Density functional theory and optimal transportation with Coulomb and Riesz cost.

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

May 02, 2017

Outline

Informal introduction to Quantum mechanics/DFT

- 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 a molecule with N electrons reduces to a PDE (called Schroedinger equation) for a function $\Psi \in L^2(\mathbb{R}^{3N}, \mathbb{C})$.

- The solution $\Psi(x_1, \ldots, x_N)$ is called wave function
- N - number of electrons, x_i position of electron i
-

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

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

- 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.

Density Functional Theory (DFT)

- To simulate chemical behaviour, approximations are needed.
- Curse of dimensionality: carbon atom: $N = 6$. Discretise $\mathbb{R}$ by 10 points$\rightarrow 10^{18}$ total grid points.
- DFT is a simplified version of quantum mechanics (QM), widely used in molecular simulations in chemistry, physics, materials science
- **Main idea**: describe complicated N-particle system (a probability on $\mathbb{R}^{3N}$) using only its one-particle marginal density

$$\rho(x_1) = \int_{\mathbb{R}^{3(N-1)}} |\Psi(x_1, \ldots, x_N)|^2 dx_2 \ldots dx_N$$

- Feasible system size: systems with more than a dozen or so electrons.

Some history of DFT

- Thomas-Fermi: 1920s simple model
- Hohenberg-Kohn-Sham (1963-1964): practical method based on semi-empirical functions of ρ
- Levy (1979), Lieb (1983): mathematical justification and simplified reformulation of the equation
- 1970s: popular in solid state physics, but not so accurate
- 1990s: explosion in quantum chemistry, due to increase of computational resources + discovery of efficient semi-empirical functionals of ρ
- 1998 Nobel Prize for 'founding father' Walter Kohn
- More than 15 000 papers per year with the keyword 'density functional theory'

Quantum mechanics-Formal definition

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

$$T_h[\Psi] = \frac{h^2}{2} \int_{\mathbb{R}^{3N}} |\nabla \Psi(x_1, \ldots, x_N)|^2 dx_1 dx_2 \ldots dx_N$$

- Electron-electron energy:

$$V_{ee}[\Psi] = \int_{\mathbb{R}^{3N}} \sum_{1 \leq i < j \leq N} \frac{1}{|x_i - x_j|} |\Psi|^2 dx_1 \ldots dx_N$$

- Nuclei-electron energy:

$$V_{ne}[\Psi] = \sum_{i=1}^{N} \int_{\mathbb{R}^{3N}} v(x_i) |\Psi(x_1 \ldots, x_N)|^2 dx_1 \ldots dx_N$$

- N-electrons density

$$\rho_N^{\Psi}(x_1,..,x_N) = |\Psi(x_1,\dots,x_N)|^2$$

- Pair electrons density

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

- Single electron density

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

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

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

- $$\begin{aligned} F_h[\rho] : &= \inf_{\Psi \in \mathcal{A}_N, \Psi \mapsto \rho} \left\{ T_h[\Psi] + V_{ee}[\Psi] \right\} \\ &= \inf_{\Psi \in \mathcal{A}_N, \Psi \mapsto \rho} \left\{ T_h[\Psi] + \frac{N(N-1)}{2} \int_{\mathbb{R}^6} \frac{1}{|x-y|} \rho_2^{\Psi}(dx, dy) \right\}, \end{aligned}$$

  $F_h[\rho]$ is the famous Hohenberg-Kohn functional.
- Not useful for computations (definitely still contains the big space of $\Psi(x_1, \ldots, x_N)$). 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.
- Roughly, DFT models $\approx$ semi-empirical models of the pair density ρ_2 in terms of ρ.
- Standard way out: start by assuming independence, add semi-empirical corrections to $F_h[\rho]$ accounting for correlations. Often but not always accurate/reliable.

Popular functionals

All functionals used in practice are of form

$$\text{Mean field} \ + \ \text{additive corrections.}$$

Why mean field? Interactions not weaker than single-particle terms.

- The mean field approximation:

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

- Local Density Approximation approximation:

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

- 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*)
- Most cited physicist of all time is a designer of DFT models, J.Perdew.

- ρ_N measure in $\mathbb{R}^{Nd}$, ρ measure in $\mathbb{R}^d$
- $0 < s < d$
- Minimize the transportation cost

$$\int_{\mathbb{R}^{Nd}} \left(\sum_{1 \le i < j \le N} \frac{1}{|x_i - x_j|^s} \right) d\rho_N(x_1, \ldots, x_N)$$

subject to the constraint

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

$$\int_{\mathbb{R}^{(N-1)d}} \rho_N(x_1, \ldots, x_{N-1}, x_N) dx_1 \ldots dx_{N-1} = \rho(x).$$

- We can symmetrise the measure ρ_N-finite exchangeable.

- Coulomb: $s = d - 2$ (for DFT, $s = 1, d = 3$)
- Riesz: $0 < s < d$
- For $s = 1, d = 3$

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

subject to equal marginals ρ.

Semiclassical limit

Theorem

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

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

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

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

- Bindini - De Pascale (2017): extension to $N = 3$.
- C, Friesecke, Klueppelberg (in preparation): extension to $N \geq 4$
- In physics literature: Seidl'99, Seidl/Perdew/Levy 1999, Seidl/Gori-Giorgi/Savin 2007

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 $(0,\infty)$, $l(0) = +\infty$.

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 $(0,\infty)$, $l(0) = +\infty$. Take $\rho \in \mathcal{P}(\mathbb{R}^d) \cap L^1(\mathbb{R}^d)$. Then

- *There exists a unique optimizing measure ρ_2 with minimizer of Monge form, i.e.*

$$\rho_2 = (id, T)_{\#\rho},$$

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

For Coulomb cost, we have also $T(x) = x + \frac{\nabla v(x)}{|\nabla v(x)|^{3/2}}$ for some convex potential $v : \mathbb{R}^3 \to \mathbb{R}$

- 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 pair density ρ_2.

The Method

- Adaptation of **Gangbo and McCann:** The geometry of optimal transportation, *Acta Math.* **177**, 113-161 (1996).
- Check that formula (originally for increasing costs) generalizes to decreasing costs with singularity on the diagonal.
- Explicit Solution: For $d = 1$, for all marginals. As simple example, take ρ to be the uniform measure on $[0, 1]$. Then $\rho_2 = (id, T)_{\#\rho}$, and T rigidly switches right and left half of $[0, 1]$.
- More precisely, we have $T(x) = x + 1/2$ for $x < 1/2$, and $T(x) = x - 1/2$ for $1/2 < x < 1$.

- For $d \geq 2$, explicit solution for all symmetric marginals. Then for 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 $g(+\infty) = 0$.
- g explicitly computable
- Physical interpretation of solution: 2nd electron is in the opposite direction of first.

Many-marginals optimal transport problem

- Kantorovich problem coincides with infimum over Monge states "strongly correlated electrons" (Colombo-Di Marino 2015)
- Existence and uniqueness of Monge solution for $N \geq 2$ in $d = 1$ (Colombo-De Pascale-Di Marino 2013)
- Buttazzo, De Pascale & Gori-Giorgi (2012); Pass (2013); Benamou, Carlier & Nenna (2015); Di Marino, Gerolin & Nenna (2015)

The infinite-dimensional Optimal Transportation problem

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

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

subject to the constraint

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

Theorem

(C, Friesecke, Pass - Calc Var PDEs 2015)

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

- Proof by use of de Finetti's Theorem and Fourier transforms
- De Finetti theorem: Let γ_∞ be a symm. inf. dim. Borel measure. Then there exists a unique Borel prob. measure ν such that

$$\gamma_\infty = \int_{\mathcal{P}(\mathbb{R}^d)} Q^{\otimes\infty} d\nu(Q).$$

-

$$F^\infty_{OT}[\rho] = \int_{\mathbb{R}^{2d}} \ell(x-y)\,\rho(dx)\,\rho(dy) + \int_{\mathbb{R}^d} \hat{\ell}(z)\left(\mathrm{var}_{\nu(dQ)} Re(\hat{Q}(z)) + \mathrm{var}_{\nu(dQ)} Im(\hat{Q}(z))\right) dz$$

Let γ be an infinite dimensional measure, γ symmetric (exchangeable), ρ probability measure in $\mathbb{R}^d$ and let $0 < s < d$.

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

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 positive-definiteness, we have

Theorem

(Petrache 2015)

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

C, di Marino, Lewin, Lieb, Petrache ...

$$\lim_{N\to\infty} N^{1-s/d}\left(F^N_{OT,s}[\rho] - \frac{N^2}{2}\int_{\mathbb{R}^{2d}} \frac{1}{|x-y|^s}\rho(x)\rho(y)dxdy\right) = ?$$

- Lieb-Oxford bound

$$N^{1-s/d}\left(F^N_{OT,s}[\rho] - \frac{N^2}{2}\int_{\mathbb{R}^{2d}} \frac{1}{|x-y|^s}\rho(x)\rho(y)dxdy\right) \geq -C_{LO}\int_{\mathbb{R}^d}\rho(x)^{1+s/d}dx.$$

- Trivially, we have

$$F^N_{OT,s}[\rho] - F^\infty_{OT}[\rho] \leq 0.$$

- Question: Does the limit exit?

- Exists $C(d,s,\rho) > 0$ such that

$$\lim_{N\to\infty} N^{1-s/d}\left(F^N_{OT,s}[\rho] - \frac{N^2}{2}\int_{\mathbb{R}^{2d}} \frac{1}{|x-y|^s}\rho(x)\rho(y)dxdy\right) = -C(d,s,\rho)\int_{\mathbb{R}^d}\rho(x)^{1+s/d}dx$$

- Uniform marginal: $\rho_\Omega = 1_\Omega/|\Omega|$ (uniform electron gas)
- C_{unif} independent of Ω
- Chemists conjecture for $s = 1, d = 3$: $C(d,s,\rho) = C_{unif}$ (Rasanen, Pittalis, Capelle & Proetto 2009)
- Exact value of C_{unif} is unknown, although everybody thought for decades that it is approx 1.4441, related to Epstein Zeta function
- C_{unif} is exactly known for $d = 1$ (di Marino-2017).
- Numerics by Seidl-Vuckovic-Gori Giorgi 2015 ($N = 50$). $C_{unif} \geq 1.401$.

Connection with Coulomb and Riesz gases (Jellium)

- N electrons and a neutralizing background in a domain Ω with $|\Omega| = N$.
- Minimize over x_i

$$\sum_{1\leq i<j\leq N} \frac{1}{|x_i - x_j|} - \sum_{j=1}^{N} \int_\Omega \frac{1}{|x_j - y|} dy + \frac{1}{2} \int_\Omega \int_\Omega \frac{1}{|x-y|} dxdy$$

- Let minimization be $\xi(N, \Omega)$, then the limit (Lieb & Narnhofer 1975)

$$\lim_{N\to\infty} \frac{\xi(N,\Omega)}{N^{4/3}} = -C_{jel}.$$

- Wigner crystallisation conjecture: in limit $N \to \infty$, the electrons place themselves on a BCC lattice (hexagonal lattice in $d = 2$)
- Lewin-Lieb (2015): comparison with uniform electron gas constant in $d = 3$

Sandier, Serfaty, Rougerie, Petrache ...

- Let $V : \mathbb{R}^d \to \mathbb{R}$ be a confining potential growing at infinity
- For $0 < d - 2 \leq s < d$, let

$$H_N(x_1, \dots, x_N) := \sum_{i \neq j} \frac{1}{|x_i - x_j|^s} + N \sum_i V(x_i).$$

- Let μ_V be the minimizer (among probability measures) of

$$\mathcal{E}_V(\mu) = \int \int \frac{1}{|x-y|^s} d\mu(x) d\mu(y) + \int V(x) d\mu(x)$$

- If the density ρ_V is smooth enough

$$\min_{x_1,\dots,x_N} H_N(x_1, \dots, x_N) - N^2 \mathcal{E}_V(\mu_V) = C_{Jel} N^{1+\frac{s}{d}} \int_\Sigma \rho^{1+\frac{s}{d}}(x) dx + o\left(N^{1+\frac{s}{d}}\right).$$

- For $d-2<s<d$, we have (C-Petrache (in progress))

$$C_{unif}=C_{Jel}.$$

- C_{Jel} minimizer of a limiting energy $\mathcal{W}$
- Abrikosov conjecture: in $2d$, the regular triangular lattice is a minimizing configuration for $\mathcal{W}$.
- Known for Coulomb case in $d=2$ (Sandier, Serfaty 2012)
- For general dimension, the conjecture is that the minimum of $\mathcal{W}$ for all $d-2\leq s<d$ is always achieved by some lattice.

THANK YOU!

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

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