1 The discrete case.

1.1 Optimal transportation

An N-chain is a sequence $\{\theta_0, ..., \theta_N\} \subset \Omega$. It is closed if $\theta_0 = \theta_N$; a closed N-chain is also called an N-cycle. A chain is an N-chain, for some finite N. We shall denote by Θ the set of all chains, and by Θ_0 the set of all closed chains.

Proposition 1 *A bijection $x : \Omega \to X$ is optimal iff, for every closed chain $\{\theta_0, ..., \theta_N\} \in \Theta_0$, we have:*

$$\sum_{n=0}^{N-1} [u(\theta_n, x(\theta_n)) - u(\theta_n, x(\theta_{n+1}))] \geq 0 \tag{A}$$

Proof (sufficiency). Let x be a bijection satisfying condition (A), and let y be another bijection from Ω to X. Pick any point $\theta_0 \in \Omega$, and define the θ_n, $n \geq 1$, recursively by $y(\theta_n) = x(\theta_{n+1})$. Since Ω is a finite set, the θ_n cannot all be different. Let m be the lowest integer such that $\theta_m = \theta_N$ for some $N > m$. If $m \neq 0$,we have:

$$y(\theta_{m-1}) = x(\theta_m) = x(\theta_N) = y(\theta_{N-1})$$

From the definition of m, it follows that $\theta_{m-1} \neq \theta_{N-1}$, and the above equality contradicts the fact that y is a bijection. So we must have $m = 0$, and $\Gamma_1 = \{\theta_0, ..., \theta_N\}$ is a cycle. Because of condition (A), we have:

$$\sum_{\theta \in \Gamma_1} [u(\theta, x(\theta)) - u(\theta, y(\theta))] = \sum_{n=0}^{N-1} [u(\theta_n, x(\theta_n)) - u(\theta_n, x(\theta_{n+1}))] \geq 0$$

If Γ_1 is the whole of Ω, this concludes the proof. If not, pick another point $\theta'_0 \notin \Gamma_1$, and go through the same procedure. We get another cycle Γ_2, which has no common point with Γ_1. Proceeding in this way, we can partition Ω in a disjoint union of cycles Γ_k, $1 \leq k \leq K$, and:

$$\sum_{\theta \in \Omega} [u(\theta, x(\theta)) - u(\theta, y(\theta))] = \sum_{k=1}^{K} \sum_{\theta \in \Gamma_k} [u(\theta, x(\theta)) - u(\theta, y(\theta))]$$

and the right-hand side is non-negative by condition (A). This means that x is optimal. ■

Proof (necessity). Assume a bijection x is optimal. Pick any cycle $\Gamma = \{\theta_0, ..., \theta_N\}$ in Ω. Define a new map $y : \Omega \to X$ as follows:

$$\begin{aligned} y(\theta_n) &= x(\theta_{n+1}) \text{ for } 0 \leq n \leq N-1 \\ y(\theta) &= x(\theta) \text{ if } \theta \notin \Gamma \end{aligned}$$

Then y is still a bijection, and since x is optimal we must have:

$$\sum_{\theta\in\Gamma} u(\theta, x(\theta)) \geq \sum_{\theta\in\Gamma} u(\theta, y(\theta))$$

Replacing $y(\theta)$ by its values, we get exactly condition (A) $\blacksquare$

Note for future reference that condition (A) can be rewritten as follows:

$$\sum_{n=1}^{N} \left[u(\theta_n, x(\theta_n)) - u(\theta_{n-1}, x(\theta_n))\right] \geq 0$$

1.2 More about condition (A)

Proposition 2 *Suppose a map $x : \Omega \to \Omega$ (not necessarily a permutation) satisfies condition (A). Pick a point $\bar{\theta} \in \Omega$. Then the formula:*

$$V(\theta) := \inf \left\{ \sum_{n=1}^{N} \left[u(\theta_n, x(\theta_n)) - u(\theta_{n-1}, x(\theta_n))\right] \mid \{\theta_0, ..., \theta_N\} \in \Theta,\ \theta_0 = \bar{\theta},\ \theta_N = \theta \right\} \tag{1}$$

defines a real-valued function V on Ω with the following properties:

1. $V(\bar{\theta}) = 0$
2. $V(\theta') \geq V(\theta) + u(\theta', x(\theta)) - u(\theta, x(\theta))\ \forall \theta, \theta'$

Any such function V is called a potential associated with the map x.

Proof. Since any chain starting and ending at $\bar{\theta}$ must be closed, we have:

$$V(\bar{\theta}) \geq \inf \left\{ \sum_{n=1}^{N} \left[u(\theta_n, x(\theta_n)) - u(\theta_{n-1}, x(\theta_n))\right] \mid \{\theta_0, ..., \theta_N\} \in \Theta_0 \right\}$$

The right-hand side is non-negative by condition (A), so $V(\bar{\theta}) \geq 0$. On the other hand, taking the trivial $1-$path $\theta_0 = \bar{\theta} = \theta_1$ yields $V(\bar{\theta}) \leq 0$, so we have $V(\bar{\theta}) = 0$.

Formula (1) defines V as function with values in $\mathbb{R} \cup \{-\infty\}$. Pick two points θ and θ'. We shall consider special chains connecting $\bar{\theta}$ to θ, those whose last leg is $\{\theta', \theta\}$. From the definition of V, we have:

$$\begin{aligned}
V(\theta) &\leq \inf \left\{ \begin{array}{c} \sum_{n=1}^{N} \left[u(\theta_n, x(\theta_n)) - u(\theta_{n-1}, x(\theta_n))\right] \\ \mid \{\theta_0, ..., \theta_N\} \in \Theta,\ \theta_0 = \bar{\theta},\ \theta_{N-1} = \theta',\ \theta_N = \theta \end{array} \right\} \\
&= \inf \left\{ \begin{array}{c} \sum_{n=1}^{N-1} \left[u(\theta_n, x(\theta_n)) - u(\theta_{n-1}, x(\theta_n))\right] + u(\theta, x(\theta)) - u(\theta', x(\theta)) \\ \mid \{\theta_0, ..., \theta_{N-1}\} \in \Theta,\ \theta_0 = \bar{\theta},\ \theta_{N-1} = \theta' \end{array} \right\} \\
&= \inf \left\{ \sum_{n=1}^{N-1} \left[u(\theta_n, x(\theta_n)) - u(\theta_{n-1}, x(\theta_n))\right] \mid \{\theta_0, ..., \theta_{N-1}\} \in \Theta,\ \theta_0 = \bar{\theta},\ \theta_{N-1} = \theta' \right\} \\
&\quad + u(\theta, x(\theta)) - u(\theta', x(\theta)) \\
&= V(\theta') + u(\theta, x(\theta)) - u(\theta', x(\theta))\ \forall \theta, \theta'
\end{aligned}$$

Applying this inequality to $\theta = \bar{\theta}$, we get:

$$\begin{aligned} V(\theta') &\geq u\left(\theta', x\left(\bar{\theta}\right)\right) - u\left(\bar{\theta}, x\left(\bar{\theta}\right)\right), \ \ \forall \theta' \\ &\geq \inf\left\{u\left(\theta', x\left(\bar{\theta}\right)\right) - u\left(\bar{\theta}, x\left(\bar{\theta}\right)\right) \mid \theta \in \Omega\right\} \end{aligned}$$

and since the set Ω is finite, the right-hand side is bounded. So the function V is real-valued, and satisfies the desired inequality ■

Note for future use that equality is achieved if $\theta = \theta'$, so that inequality (2) can be rewritten:

$$\begin{aligned} V(\theta') &= \sup_{\theta}\left\{u(\theta', x(\theta)) - u(\theta, x(\theta) + V(\theta))\right\} \qquad (2) \\ &= \sup_{\theta}\left\{u(\theta', x(\theta)) + t(\theta)\right\} \end{aligned}$$

1.3 Incentive-compatible allocations

Ω is a set of types, and X is a set of tasks. An individual of type θ perfoming task x and getting paid a sum t will have a total utility of

$$u(\theta, x) + t$$

The principal wishes individuals of type θ to perform a certain task $x(\theta) \in \Omega$, and he will reward them with a payment $t(\theta)$. The map

$$(x, t) : \Omega \to X \times \mathbb{R}$$

is called a *contract*. Such a contract will be called *incentive-compatible* (henceforth IC) if there is no incentive for individuals to lie about their type, that is, if:

$$u(\theta, x(\theta)) + t(\theta) \geq u(\theta, x(\theta')) + t(\theta') \ \ \forall (\theta, \theta')$$

An allocation $x : \Omega \to X$ will be called *incentive-compatible* if there is a map $t : \Omega \to \mathbb{R}$ such that the contract (x, t) is IC.

Proposition 3 *A map $x : \Omega \to \Omega$ is an incentive-compatible allocation iff it satisfies condition (A)*

Proof (necessity). Assume x is IC, and let t be the corresponding payment. Pick any cycle $(\theta_0, ..., \theta_N)$, so that $.\theta_0 = \theta_N$. We have:

$$\begin{aligned} u(\theta_0, x(\theta_0)) + t(\theta_0) &\geq u(\theta_0, x(\theta_1)) + t(\theta_1) \\ u(\theta_1, x(\theta_1)) + t(\theta_1) &\geq u(\theta_1, x(\theta_2)) + t(\theta_2) \\ &\ldots \\ u(\theta_{N-1}, x(\theta_{N-1})) + t(\theta_{N-1}) &\geq u(\theta_{N-1}, x(\theta_N)) + t(\theta_N) \end{aligned}$$

Summing up, and remembering that $t(\theta_0) = t(\theta_N)$, we find that all the terms in t cancel out, and we are left with condition (A) $\blacksquare$

Proof (sufficiency). Assume x satisfies condition (A). Let V be the potential associated with x. Set:

$$t(\theta) = V(\theta) - u(\theta, x(\theta))$$

Writing this into inequality (2) in Proposition 2 we get exactly condition (A) $\blacksquare$

2 A formal approach to the continuous case

The continuous version of the optimal transportation problem reads as follows. Given two subsets Ω and X of $\mathbb{R}^n$, endowed with the Borelian tribe and positive measures μ and ν, find, among all measure-preserving maps $x : \Omega \to X$, the one(s) which minimize the integral:

$$\int_\Omega u(\theta, x(\theta))\, d\theta$$

Going to the limit in condition (A), we obtain that an optimal map x must be such that:

$$\oint \sum u_\theta(\theta, x(\theta))\, d\theta_i \geq 0$$

along every closed loop. Running the loop in the other direction, we see that the inequality is in fact an equality. By the Poincaré lemma, this means that there is a function $V : \Omega \to \mathbb{R}$ such that:

$$\frac{\partial u}{\partial \theta_i}(\theta, x(\theta)) = \frac{\partial V}{\partial \theta_i}$$

So we find the potential again. Note that if the map $x \to u_\theta(\theta, x)$ can be inverted, then the map x can be deduced from the function V. Looking for V instead of x is a tremendous simplification, and this is what we are going to do.

3 u-convex analysis

We are given sets Ω and X (no longer finite) and a map $u : \Omega \times X \to \mathbb{R}$. Points in Ω will be denoted by θ, and points in X by x

3.1 u-convex functions

We will be dealing with function taking values in $\mathbb{R}\cup\{+\infty\}$. Such a function will be called *proper* if it is not identically $\{+\infty\}$.

A function $f : \Omega \to \mathbb{R}\cup\{+\infty\}$ will be called *u-convex* iff there exists a non-empty subset $A \subset X \times \mathbb{R}$ such that:

$$f(\theta) = \sup_{(x,t)\in A} \{u(\theta.x) + t\} \tag{3}$$

A function $g : x \to \mathbb{R}\cup\{+\infty\}$ will be called *u-convex* iff there exists a non-empty subset $B \subset \Omega \times \mathbb{R}$ such that:

$$g(x) = \sup_{(\theta,t)\in B} \{u(\theta.x) + s\}$$

3.2 u-conjugates

Let $f : \Omega \to \mathbb{R}\cup\{+\infty\}$ be a proper function (not necessarily u-convex). We define its u-conjugate f^* by:

$$f^*(x) = \sup_{\theta} \{u(\theta,x) - f(\theta)\} \tag{4}$$

Let $g : X \to \mathbb{R}\cup\{+\infty\}$ be a proper function (not necessarily u-convex). We define its u-conjugate g^* by:

$$g^*(\theta) = \sup_{x} \{u(\theta.x) - g(x)\}$$

If f is proper, then f^* takes values in $\mathbb{R}\cup\{+\infty\}$ and is a *u-convex* function on X.

If g is proper, then g^* takes values in $\mathbb{R}\cup\{+\infty\}$ and is a *u-convex* function on Ω

Example 4 *Set $\varphi(\theta) = u(\theta,\bar{x}) + \bar{t}$. Then*

$$\varphi^*(\bar{x}) = \sup_{\theta} \{u(\theta,\bar{x}) - u(\theta,\bar{x}) - \bar{t}\} = -\bar{t}$$

Example 5 *Let $\theta \to x(\theta)$ be a map satisfying condition (A). The associated potential V is u-convex: this follows from condition (2)*

Conjugation reverses ordering: if $f_1 \leq f_2$, then $f_1^* \geq f_2^*$. As a consequence, if f is u-convex, then f^* is proper. Indeed, since f is u-convex, we have $f \geq \varphi$ for some function φ of this type, and then $f^*(\bar{x}) \leq -\bar{t} < \infty$.

Proposition 6 (the Fenchel inequality) *For any proper functions $f : \Omega \to \mathbb{R}\cup\{+\infty\}$ and $g : X \to \mathbb{R}\cup\{+\infty\}$, we have:*

$$\begin{aligned} f(\theta) + f^*(x) &\geq u(\theta,x) \quad \forall(\theta,x) \\ g(x) + g^*(\theta) &\geq u(\theta,x) \quad \forall(\theta,x) \end{aligned}$$

3.3 u-subgradients

Let $f : \Omega \to \mathbb{R} \cup \{+\infty\}$ be a proper u-convex function. Take some point $\theta \in \Omega$. We shall say that a point $x \in X$ is a *u-subgradient* of f at θ if:

$$f(\theta) + f^*(x) = u(\theta, x) \tag{5}$$

The set of subgradients of f at θ will be called the subdifferential of f at θ and denoted by $\partial_u f(\theta)$.

Similarly, let $g : X \to \mathbb{R} \cup \{+\infty\}$ be a proper convex function. Take some point $x \in X$. We shall say that a point $\theta \in \Omega$ is a *u-subgradient* of g at θ if:

$$g^*(\theta) + g(x) = u(\theta, x) \quad \forall (\theta, x)$$

Proposition 7 *$x \in \partial_u f(\theta)$ iff :$f(\theta') \geq f(\theta) + u(\theta', x) - u(\theta, x) \quad \forall (\theta, x)$*

Proof (necessity). Assume $x \in \partial_u f(\theta)$. Then, by (5), we have:

$$f(\theta') \geq u(\theta', x) - f^*(x) = u(\theta', x) - [u(\theta, x) - f(\theta)]$$

■

Proof (sufficiency). We have:

$$\begin{aligned} f^*(x) &= \sup_{\theta'} \{u(\theta', x) - f(\theta')\} \\ &\leq \sup_{\theta'} \{u(\theta', x) - f(\theta) - u(\theta', x) + u(\theta, x)\} \\ &= u(\theta, x) - f(\theta) \end{aligned}$$

so $f(\theta) + f^*(x) \leq u(\theta, x)$. We have the converse by the Fenchel inequality, so equality holds ■

Example 8 *Consider a map $\theta \to x(\theta)$. Then it satisfies condition (A) iff there is a u-convex function V such that $x(\theta)$ is a subgradient of V at θ.*

3.4 u-biconjugates

It follows from the Fenchel inequality that:

$$f^{**}(\theta) = \sup_x \{u(\theta, x) - f^*(x)\} \leq f(\theta)$$

Example 9 *Set $\varphi(\theta) = u(\theta, \bar{x}) + \bar{t}$. Then*

$$\varphi^{**}(\theta) = \sup_x \{u(\theta, x) - \varphi^*(x)\} \geq u(\theta, \bar{x}) + \bar{t} = \varphi(\theta)$$

and hence $\varphi^{**}(\theta) = \varphi(\theta)$

Proposition 10 *For every proper function $f : \Omega \to \mathbb{R}$, we have*

$$f^{**}(\theta) = \sup_{\varphi} \{\varphi(\theta) \mid \varphi \le f,\ \varphi \in \mathbb{C}_u(\Omega)\}$$

Proof. Denote by $\bar{f}$ the right-hand side of the above formula. We want to show that $f^{**} = \bar{f}$.

Since $f^{**} \le f$ and f^{**} is u-convex, we must have $f^{**} \le \bar{f}$.

To show that $f^{**} \ge \bar{f}$, since $\bar{f}$ is u-convex, it is enough to show that every u-affine function which is less than $\bar{f}$ is also less that f^{**}. Let $(\bar{x}, \bar{t})$ be such that $u(\theta, \bar{x}) + \bar{t} \le \bar{f}(\theta)$. Since $\bar{f} \le f$, we also have $u(\theta, \bar{x}) + \bar{t} \le f(\theta)$. Taking biconjugates, we get $u(\theta, \bar{x}) + \bar{t} \le f^{**}(\theta)$ as well. ■

Corollary 11 *If f is convex, then $f^{**} = f$. and we have:*

$$f(\theta) = \sup_{x} \{u(\theta, x) - f^*(x)\} \tag{6}$$

Corollary 12 *The following are equivalent:*

1. $x \in \partial_u f(\theta)$
2. $\theta \in \partial_u f^*(x)$
3. $f(\theta) + f^*(x) = u(\theta, x)$

3.5 Smoothness

From now on, we shall assume the following:

Condition 13 (c1) *Ω is an open bounded subset of $\mathbb{R}^n$ and X is a compact subset of $\mathbb{R}^m$. The functions u and u_θ are well-defined and continuous, and that they extend continously to $\bar{\Omega} \times X$.*

So the family $\{u(\bullet, x) \mid x \in X\}$ is uniformly Lipschitz, and therefore all u-convex functions on Ω are Lipschitz on Ω. By a theorem of Rademacher, they are differentiable almost everywhere with respect to the Lebesgue measure..

The family $\{u(\theta, \bullet) \mid \theta \in \bar{\Omega}\}$ is equicontinuous on $\dot{X}$. So all u-convex functions on X are continuous on X. Because of the compactness of X, the supremum is attained in condition (6), so that they all u-convex functions on Ω are subdifferentiable everywhere on Ω.

If a function f is u-differentiable at a point $\bar{\theta}$, and $\bar{x} \in \partial_u f(\bar{\theta})$,we have:

$$f_\theta(\bar{\theta}) = \frac{\partial u}{\partial \theta}(\bar{\theta}, \bar{x})$$

Proposition 14 *For every u-convex function f, there is a map $\theta \to x(\theta)$ such that*

$$f_\theta(\theta) = \frac{\partial u}{\partial \theta}(\theta, x(\theta)) \text{ Lebesgue-a.e.}$$

3.6 Generalized Spence-Mirrlees

Definition 15 *We shall say that u satisfies the generalized Spence-Mirrlees condition (henceforth GSM) if, for Lebesgue-a.e. $\theta \in \Omega$, the map $u_\theta(\theta, \bullet)$ is one-to-one:*

$$u_\theta(\theta, x_1) = u_\theta(\theta, x_2) \Longrightarrow x_1 = x_2$$

Example 16 *If $\Omega = X = \mathbb{R}^n$, then $u(\theta, x) = \|\theta - x\|^\alpha$ satisfies (GSM) for $\alpha \neq 0$ and $\alpha \neq 1$.*

Proposition 17 *Assume condition (**??**) is satisfied, as well as GSM. Then, for every u-convex function f, the map $\theta \to x(\theta)$ defined in Proposition 14 is unique up to equality almost everywhere:*

$$f_\theta(\theta) = \frac{\partial u}{\partial \theta}(\theta, x_1(\theta)) = \frac{\partial u}{\partial \theta}(\theta, x_2(\theta)) \Rightarrow x_1(\theta) = x_2(\theta) \ \textit{Lebesgue-a.e.}$$

We shall refer to it as the u-subgradient map of f.

4 Optimal transportation

4.1 A bizarre optimization problem

We are given two borelian subsets $\Omega \subset \mathbb{R}^{d_1}$ and $X \subset \mathbb{R}^{d_2}$, endowed with positive and finite measures μ and ν. We posit the following optimization problem:

$$\inf \left[\int_\Omega f(\theta)\, d\mu + \int_X g(x)\, d\nu\right] \tag{7}$$

$$f(\theta) + g(x) \geq u(\theta, x) \quad \forall (\theta, x) \tag{8}$$

Changing f to $f + a$ and g to $g - a$, for any constant a, does not affect the constraint, but changes the criterion by $[\mu(\Omega) - \nu(X)]\, a$. It follows that, for the problem to be meaningful, we must require that:

$$\mu(\Omega) = \nu(X)$$

and that, even then, optimal solutions, if they exist, are not unique. If (f, g) is optimal, so is $(f + a, g - a)$.

Proposition 18 *Assume Ω, X and u satisfy condition (**??**) and that $\mu(\Omega) = \nu(X)$. Then the problem (7),(8) has an optimal solution (f, g). If (f, g) is an optimal solution, then there is a u-convex function $\bar{f}$ such that $f = \bar{f}$ μ-a.e. and $g = \bar{f}^*$ ν-a.e..*

Proof. Let (f_n, g_n) be a minimizing sequence:

$$\int_\Omega f(\theta)\, d\mu + \int_X g(x)\, d\nu \to \inf \tag{9}$$

Then $f_n^* \leq g_n$ and (f, f_n^*) still satisfies the constraints, so (f, f_n^*) still is a minimizing sequence. Then $f_n^{**} \leq f_n$ and (f_n^{**}, f_n) still satisfies the constraints, so (f_n^{**}, f_n^*) is a minimizing sequence. The sequences f_n^{**} and f_n^* both are equicontinuous, and they will have uniformly convergent subsequences provided they are bounded (Ascoli's theorem).

Set $a_n = \min_\Omega f_n^{**}.$, and consider the functions $h_n(\theta) = f_n^{**}(\theta) - a_n$, so that the h_n are u-convex and $h_n \geq 0$. Since Ω is compact, there is a point θ_n where $h_n(\theta_n) = 0$. We have:

$$\begin{aligned} h_n^*(x) &= \sup_\theta \{u(\theta, x) - h_n(\theta)\} \geq u(\theta_n.x) \geq \inf_{\Omega\times X}\{u(\theta, x)\} \\ h_n(x) &= \sup_x \{u(\theta, x) - h_n^*(x)\} \leq \sup_{\Omega\times X}\{u(\theta, x)\} - \inf_{\Omega\times X}\{u(\theta, x)\} \end{aligned}$$

So the sequence h_n is uniformly bounded. A similar argument shows that the sequence h_n^* is uniformly bounded. By Ascoli's theorem, we can extract subsequences which converge uniformly to u-convex functions $\bar{f}$ and $\bar{f}^*$. Taking limits in (9), we get the result. ∎

Theorem 19 *Assume moreover that μ is absolutely continuous with respect to the Lebesgue measure and that u satisfies GSM. Then the subgradient map x of f is a measure-preserving map from Ω to X :*

$$x(\mu) = \nu$$

and if (f_1, g_1) and (f_2, g_2) are two optimal solutions, then $x_1(\theta) = x_2(\theta)$ μ-a.e.

Proof. Let $(f = g^*, g)$ be an optimal solution of problem (7),(8). Take any continuous function φ on X. For any $h > 0$, we have:

$$\int_\Omega (g + h\varphi)^*\, d\mu + \int_X (g + h\varphi)\, d\nu \geq \int_\Omega g^* d\mu + \int_X g d\nu$$

which yields immediately:

$$\frac{1}{h}\int_\Omega \left[(g + h\varphi)^* - g^*\right] d\mu + \int_X \varphi d\nu \geq 0 \tag{10}$$

Now, let $x(\theta)$ and $x_h(\theta)$ be measurable selections of $\partial_u g$ and $\partial_u(g + h\varphi)$, that is, maps $\theta \to x(\theta)$ and $\theta \to x_h(\theta)$ such that

$$\begin{aligned} u_\theta(\theta, x(\theta)) &= \frac{\partial g^*}{\partial\theta}(\theta) \quad \mu\text{-a.e.} \\ u_\theta(\theta, x_h(\theta)) &= \frac{\partial}{\partial\theta}(g + h\varphi)^*(\theta) \quad \mu\text{-a.e.} \end{aligned}$$

From the definition of g^*, we have:

$$u(\theta, x_h(\theta)) - g(x_h(\theta)) \leq g^*(\theta) = u(\theta, x(\theta)) - g(x(\theta))$$

so that:

$$u(\theta, x_h(\theta)) - g(x_h(\theta)) - u(\theta, x(\theta)) + g(x(\theta)) \leq 0 \tag{11}$$

From the definition of $(g + h\varphi)^*$, we have:

$$u(\theta, x(\theta)) - g(x(\theta)) - h\varphi(x(\theta)) \leq (g + h\varphi)^*(\theta) = u(\theta, x_h(\theta)) - g(x_h(\theta)) - h\varphi(x_h(\theta))$$

from which we deduce, taking into account the fact that $x(\theta) \in \partial_u g^*(\theta)$:

$$\begin{aligned} h\varphi(x_h(\theta)) - h\varphi(x(\theta)) &\leq u(\theta, x_h(\theta)) - g(x_h(\theta)) - u(\theta, x(\theta)) + g(x(\theta)) \qquad (12) \\ &= (g + h\varphi)^*(\theta) + h\varphi(x_h(\theta)) - g^*(\theta) \qquad (13) \end{aligned}$$

Finally, comparing (11), (12), and (13), we have::

$$-h\varphi(x(\theta)) \leq (g + h\varphi)^*(\theta) - g^*(\theta) \leq -h\varphi(x_h(\theta)) \tag{14}$$

Let us now use the assumptions. By Rademacher's theorem and GSM, the maps $x(\theta)$ and $x_h(\theta)$ are uniquely defined, up to a.e. equivalence. Letting $h \to 0$, and fixing θ, we find that, because of the compactness of X, the sequence $x_h(\theta)$ must have cluster points. Any cluster point $x(\theta)$ of $x_h(\theta)$ must satisfy:

$$u_\theta(\theta, x(\theta)) = g^*_\theta(\theta)$$

which defines $x(\theta)$ uniquely, by GSM. So, $x_h(\theta)$ converges to $x(\theta)$ almost everywhere[1].

Dividing by h, taking the limit as $h \to 0$ and using the continuity of φ yields:

$$\lim_{h \to 0} \frac{1}{h} \left[(g + h\varphi)^*(\theta) - g^*(\theta) \right] = \varphi(x(\theta))$$

Taking limits in equation (10) yields:

$$\int_\Omega \varphi(x(\theta)) \, d\mu + \int_X \varphi d\nu \geq 0$$

and since this inequality holds both for φ and -φ, it must be an equality. So x sends μ on ν, as announced.

As for uniqueness, assume that (f_1, g_1) and (f_2, g_2) both are optimal solutions to problem (7),(8), with x_1 and x_2 being the corresponding subgradient maps. Then:

$$f_1(\theta) + g_1(x_2(\theta)) \geq u(\theta, x_2(\theta)) \quad \mu\text{-a.e.} \tag{15}$$

Integrating, and remembering that x_2 sends μ to ν, we get:

[1] There is a nicety here. We are applying Rademacher's theorem to each $(g + h\varphi)^*$, leaving aside a negligible set N_h which depends on h. We should take $h = 1/k$, with $k \to \infty$ an integer, so that the union $\cup_h N_h =: N$ is still negligible

$$\begin{aligned}
\int_\Omega f_1 d\mu + \int_X (g_1 \circ x_2)\, d\mu &= \int_\Omega f_1 d\mu + \int_X g_1 d\nu \\
&\geq \int_\Omega u(\theta, x_2(\theta))\, d\mu(\theta)
\end{aligned}$$

Since (f_2, g_2) is another optimal solution, we also have:

$$\begin{aligned}
\int_\Omega f_1 d\mu + \int_X g_1 d\nu &= \int_\Omega f_2 d\mu + \int_X g_2 d\nu \\
&= \int_\Omega u(\theta, x_2(\theta))\, d\mu(\theta)
\end{aligned}$$

Comparing this with (15) we find that $f_1(\theta) + g_1(x_2(\theta)) = u(\theta, x_2(\theta))$ μ-a.e. So $x_1(\theta) = x_2(\theta)$ μ-a.e. ■

Corollary 20 *If Ω is connected, if (f_1, g_1) and (f_2, g_2) are two solutions of problem(7),(8) then there is a constant a such that $f_2 = f_1 + a$ and $g_2 = g_1 - a$.*

Note that, if GSM does not hold, then Theorem 19 does not apply. The proof carries over up to equation (14). But now there is no reason why the functions $x_h(\theta)$ should converge a.e., and no good way to pick $x(\theta)$. In fact, the original Monge problem, with $u(\theta, x) = \|\theta - x\|$, does have an optimal solution, although u does not satisfy GSM but there is great difficulty in proving it. Only recently did we get a satisfactory proof.

5 The optimal transportation problem

5.1 Main results

Theorem 21 *Assume that Ω, X and u satisfy condition (**??**), that Ω is endowed with a positive measure μ, absolutely continuous with respect to the Lebesgue measure, that X is endowed with a positive measure ν such that:*

$$\mu(\Omega) = \nu(X) < \infty$$

and that u satisfies GSM. Then the problem:

$$\min \int_\Omega u(\theta, x(\theta))\, d\mu(\theta) \tag{16}$$

$$x : \Omega \to X,\ x(\mu) = \nu \tag{17}$$

has a solution, and any two solutions are equal.μ-a.e..

Proof. We just consider the problem (7),(8) and apply theorem 19. Let (f, g) be an optimal solution, and x be the subgradient map of f. Then x preserves measure, and if $\bar{x} : \Omega \to X$ is another measure-preserving map, we have, by the Fenchel inequality::

$$\begin{aligned}
\int_\Omega u(\theta, \bar{x}(\theta))\, d\mu(\theta) &\leq \int_\Omega f(\theta)\, d\mu(\theta) + \int_X g(\bar{x}(\theta))\, d\mu(\theta) \\
&= \int_\Omega f(\theta)\, d\mu(\theta) + \int_\Omega g(x)\, d\nu(x) \\
&= \int_\Omega u(\theta, x(\theta))\, d\mu(\theta)
\end{aligned}$$

the latter equality expressing the fact that $x(\theta) \in \partial_u f(\theta)$ ■

We shall now give a theorem where θ and x play fully symmetric roles: GSM will be satisfied, not only with respect to θ, but also with respect to x. In that case the mapping x can be inverted (up to a.e. equivalence)

Theorem 22 *Let Ω be an open bounded subset of $\mathbb{R}^{d_1}$ and X_0 be an open bounded subset of $\mathbb{R}^{d_2}$. Let μ and ν be positive measures on Ω and X respectively, both absolutely continuous with respect to the Lebesgue measure, and such that:*

$$\mu(\Omega) = \nu(X_0) < \infty$$

Let $u : \Omega \times X_0 \to \mathbb{R}$ be continously differentiable, and assume that all derivatives extend continuously to the closure $\bar{\Omega} \times \overline{X_0}$. Assume finally that u satisfies GSM with respect to the variables θ and x:

$$\begin{aligned}
u_\theta(\theta, x_1) &= u_\theta(\theta, x_2) \Longrightarrow x_1 = x_2 \\
u_x(\theta_1, x) &= u_x(\theta_2, x) \Longrightarrow \theta_1 = \theta_2
\end{aligned}$$

Then the problem

$$\min \int_\Omega u(\theta, s(\theta))\, d\mu(\theta) \tag{18}$$

$$s : \Omega \to X,\ s(\mu) = \nu \tag{19}$$

has a solution s, and the problem

$$\min \int_X u(t(x), x)\, d\nu(x) \tag{20}$$

$$t : X \to \Omega,\ t(\nu) = \mu \tag{21}$$

has a solution t. These solutions are given by the formulas:

$$\begin{aligned}
f(s(\theta)) &= \max_x \{u(\theta, x) - g(x)\} \quad a.e. \\
g(t(x)) &= \max_\theta \{u(\theta, x) - f(\theta)\} \quad a.e.
\end{aligned}$$

where $f = g^*$ and $g = f^*$ are a suitable pair of u-convex functions. If s_1 and s_2 are two solutions of (18),(19), then $s_1 = s_2$ a.e..If $t_1 = t_2$ are two solutions of (20),(21), then $t_1 = t_2$ a.e..If s is a solution of (18),(19), and t is a solution of (20),(21), then

$$\begin{aligned} (s \circ t)(x) &= x \ \ a.e. \\ (t \circ s)(\theta) &= \theta \ \ a.e. \end{aligned}$$

Proof. The existence and uniqueness of s and t follow from the preceding theorem. For the same reason, we have$\{s(\theta)\} = \partial_u f(\theta)$ a.e. and $\{t(x)\} = \partial_u g(x)$ a.e. (the brackets mean that the sets are singletons), with $g = f^*$. By corollary 12, it follows that s and t inverse of each other. $\blacksquare$

The assumptions will be satisfied if, for instance, Ω and X_0 are open bounded subsets of $\mathbb{R}^d$, endowed with positive measures μ and ν, absolutely continuous with respect to the Lebesgue measure, and if $u(\theta, x) = c(\theta - x)$,with:

- either $u(\theta, x) = c(\theta - x)$. where $c : \mathbb{R}^d \to \mathbb{R}$ is $C^1$and $c' : \mathbb{R}^d \to \mathbb{R}$ is one-to-one
- or $\bar{\Omega}$ and $\overline{X_0}$ are disjoint, and $u(\theta, x) = c(\|\theta - x\|)$, where $c' : (0, \infty) \to \mathbb{R}$ is one-to-one

For instance, if $u(\theta, x) := \|\theta - x\|^\alpha$, the case $\alpha > 1$ falls into the first category, and the case $\alpha < 1$ falls into the second category (and so requires $\bar{\Omega}$ and $\overline{X_0}$ to be disjoint for the optimal transportation problem to have a solution)

5.2 Brenier's theorems

All these results originate with Yann Brenier, who investigated the optimal transportation problem for the special case when $u(\theta, x) = \|\theta - x\|^2$, yielding the problem:

$$\min_s \int_\Omega \frac{1}{2} \|\theta - s(\theta)\|^2 d\mu(\theta) \tag{22}$$

$$s : \Omega \to X, \ s(\mu) = \nu \tag{23}$$

Note that this is equivalent to the seemingly different problem:

$$\max_s \int_\Omega \sum_{i=1}^n \theta^i s_i(\theta) d\mu(\theta) \tag{24}$$

$$s : \Omega \to X, \ s(\mu) = \nu \tag{25}$$

corresponding to $u(\theta, x) = \theta' x$. Indeed, we have:

$$\begin{aligned} \int_\Omega \frac{1}{2} \|\theta - s(\theta)\|^2 d\mu(\theta) &= \int_\Omega \frac{1}{2} \|\theta\|^2 d\mu(\theta) + \int_\Omega \frac{1}{2} \|s(\theta)\|^2 d\mu(\theta) - \int_\Omega \frac{1}{2} \theta' s(\theta) d\mu(\theta) \\ &= \int_\Omega \frac{1}{2} \|\theta\|^2 d\mu(\theta) + \int_\Omega \frac{1}{2} \|x\|^2 d\nu(x) - \int_\Omega \frac{1}{2} \theta' s(\theta) d\mu(\theta) \end{aligned}$$

because s preserves measure. The first two terms on the right-hand side are constants, and we are left with the third one to optimize.

As above, we assume that Ω and X are subsets of $\mathbb{R}^d$, that condition (**??**) is satisfied, and that μ is a positive measure on Ω,absolutely continuous with respect to the Lebesgue measure, and finite: $\mu(\Omega) < \infty$.We then have a series of results, which are all due to Brenier.

Proposition 23 *Let ν be a positive measure on X, such that $\mu(\Omega) = \nu(X)$. Then there is a standard convex function f on $\mathbb{R}^d$, and a μ-negligible subset $N \subset \Omega$, such that f is differentiable on $\Omega \backslash N$, and the gradient f' maps $\Omega \backslash N$ into X and μ on ν :*

$$f'(\mu) = \nu$$

If g is another standard convex function with the same property, then $f' = g'$ a.e.

This is a remarkable theorem, because the set Ω and X are not required to be convex.In fact, they can have any shape at all. Note that the definition of a standard convex function as a supremum of affine functions over Ω implies that it is defined, not only over Ω, but over $\mathbb{R}^d$.

Proposition 24 *Given any Borelian map $x : \Omega \to X$, there is a standard convex function f on $\mathbb{R}^d$, and a μ-negligible subset $N \subset \Omega$, such that f is differentiable on $\Omega \backslash N$, and the gradient f' has the same distribution as x:*

$$f'(\mu) = x(\mu)$$

If g is another standard convex function with the same property, then $f' = g'$ a.e.

This follows from the preceding proposition by taking $\nu := x(\mu)$ .The map $f' : \Omega \backslash N \to X$ is called the *increasing rearrangement* of x (see the one-dimensional case). It satisfies the following inequalities:

$$\begin{aligned} \int_\Omega \|\theta - f'(\theta)\|^2 \, d\mu &\leq \int_\Omega \|\theta - x(\theta)\|^2 \, d\mu \\ \int_\Omega \sum \theta_i \frac{\partial f}{\partial \theta_i} d\mu &\geq \int_\Omega \sum \theta_i x^i(\theta) \, d\mu \end{aligned}$$

Proposition 25 *Let $X_0 \subset R^d$ be an open bounded set and $x : \Omega \to X_0$ be a Borelian map such that $x(\mu)$ is absolutely continuous with respect to the Lebesgue measure. Then we have $x = f' \circ \varphi$ a.e., where. f is a standard convex function on Ω and $\varphi : \Omega \to \Omega$ satisfies $\varphi(\mu) = \mu$*

Proof. There is a standard convex function f such that $f'(\mu) = x(\mu)$. There is an inverse $g' : X \to \Omega$, which also preserves measure. Setting $g' \circ x := \varphi$ and writing $x = f' \circ g' \circ x$ gives the desired result. ■

We begin by recalling the standard polar decomposition theorem for matrices.

Proposition 26 *Any invertible real matrix M can be written as $M = UA$, where U is orthogonal and A is symmetric and positive definite.*

Proof. Take M an $n \times n$ matrix with $\det M \neq 0$. Then $A := (M^* M)^{1/2}$ is a symmetric matrix, positive definite. Set $M = UA$. Then $U = MA^{-1}$ and $U^* = A^{-1} M^*$, so that $UU^* = MA^{-2} M^* = I$. ■

Theorem 27 *Let Ω be an open bounded subset of $\mathbb{R}^d$, and μ a positive measure on Ω, absolutely continuous with respect to the Lebesgue measure. Then any borelian map φ from Ω into another open bounded subset X of $\mathbb{R}^d$ can be written as $\varphi = \alpha \psi$, where $\alpha(\theta) = f'(\theta)$ a.e,. f is a standard convex function on Ω and $\psi : \Omega \to \Omega$ preserves μ*

6 The principal-agent problem

6.1 Main results

Let Ω be the space of parameters characterizing agents, and X be the space of actions which the principal wishes the agents to undertake. The distribution μ of parameters is known to the principal. An agent with parameter θ performing action x and paying p to do so has utility $u(\theta, x) - p$. If this is less than a certain quantity $\bar{u}(\theta)$ (his reservation utility), he will not undertake the action. A contract is a map $x \to p(x)$, meaning that the agent will pay $p(x)$ to the principal for performing action x. This contract will be incentive-compatible if:

$$\begin{aligned} u(\theta, x(\theta)) - p(x(\theta)) &\geq u(\theta, x(\theta')) - p(x(\theta')) \quad \forall (\theta, \theta') \\ u(\theta, x(\theta)) - p(x(\theta)) &\geq \bar{u}(\theta) \quad \forall \theta \end{aligned}$$

There is no cost of production. The utility which the principal derives from a contract $\theta \to (x(\theta), t(\theta))$ is

$$\int_A [p(x(\theta)) - c(x(\theta))] \, d\mu(\theta)$$

where $A \subset \Omega$ is the set of agents which actually buy. The principal-agent problem consists in maximizing this integral over all incentive-compatible contracts.

Introduce the function

$$V(\theta) := \max_x \{u(\theta, x) - p(x)\}$$

which is the indirect utility which consumer θ derives from the contract p. If GSM holds, and μ is absolutely continous with respect to the Lebesgue measure, the subgradient map $\nabla_u V$, given by $\{\nabla_u V(\theta)\} = \partial_u V(\theta)$, is well-defined a.e., and the problem can be rewritten as follows:

$$\sup \int_{V(\theta) \geq \bar{u}(\theta)} [u(\theta, \nabla_u V(\theta)) - V(\theta)] \, d\mu$$

There is an existence theory for such problems, which was developed by Guillaume Carlier. We will not give it here, and we will focus instead on the standard convex case. Note, however, the following very general result, which is an economic version of the rearrangement theorem:

Proposition 28 *Let $x : \Omega \to X$ be an allocation such that $u(\theta, x(\theta)) - p(x(\theta)) \geq \bar{u}(\theta) \;\; \forall \theta$. Then there is an incentive-compatible allocation y with the same distribution*

6.2 The Rochet-Choné problem

Set $\Omega = [a_1,\, b_1] \times [a_2,\, b_2]$, with μ the Lebesgue measure, $X = R_+^2$ and $c = 0$.

$$u(\theta, x) = -\frac{1}{2}\left(x_1^2 + x_2^2\right) + \theta_1 x_1 + \theta_2 x_2$$

Note that, since x_1 and x_2 are positive, utility is increasing in the parameters θ. The higher θ_1 and θ_2, the more interested the agent is in the action θ. Let all agents have the same reservation utility u_0. The principal-agent problem then becomes:

$$\sup \int_\Omega \left[-\frac{c}{2}\left(\frac{\partial V}{\partial x_1}^2 + \frac{\partial V}{\partial x_2}^2\right) + \theta_1 \frac{\partial V}{\partial x_1} + \theta_2 \frac{\partial V}{\partial x_2} - V(\theta)\right] d\theta_1 d\theta_2$$
$$V \in H^1(\Omega), \text{ convex, } V(x) \geq u_0 \text{ a.e.}$$

Here is the state of the art on that problem:

- The problem has a unique solution.
- Necessary conditions for optimality are difficult to find..
- Regularity: it has been proved by Carlier and Lachand-Robert that the solution V is C^1
- Numerical computations are difficult as well. Note that a straight finite element methods will fail.
- Qualitative properties are obvious from the economic point of view: Ω is divided in three regions, from the bottom left corner to the top right corner, $\Omega_0, \Omega_1, \Omega_2$, and the Hessian of V has i positive eigenvalues in Ω_i. One would expect such a feature to be robust, but the mathematics are not there.

Proposition 29 *The Rochet-Chone problem has a unique solution.*

Proof. Let $C \subset H^1$ be the set of convex functions, $C_0 \subset C$ the set of functions $u \geq u_0$ a.e., and

$$F(u) = \int_\Omega \left[-\frac{1}{2} \left(\frac{\partial V}{\partial x_1}^2 + \frac{\partial V}{\partial x_2}^2 \right) + \theta_1 \frac{\partial V}{\partial x_1} + \theta_2 \frac{\partial V}{\partial x_2} - V(\theta) \right] d\theta_1 d\theta_2$$

C_0 is a closed convex set and F is a concave lsc function on H^1. By the Hahn-Banach theorem, F is weakly upper semi-continuous and C_0 is weakly closed.

Let u_n be a maximizing sequence in C_0:

$$F(u_n) \to \sup_{C_0} F$$

Since the u_n are convex, they are lower semi-continous on $\bar{\Omega}$. Set $a_n = \min u_n$. If $a_n > u_0$, the function $v_n = u_n - (a_n - u_0)$ still belongs to C_0, and $F(u_n) < F(v_n)$, so the sequence v_n is maximizing. So, without loss of generality, we may assume that $\min u_n = u_0$.

Set $u_n(\theta_n) = u_0$. We have:

$$u_n(\theta) = u_0 + \int_0^1 (\theta - \theta_n)' \nabla u_n (\theta_n + t(\theta - \theta_n)) \, dt$$

and hence

$$\begin{aligned}
|u_n(\theta) - u_0|^2 &\leq ||\theta - \theta_n||^2 \int_0^1 ||\nabla u_n (\theta_n + t(\theta - \theta_n))||^2 \, dt \\
||u_n - u_0||_{L^2}^2 &\leq (\text{diameter } \Omega)^2 \int_\Omega \int_0^1 ||\nabla u_n (\theta_n + t(\theta - \theta_n))||^2 \, dt d\theta \\
&= (\text{diameter } \Omega)^2 \int_0^1 \int_\Omega ||\nabla u_n (\theta_n + t(\theta - \theta_n))||^2 \, d\theta dt \\
&\leq (\text{diameter } \Omega)^2 \int_0^1 \int_\Omega ||\nabla u_n(\theta)||^2 \, d\theta dt \\
&= (\text{diameter } \Omega)^2 ||\nabla u_n||_{L^2}^2
\end{aligned}$$

(the Poincaré inequality). Therefore,

$$F(u_n) \geq -c ||\nabla u_n||_{L^2}^2 - \left(\int_\Omega ||\theta||^2 \, d\theta \right)^{1/2} ||\nabla u_n||_{L^2} - \left(\int_\Omega ||\theta||^2 \, d\theta \right)^{1/2} (|u_0| + (\text{diameter } \Omega) ||\nabla u_n||_{L^2})$$

It follows that $||\nabla u_n||_{L^2}$ is bounded, and so ∇u_n converges weakly to some ξ in L^2. From standard arguments if follows that $\xi = \nabla u$, with $u \in C_0$ and $\min u = u_0$. Since F is weakly upper semi-continuous, we have:

$$F(u) \geq \limsup F(u_n) = \sup_{C_0} F$$

Si u is a maximizer. Since F is strictly concave, it is the only one. ■

We will first give necessary conditions, not for the Rochet-Chone problem, but for related ones.

6.3 The projection problem

Let Ω be a bounded open convex subset of R^d. Introduce the set

$$K = \left\{ \nabla u \mid u \in C,\ \nabla u \in L^2 \right\} \subset L^2$$

and let us consider the problem:

$$\min_{y \in K} ||x - y||_{L^2}^2 = \min_{u \in C} \int_\Omega ||x - \nabla u||^2 \, d\theta \tag{26}$$

the solution of which, $y = p(x)$, is the projection of v on K. It is characterized by the set of inequalities:

$$y \in K \text{ and } (x - y, z - y)_{L^2} \leq 0 \ \ \forall z \in K \tag{27}$$

This can be rewritten as $(x - y) \in K^-$, where $K^- = \{\xi \mid (\xi, z) \leq 0, \forall z \in K\}$ is the *polar cone* of K. Note that inequality (27) can be rewritten as two separate conditions::

$$(x - y) \in K^- \quad \text{and} \quad (x - y, y)_{L^2} = 0$$

The problem now is to characterize the functions $\xi \in K^-$. Define a *bistochastic measure* on $\Omega \times \Omega$ to be a Radon measure μ on $\Omega \times \Omega$ such that:

$$\begin{aligned} \int\int_{\Omega \times \Omega} f(\theta_1)\, d\mu(\theta_1, \theta_2) &= \int_\Omega f(\theta)\, d\theta \quad \forall f \in C^0\left(\bar{\Omega}\right) \\ \int\int_{\Omega \times \Omega} g(\theta_2)\, d\mu(\theta_1, \theta_2) &= \int_\Omega g(\theta)\, d\theta \quad \forall f \in C^0\left(\bar{\Omega}\right) \end{aligned}$$

In other words, the horizontal and vertical projections of μ coincide with the Lebesgue measure on Ω. We can then write:

$$\begin{aligned} \int\int_{\Omega \times \Omega} f(\theta_1, \theta_2)\, d\mu(\theta_1, \theta_2) &= \int_\Omega \left[\int_\Omega f(\theta_1, \theta_2)\, d\mu^v[\theta_2 \mid \theta_1] \right] d\theta_1 \\ &= \int_\Omega \left[\int_\Omega f(\theta_1, \theta_2)\, d\mu^h[\theta_1 \mid \theta_2] \right] d\theta_2 \end{aligned}$$

where $\mu^v[\cdot \mid \theta_1]$ and $\mu^h[\cdot \mid \theta_2]$ are probabilities on Ω depending on θ_1 and θ_2 respectively.

Denote by M the set of bistochastic measures. Note that it is a convex set (but not a cone), and it is closed in the weak topology.

Example 30 *Let $s : \Omega \to \Omega$ be a map which preserves the Lebesgue measure. Then the measure μ defined by:*

$$\int\int_{\Omega \times \Omega} f(\theta_1, \theta_2)\, d\mu(\theta_1, \theta_2) = \int_\Omega f(\theta_1, s(\theta_1))\, d\theta_1 \quad \forall f \in C^0\left(\bar{\Omega} \times \bar{\Omega}\right) \tag{28}$$

is bistochastic. We have

$$\begin{aligned} \mu^v[\cdot \mid \theta_1] &= \delta_{s(\theta_1)} \\ \mu^h[\cdot \mid \theta_2] &= \delta_{t(\theta_2)} \end{aligned}$$

where δ is the Dirac mass and $t : \Omega \to \Omega$ is the inverse of s

Proposition 31 *The following statements are equivalent:*

1. $\xi \in K^-$
2. $\xi \in L^2(\Omega)$ *and there are a bistochastic measure* μ *and a number* $a > 0$ *such that*

$$\xi(\theta_1) = a\left(\int_\Omega \theta_2 d\mu^v [\theta_2 \mid \theta_1] - \theta_1\right) \quad \text{for a.e. } \theta_1 \tag{29}$$

Proof. Let $\widetilde{K}$ be the set of functions ξ defined by (29). Note that $\widetilde{K}$ is a closed convex cone. We shall prove that $(\widetilde{K})^- = K$. By the bipolar theorem, $\widetilde{K} = \left((\widetilde{K})^-\right)^- = K^-$, and the proposition will follow.

We have to prove the following. Given $z \in \left[L^2(\Omega)\right]^d$, we have $z \in K$ if and only if

$$\int_\Omega z(\theta)' \xi(\theta)\, d\theta = 0 \tag{30}$$

for all $\xi \in.\widetilde{K}$ Note that it is enough to prove this for $a = 1$, which we shall do henceforth.

Let us assume first that $z \in K$, that is, $z = \nabla u$ for some convex u. Then the integral becomes:

$$\begin{aligned}\int_\Omega z(\theta)' \xi(\theta)\, d\theta &= \int_\Omega \nabla u(\theta)' \xi(\theta)\, d\theta \\ &\leq \int_\Omega u(\theta + \xi(\theta))\, d\theta - \int_\Omega u(\theta)\, d\theta\end{aligned}$$

since u is convex. Using formula (29), and the Jensen inequality, this becomes:

$$\begin{aligned}\int_\Omega z(\theta)' \xi(\theta)\, d\theta &\leq \int_\Omega u\left(\int_\Omega \theta_2 d\mu^v [\theta_2 \mid \theta_1]\right) d\theta_1 - \int_\Omega u(\theta)\, d\theta \\ &\leq \int_\Omega u(\theta_1)\, d\theta_1 - \int_\Omega u(\theta)\, d\theta = 0\end{aligned}$$

Conversely, let us assume that (30) holds for all $\xi \in.\widetilde{K}$. Then it certainly holds for all ξ of type (28), so that, for all maps $\Omega \to \Omega$ which preserve Lebesgue measure, we have:

$$\int_\Omega z(\theta)' (s(\theta) - \theta)\, d\theta \leq 0$$

Pick a cycle $\theta_0, ..., \theta_N = \theta_0$ in Ω, and around each θ_i a ball $B_i \subset \Omega$ of radius ε. We then define a measure-preserving map by s by switching the B_i around: $s(B_i) = B_{i+1}$. Writing the preceding inequality for s, and taking the limit as $\varepsilon \to 0$. we get the inequality:

$$\sum_{i=0}^{N-1} z(\theta_i)' (\theta_{i+1} - \theta_i) \leq 0$$

which is a necessary and sufficient condition for z to be the gradient of a convex function. ■

Lemma 32 *The polar cone K^- of K is the closed convex cone generated by S in L^2, where:*

$$S = \left\{ s - I \mid s : \Omega \to \Omega,\ s \in \left(L^2\right)^n,\ s\,(d\theta) = d\theta \right\}$$

We can then apply the projection theorem in Hilbert space, and we get:

Proposition 33 *For any $x \in \left[L^2(\Omega)\right]^d$, the projection problem (26) has a unique sol≤ution $p(x) = \nabla u$. Let u be a convex function on Ω. Then ∇u is the optimal solution if and only if there exists a bistochastic measure μ on $\Omega \times \Omega$ and a number $a > 0$ such that:*

$$\begin{aligned} x &= \nabla u + \xi, && (31)\\ 0 &= \int_\Omega \xi(\theta)' \nabla u(\theta)\, d\theta && (32)\\ \xi(\theta_1) &= a\left(\int_\Omega \theta_2 d\mu^v[\theta_2 \mid \theta_1] - \theta_1\right) \quad \textit{for a.e. } \theta_1 \in \Omega && (33) \end{aligned}$$

Corollary 34 *Condition () is equivalent to the following:*

$$u\left(\int_\Omega \theta_2 d\mu^v[\theta_2 \mid \theta_1]\right) - u(\theta_1) - \nabla u(\theta_1)'\left(\int_\Omega \theta_2 d\mu^v[\theta_2 \mid \theta_1] - \theta_1\right) = 0 \ \textit{for a.e. } \theta_1.$$
$$\forall \theta_2 \in \mathit{Supp}\,(\mu^v[\cdot \mid \theta_1]),\ \ u(\theta_2) - u(\theta_1) = \nabla u(\theta_1)'(\theta_2 - \theta_1) \quad \textit{for a.e. } \theta_1 \in \Omega,$$
$$\int_\Omega u\left(\int_\Omega \theta_2 d\mu^v[\theta_2 \mid \theta_1]\right) d\theta_1 = \int_\Omega u(\theta)\, d\theta$$

and in addition we have:

$$.\quad \forall \theta_2 \in \ \mathit{Supp}\,(\mu^v[\cdot \mid \theta_1]),\ \ u(\theta_2) - u(\theta_1) = \nabla u(\theta_1)'(\theta_2 - \theta_1) \quad \textit{for a.e. } \theta_1 \in \Omega,$$
$$\int_\Omega u\left(\int_\Omega \theta_2 d\mu^v[\theta_2 \mid \theta_1]\right) d\theta_1 = \int_\Omega u(\theta)\, d\theta$$

Proof. Because of the Jensen inequality, we have:

$$\int_\Omega u(\theta_2)\, d\mu^v[\theta_2 \mid \theta_1] \geq u\left(\int_\Omega \theta_2 d\mu^v[\theta_2 \mid \theta_1]\right) \ \ \forall \theta_1$$

Integrating with respect to $d\theta_1$, we get:

$$\int_\Omega u(\theta_1)\, d\theta_1 = \int_\Omega u(\theta_2)\, d\mu^v[\theta_2 \mid \theta_1]\, d\theta_1 \geq \int_\Omega u\left(\int_\Omega \theta_2 d\mu^v[\theta_2 \mid \theta_1]\right) d\theta_1$$

Writing in relation (34), we get:

$$\int_\Omega \left[u\left(\int_\Omega \theta_2 d\mu^v[\theta_2 \mid \theta_1]\right) - u(\theta_1) - \nabla u(\theta_1)'\left(\int_\Omega \theta_2 d\mu^v[\theta_2 \mid \theta_1] - \theta_1\right)\right] d\theta_1 \leq 0$$

But the integrand is always ≥ 0 by Fenchel's inequality, so it must vanish:

$$u\left(\int_\Omega \theta_2 d\mu^v\left[\theta_2 \mid \theta_1\right]\right) - u\left(\theta_1\right) - \nabla u\left(\theta_1\right)'\left(\int_\Omega \theta_2 d\mu^v\left[\theta_2 \mid \theta_1\right] - \theta_1\right) = 0 \text{ for a.e. } \theta_1.$$

and the two other results follow. ■

Given a convex function u on Ω, and a point θ_1 in Ω, define the facet of u containing θ_1 to be the set:

$$\text{Facet}\left\{u \mid \theta_1\right\} = \left\{\theta_2 \mid u\left(\theta_2\right) - u\left(\theta_1\right) = \nabla u\left(\theta_1\right)'\left(\theta_2 - \theta_1\right)\right\}$$

If u is strictly convex, $\text{Facet}\left\{u \mid \theta_1\right\} = \left\{\theta_1\right\}$ for every θ_1. In the general case, $\text{Facet}\left\{u \mid \theta_1\right\}$ is a closed convex subset of Ω. Corollary **??** states that the probability $\mu^v\left[\cdot \mid \theta_1\right]$ is carried by $\text{Facet}\left\{u \mid \theta_1\right\}$.

So, if there is a convex subset $C \subset \Omega$ such that the restriction of u to C is strictly convex, then $\text{Facet}\left\{u \mid \theta\right\} = \left\{\theta\right\}$ for every $\theta \in C$. so that $\mu^v\left[\cdot \mid \theta\right]$ must be the Dirac mass δ_θ, and $\xi\left(\theta\right) = 0$ on C.

6.4 Calculus of variations with convexity constraints

Consider the problem:

$$\begin{aligned}
&\inf_u \int_\Omega L\left(\theta, u\left(\theta\right), \nabla u\left(\theta\right)\right) d\theta \\
&u : \Omega \to \Omega,\ u \text{ convex}
\end{aligned}$$

Assume that $L\left(\theta, u, y\right)$ is C^1 and there is a C^1 solution $\bar{u}$, such that $\bar{u}$ and $\nabla\bar{u}$ extend continuously to $\bar{\Omega}$.

Comparing $\bar{u}$ with $u = \bar{u} + h$, where h is a constant, we derive:

$$\int_\Omega \frac{\partial L}{\partial u}\left(\theta, \bar{u}\left(\theta\right), \nabla\bar{u}\left(\theta\right)\right) d\theta = 0 \tag{34}$$

Comparing $\bar{u}$ with $u = \bar{u} + h\left(v - \bar{u}\right)$, where v is another convex function, and letting $h \to 0$ we get:

$$\int_\Omega \left[\frac{\partial L}{\partial u}\left(v - \bar{u}\right) + \sum_i \frac{\partial L}{\partial y_i}\left(\frac{\partial v}{\partial \theta_i} - \frac{\partial \bar{u}}{\partial \theta_i}\right)\right] d\theta \geq 0 \tag{35}$$

Introduce the solution ϕ of the Neumann problem:

$$\begin{aligned}
\triangle\phi &= \frac{\partial L}{\partial u}\left(\theta, \bar{u}\left(\theta\right), \nabla\bar{u}\left(\theta\right)\right) \text{ in } \Omega \\
\frac{\partial \phi}{\partial n} &= 0 \text{ on } \partial\Omega
\end{aligned}$$

This solution exists and is unique, because of condition (34). Substituting into the variational inequality (35) and integrating by parts yields:

$$\int_\Omega \sum_i \left(\frac{\partial L}{\partial y_i} - \frac{\partial \phi}{\partial \theta_i}\right)\left(\frac{\partial v}{\partial \theta_i} - \frac{\partial \bar{u}}{\partial \theta_i}\right) d\theta \geq 0 \text{ for all convex } v$$

Applying the results of the preceding subsection, we see that this amounts to:

$$\begin{aligned}\left[\frac{\partial L}{\partial y_i}\left(\theta,\bar{u}\left(\theta\right),\nabla\bar{u}\left(\theta\right)\right)-\frac{\partial\phi}{\partial\theta_i}\right]_{1\leq i\leq d} &\in -K^{-}\\ \int_{\Omega}\sum_i\left(\frac{\partial L}{\partial y_i}\left(\theta,\bar{u}\left(\theta\right),\nabla\bar{u}\left(\theta\right)\right)-\frac{\partial\phi}{\partial\theta_i}\right)\frac{\partial\bar{u}}{\partial\theta_i}d\theta &= 0\end{aligned}$$

Rewrite this as follows:

$$\begin{aligned}\frac{\partial L}{\partial y}\left(\theta,\bar{u}\left(\theta\right),\nabla\bar{u}\left(\theta\right)\right) &= \nabla\phi\left(\theta\right)-\xi\left(\theta\right)\\ \int_{\Omega}\xi\left(\theta\right)'\nabla\bar{u}\left(\theta\right)d\theta &= 0\end{aligned}$$

where $\xi\left(\theta_1\right)=a\left(\int_{\Omega}\theta_2 d\mu^v\left[\theta_2\mid\theta_1\right]-\theta_1\right)$ for some bistochastic measure μ. This is the integrated version of the Euler-Lagrange equation in the presence of global convexity constraints. Differentiating the first equation, we get the familiar form:

$$\sum_i\frac{\partial}{\partial\theta_i}\left[\frac{\partial L}{\partial y_i}\left(\theta,\bar{u}\left(\theta\right),\nabla\bar{u}\left(\theta\right)\right)+\xi\left(\theta\right)\right]-\frac{\partial L}{\partial u}\left(\theta,\bar{u}\left(\theta\right),\nabla\bar{u}\left(\theta\right)\right)=0$$

6.5 The one-dimensional case

One would like to have a direct characterization of the maps $\xi\in K^{-}$.

This is easy in the case when $\Omega=[a,b]$. Then $f\in K$ iff $f:[a,b]\to R$ is non-decreasing, and $\xi:[a,b]\to R$ belongs to K iff

$$\int_a^b f\left(\theta\right)\xi\left(\theta\right)d\theta\leq 0 \qquad (36)$$

for all f in K. Changing f to $f+c$, where c is any constant, we find that:

$$\int_a^b\xi\left(\theta\right)d\theta=0$$

Set $X\left(\theta\right)=\int_a^{\theta}\xi\left(t\right)dt$. If f is increasing, we have $f\left(\theta\right)=f\left(a\right)+\int_a^{\theta}df$, where df is a positive measure. Integrating by parts inequality (36) then yields:

$$\int_a^b X\left(\theta\right)df\geq 0$$

and since df can be any finite positive measure on $[a,b]$, we find that X is continuous, with $X\left(\theta\right)\geq 0\ \ \forall\theta$ and $X\left(a\right)=X\left(b\right)=0$

So, in the one-dimensional case, we have:

$$K^{-}=\left\{\frac{dX}{d\theta}\mid X\in C^0\left([a,b]\right),\ X\left(a\right)=X\left(b\right)=0,\ X\left(\theta\right)\geq 0\ \ \forall\theta\right\}$$

As an immediate consequence, we have:

Proposition 35 *If $\Omega = [a, b]$, the solution of the projection problem with $x = \frac{df}{dt}$ is $y = \frac{dg}{dt}$ with $g = f^{**}$, the convex hull of f.*

Proof. We just write $x = y + \frac{d}{d\theta}(f - g)$, and this is the required decomposition. ∎

This is no longer true in higher dimension.