Density functional theory and optimal transportation with Coulomb cost.

Codina Cotar
(joint works with G. Friesecke, C. Klueppelberg, C. Mendl and B. Pass)

April 02, 2014

Outline

Informal introduction to Quantum mechanics

- All materials systems we study essentially consist of electrons and nuclear charge.
- Mechanical, electronic, magnetic etc. properties are due to electrons and their interaction with other electrons.
- In order to define electrons and their interaction we use Schrodinger equation.
- It allows to predict, e.g., binding energies, equilibrium geometries, intermolecular forces
- Quantum mechanics for electrons reduces to a PDE (the Schroedinger equation)

Density Functional Theory (DFT)

- If Schrodinger equation for the many electrons problem could be solved accurately and efficiently then almost any property of the materials could be determined determined accurately.
- Unfortunately, there is neither an accurate nor an efficient method to solve these problems.
- DFT is a simplified version of quantum mechanics (QM), widely used in molecular simulations in chemistry, physics, materials science
- Introduced by Hohenberg-Kohn-Sham in the 1960s
- Feasible system size: up to a million atoms
- 1998 Nobel Prize for ‘founding father’ Walter Kohn

Quantum mechanics-Formal definition

- The solution for this PDE is the wave function $\Psi(x_1, s_1, \ldots, x_N, s_N) \in L^2((\mathbb{R}^3 \times \mathbb{Z}_2)^N; \mathbb{C})$
- N - number of electrons, x_i position of electron i, s_i spin of electron i
-

$$|\Psi(x_1, s_1, \ldots, x_N, s_N)|^2 = \text{probability density that the electrons are at positions } x_i \text{ with spins } s_i.$$

Ψ is an anti-symmetric function, which makes $|\Psi|^2$ a symmetric (N-exchangeable) probability measure.

- $$\mathcal{A}_N = \{\Psi \in L^2((\mathbb{R}^3 \times \mathbb{Z}_2)^N; \mathbb{C}) \,|\, \nabla\Psi \in L^2,\ \Psi \text{ antisymmetric},\ ||\Psi||_{L^2} = 1\}$$

- Key quantum mechanics quantity is the ground state energy E_0

$$E_0 = \inf_{\Psi \in \mathcal{A}_N} E[\Psi]$$

where

$$E[\Psi] = T_h[\Psi] + V_{ee}[\Psi] + V_{ne}[\Psi]$$

- Kinetic energy:

$$T_h[\Psi] = \frac{h^2}{2} \int |\nabla \Psi(x_1, s_1, \ldots, x_N, s_N)|^2 dz_1 dz_2 \ldots dz_N$$

- Electron-electron energy:

$$V_{ee}[\Psi] = \int \sum_{1 \leq i < j \leq N} \frac{1}{|x_i - x_j|} |\Psi|^2 dz_1 \ldots dz_N$$

- Nuclei-electron energy:

$$V_{ne}[\Psi] = \int v(x_i) |\Psi(z_1, z_2, \ldots, z_N)|^2 dz_1 \ldots dz_N$$

- N-electrons density

$$\rho_N^{\Psi}(x_1,..,x_N) = \sum_{s_1,..,s_N \in \mathbb{Z}_2} |\Psi(x_1,s_1,\ldots,x_N,s_N)|^2.$$

- Pair electrons density

$$\rho_2^{\Psi}(x_1,x_2) = \binom{N}{2} \int_{\mathbb{R}^{3(N-2)}} \rho_N^{\Psi}(x_1,\ldots,x_N)dx_3\ldots dx_N$$

- Single electron density

$$\rho^{\Psi}(x_1) = N \int_{\mathbb{R}^{3(N-1)}} \rho_N^{\Psi}(x_1,\ldots,x_N)dx_2\ldots dx_N.$$

- $\mathcal{R}_N := \{\rho : \mathbb{R}^3 \to \mathbb{R} \,|\, \rho \text{ is the density of some } \Psi \in \mathcal{A}_N\}$

Full Scrod. eqn. can be reformulated as a hierarchy of eqn: for ρ in terms of of the pair electrons density ρ_2, for ρ_2 in terms of ρ_3 etc.

Variational formulation of density functional theory

(Hohenberg/Kohn 1964, M. Levy 1979, E. Lieb 1983)
For any external potential v, the exact Schroedinger eqn. satisfies

- $$E_0 = \inf_{\rho \in \mathcal{R}_N} \left\{ F_h[\rho] + N \int_{\mathbb{R}^3} v(x)\, \rho(x) dx \right\}$$

  with

- $$\begin{aligned} F_h[\rho] : &= \inf_{\Psi \in \mathcal{A}_N, \Psi \mapsto \rho} \left\{ T_h[\Psi] + V_{ee}[\Psi] \right\} \\ &= \inf_{\Psi \in \mathcal{A}_N, \Psi \mapsto \rho} \left\{ T_h[\Psi] + \int_{\mathbb{R}^6} \frac{1}{|x-y|} \rho_2^{\Psi}(dx, dy) \right\}, \end{aligned}$$

  $F_h[\rho]$ is the famous Hohenberg-Kohn functional.
- Not useful for computations (definitely still contains the big space of $\Psi(x_1, s_1, \ldots, x_N, s_N)$). But useful starting point for model reduction in asymptotic limits.

Correlations in DFT

- Mathematical structure: Minimize an approximate energy functional $F[\rho]$ which depends on the electron density $\rho(x)$, a function on $\mathbb{R}^3$.
- Catch: exact QM energy requires knowledge of electron-pair density $\rho_2(x,y)$, a function on $\mathbb{R}^6$, which entails correlations.
- Standard way out: start by assuming independence, add semi-empirical corrections to $F[\rho]$ accounting for correlations. Often but not always accurate/reliable.

Popular functionals

- The mean field approximation:

$$\int_{\mathbb{R}^6} \frac{1}{|x-y|} \rho_2^{\Psi}(dx,dy) = \frac{1}{2} \int_{\mathbb{R}^6} \frac{1}{|x-y|} \rho(dx)\rho(dy) =: J[\rho].$$

- Local Density Approximation approximation:

$$\int_{\mathbb{R}^6} \frac{1}{|x-y|} \rho_2^{\Psi}(dx,dy) = J[\rho] - \frac{4}{3}(3/\pi)^{1/3} \int_{\mathbb{R}^3} \rho(x)^{4/3} dx.$$

- Quantum mechanics is becoming so unbelievably complex that it is taking longer and longer to train a quantum theorist.
- It is taking so long, in fact, to train him to the point where he understands the nature of physical problems that he is already too old to solve them. (*Eugene Wigner*)

DFT Results-Our approach

Let

- $$E_0 = \inf_{\Psi \in \mathcal{A}_N} \left\{ T_h[\Psi] + V_{ne}[\rho^\Psi] + V_{ee}[\rho_2^\Psi] \right\}$$

  and

- $$E_0^{OT} = \inf_{\Psi \in \mathcal{A}_N} \left\{ T[\Psi] + V_{ne}[\rho^\Psi] + E_{OT}^N[\rho^\Psi] \right\},$$

  where

$$E_{OT}^N[\rho^\Psi] = \frac{1}{\binom{N}{2}} \inf_\gamma \sum_{1 \le i < j \le N} \int \frac{1}{|x_i - x_j|} d\gamma(x_1, x_2, \ldots, x_N),$$

subject to equal marginals ρ^Ψ.

Theorem

(C, Friesecke, Klueppelberg - CPAM 2013) *Fix $\rho \in \mathcal{R}_N$. Let $N \geq 2$. Then*

$$\lim_{h \to 0} F_h[\rho] = E^N_{OT}[\rho]$$

for every $\rho \in \mathcal{R}_N$, where recall that

$$F_h[\rho] := \inf_{\Psi \in \mathcal{A}_N, \Psi \mapsto \rho} \left\{ T_h[\Psi] + V_{ee}[\Psi] \right\}.$$

Theorem

(C,Friesecke, Klueppelberg - CPAM 2013) *For every N and every v*

$$E_0 \geq E^{OT}_0.$$

We are the first to link electronic structure to optimal transportation. Seidl'99, Seidl/Perdew/Levy'99, Seidl/Gori-Giorgi/Savin'07

Optimal transportation

- γ measure in $\mathbb{R}^{2d}$, ρ, ρ' measures in $\mathbb{R}^d$
- The Cost Function $c : \mathbb{R}^d \times \mathbb{R}^d \to \mathbb{R} \cup \{+\infty\}$
- Prototype problem: transport mass from a given pile ρ into a given hole ρ' so as to minimize the transportation cost

$$\int_{\mathbb{R}^{2d}} c(x,y) d\gamma(x,y)$$

subject to the constraint

$$\int_{\mathbb{R}^d} \gamma(x,y) dy = \rho(x) \quad \text{and} \quad \int_{\mathbb{R}^d} \gamma(x,y) dx = \rho'(y).$$

- $\gamma(x,y) =$ amount of mass transported from x to y
- $c(x,y) =$ cost of transporting one unit of mass from x to y, e.g. $|x-y|$, $|x-y|^2$

Issues

- Can we find an optimal measure γ which minimizes

$$\int_{\mathbb{R}^{2d}} c(x,y)d\gamma(x,y)?$$

- Under what conditions will the solution γ be unique?
- Can the optimal measure γ be characterized geometrically?
- Can we find γ explicitly?

Some known results

Optimal transport goes back to Monge (1781), Kantorovich (1942) and has recently become a very active area of mathematics,e.g. Villani (2009).

- $c(x,y) = |x-y|^2$: an optimal measure exist which is unique and it is characterized through the gradient of a convex function (Brenier, Knott and Smith, Cuesta-Albertos, Rüschendorf and Rachev)
- $c(x,y) = h(x-y)$ with h strictly convex, or $c(x,y) = l(|x-y|)$ with $l \geq 0$ strictly concave and increasing (Gangbo and McCann-1996)

Many-marginals Optimal Transportation

- γ measure in $\mathbb{R}^{Nd}$, $\rho_1, \rho_2, \ldots, \rho_N$ measures in $\mathbb{R}^d$
- The Cost Function $c : \mathbb{R}^d \times \mathbb{R}^d \ldots \times \mathbb{R}^d \to \mathbb{R} \cup \{+\infty\}$
- We want to transport mass from a given pile ρ_1 into a number of given holes $\rho_2, \rho_3, \ldots, \rho_N$, so as to minimize the transportation cost

$$\int c(x_1, x_2, \ldots, x_N) d\gamma(x_1, x_2, \ldots, x_N).$$

subject to the constraints

$$\int_{\mathbb{R}^{(N-1)d}} \gamma(x_1, x_2, \ldots, x_N) dx_2 \ldots dx_N = \rho(x_1), \ldots$$

$$\int_{\mathbb{R}^{(N-1)d}} \gamma(x_1, x_2, \ldots, x_N) dx_1 \ldots dx_{N-1} = \rho(x_N),$$

- Results by Carlier, Gangbo and Swietch, Pass

The 2-marginal Optimal Transport Problem with Coulomb Cost

- ρ_2 measure in $\mathbb{R}^{2d}$, ρ measure in $\mathbb{R}^d$
- Minimize the transportation cost

$$\int_{\mathbb{R}^{2d}} \frac{1}{|x-y|} d\rho_2(x,y)$$

subject to the constraint

$$\int_{\mathbb{R}^d} \rho_2(x,y)dy = \rho(x) \quad \text{and} \quad \int_{\mathbb{R}^d} \rho_2(x,y)dx = \rho(y).$$

- General pattern: $c : \mathbb{R}^d \times \mathbb{R}^d \to \mathbb{R} \cup \{+\infty\}$,with $c(x,y) := l(|x-y|)$, such that $l \geq 0$ is strictly convex, strictly decreasing and C^1 on $(0,\infty)$, $l(0) = +\infty$.

Optimal Transport Results

Theorem

(C, Friesecke, Klueppelberg - CPAM 2013) *Let $c(x,y) := l(|x-y|)$, such that $l \geq 0$ is strictly convex, strictly decreasing and C^1 on $(0,\infty)$, $l(0) = +\infty$, ρ absolutely continuous with respect to the Lebesgue measure. Then*

- *There exists a unique optimizing measure ρ_2 with*

$$\rho_2(x,y) = \rho(x)\delta_{T(x)}(y),$$

where the optimal map $T : \mathbb{R}^d \to \mathbb{R}$ is unique. Moreover $\rho \circ T^{-1} = \rho$.

- Physical meaning 1: $T(x) =$ position of the 2nd electron if the first electron is at x.
- Physical meaning 2: the graph of T is the support of the electron

The Method

- Adaptation of **W. Gangbo, R. McCann:** The geometry of optimal transportation, *Acta Math.* **177**, 113-161 (1996).
- Explicit Solution: ρ_1 and ρ_2 densities of μ, ν, with $\rho_1(x) = \lambda_1(|x|)$ and $\rho_2(x) = \lambda_2(|x|), x \in \mathbb{R}^d$.
- Then T is of form: $T(x) = x\frac{g(|x|)}{|x|}, x \in \mathbb{R}^d$, with $g : [0, \infty) \to \mathbb{R}$. Moreover $g \leq 0$, and g is an increasing function with $g(0_+) = -\infty$ and $g(+\infty) = 0$.
- Physical interpretation: 2nd electron is in the opposite direction of first.

Theorem

(C, Frieescke, Klueppelberg - CPAM 2013) *Suppose that* $\mu = \nu$. *Let* $t \in (0, \infty)$ *and let*

$$F_1(t) = |S^{d-1}| \int_0^t \lambda(s) s^{d-1} ds$$

and

$$F_2(-t) = |S^{d-1}| \int_t^\infty \lambda(s) s^{d-1} ds.$$

Then

$$g(t) = F_2^{-1}(F_1(t)).$$

The infinite Optimal Transportation marginal problem

Let γ be an infinite dimensional measure, γ symmetric (exchangeable), ρ probability measure in $\mathbb{R}^d$.

$$F_{OT}^{\infty}[\rho] = \inf_{\gamma} \lim_{N\to\infty} \frac{1}{\binom{N}{2}} \int_{\mathbb{R}^{dN}} \sum_{1\le i<j\le N} \frac{1}{|x_i - x_j|} d\gamma(x_1, .., x_N),$$

subject to the constraint

$$\int_{\mathbb{R}\times\mathbb{R}\times\ldots} \gamma(x_1, x_2, \ldots, x_N, \ldots) dx_2 dx_3 \ldots = \rho(x_1).$$

Then by de Finetti's Theorem and new results on Fourier transforms

Theorem

(C, Friesecke, Pass - 2013)

$$\lim_{N\to\infty} F_{OT}^{N}[\rho] = F_{OT}^{\infty}[\rho] = \frac{1}{2} \int_{\mathbb{R}^6} \frac{1}{|x-y|} \rho(x)\rho(y) dxdy.$$

Recap-Our Current Project

Fresh look at the DFT correlation problem from the point of view of recent optimal transport/exchangeable processes methods

- C.C., G. Friesecke, C. Klueppelberg CPAM (2013): Exact $F_h[\rho]$ in semi-classical limit ($h \to 0$) for $N = 2$
  - Novel functional form, complete anticorrelation
  - Opposite starting point for designing approximations than usual
- C.C., G. Friesecke, C. Klueppelberg, B. Pass, J. Chem. Phys. (2013)
- C.C., G. Friesecke, B. Pass Calc. Var. and PDEs., under revision (2013): Limit of the $F_h[\rho]$ for large N and small h

THANK YOU!