,pdfcreator=HAL,pdfproducer=PDFLaTeX,pdfsubject=Mathematics [math]/Analysis of PDEs [math.AP], Humanities and Social Sciences/Economics and Finance

Optimal transport and Cournot-Nash equilibria

Adrien Blanchet, Guillaume Carlier

To cite this version:

Adrien Blanchet, Guillaume Carlier. Optimal transport and Cournot-Nash equilibria. Mathematics of Operations Research, 2015, 41 (1), pp.125-145. 10.1287/moor.2015.0719 . hal-00712488

HAL Id: hal-00712488

https://hal.science/hal-00712488v1

Submitted on 27 Jun 2012

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

OPTIMAL TRANSPORT AND COURNOT-NASH EQUILIBRIA

Adrien Blanchet[1] and Guillaume Carlier[2]

Abstract We study a class of games with a continuum of players for which Cournot-Nash equilibria can be obtained by the minimisation of some cost, related to optimal transport. This cost is not convex in the usual sense in general but it turns out to have hidden strict convexity properties in many relevant cases. This enables us to obtain new uniqueness results and a characterisation of equilibria in terms of some partial differential equations, a simple numerical scheme in dimension one as well as an analysis of the inefficiency of equilibria.

Keywords: Cournot-Nash equilibria, mean-field games, optimal transport, externalities, Monge-Ampère equations, convexity along generalised geodesics.

1. INTRODUCTION

Since Aumann's seminal works Aumann [1964, 1966], models with a continuum of agents have occupied a distinguished position in economics and game theory. Schmeidler [1973] introduced a notion of non-cooperative equilibrium in games with a continuum of agents and established several existence results. In Schmeidler's own words: *Non-atomic games enable us to analyze a conflict situation where the single player has no influence on the situation but the aggregative behavior of "large" sets of players can change the payoffs. The examples are numerous: Elections, many small buyers from a few competing firms, drivers that can choose among several roads, and so on.*

Following the approach of Kohlberg et al. [1974] to Walras equilibrium analysis, Mas-Colell [1984] reformulated Schmeidler's analysis in terms of joint distributions over agents' actions and characteristics and, in particular, the concept of Cournot-Nash equilibrium distributions. Not only Mas-Colell's reformulation enabled him to obtain general existence results in an easy and elegant way but it is flexible enough to accommodate quite weak assumptions on the data (which is relevant in the framework of games with incomplete information and a continuum of players for instance). Roughly speaking, in analysing Cournot-Nash equilibria in the sense of Mas-Colell [1984] one can take great advantage of (topological but also geometric) properties of spaces of probability measures. With this respect, it is natural to expect that optimal transport theory (which is an extremely active field in research in mathematics both from an applied and fundamental point, as illustrated by the monumental textbook Villani [2009]) may be useful.

[1]TSE (GREMAQ, Université de Toulouse), Manufacture des Tabacs, 21 allée de Brienne, 31000 Toulouse, FRANCE. Adrien.Blanchet@univ-tlse1.fr

[2]CEREMADE, UMR CNRS 7534, Université Paris-Dauphine, Pl. de Lattre de Tassigny, 75775 Paris Cedex 16, FRANCE. carlier@ceremade.dauphine.fr

Even though there are very general existence results for Cournot-Nash equilibria (see for instance Kahn [1989]) in the literature, we are not aware of classes of problems where there is uniqueness and a full characterisation of such equilibria which is tractable enough to obtain close-form solutions or efficient numerical computation schemes. One of our goals is precisely to go one step beyond abstract existence results (in mixed or pure strategies) and to identify classes of non-atomic games where Cournot-Nash equilibria are unique and can be fully characterised or numerically computed.

Given a space of players types X endowed with a probability measure $\mu \in \mathcal{P}(X)$ (which gives the exogenous distribution of the type of the agents), an action space Y and a cost Φ: $X \times Y \times \mathcal{P}(Y) \to \mathbb{R}$, x-type agents taking action y pay the cost $\Phi(x, y, \nu)$ where $\nu \in \mathcal{P}(Y)$ represents the action distribution. The fact that this cost depends on the other agents actions only through the distribution ν means that who plays what does not matter *i.e.* the game is *anonymous*. A *Cournot-Nash equilibrium* is a joint probability measure $\gamma \in \mathcal{P}(X \times Y)$ with first marginal μ such that

$$\gamma(\{(x, y) \in X \times Y \ : \ \Phi(x, y, \nu) = \min_{z \in Y} \Phi(x, z, \nu)\}) = 1 \tag{1.1}$$

where ν represents γ's second marginal. The probability γ is naturally interpreted by saying that $\gamma(A \times B)$ is the probability that agents have their type in A and an action in B. The equilibrium is called *pure* if, in addition, γ is carried by a graph *i.e.* μ-a.e. the agents play in pure strategy. Condition (1.1) means that agents choose cost minimising strategies given their type and ν so that, finally, imposing that ν is the second marginal of γ is a simple self-consistency requirement.

In the sequel, we will restrict ourselves to the additively separable case where $\Phi(x, y, \nu) = c(x, y) + \mathcal{V}[\nu](y)$, which seems to be a necessary limitation for the optimal transport approach we will develop. Under this separability specification, the connection with optimal transport is almost obvious: if γ is a Cournot-Nash equilibrium it necessarily minimises the average of c among probability measures having μ and ν (which is *a priori* unknown) as marginals *i.e.* it solves the optimal transport problem:

$$\mathcal{W}_c(\mu, \nu) := \inf_{\gamma \in \Pi(\mu, \nu)} \iint_{X \times Y} c(x, y) \, \mathrm{d}\gamma(x, y) \tag{1.2}$$

where $\Pi(\mu, \nu)$ is the set of joint probabilities having μ and ν as marginals. In an euclidean setting, there are well-known conditions on c and μ which guarantee that such an optimal γ necessarily is pure whatever ν is and this of course implies purity of equilibria.

If we go one step further and assume that $\mathcal{V}[\nu]$ is the differential of some functional $\mathcal{E}$ (see Section 3 for a precise definition), it turns out that if ν is a minimiser of

$\mathcal{E}[\nu]+\mathcal{W}_c(\mu,\nu)$ and γ solves (1.2) then it is a Cournot-Nash equilibrium. This gives a variational device to find equilibria: first find ν by minimising $\mathcal{E}[\nu]+\mathcal{W}_c(\mu,\nu)$ and then find γ by solving the optimal transport problem (1.2) between μ and ν. This variational approach actually gives new existence results. To the best of our knowledge, usual general existence proofs are via fixed-point arguments and thus require a lot of regularity for the dependence of $\mathcal{V}[\nu]$ with respect to ν, in a compact metric setting, it is typically asked that $\nu \mapsto \mathcal{V}[\nu]$ is continuous (or at least upper-semi continuous in some sense) from the set of probabilities equipped with the weak-$*$ topology to the set of continuous functions equipped with the supremum norm. This is harmless if Y is finite but extremely restrictive in general, in particular it excludes the case of a purely local dependence which is relevant to capture congestion effects (actions that are frequently played are more costly). In contrast, the variational approach will enable us to treat such local congestion effects. If $\mathcal{E}$ (the primitive of $\mathcal{V}$ in some sense) is convex then equilibria and minimisers coincide and strict convexity gives uniqueness of the second marginal, ν, of the equilibrium. Such a convexity is quite demanding in applications but we shall prove that in an euclidean setting and for a quadratic c (and more generally strictly convex c's in dimension one), there is some hidden convexity (in the spirit of the seminal results of McCann [1997]) in the problem from which one can deduce uniqueness of equilibria but also a characterisation in terms of a nonlinear partial differential equation of Monge-Ampère type. This partial differential equation cannot be solved explicitly in general but, in dimension one, it is easy to solve the variational problem numerically in an efficient way as we shall illustrate on several examples. Another advantage of the variational approach is that it allows for an elementary (in-)efficiency analysis of the equilibrium and the design of a tax system to restore the efficiency of the equilibrium (see Section 5). Of course, the variational approach described above presents strong similarities with the potential games of Monderer and Shapley [1996] and our framework is very close to that of Konishi et al. [1997] or LeBreton and Weber [2011] in the case of a finite number of players; however we are not aware of any extension of the analysis of Monderer and Shapley [1996] to the case of a continuum of players.

Apart from our results on Cournot-Nash equilibria, another objective of the paper is to contribute to popularise the use of optimal transport in economics. Several recent papers have fruitfully used optimal transport arguments in such different fields as hedonics and matching problems (Chiappori et al. [2010], Ekeland [2010]), multidimensional screening (Carlier [2003], Figalli et al. [2011]) or urban economics (Blanchet et al. [2012], Carlier and Ekeland [2007]). We believe that cross-fertilisation between economics and optimal transport will rapidly develop. This is why we have included in Appendix A some basic results from optimal transport theory which we hope can serve as a comprehensive introduction to this vast subject to an economists readership.

The present introduction would neither be complete nor fair without an explicit reference to the mean-field games theory of Lasry and Lions [2006a,b, 2007]. Indeed, our variational approach is largely inspired by the Lasry and Lions optimal control approach to mean-field games (that has some similarities with optimal transport), but also mean-field games theory enables to treat considerably richer situations than the somehow static one we treat here. Another line of research we would like to mention concerns congestion games (another example of potential games) and the literature on the cost of anarchy (see Roughgarden [2005] and the references therein), indeed the variational approach we develop presents some similarities with the variational approach to Wardrop equilibria on congested networks and in both cases equilibria are socially inefficient.

The paper is organised as follows. In Section 2, we introduce the model, define equilibria and emphasise some connections with optimal transport. In Section 3, we adopt a variational approach and prove that for a large class of interactions, equilibria naturally arise as local minimisers of a certain functional. Section 4 is devoted to further uniqueness and variational characterisation of equilibria results thanks to notions of displacement convexity arising in optimal transport, we also characterise the equilibrium via a certain nonlinear partial differential equation and compute numerically the equilibrium in dimension one. Section 5 concludes. The proofs as well as well as a presentation of various results from optimal transport theory which are used throughout the paper are gathered in the Appendix.

2. THE EQUILIBRIUM MODEL

The model consists of a compact metric type space X equipped with a Borel probability measure $\mu \in \mathcal{P}(X)$, giving the distribution of types, a compact metric action space Y, a reference[1] Borel non-negative measure m_0, a continuous function $c \in \mathcal{C}(X \times Y)$ and interactions are captured by a map which to every action distribution $\nu \in \mathcal{P}(Y) \cap \mathcal{L}^1(m_0)$ associates a function $\mathcal{V}[\nu]$ defined m_0-almost-everywhere. Given an action distribution ν, x-type agents taking action y then incur the additively separable cost $c(x, y) + \mathcal{V}[\nu](y)$. The unknown is a probability distribution $\gamma \in \mathcal{P}(X \times Y)$, with the interpretation that $\gamma(A \times B)$ is the probability that an agent has her type in A and takes an action in Y, such a γ induces as action distribution ν, its second marginal which we denote $\nu = \pi_{Y\#}\gamma$. By construction, the first marginal of γ, $\pi_{X\#}\gamma$ should be equal to μ. Since we will be interested by efficiency (or rather inefficiency) properties of equilibria, we will also impose that

[1]The role of the reference measure m_0 is here to capture purely local congestion effects as in the examples below. In other words, we will require action distributions to be absolutely continuous with respect to m_0. This departs from the common assumption that the cost is well defined for every action distribution and satisfies some strong continuity/semi-continuity with respect to the weak $*$ topology of measures as in Mas-Colell [1984] or Kahn [1989].

γ has finite social cost, where the latter is given by

$$\begin{aligned}\mathrm{SC} &= \iint_{X\times Y} (c(x,y) + \mathcal{V}[\nu](y))\ \mathrm{d}\gamma(x,y)\\ &= \iint_{X\times Y} c(x,y)\ \mathrm{d}\gamma(x,y) + \int_Y \mathcal{V}[\nu](y)\ \mathrm{d}\nu(y)\end{aligned}$$

since c is continuous the first term is finite for every γ, but the second requires the action marginal ν to belong to the domain

$$\mathcal{D} := \{\nu \in \mathcal{L}^1(m_0)\ :\ \mathcal{V}[\nu] \in \mathcal{L}^1(\nu)\} = \{\nu \in \mathcal{L}^1(m_0)\ :\ \int_Y |\mathcal{V}[\nu]|\ \mathrm{d}\nu < +\infty\}. \tag{2.1}$$

Cournot-Nash equilibria are then defined as follows.

Definition 2.1 *$\gamma \in \mathcal{P}(X\times Y)$ is a* Cournot-Nash equilibria *if its first marginal is μ, its second marginal, ν, belongs to $\mathcal{D}$ and there exists $\varphi \in \mathcal{C}(X)$ such that*

$$c(x,y) + \mathcal{V}[\nu](y) \geq \varphi(x) \text{ for all } x \in X \text{ and } m_0\text{-a.e. } y \text{ with equality } \gamma\text{-a.e.}. \tag{2.2}$$

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

The previous definition is slightly different from that of Mas-Colell [1984] because we require the action distribution to be absolutely continuous with respect to m_0, so as to take into account congestion effects as explained in the examples below. This makes the existence of equilibria nontrivial, indeed, when $\nu \mapsto \mathcal{V}[\nu]$ is continuous from $(\mathcal{P}(Y), w{-}*)$ to $(\mathcal{C}(Y), \|.\|_\infty)$ (as is the case for instance when $\mathcal{V}[\nu](y) = \int_Y \phi(y,z)\ \mathrm{d}\nu(z)$ with ϕ continuous) standard fixed-point arguments immediately give the existence of Cournot-Nash equilibria but here, we do not have such regularity.

2.1. *Examples*

Holiday choice

Let us consider a population of agents whose location is distributed according to some probability distribution $\mu \in \mathcal{P}(X)$ where X is some compact subset of $\mathbb{R}^2$ (say). These agents have to choose their holidays destination (possibly in mixed strategy). The set of possible holiday destinations is some compact subset of the plane Y (it can be X, a finite set, ...). The commuting cost from x to y is $c(x,y)$. In addition to the commuting cost, agents incur costs resulting from interactions with other agents, this is captured by a map $\nu \mapsto \mathcal{V}[\nu]$ that can be modelled as follows. A natural effect that has to be taken into account is congestion, *i.e.* the fact that

more crowded location results in more disutility for the agents. Congestion thus requires to consider local effects and actually imposes that ν is not too concentrated; a way to capture this is to impose that ν is absolutely continuous with respect to some reference probability measure m_0. Still denoting by ν the Radon-Nikodym derivative of ν, a natural congestion cost is of the form $y \mapsto f(\nu(y))$ with f non-decreasing. In addition to the negative externality due to congestion effect, there may be a positive externality effect due to the positive social interactions between agents which can be captured through a non-local term of the form $y \mapsto \int_Y \phi(y,z)\ \mathrm{d}\nu(z)$ where for instance $\phi(y,.)$ is minimal for $z = y$ so that the previous term represents a cost for being far from the rest of the population. Finally, the presence of purely geographical factors (e.g. distance to the sea) can be reflected by a term of the form $y \mapsto v(y)$. The total externality cost generated by the distribution ν combines the three effects of congestion, positive interactions and geographical factors and can then be taken of the form

$$\mathcal{V}[\nu](y) = f(\nu(y)) + \int_Y \phi(y,z)\ \mathrm{d}\nu(z) + v(y).$$

Technological choice

Consider now a simple model of technological choice in the presence of externalities. There is a set of consumers indexed by a type $x \in X$ drawn according to the probability μ, and a set of technologies Y for a certain good (cell-phone, computer, tablet...). On the supply side, assume there is a single profit maximising profit firm with convex production cost $F(y,.)$ producing technology y, the supply (equals demand at equilibrium) of this firm is thus determined by the marginal pricing rule $p(y) = \partial_\nu F(y,\nu(y))$. Agents aim to minimise with respect to y a total cost which is the sum of their individual purchasing cost $c(x,y)+p(y) = c(x,y)+\partial_\nu F(y,\nu(y))$ and an additional usage/maintenance or accessibility cost which is positively affected by the number of consumers having purchased similar technologies *i.e.* a term of the form $\int_Y \phi(y,z)\ \mathrm{d}\nu(z)$ where ϕ is increasing in the distance between technologies y and z.

2.2. *Connection with optimal transport and purity of equilibria*

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 (nonempty) 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 2.2 *If γ is a Cournot-Nash equilibrium and ν denotes its second marginal then $\gamma \in \Pi_o(\mu,\nu)$.*

PROOF: Indeed, let $\varphi \in \mathcal{C}(X)$ be such that (2.2) holds and let $\eta \in \Pi(\mu,\nu)$ then we have

$$\begin{aligned}\iint_{X\times Y} c(x,y)\ \mathrm{d}\eta(x,y) &\geq \iint_{X\times Y} (\varphi(x) - \mathcal{V}[\nu](y))\ \mathrm{d}\eta(x,y)\\ &= \int_X \varphi(x)\ \mathrm{d}\mu(x) - \int_Y \mathcal{V}[\nu](y)\ \mathrm{d}\nu(y) = \iint_{X\times Y} c(x,y)\ \mathrm{d}\gamma(x,y)\end{aligned}$$

so that $\gamma \in \Pi_o(\mu,\nu)$. *Q.E.D.*

The previous proof also shows that φ solves the dual of $\mathcal{W}_c(\mu,\nu)$ (see Appendix (A.3)) *i.e.* maximises the functional

$$\int_X \varphi(x)\ \mathrm{d}\mu(x) + \int_Y \varphi^c(y)\ \mathrm{d}\nu(y)$$

where φ^c denotes the c-transform of φ *i.e.*

$$\varphi^c(y) := \min_{x\in X}\{c(x,y) - \varphi(x)\}\ . \tag{2.3}$$

In an euclidean setting, there are well-known conditions on c and μ which guarantee that such an optimal γ necessarily is pure whatever ν is. It is the case for instance if μ is absolutely continuous with respect to the Lebesgue measure, $c(x,y)$ is a smooth and strictly convex function of $x-y$ (see McCann and Gangbo [1996] who extended the seminal results of Brenier [1991] in the quadratic cost case), or more generally, when it satisfies a generalised Spence-Mirrlees condition (see Carlier [2003] for details):

COROLLARY 2.3 *Assume that $X = \overline{\Omega}$ where Ω is some open connected bounded subset of $\mathbb{R}^d$ with negligible boundary, that μ is absolutely continuous with respect to the Lebesgue measure, that c is differentiable with respect to its first argument, that $\nabla_x c$ is continuous on $\mathbb{R}^d \times Y$ and that it satisfies the generalised Spence-Mirrlees condition:*

for every $x \in X$, the map $y \in Y \mapsto \nabla_x c(x,y)$ is injective,

then for every $\nu \in \mathcal{P}(Y)$, $\Pi_o(\mu,\nu)$ consists of a single element and the latter is of the form $\gamma = (\mathrm{id}, T)_{\#}\mu$ hence every Cournot-Nash equilibrium is pure.

3. A VARIATIONAL APPROACH

In this section, we will see that in many relevant cases, one may obtain equilibria by the minimisation of some functional over a set of probability measures[2]. The main assumption for this variational approach to be valid is that the interaction map $\mathcal{V}[\nu]$ has the structure of a differential *i.e.* that $\mathcal{V}[\nu]$ can be seen as the first variation of some function $\nu \mapsto \mathcal{E}[\nu]$. In this case, the variational approach is based on the observation that the equilibrium condition is the first-order optimality condition for the minimisation of $\mathcal{W}_c(\mu,\nu) + \mathcal{E}[\nu]$.

3.1. *Interaction maps which are differentials*

The main assumption for the variational approach to be valid is that $\nu \mapsto \mathcal{V}[\nu]$ is a differential in the following sense:

Definition 3.1 (Differential) *Let $\mathcal{D}$ be defined by (2.1). The map $\nu \in \mathcal{D} \mapsto \mathcal{V}[\nu]$ is a* differential *on $\mathcal{D}$ if $\mathcal{D}$ is convex and there exists $\mathcal{E}\colon \mathcal{D} \to \mathbb{R}$ such that for every $(\rho,\nu) \in \mathcal{D}^2$, $\mathcal{V}[\nu] \in \mathcal{L}^1(\rho)$ and*

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

i.e. $\mathcal{V}[\nu]$ is the first variation of $\mathcal{E}$ which we denote $\mathcal{V}[\nu] = \dfrac{\delta \mathcal{E}}{\delta \nu}$.

Before going any further, let us consider some examples to illustrate the previous definition.

Local term

Let us consider first the case of a local dependence, again m_0 is our reference measure and for $\nu \in \mathcal{D} := \mathcal{P}(Y) \cap \mathcal{L}^1(m_0)$ and m_0-a.e. y:

$$\mathcal{V}[\nu](y) = f(y, \nu(y))$$

for some continuous f. Assume first that f is bounded and define, for all $\nu \in \mathcal{D}$:

$$F(y,\nu) = \int_0^\nu f(y,s) \, \mathrm{d}s, \quad \mathcal{E}[\nu] = \int_Y F(y,\nu(y)) \, \mathrm{d}m_0(y) \, . \tag{3.1}$$

[2]Note the analogy with the variational approach of Monderer and Shapley [1996] for potential games, *i.e.* games whose equilibria can be obtained by minimising some potential function.

Then since F is Lipschitz in ν uniformly in y, it easily follows from Lebesgue's dominated convergence theorem that $\mathcal{V}[\nu]$ is the differential of $\mathcal{E}$ on $\mathcal{D}$. Now, rather assume that f satisfies the growth condition

$$a(\nu^\alpha - 1) \leq f(y,\nu) \leq b(\nu^\alpha + 1) \tag{3.2}$$

for some $a \geq 0$, $b > 0$, $\alpha > 0$, m_0-a.e. y and every $\nu \geq 0$. For $p = \alpha + 1$, the corresponding energy functional $\mathcal{E}$ is then defined for all $\nu \in \mathcal{D} := \mathcal{P}(Y) \cap \mathcal{L}^p(m_0)$ as above by (3.1). Thanks to (3.2) and Lebesgue's dominated convergence theorem, $\mathcal{V}$ is the differential of $\mathcal{E}$ on $\mathcal{D}$, $\mathcal{V}[\nu] \in \mathcal{L}^{p'}(m_0)$ (with p' the conjugate exponent of p *i.e.* $p' = p/(p-1)$) as soon as $\nu \in \mathcal{D}$. Apart from the technical growth condition (which is useful to apply Lebesgue's theorem and guarantee that $\mathcal{V}[\nu]\nu$ is integrable) we therefore see that local $\mathcal{V}$'s are differentials. In Section 3.3, we will treat local $\mathcal{V}$'s under a different Inada-like condition on f which is more customary in economics and will ensure that ν remains positive hence simplifying the equilibrium/optimality condition.

Non-local interaction term and the role of symmetry

Let us now consider the case of (pairwise) interactions where $\mathcal{V}[\nu]$ is defined by

$$\mathcal{V}[\nu](y) = \int_Y \phi(y,z)\, \mathrm{d}\nu(z)$$

for some $\phi \in \mathcal{C}(Y \times Y)$. It is then natural to define the quadratic functional

$$\mathcal{E}[\nu] = \frac{1}{2} \iint_{Y\times Y} \phi(y,z)\, \mathrm{d}\nu(y)\, \mathrm{d}\nu(z)\ .$$

By expanding in ε, $\mathcal{E}[\nu + \varepsilon(\rho - \nu)]$, its differential is immediate to compute

$$\begin{aligned}\lim_{\varepsilon\to 0} &\frac{\mathcal{E}[\nu + \varepsilon(\rho-\nu)] - \mathcal{E}[\nu]}{\varepsilon} \\ &= \frac{1}{2}\iint \phi(y,z)[\mathrm{d}\nu(y)\, \mathrm{d}(\rho-\nu)(z) + \mathrm{d}\nu(z)\, \mathrm{d}(\rho-\nu)(y)] \\ &= \frac{1}{2}\iint [\phi(y,z) + \phi(z,y)]\, \mathrm{d}\nu(z)\, \mathrm{d}(\rho-\nu)(y)\ .\end{aligned}$$

So that

$$\frac{\delta\mathcal{E}}{\delta\nu}(y) = \int_Y \phi^{\mathrm{sym}}(y,z)\, \mathrm{d}\nu(z)\ :\ \phi^{\mathrm{sym}}(y,z) = \frac{\phi(y,z) + \phi(z,y)}{2}.$$

Hence $\mathcal{V}$ is the differential of $\mathcal{E}$ on $\mathcal{P}(Y)$ as soon as ϕ is symmetric[3] *i.e.* $\phi(y,z) = \phi(z,y)$ (which is the case for instance if ϕ is the function of the distance between y

[3]Let us remark that in the case of a finite number of players, the role of symmetry for the potential approach to work was already pointed out in LeBreton and Weber [2011].

and z). Note that the assumption that $\mathcal{V}$ is a differential requires ϕ to be symmetric.[4] Of course, one can combine the previous examples and consider a $\mathcal{V}$ which is the sum of a symmetric interaction term and a local term, such $\mathcal{V}$'s still have the structure of a differential.

3.2. *Minimisers are equilibria*

Throughout this paragraph, we assume that

$$\mathcal{V}[\nu] = \frac{\delta \mathcal{E}}{\delta \nu} \text{ on } \mathcal{D}. \tag{3.4}$$

We then consider the variational problem

$$\inf_{\nu \in \mathcal{D}} \mathcal{J}_\mu[\nu] \quad \text{where} \quad \mathcal{J}_\mu[\nu] := \mathcal{W}_c(\mu, \nu) + \mathcal{E}[\nu]. \tag{3.5}$$

To prove that minimisers of (3.5) are equilibria, we first need to be able to differentiate the term $\mathcal{W}_c(\mu, \nu)$ with respect to ν, this is possible thanks to Lemma A.1 proved in Appendix A but it requires more structure on X, μ and c: in particular X is a connected subset of $\mathbb{R}^d$, c is differentiable with respect to x and μ is equivalent to the Lebesgue measure.

Theorem 3.2 (Minimisers are equilibria) *Assume that $\mathcal{V}[\nu]$ satisfies (3.4) with $\mathcal{D} = \mathcal{P}(Y) \cap \mathcal{L}^p(m_0)$ for some $p \in [1, +\infty[$ and that the assumptions of* Lemma A.1 *hold true. If ν solves (3.5) and $\gamma \in \Pi_o(\mu, \nu)$ then γ is a Cournot-Nash equilibrium.*

See Appendix B.1 for the proof. Let us mention however that the optimality condition for (3.5) is the following: there is a constant M such that

$$\begin{cases} \varphi^c + \mathcal{V}[\nu] \geq M \\ \varphi^c + \mathcal{V}[\nu] = M \qquad \nu\text{-a.e. ,} \end{cases} \tag{3.6}$$

[4] In a similar way, if we consider the case of higher-order interactions

$$\mathcal{V}[\nu](y) = \int_{Y^m} \phi(y, .) \, \mathrm{d}\nu^{\otimes m} = \int_{Y^m} \phi(y, z_1, \dots, z_m) \, \mathrm{d}\nu(z_1) \dots \, \mathrm{d}\nu(z_m)$$

where $\phi \in \mathcal{C}(Y^{m+1})$ satisfies the symmetry relations

$$\phi(y, z_1, \dots, z_m) = \phi(z_1, y, \dots, z_m) = \dots = \phi(z_m, z_1, \dots, y), \tag{3.3}$$

for all $(y, z_1, \dots, z_m) \in Y^{m+1}$, then $\mathcal{V}$ is the differential of

$$\mathcal{E}[\nu] = \frac{1}{m+1} \int_{Y^{m+1}} \phi \, \mathrm{d}\nu^{\otimes (m+1)}.$$

where φ^c is the c-transform of φ as in (2.3).

To deduce an existence result from Theorem 3.5, assume that $\mathcal{V}[\nu]$ is defined for $\nu \in \mathcal{P}(Y) \cap \mathcal{L}^1(m_0)$ by

$$\mathcal{V}[\nu](y) := f(y, \nu(y)) + \int_Y \phi(y, z) \, \mathrm{d}\nu(z) \tag{3.7}$$

where $\phi \in \mathcal{C}(Y \times Y)$ is symmetric, f is continuous and non-decreasing with respect to its second argument and satisfies the growth condition (3.2) for some $a > 0$, $b > 0$ and $\alpha > 0$. For $p = \alpha + 1$, the corresponding energy functional is then defined for all $\nu \in \mathcal{D} := \mathcal{P}(Y) \cap \mathcal{L}^p(m_0)$ by

$$\mathcal{E}[\nu] = \int_Y F(y, \nu(y)) \, \mathrm{d}m_0(y) + \frac{1}{2} \iint_{Y^2} \phi(y, z) \, \mathrm{d}\nu(y) \, \mathrm{d}\nu(z)$$

where F is defined by (3.1). The functional $F(y, .)$ is convex and satisfies the growth condition

$$a(p^{-1}\nu^p - \nu) \leq F(y, \nu) \leq b(p^{-1}\nu^p + \nu) \, .$$

Hence $\mathcal{V}[\nu] \in \mathcal{L}^{p'}(m_0)$ as soon as $\nu \in \mathcal{D}$ and thus, by Hölder's inequality, $\mathcal{V}[\nu]\rho \in \mathcal{L}^1(m_0)$ for every ρ and ν in $\mathcal{D}$.

Corollary 3.3 (Existence of equilibria by minimisation) *Assume that the assumptions of* Lemma A.1 *hold, that $\nu \mapsto \mathcal{V}[\nu]$ is of the form* (3.7) *where f and ϕ satisfy the assumptions above, then* (3.5) *admits minimisers in $\mathcal{P}(Y) \cap \mathcal{L}^p(m_0)$ so that there exists Cournot-Nash equilibria.*

The proof is given in Appendix B.2. Note that this in particular provides existence of equilibria results for the holiday and technological choice model examples above.

Remark 3.4 Under the assumptions of the previous corollary, one can prove that the minimisers are actually bounded: indeed let ν be such a minimiser either $\nu(y) = 0$ or $\nu(y) > 0$ and for m_0-a.e. such points by the optimality condition (3.6) and (3.2) one should have for some constant M

$$a(\nu(y)^\alpha - 1) \leq f(y, \nu(y)) = M - \varphi^c(y) - \int_Y \phi(y, z) \, \mathrm{d}\nu(z) \, .$$

Since φ is a c-transform, it is continuous hence bounded on Y and the integral term is bounded since ϕ is. We therefore have $\nu \in \mathcal{L}^\infty(m_0)$.

Let us now emphasise the role of convexity in the variational approach. As expected if $\mathcal{E}$ is convex, then finding equilibria and minimising $\mathcal{J}_\mu$ are equivalent:

Proposition 3.5 (Equivalence in the convex case) *Assume that the assumptions of* Theorem 3.2 *are satisfied. If $\mathcal{E}$ is convex on $\mathcal{D}$ then the following statements are equivalent:*

- *ν solves* (3.5) *and $\gamma \in \Pi_o(\mu, \nu)$,*
- *γ is an equilibrium and $\nu = \Pi_{Y\#}\gamma$.*

If moreover $\mathcal{E}$ is strictly convex the following uniqueness result holds:

Corollary 3.6 (Uniqueness in the strictly convex case) *Assume that the assumptions of* Theorem 3.2 *are satisfied. If $\mathcal{E}$ is strictly convex then all equilibria share the same second marginal ν. If in addition, the assumptions of* Corollary 2.3 *are satisfied then there is at most one Cournot-Nash equilibrium.*

As an application, let us observe that if the assumptions of Lemma A.1 are satisfied and if $\mathcal{V}[\nu](y) = f(y, \nu(y))$ with an f which is increasing in ν and satisfies (3.2) then there exists a unique minimiser so the previous uniqueness result holds. This applies naturally to the technological choice equilibrium problem as well as to the holiday choice example with pure congestion or, more generally, in the case where the congestion effects dominate as explained below.

In the case where

$$\mathcal{E}[\nu] = \int_Y F(y, \nu(y))\, \mathrm{d}m_0(y) + \frac{1}{2} \iint_{Y^2} \phi(y, z)\, \mathrm{d}\nu(y)\, \mathrm{d}\nu(z)$$

the second non-local term typically favours the concentration of ν (when $\phi(y, z)$ is increasing with the distance between y and z for instance) and it is not convex, while the congestion terms fosters dispersion and is convex. There may however be some compensation between the two terms that makes $\mathcal{E}$ convex. For instance, by Cauchy-Schwarz inequality, the quadratic form

$$\int_Y \nu^2(y)\, \mathrm{d}m_0(y) + \iint_{Y^2} \phi(y, z)\, \mathrm{d}\nu(y)\, \mathrm{d}\nu(z)$$

is positive definite hence convex as soon as

$$\int_{Y^2} \phi^2(y, z)\, \mathrm{d}m_0(y)\, \mathrm{d}m_0(z) < 1.$$

Whence in this case, the uniqueness result of Corollary 3.6 applies.

3.3. *The case of Inada's condition*

We now consider the case where $\mathcal{V}$ contains a local congestion term that satisfies an Inada-like condition:

$$\lim_{\nu \to 0^+} f(\nu) = -\infty \quad \text{and} \quad \lim_{\nu \to +\infty} f(\nu) = +\infty. \tag{3.8}$$

This will imply that minimisers of (3.5) are positive m_0-a.e.. The optimality condition $\varphi^c + \mathcal{V}[\nu] = M$, for some constant M, will therefore be satisfied m_0-a.e. which implies the regularity of ν. More precisely, let us consider the case where the interaction are given for $\nu \in \mathcal{P}(Y) \cap \mathcal{L}^1(m_0)$ by the map:

$$\mathcal{V}[\nu](y) = f(\nu(y)) + \int_Y \phi(y,z)\, \mathrm{d}\nu(z) \tag{3.9}$$

(for the sake of simplicity we have dropped the dependence in y of f) where

- $\phi \in \mathcal{C}(Y \times Y)$ is symmetric,
- $f : (0, +\infty) \mapsto \mathbb{R}$ is continuous increasing, locally integrable on $[0, +\infty)$ and satisfies the Inada condition (3.8).

We then define F by $F(0) = 0$ and $F' = f$ so that F is strictly convex, continuous on $[0, +\infty)$ and $\mathcal{C}^1(0, +\infty)$, bounded from below and coercive *i.e.* $F(\nu)/\nu \to +\infty$ as $\nu \to +\infty$. As before, for any ν in $\mathcal{P}(Y) \cap \mathcal{L}^1(m_0)$, we define the associated cost functional

$$\mathcal{E}[\nu] = \begin{cases} \displaystyle\int_Y F(\nu)\, \mathrm{d}m_0 + \frac{1}{2} \iint_{Y^2} \phi(y,z)\, \mathrm{d}\nu(y)\, \mathrm{d}\nu(z) & \text{if } \displaystyle\int_Y F(\nu)\, \mathrm{d}m_0 < +\infty \\ +\infty & \text{otherwise.} \end{cases}$$

The typical example we have in mind is $f(\nu) = \log(\nu)$ and $F(\nu) = \nu \log \nu - \nu$, or simply $F(\nu) = \nu \log \nu$ since we are only dealing with probability measures. In this case, the domain of $\mathcal{E}$ consists of absolutely continuous measures with finite entropy. Again, we look for equilibria by solving the minimisation problem (3.5). The implication of Inada's condition on the interiority of minimisers is given by the following:

Lemma 3.7 (Existence and positivity of minimisers) *Under the above assumptions, the variational problem* (3.5) *admits solutions and if ν is such a solution $\nu \geq \delta$ m_0-a.e. for some $\delta > 0$ and $\nu \in \mathcal{L}^\infty(m_0)$.*

The proof (see Appendix B.4) relies on the fact that since $f(0^+) = -\infty$ the functional $\mathcal{E}$ abhors a vacuum. Now that we know that minimisers ν exist and are bounded from above and bounded away from 0, under the assumptions of Lemma A.1 it is easy to see, as in the previous paragraph, that they necessarily are equilibria and satisfy the optimality condition (3.6):

$$f(\nu(y)) + \int_Y \phi(y,z)\, \mathrm{d}\nu(z) + \varphi^c(y) = M$$

for some constant M and where as usual φ is the Kantorovich potential between μ and ν and φ^c is its c-transform. Note that this equality is true not only ν-a.e.

but m_0-a.e., one can then invert this relation to deduce that ν coincides m_0 with the continuous function

$$\nu(y) = f^{-1}\left(M - \varphi^c(y) - \int_Y \phi(y,z)\,\mathrm{d}\nu(z)\right). \tag{3.10}$$

In particular there exists equilibria that have a continuous representative.[5] Relation (3.10) is however not very tractable in general since it involves the very indirect quantity φ^c and an integral term. We will see in Section 4 how it can be simplified and reformulated as a nonlinear partial differential equation in the case of a quadratic cost.

For the moment, the Inada condition has just enabled us to prove some further regularity properties of minimisers hence of some special equilibria. Let us summarise all this by:

THEOREM 3.8 (Main results under the Inada condition) *Let $\mathcal{V}$ be of the form (3.9) where f and ϕ satisfy the assumptions of this paragraph. If ν solves (3.5) and $\gamma \in \Pi_o(\mu,\nu)$ then γ is an equilibrium; in particular, there exists equilibria. Moreover any minimiser ν of (3.5) is bounded and bounded away from 0 and coincides m_0-a.e. with the continuous function given by (3.10).*

If, in addition, $\mathcal{E}$ is convex γ is an equilibrium if and only if $\gamma \in \Pi_o(\mu,\nu)$ where ν solves (3.5).

If, in addition, $\mathcal{E}$ is strictly convex there is a uniqueness of the equilibrium second marginal ν.

4. HIDDEN CONVEXITY AND FURTHER UNIQUENESS RESULTS

So far, our variational approach has enabled us to prove the existence of equilibria by the minimisation problem (3.5). However, the previous results are not totally satisfying since in general there might exist equilibria that are not minimisers and even if we are only interested in the special equilibria obtained by minimisation, optimality conditions like (3.10) are not tractable enough to provide a full characterisation. Under further convexity conditions that are quite stringent we have seen that equilibria necessarily are minimisers and obtained uniqueness of both. In the case where

$$\mathcal{E}[\nu] = \int_Y F(y,\nu(y))\,\mathrm{d}m_0(y) + \frac{1}{2}\iint_{Y\times Y} \phi(y,z)\,\mathrm{d}\nu(y)\,\mathrm{d}\nu(z)$$

[5] Inada's condition is actually not essential to obtain a relation of the form (3.10). Indeed, in the case of a power congestion function, $f(\nu) = \nu^\alpha$, $\alpha > 0$, using the positive part function, one obtains a similar relation

$$\nu(y) = \left(M - \varphi^c(y) - \int_Y \phi(y,z)\,\mathrm{d}\nu(z)\right)_+^{1/\alpha}.$$

there is a competition between the convexity of the congestion term that favours dispersion and the non-convexity of the interaction term so that in general nothing can be said about the convexity of $\mathcal{E}$ in the usual sense. We shall see however, that some convexity structure, more adapted to optimal transport, can be used to derive new uniqueness and characterisation results. The aim of this section is precisely to exploit some hidden convexity structure in one dimension and in higher dimensions when the cost is quadratic. This goal can be achieved thanks to the very powerful notion of displacement convexity (or some slight variant of it) due to McCann [1997]. In recent years, these notions of convexity, intimately linked to optimal transport, have proved to be an extremely useful and flexible tool in particular in the study of nonlinear diffusions, to our knowledge, this is the first time they are used in an economic context, see also Blanchet et al. [2012]. We refer to Appendix A for a very short presentation and Section 4.1 for a detailed exposition in the easier one-dimensional case. Much more on this rich subject can be found in the books Ambrosio et al. [2005], Villani [2003, 2009].

4.1. *Hidden convexity in dimension one*

Let us start with the simple one-dimensional case where the intuition is easy to understand: the functional $\mathcal{J}_\mu$ 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 $\mathcal{V}[\nu]$ takes the form:

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

and that

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

As already noted the corresponding cost

$$\mathcal{E}[\nu] := \int_0^1 F(\nu(y))\, \mathrm{d}y + \frac{1}{2} \iint_{[0,1]^2} \phi(y,z)\, \mathrm{d}\nu(y)\, \mathrm{d}\nu(z) + \int_0^1 v(y)\, \mathrm{d}\nu(y)$$

(with $F' = f$) is not convex in the usual sense in general and neither is the functional $\mathcal{J}_\mu = \mathcal{W}_c(\mu, .) + \mathcal{E}$.

However, we shall see that $\mathcal{J}_\mu$ has good convexity properties when one considers the following interpolation. 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 non-decreasing (see Villani [2003]). For $t \in [0,1]$, let us define:

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

then by construction, the curve $t \mapsto \nu_t$ connects $\nu_0 = \nu$ to $\nu_1 = \rho$.

Definition 4.1 *A functional $\mathcal{J} : \mathcal{P}(Y) \to \mathbb{R} \cup \{+\infty\}$ is called displacement convex whenever $t \in [0,1] \mapsto \mathcal{J}[\nu_t]$ is convex (for every choice of endpoints ν and ρ), it is called strictly displacement convex when, in addition $\mathcal{J}[\nu_t] < (1-t)\mathcal{J}[\nu] + t\mathcal{J}[\rho]$ when $t \in (0,1)$ and $\rho \neq \mu$.*

We claim that $\mathcal{J}_\mu$ is strictly displacement convex; indeed, take (ν, ρ) two probability measures in the domain of $\mathcal{E}$ (which is convex by convexity of F), define ν_t as above and, let us consider the four terms in $\mathcal{J}_\mu$ separately:

- By definition of $\mathcal{W}_c$, ν_t and the strict convexity of C we have
$$\begin{aligned}\mathcal{W}_c(\mu,\nu_t) &\leq \int_0^1 C(x - ((1-t)T_0(x) + tT_1(x)))\,\mathrm{d}\mu(x)\\ &\leq (1-t)\int_0^1 C(x - T_0(x))\,\mathrm{d}\mu(x) + t\int_0^1 C(x - T_1(x))\,\mathrm{d}\mu(x)\\ &= (1-t)\mathcal{W}_c(\mu,\nu) + t\mathcal{W}_c(\mu,\rho)\end{aligned}$$
with a strict inequality if $t \in (0,1)$ and $\nu \neq \rho$,
- By construction
$$\int_0^1 v\,\mathrm{d}\nu_t = \int_0^1 v(T_t(x))\,\mathrm{d}\mu(x)$$
which is convex with respect to t, by convexity of v,
- Similarly
$$\iint_{[0,1]^2} \phi(y,z)\,\mathrm{d}\nu_t(y)\,\mathrm{d}\nu_t(z) = \iint_{[0,1]^2} \phi(T_t(x), T_t(y))\,\mathrm{d}\mu(x)\,\mathrm{d}\mu(y)$$
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 non-decreasing, at least formally[6] we have $\nu_t(T_t(x))T_t'(x) = \mu(x)$, by the change of variables formula we also have
$$\int_0^1 F(\nu_t(y))\,\mathrm{d}y = \int_0^1 F(\nu_t(T_t(x)))T_t'(x)\,\mathrm{d}x = \int_0^1 F\Big(\frac{\mu(x)}{T_t'(x)}\Big)T_t'(x)\,\mathrm{d}x$$
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.

[6] see Ambrosio et al. [2005], Villani [2003, 2009] for a rigorous justification.

Under the assumptions above, $\mathcal{J}_\mu$ is therefore strictly displacement convex and thus admits at most one minimiser (indeed if ν and ρ were different minimisers, by strict displacement convexity, one would have $\mathcal{J}_\mu[\nu_{1/2}] < \frac{1}{2}\mathcal{J}_\mu[\nu] + \frac{1}{2}\mathcal{J}_\mu[\rho]$). Actually more is true (see Appendix B.5 for details): if γ is an equilibrium then its second marginal solves (3.5) and therefore is unique. Since c satisfies the generalised Spence-Mirrlees condition (see Corollary 2.3), we deduce the following uniqueness result

Theorem 4.2 (Uniqueness of an equilibrium by displacement convexity in dimension one) *Under the assumptions above, we have the equivalence*

$$\nu \text{ is a minimiser to (3.5) and } \gamma \in \Pi_o(\mu,\nu) \quad \Leftrightarrow \quad \gamma \text{ is an equilibrium}$$

and since $\mathcal{J}_\mu$ is strictly displacement convex, there is uniqueness of the equilibrium (which is actually necessarily pure).

4.2. *Hidden convexity under quadratic cost*

The arguments of the previous paragraph can be generalised in higher dimensions when the transport cost is quadratic. Throughout this section, we will assume the following:

- $X = Y = \overline{\Omega}$ where Ω is some open bounded convex subset of $\mathbb{R}^d$,
- μ is absolutely continuous with respect to the Lebesgue measure (that will be the reference measure m_0 from now on) and has a positive density on Ω,
- c is quadratic *i.e.*

$$c(x,y) := \frac{1}{2}|x-y|^2,\ (x,y) \in \mathbb{R}^d \times \mathbb{R}^d,$$

- $\mathcal{V}$ again takes the form

$$\mathcal{V}[\nu](y) = f(\nu(y)) + v(y) + \int_Y \phi(y,z)\ \mathrm{d}\nu(z)$$

  where v is convex, f satisfies the assumptions of Section 3.3 and $\phi \in \mathcal{C}(\mathbb{R}^d \times \mathbb{R}^d)$ is symmetric and $\mathcal{C}^{1,1}_{\mathrm{loc}}$ (*i.e.* $\mathcal{C}^1$ with a locally Lipschitz gradient).

Again denoting by F the primitive of f that vanishes at 0, the corresponding energy reads

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

Note that as c is quadratic, Brenier's Theorem (see Theorem A.2) implies the uniqueness and the purity of optimal plans γ between μ and an arbitrary ν. The

variational problem (3.5) then takes the form

$$\inf_{\nu\in\mathcal{P}(\overline{\Omega})} \mathcal{J}_\mu[\nu] \quad \text{where} \quad \mathcal{J}_\mu[\nu] := \frac{1}{2}\mathcal{W}_2^2(\mu,\nu) + \mathcal{E}[\nu] \tag{4.1}$$

with $\mathcal{W}_2^2(\mu,\nu)$ is the squared-2-Wasserstein distance between μ and ν i.e.:

$$\mathcal{W}_2^2(\mu,\nu) := \inf_{\gamma\in\Pi(\mu,\nu)} \int_{X^2} |x-y|^2 \, \mathrm{d}\gamma(x,y).$$

Two more structural assumptions are needed to guarantee the strict convexity of $\mathcal{J}_\mu$ along generalised geodesics with base μ (see Appendix A for details), namely McCann's condition:

$$\nu \mapsto \nu^d F(\nu^{-d}) \text{ is convex non-increasing on } (0,+\infty) \tag{4.2}$$

and that ϕ is convex. Note that McCann's condition is satisfied by power functions with an exponent larger than 1 as well as by the entropy $F(\nu) = \nu\log(\nu)$.

In contrast with Theorem 3.8 where minimisers are equilibria but the reverse is not always true, convexity along generalised geodesics ensures the converse reciprocal property:

Theorem 4.3 (Equilibria and minimisers coincide, uniqueness and regularity under generalized convexity) *Under the assumptions above, we have the equivalence*

$$\nu \textit{ is a minimiser to } (4.1) \textit{ and } \gamma\in\Pi_o(\mu,\nu) \quad \Leftrightarrow \quad \gamma \textit{ is an equilibrium}$$

Moreover, there exists a unique equilibrium (which is actually pure) and the second marginal ν of this equilibrium has a continuous density.

4.3. *A partial differential equation for the equilibrium*

In the quadratic cost framework of Paragraph 4.2, our aim now is to write the optimality condition (3.10) in the form of a nonlinear and non-local equation partial differential equation of Monge-Ampère type. For computational simplicity, we take $v = 0$ and $f(\nu) = \log(\nu)$ but any convex, $\mathcal{C}^{1,1}_{\mathrm{loc}}$, symmetric v and any increasing f satisfying McCann's condition would lead to a similar partial differential equation. Let us recall that the unique minimiser/equilibrium ν satisfies (3.10) and is actually characterised by this condition. In this equation, the less explicit term is φ^c. Thanks to Brenier's theorem (see Appendix A), this term can be made more explicit, as follows: the Brenier map T between μ and ν is the gradient of some convex function, $T = \nabla u$ with u convex, and similarly the Brenier map between ν and μ is ∇u^*

where u^* is the Legendre transform of u. In the case of a quadratic cost, u and u^* are related to the Kantorovich potential φ and its c-transform φ^c through

$$\varphi(x) = \frac{1}{2}|x|^2 - u(x), \quad \varphi^c(y) = \frac{1}{2}|y|^2 - u^*(y), \quad \forall (x,y) \in \Omega \times \Omega \,. \tag{4.3}$$

When ∇u is smooth enough, the Monge-Ampère equation, see (A.6), reads:

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

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

$$\nabla u(\Omega) = \Omega. \tag{4.4}$$

We may then rewrite the optimality condition (3.10) as

$$\nu(y) = C \exp\left(-\frac{1}{2}|y|^2 + u^*(y) - \int_Y \phi(y,z)\, \mathrm{d}\nu(z)\right) \tag{4.5}$$

where C is a normalisation constant that makes the total mass of the right hand side be 1 on Ω. Since u is defined up to an additive constant, one may actually choose $C = 1$. As $\nabla u_{\#}\mu = \nu$, we first have

$$\int_Y \phi(\nabla u(x), z)\, \mathrm{d}\nu(z) = \int_\Omega \phi(\nabla u(x), \nabla u(z))\, \mathrm{d}\mu(z) \,.$$

On the other hand, using the well-known convex analysis identity

$$u^*(\nabla u(x)) = x \cdot \nabla u(x) - u(x)$$

and performing the change of variable $y = \nabla u(x)$ in (4.5), the Monge-Ampère equation then becomes

$$\mu(x) = \det(D^2 u(x)) \exp\left(-\frac{1}{2}|\nabla u(x)|^2 + x \cdot \nabla u(x) - u(x)\right) \times \exp\left(-\int_\Omega \phi(\nabla u(x), \nabla u(z))\, \mathrm{d}\mu(z)\right) \,. \tag{4.6}$$

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

This kind of partial differential equation is rather complicated. However, in dimension 1, *i.e.* when Ω is an open interval, which we can assume to be $(0,1)$, the boundary condition (4.4) is $u'(0) = 0$, $u'(1) = 1$ and the Monge-Ampère equation (4.6) simplifies to

$$\mu(x) = u''(x) \exp\left(-\frac{1}{2}u'(x)^2 + x \cdot u'(x) - u(x) - \int_{(0,1)} \phi(u'(x), u'(z))\, \mathrm{d}\mu(z)\right).$$

Note that this differential equation automatically implies the strict convexity of u. We do not know whether this equation can be solved numerically in an iterative way but will see how the equilibrium can be computed in one dimension thanks to a convenient reformulation of (4.1) as explained below.

4.4. *Numerical computations in dimension one*

Let $\Omega = (0,1)$, m_0 be the Lebesgue measure on $(0,1)$, μ be absolutely continuous with respect to m_0 with a positive density still denoted μ. Consider again[7] the variational problem

$$\inf_\nu \mathcal{J}_\mu[\nu] \tag{4.7}$$

$$\text{where} \quad \mathcal{J}_\mu[\nu] := \frac{1}{2}\mathcal{W}_2^2(\mu,\nu) + \int_0^1 \nu \log \nu + \frac{1}{2}\iint_{[0,1]^2} \phi(y,z)\, \mathrm{d}\nu(y)\, \mathrm{d}\nu(z)$$

where ϕ is $\mathcal{C}^{1,1}_{\mathrm{loc}}$, convex and symmetric. Looking for ν amounts to look for its rearrangement or quantile function:

$$G(x) := \inf\{\lambda \ : \ \nu([0,\lambda]) \geq x\}, \ \forall x \in (0,1).$$

Note that G is non-decreasing and $G_\# m_0 = \nu$. We also denote by H the quantile of μ. It is then well known (see [Villani, 2003, Section 2.2] or [Ambrosio et al., 2005, Theorem 6.0.2]) that

$$\mathcal{W}_2^2(\mu,\nu) = \int_0^1 |G(x) - H(x)|^2 \ \mathrm{d}x.$$

Moreover since $G_\# m_0 = \nu$, we have

$$\iint_{[0,1]^2} \phi(y,z)\, \mathrm{d}\nu(y)\, \mathrm{d}\nu(z) = \iint_{[0,1]^2} \phi(G(x), G(\theta))\, \mathrm{d}x\, \mathrm{d}\theta\ .$$

And since ν is regular the change of variable formula yields

$$\int_0^1 \nu(x) \log(\nu(x))\, \mathrm{d}x = \int_0^1 \nu(G(x))\, \log(\nu(G(x)))\, G'(x)\, \mathrm{d}x = -\int_0^1 \log(G'(x))\, \mathrm{d}x\,,$$

as, by definition of G, $\nu(G(x))\, G'(x) = 1$ a.e.

Therefore reformulating (4.7) in terms of quantile consists in minimising the strictly convex functional

$$\frac{1}{2}\int_0^1 |G-H|^2 - \int_0^1 \log(G'(x))\, \mathrm{d}x + \frac{1}{2}\iint_{[0,1]^2} \phi(G(x), G(\theta))\, \mathrm{d}x\, \mathrm{d}\theta$$

[7] Again the choice of a logarithmic congestion function is not so essential, power congestion functions can be considered as well.

subject to the boundary conditions $G(0) = 0$, $G(1) = 1$. The discretisation of this variational problem is easy to solve using standard gradient descent methods, see Figure 1[8].

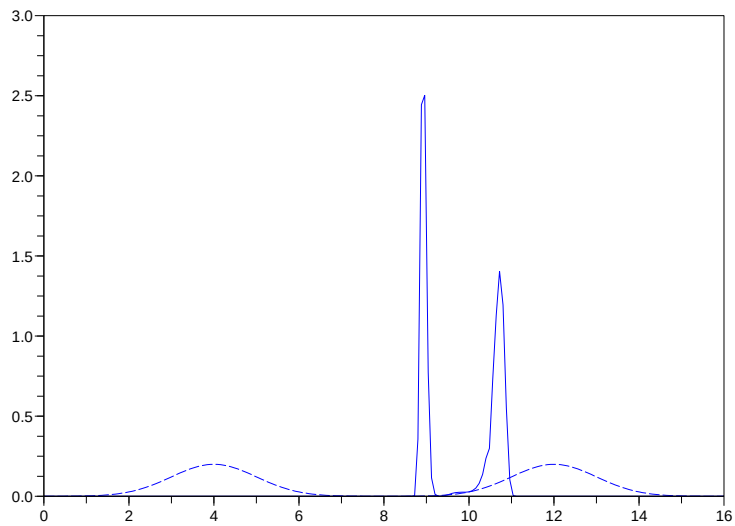

FIGURE 1.— The distribution μ of the agents is dash line and the solution ν to (4.7) in the case $f(x) = x^8$, $\phi(z) = 10^{-4}|z|^2$ and $v = (x - 10)^4$.

5. CONCLUDING REMARKS

We conclude the paper, by two remarks, for the sake of simplicity, we adopt exactly the same framework and notations, as in Section 4.2.

5.1. *Implementation by taxes*

By Theorem 4.3 the unique equilibrium is the unique minimiser of the functional $\mathcal{J}_\mu$. It would therefore be tempting to interpret this result as a kind of welfare theorem. A simple comparison between $\mathcal{J}_\mu$ and the total social cost tells us however that the equilibrium is not efficient. Indeed, the total social cost $\mathrm{SC}[\nu]$ is the sum

[8]Actually, in our simulations, we have relaxed the condition that the support of ν is $[0, 1]$ *i.e.* the boundary conditions $G(0) = 0$, $G(1) = 1$.

of the transport cost $\mathcal{W}_2^2(\mu,\nu)/2$ and the additional cost $\int_Y \mathcal{V}[\nu](y)\,\nu(y)\,\mathrm{d}y$ *i.e.*

$$\mathrm{SC}[\nu] = \frac{1}{2}\mathcal{W}_2^2(\mu,\nu) + \int_Y f(\nu(y))\,\mathrm{d}\nu(y) + \int_Y v\,\mathrm{d}\nu + \iint_{Y^2} \phi(y,z)\,\mathrm{d}\nu(y)\,\mathrm{d}\nu(z)\ .$$

The second term represents the total congestion cost and the last one the total interaction cost. The functional $\mathcal{J}_\mu$ whose minimiser is the equilibrium has a similar form, except that in its second term $f(\nu)\nu$ is replaced by $F(\nu)$ (with $F' = f$) and the interaction term is divided by 2. The equilibrium corresponds indeed to the case where agents selfishly minimise their own cost

$$c(x,.) + \mathcal{V}[\nu] = c(x,.) + f(\nu(.)) + v + \int_Y \phi(y,z)\,\mathrm{d}\nu(z)\ .$$

This individual minimisation has of course no reason to correctly estimate the marginal effect of individual behaviour on the total social cost. In other words, there is some gap between the equilibrium and the efficient (social-cost minimising) configurations, and, since we are dealing with a situation with externalities, this is actually not surprising. The computation of the equilibrium and the optimum can be done numerically in dimension 1 by using the same kind of numerical computations as explained in Section 4.4, see Figure 2.

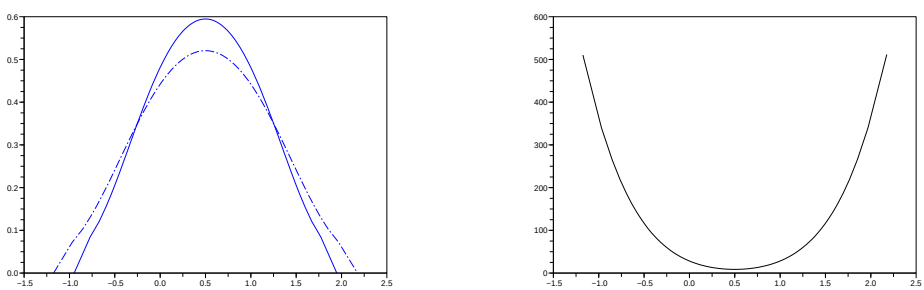

FIGURE 2.— The optimum in continuous line and the equilibrium in dash line on the left. The corresponding taxes on the right.

The natural way to restore efficiency of the equilibrium is the design by some social planner of a proper system of tax/subsidies which, added to $\mathcal{V}[\nu]$, will implement the efficient configuration (or at least a stationary point of the social cost). Thanks to our variational approach, a tax system that restores the efficiency is easy to compute (up to an additive constant):

$$\mathrm{Tax}[\nu](y) = f(\nu(y))\,\nu(y) - F(\nu(y)) + \int_Y \phi(y,z)\,\mathrm{d}\nu(z).$$

The two terms in $\mathrm{Tax}[\nu]$ represent respectively a correction to the individual estimation of congestion cost and to the individual estimation of interaction cost. A similar inefficiency of equilibria, arises in the slightly different framework of congestion games, where it is usually referred under the name *cost of anarchy*, which has been extensively studied in recent years (see Roughgarden [2005] and the references therein). In our Cournot-Nash context, we may similarly define the cost of anarchy as the ratio of the worst social cost of an equilibrium to the minimal social cost value:

$$\text{Cost of anarchy} := \frac{\max\{\mathrm{SC}[\nu^e] \ : \ \nu^e \text{ equilibrium}\}}{\min_\nu \mathrm{SC}[\nu]} \ .$$

In the previous numerical example of Figure 2, both the equilibrium and the optimum are unique and the cost of anarchy can be numerically computed as being approximately 1.8.

5.2. *A dynamical perspective*

Instead of minimising $\mathcal{J}_\mu$ directly, we may think that agents start with some distribution of strategies (that is not an equilibrium) and adjust it with time by a sort of gradient descent dynamics to decrease their individual cost dynamically. At least formally, a way to reach the equilibrium (or minimiser of $\mathcal{J}_\mu$) is then to put it into the dynamical perspective of the minimising movement scheme as follows. Fix a time step $\tau > 0$ and start with an initial configuration of strategies ν_0. The first step of the minimising movement scheme selects a new distribution of strategies ν_1 close to ν_0 (in $\mathcal{W}_2$) but also decreasing $\mathcal{J}_\mu$ by

$$\nu_1 \in \mathrm{argmin}_\nu \left\{ \frac{1}{2\tau} \mathcal{W}_2^2(\nu_0, \nu) + \mathcal{J}_\mu[\nu] \right\} \ .$$

And then it iterates the process by choosing

$$\nu_{k+1} \in \mathrm{argmin}_\nu \left\{ \frac{1}{2\tau} \mathcal{W}_2^2(\nu_k, \nu) + \mathcal{J}_\mu[\nu] \right\} \ . \tag{5.1}$$

This sort of Wasserstein Euler scheme was first introduced in Jordan et al. [1998] for the Fokker-Planck equation. Under suitable conditions it is possible to pass the continuous limit $\tau \to 0^+$ in the minimising scheme (5.1) and prove that the solution converges in some sense to the continuous evolution equation

$$\begin{cases} \partial_t \nu + \mathrm{div}\left(-\nu \nabla \left(\dfrac{\delta \mathcal{J}_\mu}{\delta \nu} \right) \right) = 0, \\ \nu_{t=0} = \nu_0 \end{cases}$$

which is the gradient flow of $\mathcal{J}_\mu$ in the Wasserstein space (see Ambrosio et al. [2005] for a detailed exposition of the theory). By construction $\mathcal{J}_\mu$ is a Lyapunov function of this equation and even though the equation may have non-unique solutions, by Lyapunov theory, it can be shown under appropriate conditions that its trajectories converge in large time to the unique minimiser of $\mathcal{J}_\mu$ *i.e.* the equilibrium. If we go back to the individual level, it can be shown that the equation above corresponds to the fact that each agent modifies her strategy according to the gradient flow of her individual cost.

We obtain a sequence of densities which converges to the equilibrium, see Figure 3. The descent algorithm is very fast and the computed equilibrium is very stable with respect to the initial density for the gradient descent as shown in the left hand figure of Figure 4.

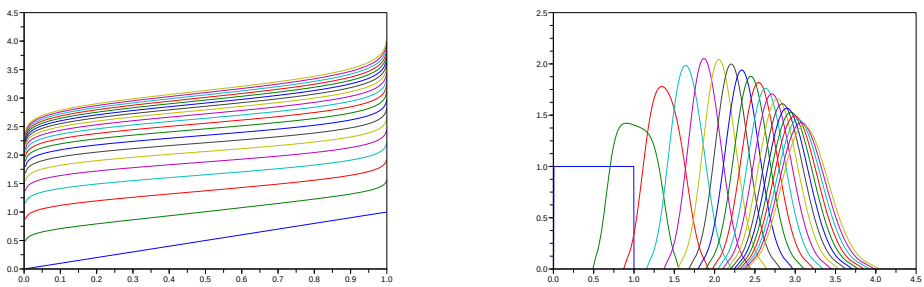

FIGURE 3.— Convergence and stabilisation toward the equilibrium in the case of a logarithmic congestion, cubic interaction, and a potential $v(x) := (x-5)^3$ with $\mathbb{1}_{[0,1]}$ as initial guess. The inverse of the cumulative function on the left and the corresponding density on the right.

Acknowledgements The authors acknowledge the support of the Agence Nationale de la Recherche through the projects ANR-09-JCJC-0096-01 EVaMEF and ANR-07-BLAN-0235 OTARIE. The authors wish to thank Jocelyn Donze, André Grimaud, Michel Le Breton, Jérôme Renault, François Salanié and the participants to the IAST LERNA - Eco/Biology and Chicago University Seminars for many interesting and fruitful discussions about the present work.

APPENDIX A: THE OPTIMAL TRANSPORT TOOLBOX

This appendix just gives some basic results from optimal transport theory that we have used in the paper, for a detailed exposition of this rich and rapidly developing subject, we refer the interested reader to the very accessible textbook Villani [2003] or Ambrosio et al. [2005], Villani [2009] or, the more probability-oriented textbook Rachev and Rüschendorf [1998].

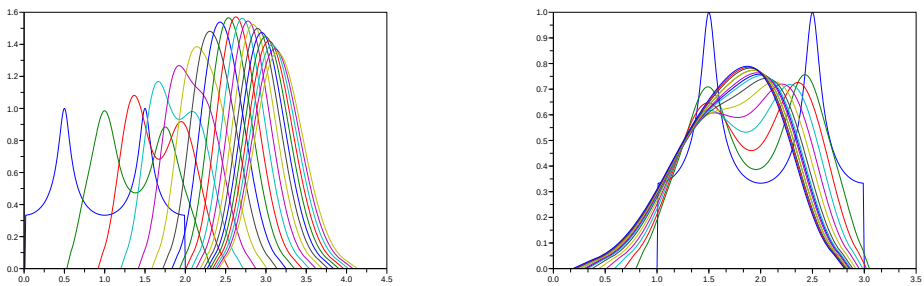

FIGURE 4.— Convergence and stabilisation toward the equilibrium in the case of a logarithmic congestion, cubic interaction with an initial guess made of two bumps. The potential is $v(x) := (x-5)^3$ on the left and $v(x) := (x-1/2)^3$ on the right.

Kantorovich duality

Let X and Y be two compact spaces equipped respectively with the Borel probability measures $\mu \in \mathcal{P}(X)$ and $\nu \in \mathcal{P}(Y)$. For $\mu \in \mathcal{P}(X)$ and T, Borel: $X \to Y$, $T_\#\mu$ denotes the *push forward* (or image measure) of μ through T which is defined by $T_\#\mu(B) = \mu(T^{-1}(B))$ for every Borel subset B of Y or equivalently by the change of variables formula

$$\int_Y \varphi \, \mathrm{d}T_\#\mu = \int_X \varphi(T(x)) \, \mathrm{d}\mu(x), \ \forall \varphi \in \mathcal{C}(X). \tag{A.1}$$

A transport map between μ and ν is a Borel map such that $T_\#\mu = \nu$. Now, let $c \in \mathcal{C}(X \times Y)$ be some transport cost function, the *Monge optimal transport* problem for the cost c consists in finding a transport T between μ and ν that minimises the total transport cost $\int_X c(x, T(x)) \, \mathrm{d}\mu(x)$. A minimiser is then called an *optimal transport*. Monge problem is in general difficult to solve (it may even be the case that there is no transport map, for instance it is impossible to transport one Dirac mass to a sum of distinct Dirac masses), this is why Kantorovich relaxed Monge's formulation as

$$\mathcal{W}_c(\mu, \nu) := \inf_{\gamma \in \Pi(\mu,\nu)} \int_{X \times Y} c(x, y) \, \mathrm{d}\gamma(x, y) \tag{A.2}$$

where $\Pi(\mu, \nu)$ is the set of transport plans between μ and ν *i.e.* Borel probability measures on $X \times Y$ having μ and ν as marginals. Since $\Pi(\mu, \nu)$ is weakly $*$ compact and c is continuous, it is easy to see that the infimum of the linear program defining $\mathcal{W}_c(\mu, \nu)$ is attained at some γ, such optimal γ's are called *optimal transport plans* (for the cost c) between μ and ν. If there is an optimal γ which is induced by a *transport map i.e.* is of the form $\gamma = (\mathrm{id}, T)_\#\mu$ for some transport map T then T is obviously an optimal solution to Monge's problem. Another advantage of the linear relaxation is that it possesses a dual formulation that can be very useful. This dual formulation consists in maximising the linear form $\int_X \varphi \, \mathrm{d}\mu + \int_Y \psi \, \mathrm{d}\nu$ among all pairs $(\varphi, \psi) \in \mathcal{C}(X) \times \mathcal{C}(Y)$ such that $\varphi(x) + \psi(y) \leq c(x, y)$, it is easy to see that this can be reformulated as a maximisation over φ only:

$$\mathcal{W}_c(\mu, \nu) := \sup_{\varphi \in \mathcal{C}(X)} \Big\{ \int_X \varphi \, \mathrm{d}\mu + \int_Y \varphi^c \, \mathrm{d}\nu \Big\} \tag{A.3}$$

where φ^c is the c-concave transform of φ *i.e.*

$$\varphi^c(y) := \min_{x\in X}\{c(x,y) - \varphi(x)\},\ \forall y \in Y.$$

Formula (A.3) is usually called *Kantorovich duality formula* and a maximiser φ in (A.3) is called a *Kantorovich potential* between μ and ν for the cost c. The existence of Kantorovich potentials under our assumptions is well-known (see Rachev and Rüschendorf [1998], Villani [2003, 2009]) and we observe that if φ is a Kantorovich potential then so is $\varphi + C$ for every constant C.

We have used in Section 3 the following result on the uniqueness of the Kantorovich potential and the differentiability of $\mathcal{W}_c(\mu,\nu)$ with respect to ν:

Lemma A.1 *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 $X \times Y$, let $\nu \in \mathcal{P}(Y)$ then there exists a unique (up to an additive constant) Kantorovich potential φ between μ and ν and for every $\rho \in \mathcal{P}(Y)$ one has*

$$\lim_{\varepsilon\to 0^+} \frac{\mathcal{W}_c(\mu,\nu+\varepsilon(\rho-\nu)) - \mathcal{W}_c(\mu,\nu)}{\varepsilon} = \int_Y \varphi^c \,\mathrm{d}(\rho-\nu).$$

Proof: The proof of the uniqueness of the Kantorovich potential φ between μ and ν up to an additive constant can be found for instance in [Carlier and Ekeland, 2007, Proposition 6.1]. As a normalisation we choose the potential φ such that $\varphi(x_0) = 0$ where x_0 is some given point of X. To shorten notations, set $\nu_\varepsilon = \nu + \varepsilon(\rho - \nu)$, thanks to Kantorovich duality formula (A.3) we have

$$\varepsilon^{-1}[\mathcal{W}_c(\mu,\nu_\varepsilon) - \mathcal{W}_c(\mu,\nu)] \geq \int_Y \varphi^c \,\mathrm{d}(\rho-\nu) \tag{A.4}$$

and similarly if φ_ε denotes the Kantorovich potential between μ and ν_ε such that $\varphi_\varepsilon(x_0) = 0$, we have

$$\varepsilon^{-1}[\mathcal{W}_c(\mu,\nu_\varepsilon) - \mathcal{W}_c(\mu,\nu)] \leq \int_Y \varphi_\varepsilon^c \,\mathrm{d}(\rho-\nu).$$

Now it is well-known that $(\varphi_\varepsilon)_\varepsilon$ is bounded and uniformly equi-continuous uniformly with respect to ε hence, thanks to Ascoli's Theorem, up to a sub-sequence, it converges uniformly to some $\overline{\varphi} \in \mathcal{C}(X)$ such that $\overline{\varphi}(x_0) = 0$ and it is easy to see that $\overline{\varphi}$ is a Kantorovich potential between μ and ν so that $\varphi = \overline{\varphi}$ and φ_ε^c converges to φ^c. We then have

$$\limsup_{\varepsilon\to 0^+} \varepsilon^{-1}[\mathcal{W}_c(\mu,\nu_\varepsilon) - \mathcal{W}_c(\mu,\nu)] \leq \int_Y \varphi^c \,\mathrm{d}(\rho-\nu).$$

The desired result thus follows from (A.4). *Q.E.D.*

When $X = Y$ and denoting by d the distance on Y, for $p \in [1,+\infty[$, the p-Wasserstein distance between $\mu \in \mathcal{P}(X)$ and $\nu \in \mathcal{P}(X)$ is by definition

$$\mathcal{W}_p(\mu,\nu) := \Big(\inf_{\gamma\in\Pi(\mu,\nu)} \Big\{\int_{X\times Y} d(x,y)^p \,\mathrm{d}\gamma(x,y)\Big\}\Big)^{1/p} \tag{A.5}$$

The Wasserstein distances are indeed distances and they metrise the weak $*$ topology of $\mathcal{P}(Y)$. For $p = 1$, it is well-known that the Kantorovich duality formula can be rewritten as

$$\mathcal{W}_1(\mu,\nu) = \sup\Big\{\int_X \varphi \,\mathrm{d}(\mu-\nu)\ :\ \varphi \text{ 1-Lipschitz}\Big\}$$

so that for every Lipschitz continuous function φ on X, one has

$$\Big|\int_X \varphi \,\mathrm{d}(\mu-\nu)\Big| \le \mathrm{Lip}(\varphi, X)\mathcal{W}_1(\mu,\nu),$$

an inequality we will use several times later on. As a simple illustration of the interest of the distance $\mathcal{W}_1$, let us equip Y^m with the distance $(x,y) = (x_1,\dots,x_m,y_1,\dots y_m) \mapsto d_m(x,y) := \sum_{k=1}^m d(x_k,y_k)$ for ν and θ in $\mathcal{P}(Y)$, let γ be an optimal transport plan between μ and ν for $\mathcal{W}_1$ then since $\gamma^{\otimes m} := \gamma\otimes\cdots\otimes\gamma$ has marginals $\nu^{\otimes m}$ and $\theta^{\otimes m}$, we have

$$\mathcal{W}_1(\nu^{\otimes m},\theta^{\otimes m}) \le \int_{Y^m\times Y^m} d_m \,\mathrm{d}\gamma^{\otimes m} = m\mathcal{W}_1(\nu,\theta)\ .$$

Which shows in particular that if $(\nu_n)_n$ weakly $*$ converges to ν then $(\nu_n^{\otimes m})_n$ weakly $*$ converges to $\nu^{\otimes m}$ *i.e.* $\int_{Y^m}\varphi\,\mathrm{d}\nu_n^{\otimes m}$ converges to $\int_{Y^m}\varphi\,\mathrm{d}\nu^{\otimes m}$ as $n\to\infty$ for every $\phi\in\mathcal{C}(Y^m)$.

Of particular interest is also the quadratic case $p=2$ in an euclidean setting for which a brief summary of the main results used in the paper is given in the next paragraphs.

The quadratic case and Monge-Ampère equation

We now restrict ourselves to the quadratic case, the solution of the quadratic optimal transport problem is due to Yann Brenier whose path-breaking paper Brenier [1991] totally renewed the field of optimal transport and was the starting point of an extremely active stream of research since the 90's.

THEOREM A.2 (Brenier's theorem) *Let $\mu\in\mathcal{P}(\mathbb{R}^d)$ be absolutely continuous with respect to the Lebesgue measure and compactly supported and $\nu\in\mathcal{P}(\mathbb{R}^d)$ be compactly supported, then the quadratic optimal transport problem*

$$W_2^2(\mu,\nu) := \inf_{\gamma\in\Pi(\mu,\nu)} \iint_{\mathbb{R}^d\times\mathbb{R}^d} |x-y|^2 \,\mathrm{d}\gamma(x,y)$$

possesses a unique solution γ which is in fact a Monge solution $\gamma = (\mathrm{id},T)_\#\mu$. Moreover $T=\nabla u$ μ-a.e. for some convex function u and ∇u is the unique (up to μ-a.e. equivalence) gradient of a convex function transporting μ to ν; $T=\nabla u$ is called the Brenier map between μ and ν.

In fact the previous theorem holds under much more general assumptions (it is enough that μ and ν have finite second moments and that μ does not charge sets of Hausdorff dimension less than $d-1$, see McCann [1995] or Villani [2003]). Brenier's theorem roughly says that there is a unique optimal transport for the quadratic cost and that it is characterised by the fact that it is of the form ∇u with u convex, in other words, solving $\nabla u_\#\mu=\nu$ with u convex determines ∇u uniquely μ-a.e.. When we have additional regularity, *i.e.* when μ and ν have regular densities (still denoted μ and ν) and ∇u is a diffeomorphism between the support of μ and that of ν, thanks to the change of variables formula, we find that u solves the Monge-Ampère partial differential equation:

$$\mu = \nu(\nabla u)\det(D^2u). \tag{A.6}$$

A deep regularity theory due to Luis Caffarelli: Caffarelli [1992a,b] implies that the Brenier map is a smooth diffeomorphism when in addition μ and ν are smooth, bounded away from 0 and have convex supports, in particular the Monge-Ampère equation is satisfied in this case which justifies the computations of Section 4.3.

Convexity along generalised geodesics

The last ingredient from optimal transport theory that we have used (in Section 4) is the powerful notion of displacement convexity along generalised geodesics due to Ambrosio, Gigli-Savaré Ambrosio et al. [2005][9]. As in Section 4.2, we assume that $X = Y = \overline{\Omega}$ where Ω is some open bounded convex subset of $\mathbb{R}^d$, that the cost is quadratic, that m_0 is the Lebesgue measure on X and that μ is absolutely continuous with respect to m_0 and has a positive density on Ω. In particular for every $\nu \in \mathcal{P}(X)$, the Brenier's map between μ and ν is well-defined. Generalised geodesics with base μ for the Wasserstein distance $\mathcal{W}_2$ and the corresponding notion of convexity are defined as follows

Definition A.3 (Convexity along generalised geodesics) *Let $\nu \in \mathcal{P}(X)$, $\rho \in \mathcal{P}(X)$, let T_0 be the Brenier's map between μ and ν and let T_1 be Brenier's map between μ and ρ, the generalised geodesic with base μ between ν and ρ is the curve of measures $t \in [0,1] \mapsto \nu_t := ((1-t)T_0+tT_1)_{\#}\mu$. The functional $\mathcal{J}\colon \mathcal{P}(X) \to \mathbb{R} \cup \{+\infty\}$ is called* convex along generalised geodesics with base μ *if for every pair of endpoints ν and ρ in $\mathcal{P}(X)$ and for every $t \in [0,1]$, one has*

$$\mathcal{J}[\nu_t] \leq (1-t)\mathcal{J}[\nu] + t\mathcal{J}[\rho].$$

If, in addition, the previous inequality is strict for $t \in (0,1)$ and $\rho \neq \nu$, $\mathcal{J}$ is called strictly convex along generalised geodesics with base μ.

In Section 4.2, we were interested in the strict convexity along generalised geodesics with base μ of the functional $\mathcal{J}_\mu$ defined by (4.1). As in Paragraph 4.1, the convexity of

$$t \mapsto \int_{Y\times Y} \phi(y,z)\,\mathrm{d}\nu_t(y)\,\mathrm{d}\nu_t(z) \quad \text{and} \quad t \mapsto \int_Y v\,\mathrm{d}\nu_t$$

directly follows from the convexity of ϕ and v respectively. As for the convexity of $t \mapsto \int_Y F(\nu_t(y))$ under McCann's condition:

$$\nu \mapsto \nu^d F(\nu^{-d}) \text{ is convex non-increasing on } (0,+\infty), \tag{A.7}$$

it follows from [Ambrosio et al., 2005, Proposition 9.3.9]. Finally, in the functional $\mathcal{J}_\mu$ defined by (4.1), we had the term $\mathcal{W}_2^2(\mu,.)$, to see that it is strictly convex along generalised geodesics with base μ, we can proceed exactly as we did for $\mathcal{W}_c(\mu,.)$ in dimension one in Paragraph 4.1.

APPENDIX B: PROOFS OF THE RESULTS

B.1. *Proof of Theorem 3.2*

Let ν be a solution of (3.5), $\rho \in \mathcal{D}$ and $\varepsilon \in (0,1)$, we then have $\varepsilon^{-1}(\mathcal{J}_\mu[\nu+\varepsilon(\rho-\nu)]-\mathcal{J}_\mu[\nu]) \geq 0$. Using the fact that $\mathcal{V}[\nu]$ is the first variation of $\mathcal{E}$ and Lemma A.1, we thus get

$$\int_Y (\varphi^c + \mathcal{V}[\nu])\,\mathrm{d}\rho \geq \int_Y (\varphi^c + \mathcal{V}[\nu])\,\mathrm{d}\nu \tag{B.1}$$

[9] Actually this notion of convexity is a slight variant of the notion of displacement convexity which first appeared in the seminal work of McCann [1997]. It is known that $\mathcal{W}_2^2(\mu,\nu)$, as a function of ν is not displacement convex in the sense of McCann (see example 9.1.5. in Ambrosio et al. [2005]) and this is the very reason why, following Ambrosio, Gigli and Savaré, we consider convexity along generalised geodesics with base μ rather than the initial notion of McCann. Let us however indicate that, in dimension one, both notions coincide.

where φ is a Kantorovich potential between μ and ν (it is unique up to an additive constant by Lemma A.1 and this constant plays no role since ρ and ν have the same mass). Minimising the left-hand side of (B.1), with respect to $\mathcal{L}^p$ probabilities ρ, this yields that ν-a.e.

$$\varphi^c + \mathcal{V}[\nu] = \inf_{\rho\in\mathcal{D}} \int_Y (\varphi^c + \mathcal{V}[\nu])\, \mathrm{d}\rho = M$$

where $M := \mathrm{Essinf}(\varphi^c + \mathcal{V}[\nu])$ denotes the essential infimum of $(\varphi^c + \mathcal{V}[\nu])$ *i.e.* the largest constant that bounds $(\varphi^c + \mathcal{V}[\nu])$ from below m_0-a.e.. Since γ is an optimal transport plan we have $c(x,y) = \varphi(x) + \varphi^c(y)$ γ-a.e. whereas by definition $c(x,z) \geq \varphi(x) + \varphi^c(z)$ for all $(x,z) \in X \times Y$. We thus have

$$\begin{cases} c(x,z) + \mathcal{V}[\nu](z) \geq M + \varphi(x) & \text{for all } x \in X \text{ and } m_0\text{-a.e. } z \in Y \\ c(x,y) + \mathcal{V}[\nu](y) = M + \varphi(x) & \text{for } \gamma\text{-a.e. } (x,y). \end{cases}$$

This proves that γ is a Cournot-Nash equilibrium.

B.2. *Proof of Corollary 3.3*

Thanks to Theorem 3.2, it is enough to prove that (3.5) admits solutions and to recall that the set $\Pi_o(\mu,\nu)$ is nonempty. Let $(\nu_n)_n$ be a minimising sequence of (3.5). Thanks to the growth condition (3.2), $(\nu_n)_n$ is bounded in $\mathcal{L}^p(m_0)$. It thus admits a (not relabelled) sub-sequence that converges weakly in $\mathcal{L}^p(m_0)$ (and thus in particular weakly $*$ in $\mathcal{P}(Y)$) to some $\nu \in \mathcal{L}^p(m_0)$. By the convexity of $F(y,.)$, the first term in $\mathcal{E}$ is lower-semi continuous for the weak topology of $\mathcal{L}^p(m_0)$. By the continuity of ϕ the second term in $\mathcal{E}$ is continuous for the weak-$*$ topology of $\mathcal{P}(Y)$. Finally, the lower-semi continuity of $\mathcal{W}_c(\mu,.)$ for the weak-$*$ topology straightforwardly follows from the Kantorovich duality formula (A.3). We thus have

$$\inf_\nu \mathcal{J}_\mu(\nu) = \liminf_n \{\mathcal{W}_c(\mu,\nu_n) + \mathcal{E}[\nu_n]\} \geq \mathcal{W}_c(\mu,\nu) + \mathcal{E}[\nu]\ .$$

So that ν solves (3.5).

B.3. *Proof of Proposition 3.5*

Assume that γ is an equilibrium and let ν be its second marginal. Let then φ be a Kantorovich potential between μ and ν such that

$$\begin{cases} \varphi^c + \mathcal{V}[\nu] \geq 0 & m_0\text{-a.e.} \\ \varphi^c + \mathcal{V}[\nu] = 0 & \nu\text{-a.e.} \end{cases} \tag{B.2}$$

Let $\rho \in \mathcal{D}$, thanks to the Kantorovich duality formula (A.3), we first have

$$\mathcal{W}_c(\mu,\rho) - \mathcal{W}_c(\mu,\nu) \geq \int_Y \varphi^c\, \mathrm{d}(\rho - \nu)\ .$$

By convexity of $\mathcal{E}$ and (3.4), we obtain

$$\mathcal{E}[\rho] - \mathcal{E}[\nu] \geq \int_Y \mathcal{V}[\nu]\, \mathrm{d}(\rho - \nu)$$

hence, finally using (B.2) and the fact that ρ is absolutely continuous with respect to m_0, we get

$$\mathcal{J}_\mu[\rho] - \mathcal{J}_\mu[\nu] \geq \int_Y (\varphi^c + \mathcal{V}[\nu])\, \mathrm{d}(\rho - \nu) \geq 0$$

which means that ν solves (3.5).

B.4. *Proof of Lemma 3.7*

The existence of a minimiser is similar to the proof of Corollary 3.3 since the coercivity of F ensures that minimising sequences are uniformly integrable and its convexity guarantees sequential weak lower semi continuity of $\mathcal{J}_\mu$.

Let ν solve (3.5) and let us prove that it is bounded away from zero. Let us assume by contraction that $m_0(\{\nu \leq \lambda\}) > 0$ for every $\lambda > 0$. Let $\delta_0 > 0$, $\delta \in (0, \delta_0)$ (to be chosen later on). Let $A := \{x : \delta_0 \leq \nu(x) \leq M\}$ where $M > 0$ is large enough so that $m_0(A) > 0$. For small $\varepsilon > 0$ such that $\varepsilon < \delta m_0(A)/2$ then define

$$\nu_\varepsilon := \nu + \varepsilon(u_{A_\delta} - u_A)$$

where $A_\delta := \{\nu \leq \delta\}$ and for $m_0(B) > 0$, u_B denotes the (sort of uniform probability on B) $u_B := m_0(B)^{-1}\mathbb{1}_B$. Since $\varepsilon < \delta m_0(A)/2$ and $\nu > \delta$ on A, ν_ε is a probability measure. By optimality of ν, we then have

$$0 \leq \mathcal{J}_\mu[\nu_\varepsilon] - \mathcal{J}_\mu[\nu] = \mathcal{W}_c(\mu, \nu_\varepsilon) - \mathcal{W}_c(\mu, \nu) + \mathcal{E}[\nu_\varepsilon] - \mathcal{E}[\nu]. \tag{B.3}$$

Denoting by φ_ε a Kantorovich potential between μ and ν_ε, we first have

$$\mathcal{W}_c(\mu, \nu_\varepsilon) - \mathcal{W}_c(\mu, \nu) \leq \varepsilon \int_Y \varphi_\varepsilon^c \, \mathrm{d}(u_{A_\delta} - u_A) \, .$$

And since φ_ε^c has a modulus of continuity that is uniform with respect to ε (that of c) and φ_ε^c can be normalised so as to vanish at the same point, φ_ε^c is uniformly bounded independently of ε and δ. So that $\mathcal{W}_c(\mu, \nu) - \mathcal{W}_c(\mu, \nu_\varepsilon) \leq C_1\varepsilon$ for some constant C_1. In a similar way, one finds a constant C_2 such that for ε small enough and uniformly in δ one has

$$\frac{1}{2} \iint_{Y^2} \phi(y, z) \, \mathrm{d}\nu_\varepsilon(y) \, \mathrm{d}(\nu_\varepsilon - \nu)(z) \leq C_2\varepsilon.$$

Now it remains to estimate the last term namely

$$\begin{aligned} \int_Y [F(\nu_\varepsilon) - F(\nu)] \, \mathrm{d}m_0 = &\int_{A_\delta} \left[F(\nu + \varepsilon m_0(A_\delta)^{-1}) - F(\nu)\right] \, \mathrm{d}m_0 \\ &+ \int_A \left[F(\nu - \varepsilon m_0(A)^{-1}) - F(\nu)\right] \, \mathrm{d}m_0 \end{aligned}$$

since F is Lipschitz on $[\delta_0/2, M]$ the second term can be bounded from above by $C_3\varepsilon$ for a constant C_3 again independent of δ and ε. Now let $C := C_1 + C_2 + C_3$ thanks to Inada's condition there is some $\delta_1 < \delta_0/2$ such that $f \leq -C - 1$ on $(0, 2\delta_1]$. Choosing $\delta \leq \delta_1$ and ε small enough so that $\varepsilon m_0(A_\delta)^{-1} \leq \delta_1$, we then have

$$\int_{A_\delta} \left[F(\nu + \varepsilon m_0(A_\delta)^{-1}) - F(\nu)\right] \, \mathrm{d}m_0 \leq (-C - 1)\varepsilon.$$

Putting everything together, the latter inequality gives the desired contradiction to (B.3).

The proof of the upper bound is similar: one assumes that $\nu \notin \mathcal{L}^\infty(m_0)$ and then considers a perturbation of the form $\nu_\varepsilon := \nu + \varepsilon(u_C - u_{C_M})$ with $C_M := \{\nu \geq M\}$, M large and C well chosen, the computations are the same as before and the contradiction comes from the Inada condition at $+\infty$: $F'(\nu) = f(\nu) \to +\infty$ as $\nu \to +\infty$.

B.5. *Proof of Theorem 4.3*

Uniqueness of a minimiser follows directly from the strict convexity along generalised geodesics with base μ of $\mathcal{J}_\mu$ which follows from McCann's condition (4.2), the convexity of v and ϕ and the strict convexity along generalised geodesics with base μ of $\mathcal{W}_2^2(\mu,.)$.

Let us now assume that γ is an equilibrium and that ν is its second marginal, then, for some constant M, we have:

$$f(\nu) + v(y) + \int_Y \phi(y,z)\,\mathrm{d}\nu(z) + \varphi^c \geq M \text{ a.e. with an equality } \nu\text{-a.e.} \tag{B.4}$$

Thanks to Inada's condition and the fact that the right hand side is continuous, this implies ν is bounded away from zero so that (B.4) actually is an equality (Lebesgue) almost everywhere $\mathcal{V}[\nu] + \varphi^c = M$ and thus ν satisfies

$$\nu(y) = f^{-1}\left(M - v(y) - \int_Y \phi(y,z)\,\mathrm{d}\nu(z) - \varphi^c(y)\right)$$

and is therefore continuous.

Let us now prove that ν solves (4.1), let ρ be another probability measure (which we can assume to have a positive and continuous density as well), and let $t \in [0,1] \to \nu_t$ denote the generalised geodesic with base μ joining ν and ρ, *i.e.* $\nu_t = T_{t\#}\mu := ((1-t)T_0 + tT_1)_\#\mu$ where T_0 (resp. T_1) denotes the Brenier map between μ and ν (resp. ρ). Since $(T_0, T_0 + t(T_1 - T_0))_\#\mu$ has marginals ν and ν_t we have

$$\mathcal{W}_1(\nu,\nu_t) \leq \int_Y |T_t - T_0|\,\mathrm{d}\mu \leq t\,\mathrm{diam}(Y)\ . \tag{B.5}$$

By the convexity of $\mathcal{J}_\mu$ along generalised geodesics with base μ, setting $g(t) = \mathcal{J}_\mu[\nu_t]$ and using $g(t) \leq (1-t)g(0) + tg(1)$ for all $t \in (0,1)$ we have:

$$\mathcal{J}_\mu[\rho] - \mathcal{J}_\mu[\nu] = g(1) - g(0) \geq \frac{1}{t}[g(t) - g(0)] = \frac{1}{t}(\mathcal{J}_\mu[\nu_t] - \mathcal{J}_\mu[\nu]).$$

Let us write ν_t as $\nu_t := \nu + th_t$, by the (usual) convexity of F and that of $\mathcal{W}_2^2(\mu,.)$, we first have

$$\frac{1}{t}\left(\frac{1}{2}\mathcal{W}_2^2(\mu,\nu_t) + \int_Y F(\nu_t)\,\mathrm{d}m_0 - \frac{1}{2}\mathcal{W}_2^2(\mu,\nu) - \int_Y F(\nu)\,\mathrm{d}m_0\right) \geq \int_Y [f(\nu) + \varphi^c]h_t\,\mathrm{d}m_0$$

Let us now expand $\iint_{Y^2} \phi(y,z)\,\mathrm{d}\nu_t(y)\,\mathrm{d}\nu_t(z)$ in powers of t as

$$\frac{1}{2t}\iint_{Y^2} \phi(y,z)\,\mathrm{d}\left[(\nu + th_t)(y)(\nu + th_t)(z) - \nu(y)\nu(z)\right] = \iint_{Y^2} \phi(y,z)\,\mathrm{d}h_t(y)\,\mathrm{d}\nu(z) + R_t$$

where

$$2|R_t| = \frac{1}{t}\left|\iint \phi(y,z)\,\mathrm{d}(\nu_t - \nu)(y)\,\mathrm{d}(\nu_t - \nu)(z)\right| \leq \frac{1}{t}\mathcal{W}_1(\nu_t,\nu)\times\mathrm{Lip}(\psi_t, Y) \leq \mathrm{diam}(Y)\,\mathrm{Lip}(\psi_t, Y)$$

where the last inequality follows from (B.5) and ψ_t is defined by

$$\psi_t(y) := \int_Y \phi(y,z)\,\mathrm{d}(\nu_t - \nu)(z) = \int_Y (\phi(y, T_t(x)) - \phi(y, T_0(x)))\,\mathrm{d}\mu(x)\ .$$

Since $\nabla\phi$ is locally Lipschitz and $T_t - T_0$ is uniformly bounded by $t\,\mathrm{diam}(Y)$ we find that there is a constant C such that $\mathrm{Lip}(\psi_t, Y) \leq Ct$ so that $R_t = O(t)$. Putting everything together and using (B.4) we get

$$\begin{aligned}\mathcal{J}_\mu[\rho] - \mathcal{J}_\mu[\nu] &\geq \int_Y (f(\nu(y)) + v(y) + \int_Y \phi(y,z)\,\mathrm{d}\nu(z) + \varphi^c(y))\,\mathrm{d}h_t(y) + R_t \\ &= \int_Y (\mathcal{V}[\nu] + \varphi^c)\,\mathrm{d}h_t + R_t \geq M\int_Y \mathrm{d}h_t + R_t = R_t \to 0 \text{ as } t \to 0^+\end{aligned}$$

which proves that ν is a minimiser.

REFERENCES

L. Ambrosio, N. Gigli, and G. Savaré. *Gradient flows in metric spaces and in the space of probability measures*. Lectures in Mathematics. Birkhäuser, 2005.

R. Aumann. Existence of competitive equilibria in markets with a continuum of traders. *Econometrica*, 32:39–50, 1964.

R. Aumann. Markets with a continuum of traders. *Econometrica*, 34:1–17, 1966.

A. Blanchet, P. Mossay, and F. Santambrogio. City equilibria via optimal transport and differential equations. Technical report, In preparation, 2012.

Y. Brenier. Polar factorization and monotone rearrangement of vector-valued functions. *Comm. Pure Appl. Math.*, 44(4):375–417, 1991.

L. A. Caffarelli. The regularity of mappings with a convex potential. *J. Amer. Math. Soc.*, 5: 99–104, 1992a.

L. A. Caffarelli. Boundary regularity of maps with convex potentials. *Comm. Pure Appl. Math.*, 45:1141–1151, 1992b.

G. Carlier. Duality and existence for a class of mass transportation problems and economic applications. *Adv. in Math. Econ.*, 5:1–21, 2003.

G. Carlier and I. Ekeland. Equilibrium structure of a bidimensional asymmetric city. *Nonlin. Anal. B*, 8(3):725–748, 2007.

P.-A. Chiappori, R. J. McCann, and L. Nesheim. Hedonic price equilibria, stable matching, and optimal transport: equivalence, topology, and uniqueness. *Econom. Theory*, 42:317–354, 2010.

I. Ekeland. Existence, uniqueness and efficiency of equilibrium in hedonic markets with multidimensional types. *Econom. Theory*, 42(2):275–315, 2010.

A. Figalli, R. J. McCann, and Y. H. Kim. When is multidimensional screening a convex program? *Econom. Theory*, 146:454–478, 2011.

R. Jordan, D. Kinderlehrer, and F. Otto. The variational formulation of the Fokker-Plank equation. *SIAM J. of Math. Anal.*, 29:1–17, 1998.

M. A. Kahn. On Cournot-Nash equilibrium distributions for games with a nonmetrizable action space and upper semi-continuous payoffs. *Trans. Amer. Math. Soc.*, 1:127–146, 1989.

E. Kohlberg, S. Hart, and W. Hildenbrand. On equilibrium allocations as distributions on the commodity space. *J. Math. Econ.*, 1:159–166, 1974.

H. Konishi, M. Le Breton, and S. Weber. Pure strategy nash equilibrium in a group formation game with positive externalities. *Games Econ. Behav.*, 21:161–182, 1997.

J.-M. Lasry and P.-L. Lions. Jeux à champ moyen. i. le cas stationnaire. *C. R. Math. Acad. Sci. Paris*, 343(9):619–625, 2006a.

J.-M. Lasry and P.-L. Lions. Jeux à champ moyen. ii. horizon fini et contrôle optimal. *C. R. Math. Acad. Sci. Paris*, 343(10):679–684, 2006b.

J.-M. Lasry and P.-L. Lions. Mean field games. *Jpn. J. Math.*, 2(1):229–260, 2007.

M. LeBreton and S. Weber. Games of social interactions with local and global externalities. *Econ. Letters*, 111:88–90, 2011.

A. Mas-Colell. On a theorem of Schmeidler. *J. Math. Econ.*, 3:201–206, 1984.

R. J. McCann. Existence and uniqueness of monotone measure-preserving maps. *Duke Math. J.*, 80:309–323, 1995.

R. J. McCann. A convexity principle for interacting cases. *Adv. Math.*, 128(1):153–179, 1997.

R. J. McCann and W. Gangbo. The geometry of optimal transportation. *Acta Math.*, 177:113–161, 1996.

D. Monderer and L.S. Shapley. Potential games. *Games and Economic Behavior*, 14:124–143, 1996.

S.T. Rachev and L. Rüschendorf. *Mass Transportation problems. Vol. I: Theory; Vol. II : Applications*. Springer-Verlag,, 1998.

T. Roughgarden. *Selfish routing and the price of anarchy*. MIT Press, 2005.

D. Schmeidler. Equilibrium points of nonatomic games. *J. Stat. Phys.*, 7:295–300, 1973.

C. Villani. *Topics in optimal transportation*, volume 58 of *Graduate Studies in Mathematics*. American Mathematical Society, Providence, RI, 2003.

C. Villani. *Optimal transport: old and new*. Grundlehren der mathematischen Wissenschaften. Springer-Verlag, 2009.