Equilibria in Games with a Continuum of Agents and Transport

Guillaume Carlier [a].

Journée EDP en sciences sociales,
clotûre de la chaire de P. Markowich, 24 janvier 2014.

[a]CEREMADE, Université Paris Dauphine

Aim of this talk: illustrate the use of optimal transport theory in equilibrium problems with a continuum of agents (not so surprising since the very roots of Kantorovich work were in decentralization of planning problems). Two parts:

- Matching for teams (joint with Ivar Ekeland, CEREMADE, Dauphine),
- Cournot-Nash equilibria (joint with Adrien Blanchet, Toulouse School of Economics).

Introduction

The basic optimal transport problem is as follows. We are given two metric (compact, say) spaces X and Y, two Borel probability measures $\mu \in \mathcal{P}(X)$ and $\nu \in \mathcal{P}(Y)$, and a transport cost function $c \in C(X \times Y)$ (or simply lsc). One looks for the cheapest way to transport μ to ν.

For a more detailed overview of optimal transport theory and applications, see the books of Villani and Ambrosio, Gigli and Savaré).

The image of μ by a Borel map $T : X \to Y$ is by definition given by

$$T_{\#}\mu(B) := \mu(T^{-1}(B)), \ \forall B.$$

or equivalently (change of variables)

$$\int_Y \varphi(y) T_{\#}\mu(dy) = \int_X \varphi(T(x))\mu(dx), \ \forall \varphi \in C(X).$$

A transport map between μ and ν is a map T such that $T_{\#}\mu = \nu$ (note that there might exist no such maps, e.g. when μ is a Dirac mass and ν is not). Monge Problem

$$\inf_{T \,:\, T_{\#}\mu = \nu} \int_X c(x, T(x))\mu(dx).$$

In general quite complicated because of the awfully nonlinear constraint (Jacobian equation in the euclidean and smooth case).

This is why the problem was relaxed by Kantorovich in the 40's. A transport plan is a joint probability $\gamma \in \mathcal{P}(X \times Y)$ having μ and ν as marginals, i.e.

$$\gamma(A \times Y) = \mu(A),\ \gamma(X \times B) = \nu(B)$$

or, for all $\varphi \in C(X)$, $\psi \in C(Y)$

$$\int_{X \times Y} (\varphi(x) + \psi(y))\gamma(dx, dy) = \int_X \varphi\mu + \int_Y \psi\nu.$$

Denote by $\Pi(\mu,\nu)$ the set of transport plans between μ and ν, the Monge Kantorovich problem reads

$$W_c(\mu,\nu) := \inf_{\gamma\in\Pi(\mu,\nu)} \int_{X\times Y} c(x,y)\gamma(dx,dy)$$

- Linearity: existence of an optimal plan is for free,
- relaxation: provided μ has no atoms, $W_c(\mu,\nu)$ coincides with the infimum in Monge's problem (which need not be attained in general),
- there is a convenient dual formulation.

Dual formulation, Kantorovich duality formula, $W_c(\mu,\nu)$ coincides with

$$\sup_{(\varphi,\psi)} \left\{ \int_X \varphi\mu + \int_Y \psi\nu \;:\; \varphi(x)+\psi(y) \le c(x,y) \right\}$$

since the constraint implies that

$$\psi(y) \le \min_{x\in X}\{c(x,y)-\varphi(x)\} := \varphi^c(y),$$

one also has the (Kantorovich) duality formula

$$W_c(\mu,\nu) = \sup_{\varphi} \left\{ \int_X \varphi\mu + \int_Y \varphi^c\nu \right\}$$

and the sup is achieved. Decentralization by prices principle (Kantorovich was awarded the Nobel prize in economics in 76).

The Wasserstein space. Let μ and ν be probability measures on $\mathbb{R}^d$ with finite second moments, the squared Wasserstein distance between μ and ν is

$$W_2^2(\mu,\nu) = \inf_{\gamma\in\Pi(\mu,\nu)} \int_{\mathbb{R}^d\times\mathbb{R}^d} |x-y|^2 \gamma(dx,dy).$$

An important result of Yann Brenier says that whenever μ does not charge Lipschitz hypersurfaces, there is a unique minimizer that is given by a transport and characterized by $\gamma = (\mathrm{id}, \nabla\phi)_{\#}\mu$ with ϕ convex. Formally, related to the Monge-Ampère equation:

$$\det(D^2\phi)\nu(\nabla\phi) = \mu, \quad \phi \text{ convex}$$

with (rather tricky) boundary conditions.

Matching for teams

Market for houses, quality $z \in Z$. For one house z to be available, need for one buyer and a team of producers (mason, plumber, electrician). We shall denote by the index $i \in \{1, \cdots, I\}$ the different populations (buyers, plumbers, electricians, masons...). Contrary to the standard OT framework, we are given measures on the populations and look for a measure on Z by equilibrium requirements.

Agents in each population i are heterererogeneous, they are characterized by a certain type $x_i \in X_i$ (metric compact) which affects their cost, $c_i \in C(X_i \times Z, \mathbb{R})$ with the interpretation that $c_i(x_i, z)$ is the cost for an agent of population i with type x_i to work in a team that produces good z. The distribution of type x_i in population i is known and given by some Borel probability measure $\mu_i \in \mathcal{P}(X_i)$.

One then looks for an equilibrium and in particular an equlibrium system of monetary transfers (paid by the buyer to the producers). A system of transfers is a collection of continuous functions $\varphi_1, \dots \varphi_I \colon Z \to \mathbb{R}$ where $\varphi_i(z)$ is the amount paid to i by the other members of the team for producing z. An obvious equilibrium requirement is that teams are self-financed i.e.

$$\sum_{i=1}^{I} \varphi_i(z) = 0, \ \forall z \in Z. \tag{1}$$

Given transfers $\varphi_1, \dots \varphi_I$, an agent from population i with type $x_i \in X_i$, gets a net minimal cost given by the so-called c_i-transform of φ_i:

$$\varphi_i^{c_i}(x_i) := \min_{z \in Z} \{c_i(x_i, z) - \varphi_i(z)\} \tag{2}$$

Agents are rational: they choose cost miniminizing qualities i.e. a $z \in Z$ such that

$$\varphi_i^{c_i}(x_i) + \varphi_i(z) = c_i(x_i, z). \tag{3}$$

The last unknown is a collection of plans $\gamma_i \in \mathcal{P}(X_i \times Z)$ such that $\gamma_i(A_i \times A)$ represents the probability that an agent in population i has a type in A_i and belongs to a team that produces a quality in A. At equilibrium the first marginal of γ_i should be μ_i (this is equilibrium on the i-th labor market) and the second marginal of γ_i should not depend on i (this is equilibrium on the quality good market), this common marginal represents the equilibrium quality line.

Equilibrium: transfer system $(\varphi_1, \dots \varphi_I) \in C(Z, \mathbb{R})^I$, probability measures $\gamma_i \in \mathcal{P}(X_i \times Z)$ and a probability measure $\nu \in \mathcal{P}(Z)$ such that

- teams are self-financed i.e. (1) holds,
- $\gamma_i \in \Pi(\mu_i, \nu)$ for $i = 1, \dots, I$ (equilibrium on the labor markets and on the good market),
- (3) holds on the support of γ_i for $i = 1, \dots, I$ (i.e. agents choose cost minimizing qualities).

Variational characterization of equilibria

$$\inf_{\nu \in \mathcal{P}(Z)} J(\nu) := \sum_{i=1}^{I} W_{c_i}(\mu_i, \nu) \tag{4}$$

and its dual (concave maximization) formulation

$$\sup \left\{ \sum_{i=1}^{I} \int_{X_i} \varphi_i^{c_i}(x_i)\mu_i(dx_i) \,:\, \sum_{i=1}^{I} \varphi_i = 0 \right\}. \tag{5}$$

Theorem 1 *$(\varphi_i, \gamma_i, \nu)$ is an equilibrium if and only if:*

- *ν solves (4),*
- *the transfers $(\varphi_1, \dots \varphi_I)$ solve (5),*
- *for $i = 1, \dots, I$, γ_i solves the Monge-Kantorovich problem $W_{c_i}(\mu_i, \nu)$.*

This implies existence of equlibria.

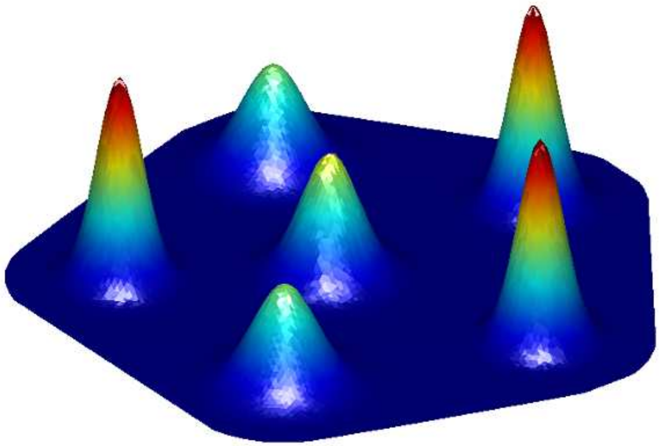
These problems can be reformulated as (large) linear programming problems. In an euclidean setting with quadratic costs $c_i(x_i, z) = \lambda_i |x_i - z|^2$, one finds equilibria as a solution of the Wasserstein barycenter problem (which also finds applications in image processing...)

$$\inf_{\nu \in \mathcal{P}_2(\mathbb{R}^d)} \sum_{i=1}^{I} \lambda_i W_2^2(\mu_i, \nu).$$

Under mild assumptions, existence, uniqueness and L^p estimates, formally related to a system of Monge-Ampère equations (C., Agueh).

Can be computed efficiently (C., Oberman, Oudet).

Wasserstein isobarycenter of 5 Gaussians:



Cournot-Nash Equilibria

Continuum of agents, each of them has to choose a strategy, individual cost depends on the distribution of strategies played by the whole population. Typical example: doctors' location choice. Typically, two competing terms: congestion (driving force for dispersion) and positive externalities (driving force for concentration).

Setting: type space X (metric compact) endowed with a probability measure $\mu \in \mathcal{P}(X)$, action space Y (metric compact). Cost: $C(x, y, \nu)$ where $\nu \in \mathcal{P}(Y)$ represents the distribution of actions (anonymous game). Unknown: $\gamma \in \mathcal{P}(X \times Y)$: $\gamma(A \times B)$ is the probability that an agent has her type in A and takes an action in B. Then define

Definition 1 *A Cournot-Nash equilibrium (CNE) is a* $\gamma \in \mathcal{P}(X \times Y)$ *such that* $\Pi_{X\#}\gamma = \mu$ *and*

$$\gamma\Big(\{(x, y)\ :\ C(x, y, \nu) = \min_{z \in Y} C(x, z, \nu)\}\Big) = 1$$

where $\nu := \Pi_{Y\#}\gamma$.

Theorem 2 (Mas-Colell, 1984) *In the regular case where*

$$\nu \mapsto C(.,.,\nu) \text{ is continuous from } (\mathcal{P}(Y), w-*) \text{ to } C(X \times Y)$$

then there exists CNE.

The assumption is extremely strong: rules out congestion/purely local effects, what about uniqueness, characterization, explicit or numerically computable solutions?

We shall restict ourselves to the separable case:

$$C(x, y, \nu) = c(x, y) + V[\nu](y) \tag{6}$$

and shall further impose that $\nu \in L^1(m_0)$ with m_0 a given reference measure on Y (congestion). Can be viewed as a simplified (static) version of the Mean-Field Games Theory of Lasry and Lions.

Benchmark: $\nu \in \mathcal{P}(Y) \cap L^1(m_0)$ (m_0: fixed reference measure according to which congestion is measured)

$$V[\nu](y) = f(y, \nu(y)) + \int_Y \phi(y, z_1, \cdots, z_m) d\nu^{\otimes m}(z_1, \cdots, z_m).$$

Due to the first term, the previous fixed-point argument does not work.

Domain

$$\mathcal{D} := \{\nu \in L^1(m_0) \ : \ \int_Y |V[\nu]| d\nu < +\infty\}.$$

Connections with optimal transport

Again $m_0 \in \mathcal{P}(Y)$ fixed reference measure, $\mathcal{D}$ domain of the cost, CNE are then defined by

Definition 2 *$\gamma \in \mathcal{P}(X \times Y)$ is a Cournot-Nash equilibria if and only if its first marginal is μ, its second marginal, ν, belongs to $\mathcal{D}$ and there exists $\varphi \in C(X)$ such that*

$$c(x,y) + V[\nu](y) \geq \varphi(x) \ \forall x \in X \text{ and } m_0\text{-a.e. } y \text{ with equality } \gamma\text{-a.e.} \tag{7}$$

A Cournot-Nash equilibrium γ is called pure whenever it is of the form $\gamma = (\mathrm{id}, T)_{\#}\mu$ for some Borel map $T : X \to Y$.

For $\nu \in \mathcal{P}(Y)$, let $\Pi(\mu, \nu)$ denote the set of probability measures on $X \times Y$ having μ and ν as marginals and let $\mathcal{W}_c(\mu, \nu)$ be the least cost of transporting μ to ν for the cost c *i.e.* the value of the Monge-Kantorovich optimal transport problem:

$$\mathcal{W}_c(\mu, \nu) := \inf_{\gamma \in \Pi(\mu,\nu)} \iint_{X \times Y} c(x, y) \, \mathrm{d}\gamma(x, y)$$

let us also denote by $\Pi_o(\mu, \nu)$ the set of optimal transport plans i.e.

$$\Pi_o(\mu, \nu) := \{\gamma \in \Pi(\mu, \nu) \ : \ \iint_{X \times Y} c(x, y) \, \mathrm{d}\gamma(x, y) = \mathcal{W}_c(\mu, \nu)\}.$$

A first link between Cournot-Nash equilibria and optimal transport is based on the following straightforward observation.

Lemma 1 *If γ is a Cournot-Nash equilibrium and ν denotes its second marginal then $\gamma \in \Pi_o(\mu, \nu)$.*

Proof. Indeed, let $\varphi \in C(X)$ be such that (7) holds and let $\eta \in \Pi(\mu, \nu)$ then we have

$$\iint_{X\times Y} c(x,y)\, \mathrm{d}\eta(x,y) \geq \iint_{X\times Y} (\varphi(x) - V[\nu](y))\, \mathrm{d}\eta(x,y)$$

$$= \int_X \varphi(x)\, \mathrm{d}\mu(x) - \int_Y V[\nu](y)\, \mathrm{d}\nu(y) = \iint_{X\times Y} c(x,y)\, \mathrm{d}\gamma(x,y)$$

so that $\gamma \in \Pi_o(\mu, \nu)$. □

Similar to what happens in mean-field games: monotonicity implies uniqueness (covers the case of pure congestion):

Theorem 3 *If $\nu \mapsto V[\nu]$ is strictly monotone in the sense that for every ν_1 and ν_2 in $\mathcal{P}(Y)$, one has*

$$\int_Y (V[\nu_1] - V[\nu_2]) d(\nu_1 - \nu_2) \geq 0$$

and the inequality is strict whenever $\nu_1 \neq \nu_2$ then all equilibria have the same second marginal ν.

Proof. Let $(\nu_1, \gamma_1, \varphi_1)$, $(\nu_2, \gamma_2, \varphi_2)$ be such that

$$V[\nu_i](y) \geq \varphi_i(x) - c(x,y), \ i = 1, 2,$$

for every x and m_0-a.e. y with an equality γ_i-a.e., using the fact that $\gamma_i \in \Pi(\mu, \nu_i)$, we get

$$\int_Y V[\nu_i] d\nu_i = \int_X \varphi_i d\mu - \int_{X \times Y} c d\gamma_i, \ i = 1, 2$$

$$\int_Y V[\nu_i] d\nu_j \geq \int_X \varphi_i d\mu - \int_{X \times Y} c d\gamma_j, \text{ for } i \neq j$$

substracting, we get $\int_Y V[\nu_1] d(\nu_1 - \nu_2) \leq \int_{X \times Y} c d(\gamma_2 - \gamma_1)$ and $\int_Y V[\nu_2] d(\nu_2 - \nu_1) \leq \int_{X \times Y} c d(\gamma_1 - \gamma_2)$ and monotonicity thus gives $\nu_1 = \nu_2$.

□

A variational approach

Take $V[\nu](y) = f(y, \nu(y)) + \int_Y \phi(y,z) d\nu(z)$ with $f(y,.)$ continuous nondecreasing (+ power or logarithm growth) and ϕ continuous and symmetric i.e. $\phi(y,z) = \phi(z,y)$. Then define $F(y,\nu) := \int_0^\nu f(y,s) ds$ and

$$E[\nu] = \int_Y F(y,\nu(y)) dm_0(y) + \frac{1}{2} \iint_{Y \times Y} \phi(y,z) \, \mathrm{d}\nu(y) \, \mathrm{d}\nu(z)$$

then $V[\nu] = \frac{\delta E}{\delta \nu}$ in the sense that for every $(\rho, \nu) \in \mathcal{D}^2$, one has

$$\lim_{\varepsilon \to 0^+} \frac{E[(1-\varepsilon)\nu + \varepsilon\rho] - E[\nu]}{\varepsilon} = \int_Y V[\nu] \, \mathrm{d}(\rho - \nu).$$

Equilibria may be obtained by solving

$$\inf_{\nu \in \mathcal{D}} J_\mu[\nu] \quad \text{where} \quad J_\mu[\nu] := \mathcal{W}_c(\mu,\nu) + E[\nu]. \tag{8}$$

Theorem 4 (Minimizers are equilibria) *Assume that $X = \overline{\Omega}$ where Ω is some open bounded connected subset of $\mathbb{R}^d$ with negligible boundary, that μ is equivalent to the Lebesgue measure on X (that is both measures have the same negligible sets) and that for every $y \in Y$, $c(.,y)$ is differentiable with $\nabla_x c$ bounded on $X \times Y$. If ν solves* (8) *and $\gamma \in \Pi_o(\mu,\nu)$ then γ is a Cournot-Nash equilibrium. In particular there exist CNE.*

Optimality condition for (8) :

$$\begin{cases} \varphi^c + V[\nu] \geq 0 & \text{with an equality } \nu\text{-a.e.} \\ \varphi^c(y) + \varphi(x) \leq c(x,y) & \text{with an equality } \gamma\text{-a.e.} \end{cases}, \tag{9}$$

If E is convex: equivalence between minimization and being an equilibrium. If E strictly convex: uniqueness (of ν). The congestion term is convex and forces dispersion whereas the interaction term is nonconvex and rather fosters concentration. It may be the case that the congestion term dominates so as to make E convex but this is more the exception than the rule. There is hidden convexity (as in McCann's displacement convexity) in the problem as we shall see now.

Hidden convexity: dimension one

Intuition is easy to understand in dimension one: the functional J_μ is not convex with respect to ν but it is with respect to T, the optimal transport map from μ to ν. Let us take $X = Y = [0,1]$, m_0 is the Lebesgue measure on $[0,1]$, μ is absolutely continuous with respect to the Lebesgue measure, and assume that $V[\nu]$ takes the form:

$$V[\nu](y) = f(\nu(y)) + V(y) + \int_{[0,1]} \phi(y,z)\, \mathrm{d}\nu(z)$$

the corresponding energy reads

$$E(\nu) := \int_0^1 F(\nu(y))\, \mathrm{d}y + \int_0^1 V(y)\, \mathrm{d}\nu(y) + \frac{1}{2}\int_{[0,1]^2} \phi\, \mathrm{d}\nu^{\otimes 2}$$

(with $F' = f$).

Assume

- the transport cost c is of the form $c(x, y) = C(x - y)$ where C is strictly convex and differentiable,
- f is convex increasing (+growth condition),
- V is convex on $[0, 1]$ and ϕ is convex, symmetric, differentiable and has a locally Lipschitz gradient.

Let $(\rho, \nu) \in \mathcal{P}([0,1])^2$ then there is a unique optimal transport map T_0 (respectively T_1) from μ to ν (respectively from μ to ρ) for the cost c and it is nondecreasing. For $t \in [0,1]$, let us define:

$$\nu_t := T_{t\#}\mu \text{ where } T_t := ((1-t)T_0 + tT_1)$$

then the curve $t \mapsto \nu_t$ connects $\nu_0 = \nu$ to $\nu_1 = \rho$. $J : \mathcal{P}(Y) \to \mathbb{R} \cup \{+\infty\}$ is called displacement convex if, for all $t \in [0,1]$, $J(\nu_t) \leq (1-t)J(\nu) + tJ(\rho)$ (for every choice of endpoints ν and ρ), it is called strictly displacement convex when, in addition $J(\nu_t) < (1-t)J(\nu) + tJ(\rho)$ when $t \in (0,1)$ and $\rho \neq \mu$.

We claim that J_μ is strictly displacement convex; take (ν, ρ) two probability measures in the domain of E (which is convex by convexity of F), define ν_t as above.

By definition of $\mathcal{W}_c$, ν_t and the strict convexity of C we have

$$\begin{aligned}\mathcal{W}_c(\mu,\nu_t) &\leq \int_0^1 C(x-((1-t)T_0(x)+tT_1(x)))\,\mathrm{d}\mu\\ &\leq (1-t)\int_0^1 C(x-T_0(x))\,\mathrm{d}\mu + t\int_0^1 C(x-T_1(x))d\mu\\ &= (1-t)\mathcal{W}_c(\mu,\nu)+t\mathcal{W}_c(\mu,\rho)\end{aligned}$$

with a strict inequality if $t\in(0,1)$ and $\nu\neq\rho$.

By construction

$$\int_0^1 V \,\mathrm{d}\nu_t = \int_0^1 V(T_t(x))d\mu(x) = \int_0^1 V((1-t)T_0(x)+tT_1(x))d\mu(x)$$

which is convex with respect to t, by convexity of V. Similarly

$$\int_{[0,1]^2} \phi \,\mathrm{d}\nu_t^{\otimes 2} = \int_{[0,1]^2} \phi(T_t(x), T_t(y)) \,\mathrm{d}\mu(x) \,\mathrm{d}\mu(y)$$

is convex with respect to t, by convexity of ϕ, The convexity of the remaining congestion term is more involved. Since $\nu_t = T_{t\#}\mu$ and T_t is nondecreasing, at least formally we have $\nu_t(T_t(x))T_t'(x) = \mu(x)$, by the change of variables formula we also have

$$\int_0^1 F(\nu_t(y))\,\mathrm{d}y = \int_0^1 F(\nu_t(T_t(x)))T_t'(x)\,\mathrm{d}x = \int_0^1 F\Big(\frac{\mu(x)}{T_t'(x)}\Big)T_t'(x)$$

and we conclude by observing that $\alpha \mapsto F(\mu(x)\alpha^{-1})\alpha$ is convex and that $T_t'(x)$ is linear in t.

All this yields:

Theorem 5 *Under the assumptions above, optima and equilibria coincide and there exists a unique equilibrium (which is actually pure).*

This can be generalized to higher dimensions when the cost is quadratic and the congestion term satisfies McCann's condition (power functions or logarithms are fine).

A PDE for the equilibrium (quadratic cost)

For computational simplicity, take $V = 0$ and $f(\nu) = \log(\nu)$ (satisfies McCann's condition and ensures that the mass remains positive everywhere). Optimality condition:

$$\log(\nu(y)) + \varphi^c(y) + \int_{Y^m} \phi(y, .)d\nu^{\otimes m} = 0 \tag{10}$$

Optimal transport map (Brenier) $T = \nabla u$ between μ and ν: Monge-Ampère equation

$$\mu(x) = \det(D^2 u(x))\, \nu(\nabla u(x)), \quad \forall x \in \Omega \tag{11}$$

which has to be supplemented with the natural sort of boundary condition

$$\nabla u(\Omega) = \Omega. \tag{12}$$

On the other hand $\varphi(x) = \frac{1}{2}|x|^2 - u(x)$, $\varphi^c(y) = \frac{1}{2}|y|^2 - u^*(y)$ so

$$\varphi^c(\nabla u) = \frac{1}{2}|\nabla u|^2 - u^*(\nabla u) = \frac{1}{2}|\nabla u|^2 - x \cdot \nabla u + u$$

substituting $y = \nabla u(x)$ in (10), using

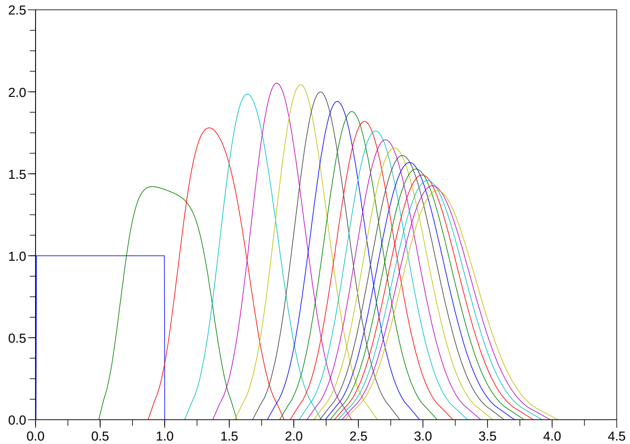
$$\int_{Y^m} \phi(\nabla u(x), .)\, \mathrm{d}\nu^{\otimes m} = \int_{\Omega^m} \phi(\nabla u(x), \nabla u(x_1), \ldots, \nabla u(x_m))\, \mathrm{d}\mu^{\otimes m} .$$

and eliminating ν thanks to (11), we get

$$\mu(x) = \det(D^2 u(x)) \exp\left(-\frac{1}{2}|\nabla u(x)|^2 + x \cdot \nabla u(x) - u(x)\right) \times$$
$$\exp\left(-\int_{\Omega^m} \phi(\nabla u(x), \nabla u(x_1), \ldots, \nabla u(x_m))\, \mathrm{d}\mu^{\otimes m}(x_1, \ldots x_m)\right) . \quad (13)$$

The equilibrium problem is therefore equivalent to a non-local and nonlinear partial differential equation.

The problem can be solved numerically in dimension 1.



Convergence and stabilisation toward the equilibrium in the case of a logarithmic congestion, cubic interaction, and a potential $V(x) := (x-5)^3$ with uniform measure on $[0,1]$ as initial guess