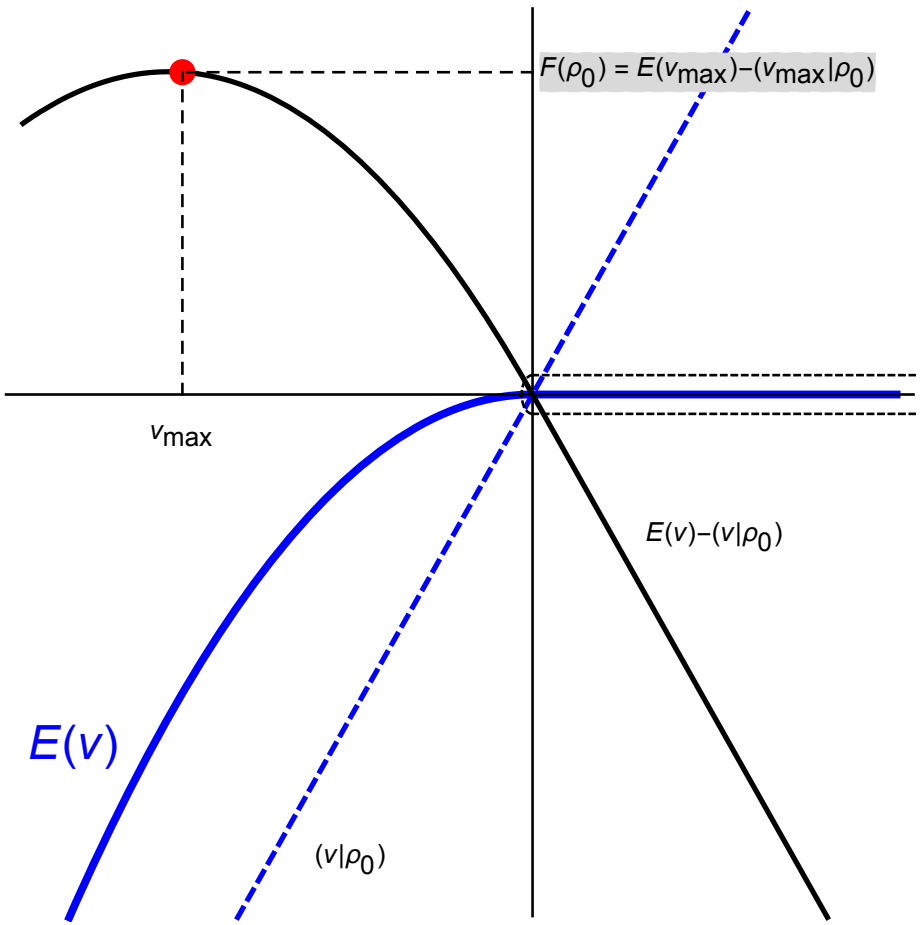
- This trivial modification of the energy ensures differentiability of the density functional

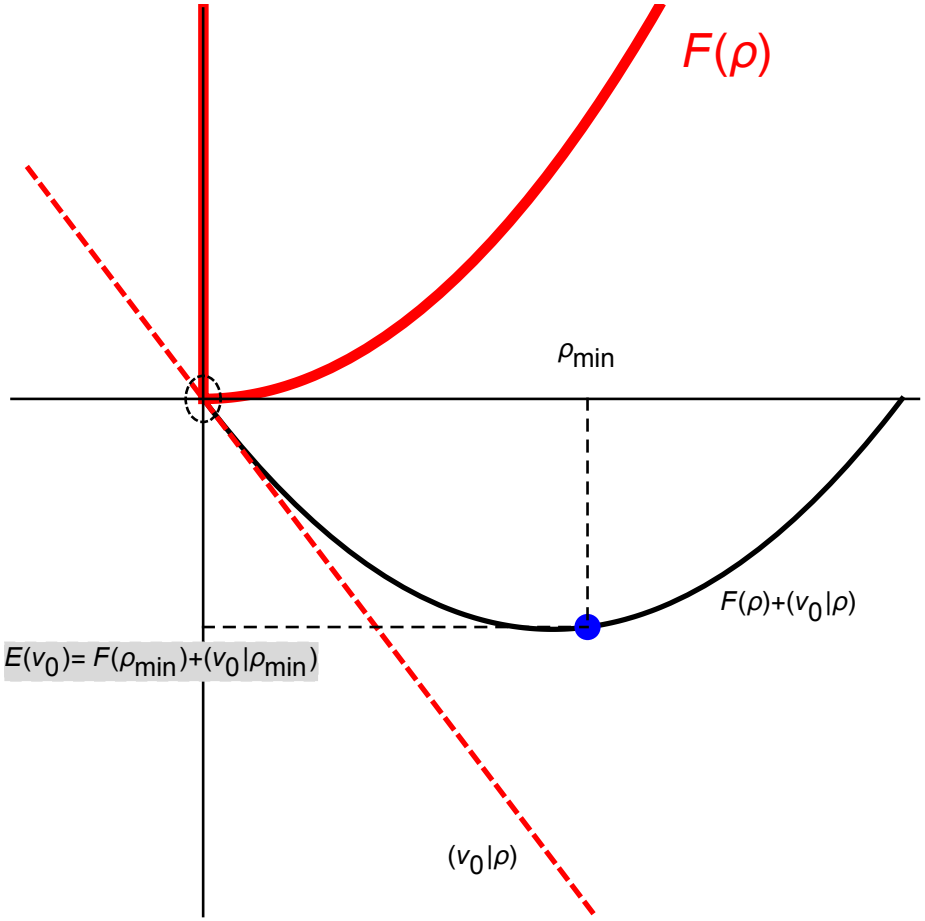
$$E_\gamma(v) = \inf_\rho \big(F_\gamma(\rho) + (v|\rho)\big) \quad \text{Hohenberg–Kohn variation principle}$$

$$F_\gamma(\rho) = \sup_v \big(E_\gamma(v) - (v|\rho)\big) \quad \text{Lieb variation principle}$$

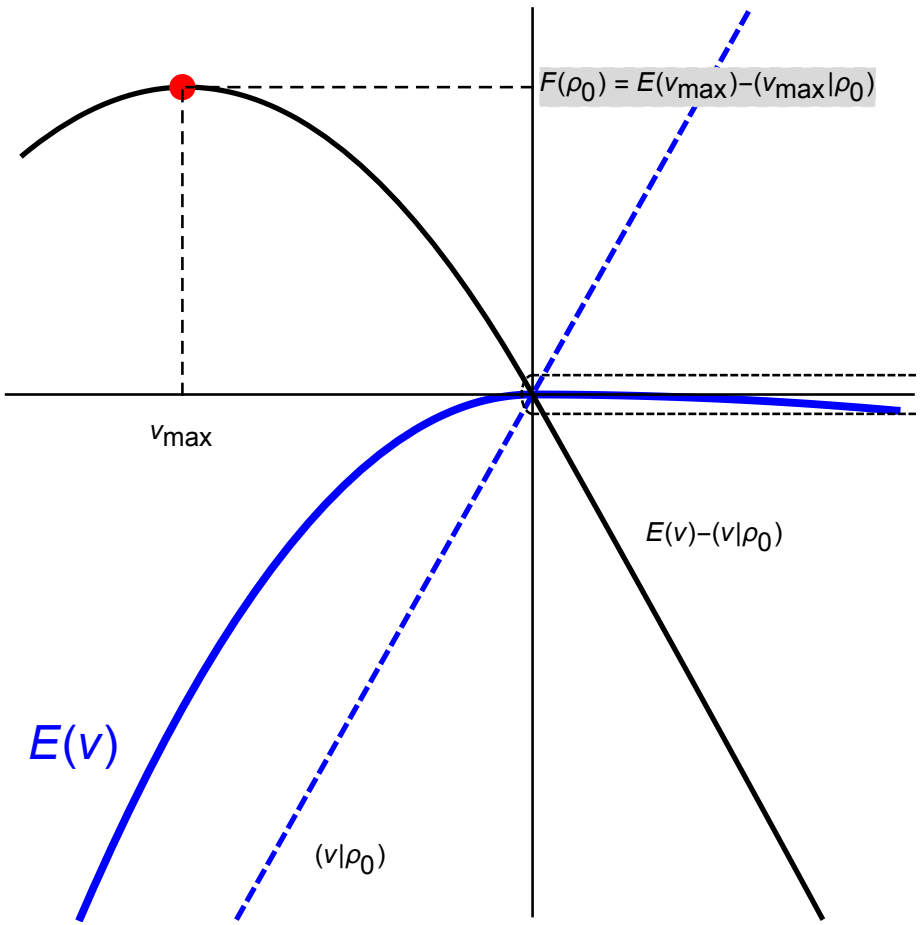
- This procedure is known as Moreau–Yosida (MY) regularisation
  - a standard technique in convex optimisation
  - it yields an alternative Moreau–Yosida DFT, with a differentiable density functional
- For Coulomb potentials v to be square integrable, the system must be in a (large) box
  - Hilbert-space methods can now be used

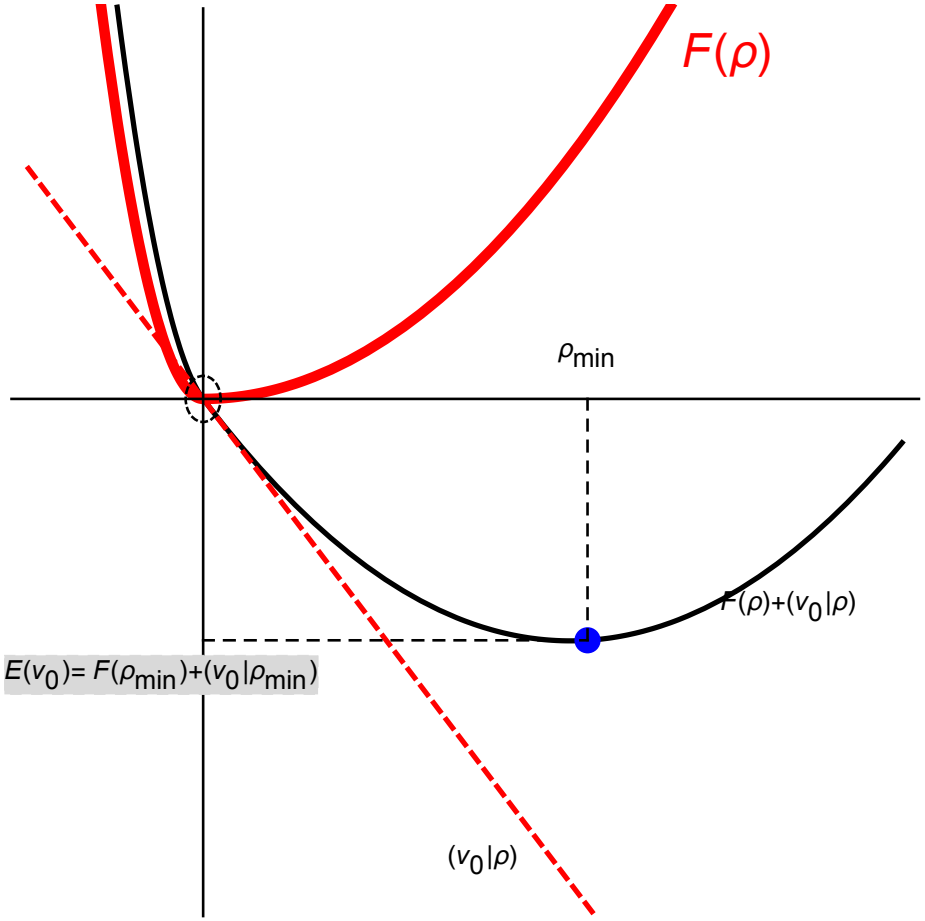
Standard DFT with conjugate functions $E \leftrightarrow F$





Moreau–Yosida DFT with conjugate functions $E_\gamma \leftrightarrow F_\gamma$





Moreau envelope F_γ and proximal density ρ_γ

- We have defined F_γ as the conjugate of E_γ by the Lieb variation principle:

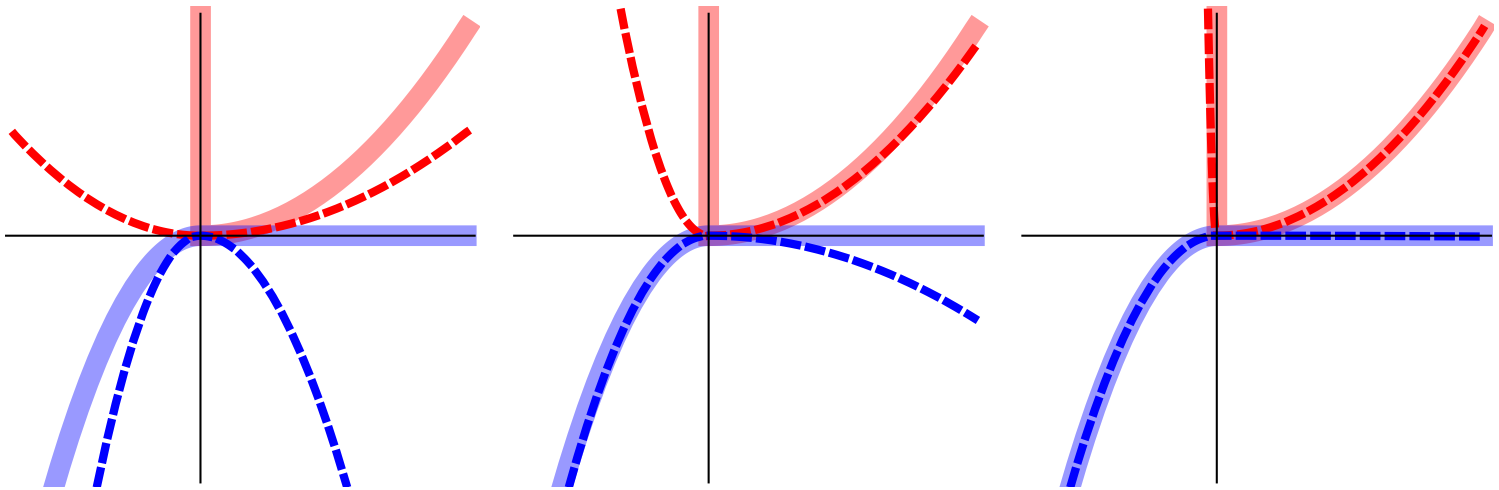
$$F_\gamma(\rho) = \sup_v \big(E_\gamma(v) - (v|\rho)\big)$$

- We may express F_γ explicitly in terms of F as

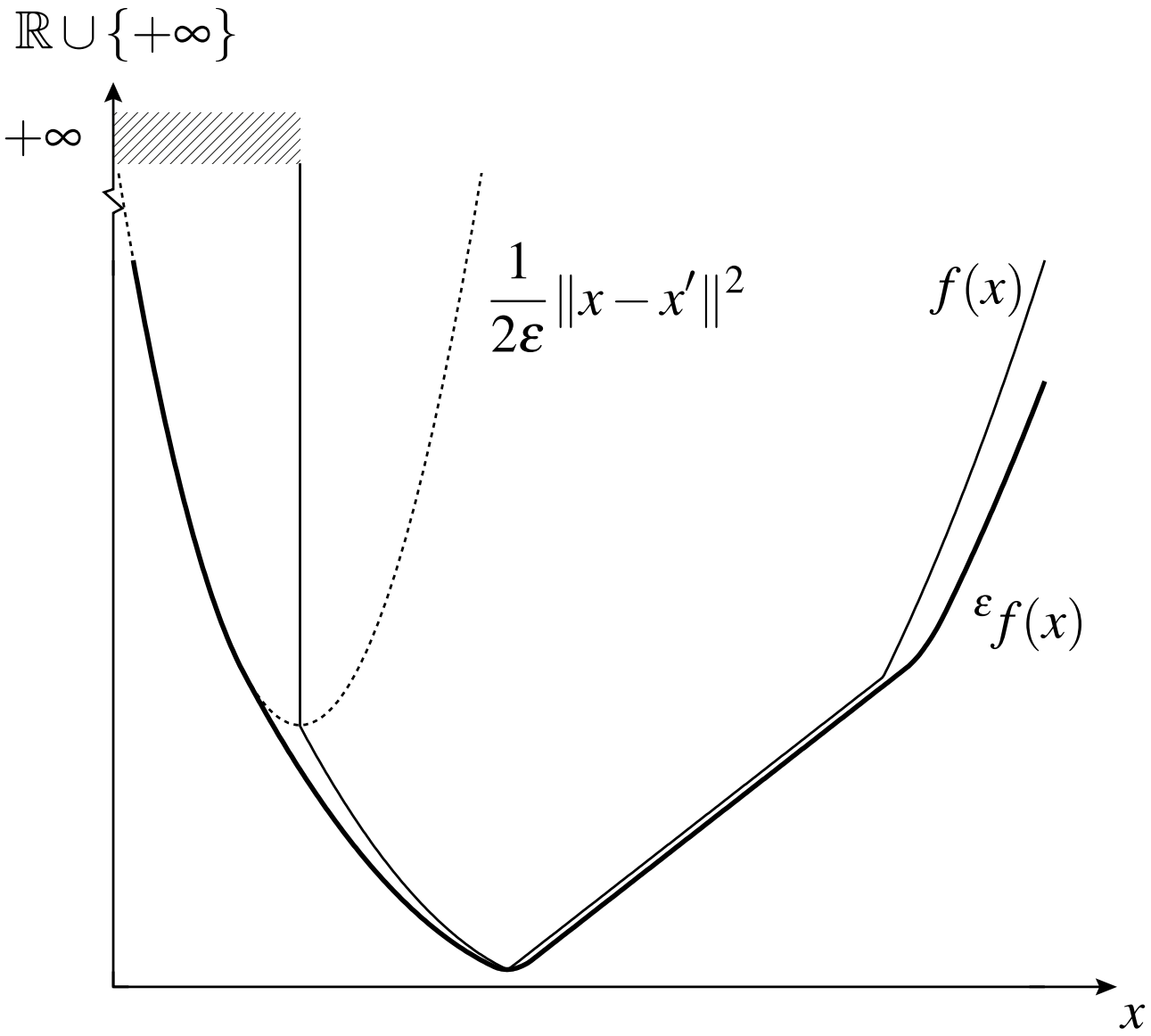
$$F_\gamma(\rho) = \min_{\tilde\rho} \left(F(\tilde\rho) + \tfrac{1}{2\gamma}\|\tilde\rho - \rho\|^2\right) \quad \leftarrow \quad \text{Moreau envelope } F_\gamma \text{ of } F$$

$$= F(\rho_\gamma) + \tfrac{1}{2\gamma}\|\rho_\gamma - \rho\|^2 \quad \leftarrow \quad \text{proximal density } \rho_\gamma \text{ of } \rho$$

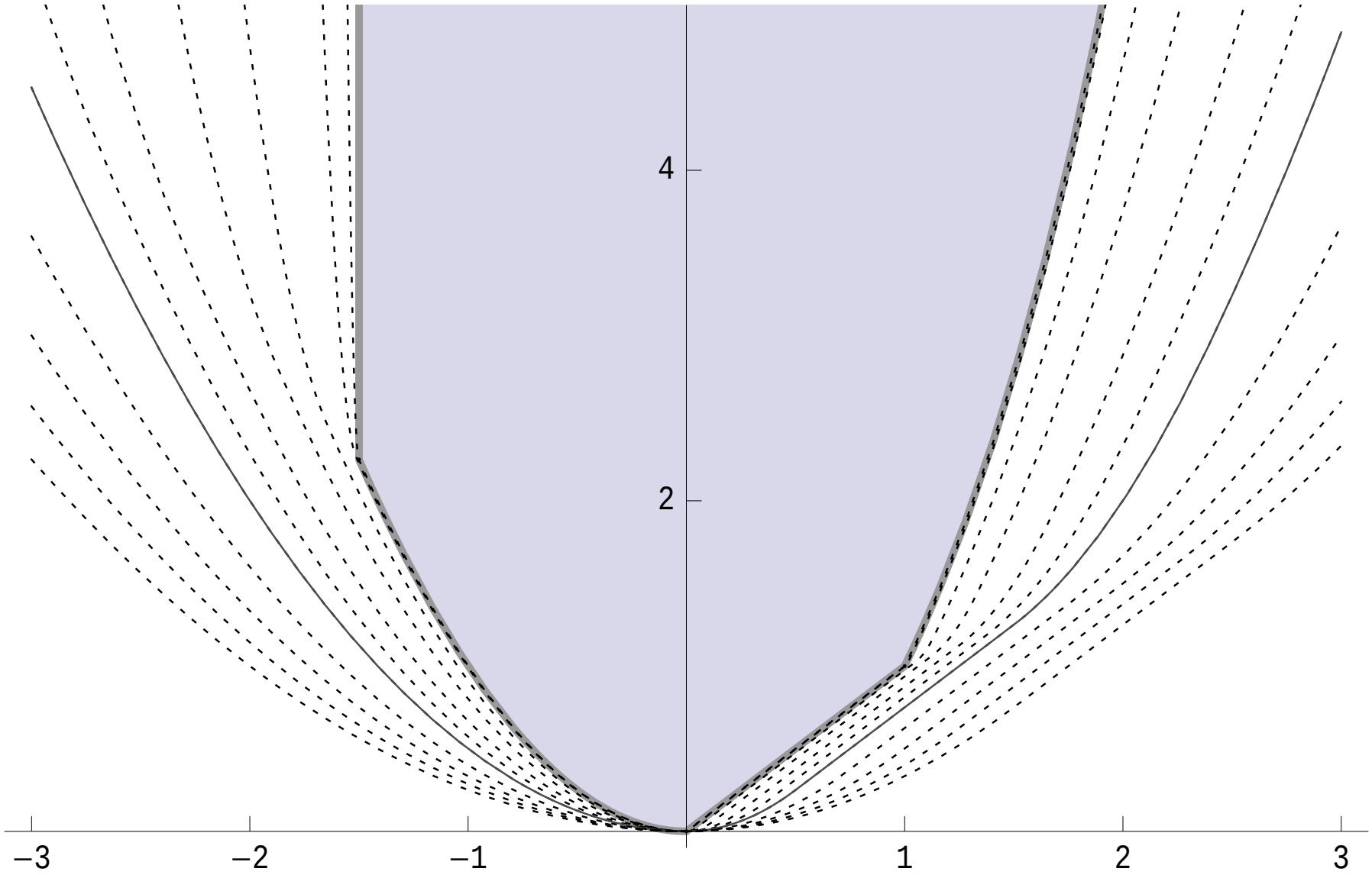
- The minimising proximal density ρ_γ is the energetically most favoured nearby density to ρ
- The Moreau envelope F_γ is a finite, continuous and differentiable approximation to F
  - it is a lower bound to F with the same minimum
  - it approaches F pointwise from below as $\gamma \to 0$



Moreau envelope illustration I



Moreau envelope illustration II



Hohenberg–Kohn variation principle in MY DFT

- We have set up MY DFT with a continuous and differentiable density functional

$$E_\gamma(v) = \min_\rho \big(F_\gamma(\rho) + (v|\rho)\big) \quad \text{Hohenberg–Kohn variation principle}$$

- Importantly, the Euler–Lagrange equation is now well defined:

$$\frac{\delta F_\gamma(\rho)}{\delta \rho(\mathbf{r})} = -v(\mathbf{r}) - \mu$$

- Its minimising solution ρ_v^γ (several may exist) yields the ground-state energy

$$E(v) = E_\gamma(v) + \tfrac{1}{2}\gamma\|v\|^2 = F_\gamma(\rho_v^\gamma) + (\rho_v^\gamma|v) + \tfrac{1}{2}\gamma\|v\|^2$$

- How is ρ_v^γ related to the regular ground-state density ρ_v?

$$E(v) = \min_\rho \big(F(\rho) + (v|\rho)\big) = F(\rho_v) + (\rho_v|v)$$

  - answer: the regular ground-state density ρ_v is the proximal density to ρ_v^γ

- Moreau–Yosida DFT approximates standard DFT arbitrarily closely as γ tends to zero:

$$\lim_{\gamma\to 0^+} \|\rho_v^\gamma - \rho_v\| = 0, \qquad \lim_{\gamma\to 0^+} F_\gamma(\rho) = F(\rho), \qquad \lim_{\gamma\to 0^+} E_\gamma(v) = E(v)$$

Significance of densities in MY DFT

- Densities in MY DFT are not physical densities but pseudo-densities or quasi-densities
  - they may be negative in some region and need not integrate to the correct number of electrons
  - they are not obtained from a wave function in the usual way
- However, each pseudo-density ρ may be uniquely decomposed in the manner

$$\rho = \rho_\gamma - \gamma v_\gamma \quad \leftarrow \text{ "depot" } = \text{ density–potential}$$

  where ρ_γ is the proximal density to ρ and the ground-state density of the potential v_γ
  - the MY "depot" ρ contains information about the density–potential pair ρ_γ and v_γ
  - they can be extracted from ρ and F_γ in the following manners:

$$\rho_\gamma = \arg\min_{\tilde{\rho}} \left(F(\tilde{\rho}) + \tfrac{1}{2\gamma} \|\tilde{\rho} - \rho\|^2 \right), \quad v_\gamma(\mathbf{r}) = -\frac{\delta F_\gamma(\rho)}{\delta \rho(\mathbf{r})}$$

- Our starting point was that derivatives of F with respect to physical densities are undefined

$$\text{ill defined} \quad \rightarrow \quad \frac{\delta F(\rho_\gamma)}{\delta \rho_\gamma(\mathbf{r})} = -v_\gamma(\mathbf{r})$$

- Instead, derivatives of F_γ with respect to pseudo-densities are well-defined:

$$\text{well defined} \quad \rightarrow \quad \frac{\delta F_\gamma(\rho)}{\delta \rho(\mathbf{r})} = -v_\gamma(\mathbf{r})$$

Kohn–Sham theory in MY DFT

- Kohn–Sham theory assumes:

  the existence of a noninteracting system with the same density as the physical system

  - this assumption is unfounded: the noninteracting representability problem

- In MY DFT, every depot is a minimising depot of some potential
  - this is true for interacting and noninteracting systems
  - the noninteracting representability problem therefore does not exist in MY DFT:

$$\frac{\delta F_\gamma(\rho)}{\rho(\mathbf{r})} = -v(\mathbf{r}) - \mu \qquad \text{interacting Euler–Lagrange equation well defined}$$

$$\frac{\delta (T_s)_\gamma(\rho)}{\rho(\mathbf{r})} = -v_s(\mathbf{r}) - \mu_s \qquad \text{noninteracting Euler–Lagrange equation well defined}$$

  - the depot ρ is shared by the systems but their ground-state densities and potentials differ:

$$\rho = \rho_{\text{phy}} - \gamma v_{\text{ext}} \qquad \text{interacting density and potential}$$

$$\rho = \rho_{\text{non}} - \gamma v_s \qquad \text{noninteracting density and potential}$$

- Kohn–Sham SCF iterations can be set up identical to standard Kohn–Sham iterations
  - an alternative SCF scheme may be based on the condition

$$\rho_{\text{phy}} = \text{prox}(\rho_{\text{phy}} - \gamma v_{\text{ext}})$$

  - for given v_{ext}, guess the ground-state density ρ_0 and iterate until self-consistency

$$\rho_{i+1} = \text{prox}(\rho_i - \gamma v_{\text{ext}})$$

Different metrics in MY DFT

- It is possible to work with a more general MY regularisation, with a weighted penalty:

$$E_\gamma(v) = E(v) - \tfrac{1}{2}\gamma\langle v|\mathbf{M}|v\rangle \quad \leftarrow \text{energy with damped potential}$$

$$F_\gamma(\rho) = \min_{\tilde\rho}\left(F(\tilde\rho) + \tfrac{1}{2\gamma}\langle\tilde\rho-\rho|\mathbf{M}^{-1}|\tilde\rho-\rho\rangle\right) \quad \leftarrow \text{energy of favoured nearby density}$$

- Two choices of metric: $\mathbf{M} = \mathbf{I}$ (overlap) and $\mathbf{M} = \mathbf{T}$ (kinetic)
  - the kinetic metric gives the regularisation of Heaton–Burgess, Bulat and Yang (2007)

$$\langle v|\mathbf{T}|v\rangle = \int |\boldsymbol{\nabla} v(\mathbf{r})|^2\,\mathrm{d}\mathbf{r} \quad \leftarrow \text{small for smooth } v$$

$$\langle \Delta\rho|\mathbf{T}^{-1}|\Delta\rho\rangle = \iint \Delta\rho(\mathbf{r}_1) r_{12}^{-1} \Delta\rho(\mathbf{r}_2)\,\mathrm{d}\mathbf{r}_1\mathrm{d}\mathbf{r}_2 \quad \leftarrow \text{small for oscillatory } \Delta\rho = \tilde\rho - \rho$$

  - in the Lieb variation principle, the potential becomes smoother as density oscillations are removed

$$F_\gamma(\rho) = \sup_v\big(E_\gamma(v) - (v|\rho)\big)$$

  - regularised optimised effective potential (OEP) at zero interaction strength (Wu–Yang)
- Exact DFT:
  - F is convex, unbounded, everywhere discontinuous and not everywhere (sub)differentiable
  - not all densities are v-representable or even N-representable
- Moreau–Yosida DFT:
  - F_γ with is convex, bounded, everywhere continuous and differentiable for $\gamma > 0$
  - all densities are N-representable and even v-representable
  - reminiscent of Lammert's coarse-grained DFT (2010)

Conclusions

- Universal density functional is a complicated function of the densities
  - it is everywhere discontinuous and nondifferentiable
  - problems with Euler–Lagrange equations and KohnSham theory
- Moreau–Yosida regularisation of DFT avoids these complications
  - it sets up an DFT with an everywhere continuous and differentiable density functional
  - it is exact and approaches standard DFT as the regularisation parameter tends to zero
- Moreau–Yosida regularisation solves the representability problem of Kohn–Sham theory
  - the noninteracting kinetic energy functional becomes differentiable for all densities
  - in particular, it is differentiable for the physical density
- Publication:
  - S. Kvaal, U. Ekström, A. Teale, and T. Helgaker, JCP **140**, 18A518 (2014)
- Acknowledgements:
  - Ulf Ekström, Simen Kvaal, and Andy Teale
  - ERC advanced grant ABACUS
  - Norwegian Research Council for Centre of Excellence CTCC