Weak optimal transport and applications to Caffarelli contraction theorem

Nathaël Gozlan*

*Université Paris Descartes

Journée Analyse et Probabilité
Institut de Mathématiques, Université de Bordeaux

November 28, 2019

Outline

Works in collaboration with N. Juillet (2018), M. Fathi et M. Prodhomme (2019) and with C. Roberto, P.-M. Samson, Y. Shu and P. Tetali (2017)

Introduction : Brenier and Strassen Theorems

I - Generalized Transport and a mixture of Brenier and Strassen Theorems

II - Link with Caffarelli contraction theorem

Introduction : Brenier and Strassen Theorems

Let $\omega : E \times E \to \mathbf{R}^+$ be a measurable function on a Polish space (E, d).

Definition

The optimal transport cost between two probability measures μ and ν is given by

$$\mathcal{T}_\omega(\nu, \mu) = \inf_{\pi \in \Pi(\mu,\nu)} \iint_{E \times E} \omega(x, y) d\pi(x, y),$$

where $\Pi(\mu, \nu)$ denotes the set of probability measures π on $E \times E$ having μ and ν as marginals (called 'transport plans between μ and ν').

Equivalently

$$\mathcal{T}_\omega(\nu, \mu) = \inf_{X \sim \mu, Y \sim \nu} \mathbb{E}[\omega(X, Y)]$$

Classical Examples : Kantorovich distances of order $p \geq 1$

$$W_p^p(\nu, \mu) = \inf_{X \sim \mu, Y \sim \nu} \mathbb{E}[d^p(X, Y)].$$

Optimal transport plans

A transport plan π° is said optimal if

$$\mathcal{T}_\omega(\nu,\mu) = \iint \omega(x,y)\, d\pi^\circ(x,y).$$

Theorem

If ω is lower-semicontinuous then there always exists at least one optimal transport plans.

Questions :

- How to characterize optimal transport plans ?
- Are they given by a transport map T : $\pi^\circ = \mathrm{Law}(X, T(X))$, $X \sim \mu$?
- Is T regular ? Main motivation of this talk : global Lispchitz continuity.
- ...

Brenier Theorem

Let $|\cdot|$ denote the standard Euclidean norm on $E = \mathbf{R}^n$.
The following result characterizes optimal transport plans for the cost function $\omega(x, y) = |y - x|^2$, $x, y \in \mathbf{R}^n$.

Theorem (Brenier (1991))

If μ is absolutely continuous with respect to Lebesgue and if $\int |x|^2 d\mu(x) < +\infty$ and $\int |y|^2 d\nu(y) < +\infty$, then there exists a unique optimal transport plan π°, such that

$$W_2^2(\nu, \mu) = \iint |y - x|^2 \, d\pi^\circ(x, y).$$

Moreover π° is deterministic : there is some map $T : \mathbf{R}^n \to \mathbf{R}^n$ such that $\pi^\circ = \mathrm{Law}(X, T(X))$ and so

$$W_2^2(\nu, \mu) = \int |T(x) - x|^2 \, d\mu(x).$$

Moreover there exists a *convex* function $\phi : \mathbf{R}^n \to \mathbf{R} \cup \{+\infty\}$ such that

$$T(x) = \nabla\phi(x), \qquad \text{for Lebesgue almost every } x \in \mathbf{R}^n.$$

Remarks on Brenier Theorem

- Many generalizations : on manifolds, for other cost functions, for more than two marginals . . .
- A necessary and sufficient condition for μ on $\mathbf{R}^n$ to be transported on *any* measure ν with finite second moment by the gradient of a convex map has been obtained by Gigli (2011).

Strassen Theorem

Notation : $\mathcal{P}_1(\mathbf{R}^n)$ the set of probability measures with a finite first moment.

Definition

Let $\mu, \nu \in \mathcal{P}_1(\mathbf{R}^n)$; μ is dominated by ν in the convex order, denoted by $\mu \leq_c \nu$, if

$$\int f\,d\mu \leq \int f\,d\nu, \qquad \text{for all convex function } f : \mathbf{R}^n \to \mathbf{R}.$$

Theorem (Strassen (1965))

Let $\mu, \nu \in \mathcal{P}_1(\mathbf{R}^n)$; the following propositions are equivalent

(1) $\mu \leq_c \nu$,

(2) there exists a martingale (X_0, X_1) such that $X_0 \sim \mu$ and $X_1 \sim \nu$.

Remarks on Strassen Theorem

- Kellerer Theorem (1972) : generalization to a continuous family of marginals.
  $\rightsquigarrow$ so called PCOC (Hirsch, Profeta, Roynette and Yor (2011))
- Optimal Transport with martingale constraints (Beiglboeck-Juillet, Henry-Labordère-Touzi, de March, Tan, Ghoussoub-Kim-Lim ...)

Remark about the assumptions of Brenier and Strassen Theorems

- If $\int x \, d\mu(x) \neq \int y \, d\nu(y)$, then there is no martingale (X_0, X_1) such that $X_0 \sim \mu$ and $X_1 \sim \nu \ldots$
- If μ has an atom and ν is diffuse then there is no map transporting μ on $\nu \ldots$

Brenier then Strassen

Elementary remark : it is always possible to compose a deterministic transport and a martingale transport to couple two arbitrary probability measures μ and ν.

Indeed if (X, Y) is an arbitrary coupling then letting $\bar{X} = \mathbb{E}[Y|X]$, the coupling $(X, \bar{X})$ is deterministic and $(\bar{X}, Y)$ is a martingale.

Definition

A coupling (X, Y) between $\mu, \nu \in \mathcal{P}_1(\mathbf{R}^n)$ is of the Brenier-Strassen type if

$$\mathbb{E}[Y|X] = \nabla\phi(X) \qquad \text{a.s}$$

with $\phi : \mathbf{R}^n \to \mathbf{R}$ a convex function of class $\mathcal{C}^1$.

Remark : the independent coupling is of the Brenier-Strassen type.

Goal of the talk : Identify in the class of Brenier-Strassen couplings a sub-class which is optimal for some *generalized transport* problem.

I - Generalized Transport and a mixture of Brenier and Strassen theorems

G.-Roberto-Samson-Tetali (2017)

Let $\pi \in \Pi(\mu, \nu)$ be a transport plan between μ and ν written in disintegrated form

$$d\pi(x, y) = d\mu(x)dp_x(y),$$

with $x \mapsto p_x$ a transition kernel (μ a.s unique).

If $\omega : E \times E \to \mathbf{R}^+$ is a cost function then

$$\iint \omega(x, y)\, d\pi(x, y) = \int \left(\int \omega(x, y)\, dp_x(y) \right) d\mu(x).$$

In other words, transports of mass coming from x are penalized through their mean cost : $\int \omega(x, y)\, dp_x(y)$.

Idea of generalized transport : introduce more general penalizations.

Generalized Transport

Let $\mathcal{P}(E)$ denote the set of all probability measures on E.

Definition

Let $c : E \times \mathcal{P}(E) \to \mathbf{R}^+ \cup \{+\infty\}$; the generalized transport cost $\mathcal{T}_c(\nu|\mu)$ is defined by

$$\mathcal{T}_c(\nu|\mu) = \inf_{p \in \mathcal{P}(\mu,\nu)} \int c(x, p_x)\, d\mu(x),$$

where $\mathcal{P}(\mu,\nu)$ is the set of all probability kernels p such that $\mu p = \nu$.

Classical transport :

$$c(x,p) = \int \omega(x,y)\, dp(y).$$

In all useful examples, the function c is convex in p.

Remarks

- First examples of these kind of transport costs appeared in K. Marton's papers on concentration of measure.
- Many applications of generalized transport in terms of dimension free concentration of measure.
- Generalized transport encompasses many variants of the transport problem : optimal transport with martingale constraints, entropic optimal transport, causal optimal transport, ...

Quadratic barycentric transport

We will denote

$$\overline{\mathcal{T}}_2(\nu|\mu) = \inf_{p \in \mathcal{P}(\mu,\nu)} \int \left| x - \int y \, dp_x(y) \right|^2 d\mu(x)$$
$$= \inf_{X \sim \mu, Y \sim \nu} \mathbb{E}[|X - \mathbb{E}[Y|X]|^2],$$

the generalized transport cost associated to the cost function

$$c(x,p) = \left| x - \int y \, dp(y) \right|^2 .$$

By Jensen,

$$\overline{\mathcal{T}}_2(\nu|\mu) \leq W_2^2(\nu,\mu).$$

A mixture of Brenier and Strassen Theorems

G.-Juillet (2018) / Alfonsi-Corbetta-Jourdain (2017)
Dimension 1 : G.-Roberto-Samson-Shu-Tetali (2015)

Let $\mathcal{P}_2(\mathbf{R}^n)$ denote the set of probability measures with a finite second moment.

Theorem 1

Let $\mu, \nu \in \mathcal{P}_2(\mathbf{R}^n)$; define $B_\nu = \{\eta \in \mathcal{P}_1(\mathbf{R}^n) : \eta \leq_c \nu\}$.
There exists a unique probability measure $\bar{\mu} \in B_\nu$ such that

$$W_2(\bar{\mu}, \mu) = \inf_{\eta \in B_\nu} W_2(\eta, \mu).$$

Moreover

$$\overline{\mathcal{T}}_2(\nu|\mu) = W_2^2(\bar{\mu}, \mu).$$

A mixture of Brenier and Strassen Theorems

G.-Juillet (2018) / Backhoff-Veraguas - Beiglboeck - Pammer (2018)

Theorem 2

Let $\mu, \nu \in \mathcal{P}_2(\mathbf{R}^n)$;

(1) There exists a convex function $\phi : \mathbf{R}^n \to \mathbf{R}$ of class $\mathcal{C}^1$ such that

$$\bar{\mu} = \nabla\phi_{\#}\mu.$$

Moreover $\nabla\phi$ is 1-Lipschitz.

(2) A coupling (X, Y) between μ and ν is optimal for $\overline{\mathcal{T}}_2(\nu|\mu)$ if and only if $\mathbb{E}[Y|X] = \nabla\phi(X)$ a.s.

Remarks

Optimal transport between μ and its projection $\bar{\mu}$ is thus more regular than in the generic case : it is automatically given by a Lipschitz continuous transport map without any particular assumption on μ.

Examples

Theorem

If $\mu \in \mathcal{P}_2(\mathbf{R}^n)$ and $\nu = \sum_{i=0}^k p_i \delta_{y_i}$ with $p_i \geq 0$ and $y_0, \ldots, y_k$ affinely independent points of $\mathbf{R}^n$, then there exists some $c \in \mathbf{R}^n$ such that

$$\bar{\mu} = T_{\#}\mu, \qquad \text{with} \qquad T(x) = \mathrm{Proj}_\Delta(x + c),$$

where Δ is the convex hull of $\{y_0, \ldots, y_k\}$ and Proj_Δ denotes the orthogonal projection on Δ.

Other example : In dimension 1, Alfonsi-Corbetta-Jourdain (2017) obtained a semi-explicit formula for the transport map T sending μ on $\bar{\mu}$.

II - Link with the Caffarelli contraction theorem

Caffarelli contraction theorem

Theorem (Caffarelli (2000))

If $\mu = \gamma$ is the standard Gaussian measure on $\mathbf{R}^n$ and $d\nu(y) = e^{-V(y)}\,dy$ is a probability measure associated to a $\mathcal{C}^2$ smooth function V on $\mathbf{R}^n$ such that $\mathrm{Hess}\,V \geq \mathrm{Id}$, then there exists a convex function $\phi : \mathbf{R}^n \to \mathbf{R}$ of class $\mathcal{C}^1$ such that $\nu = \nabla\phi_{\#}\gamma$ and such that $\nabla\phi$ is 1-Lipschitz.

In other words, the Brenier map from γ to ν is a contraction.

It will be convenient to write $d\nu(y) = e^{-W(y)}\,d\gamma(y)$, with W a convex function.

Original proof based on the Monge-Ampère equation satisfied by ϕ.

Generalizations by Kolesnikov ('10), Kim-Milman ('12), Colombo-Figalli-Jhaveri ('17).

Applications of Caffarelli contraction theorem

Numerous consequences in the field of functional inequalities.

Example : the standard Gaussian measure γ satisfies the log-Sobolev inequality (Gross (1975)) :

$$\textbf{(LSI)} \qquad \mathrm{Ent}_\gamma(f^2) \leq 2 \int |\nabla f|^2 \, d\gamma, \qquad \forall f : \mathbf{R}^n \to \mathbf{R} \ \mathcal{C}^1$$

If $d\nu(y) = e^{-V(y)}\, dy$ with $\mathrm{Hess}\, V \geq \mathrm{Id}$, then according to Caffarelli Theorem $\nu = \nabla\phi_{\#}\gamma$ with $\nabla\phi$ 1-Lispchitz.

Therefore, applying (**LSI**) to $f = g \circ \nabla\phi$ yields to

$$\begin{aligned} \mathrm{Ent}_\nu(g^2) &\leq 2 \int |\mathrm{Hess}\, \phi(x) \cdot \nabla g(\nabla\phi(x))|^2 \, d\gamma(x), \qquad \forall f : \mathbf{R}^n \to \mathbf{R} \ \mathcal{C}^1 \\ &\leq 2 \int |\nabla g(y)|^2 \, d\nu(y). \end{aligned}$$

So ν satisfies (**LSI**) : one recovers the Bakry-Emery criterion (with the good constant)

Link with Caffarelli contraction theorem

The following result is a consequence of our main results :

Corollary 1

Let $\mu, \nu \in \mathcal{P}_2(\mathbf{R}^n)$; the following propositions are equivalent

(1) There exists $\phi : \mathbf{R}^n \to \mathbf{R}$ convex and $\mathcal{C}^1$ such that $\nu = \nabla\phi_{\#}\mu$ with $\nabla\phi$ 1-Lipschitz ;

(2) $\bar{\mu} = \nu$;

(3) $W_2^2(\nu, \mu) = \overline{\mathcal{T}}_2(\nu|\mu)$.

Corollary 2

If γ is the standard gaussian measure on $\mathbf{R}^n$ and $d\nu(y) = e^{-V(y)}\,dy$, with $\mathrm{Hess}\, V \geq \mathrm{Id}$, then

$$\bar{\gamma} = \nu.$$

A new proof of Caffarelli contraction theorem

Joint work with M. Fathi et M. Prod'Homme.

Let $d\nu(y) = e^{-W(y)}\, d\gamma(y)$, with W convex.

Goal : Recover Caffarelli's theorem by showing that $\bar{\gamma} = \nu$, *i.e*

$$\eta \leq_c \nu \Rightarrow W_2(\eta, \gamma) \geq W_2(\nu, \gamma).$$

Idea : Go to the entropic level.

Relative entropy of ν_1 with respect to ν_2 :

$$H(\nu_1|\nu_2) = \int \log \frac{d\nu_1}{d\nu_2}\, d\nu_1 \qquad \text{if} \qquad \nu_1 \ll \nu_2$$

(and $H(\nu_1|\nu_2) = +\infty$ otherwise)

Heuristic calculation

Let us replace $W_2(\cdot\,,\gamma)$ with $H(\cdot|\gamma)$. Then,

$$\begin{aligned}
H(\eta|\gamma) &= \int \log \frac{d\eta}{d\gamma}\, d\eta \\
&= \int \log \frac{d\eta}{d\nu}\, d\eta + \int \log \frac{d\nu}{d\gamma}\, d\eta \\
&= H(\eta|\nu) - \int W\, d\eta \\
&\geq H(\eta|\nu) - \int W\, d\nu \qquad \text{since } \eta \leq_c \nu \\
&= H(\eta|\nu) + H(\nu|\gamma) \\
&\geq H(\nu|\gamma) \qquad \text{since } H(\eta|\nu) \geq 0
\end{aligned}$$

A new proof of Caffarelli contraction theorem

Let $R^\varepsilon = \mathrm{Law}(Z_0, Z_\varepsilon)$ where $(Z_t)_{t\geq 0}$ is a standard Ornstein-Uhlenbeck process at equilibrium.

Define the entropic regularization of the optimal transport cost :

$$\mathcal{T}^\varepsilon(\mu,\nu) = \inf_{\pi\in\Pi(\mu,\nu)} H(\pi|R^\varepsilon).$$

Also related to Schrödinger bridges - see Leonard's survey ('13).

Theorem (Fathi - G. - Prod'Homme, 2019)

If $d\nu(y) = e^{-W(y)}\,d\gamma(y)$, with W convex, then for all $\eta \leq_c \nu$ regular enough

$$\mathcal{T}^\varepsilon(\gamma,\eta) \geq \mathcal{T}^\varepsilon(\gamma,\nu).$$

Since $\varepsilon\mathcal{T}^\varepsilon(\mu,\nu) \to \frac{1}{2}W_2^2(\mu,\nu)$ (Mikami '04), one gets the desired property :

$$\eta \leq_c \nu \Rightarrow W_2(\eta,\gamma) \geq W_2(\nu,\gamma).$$

Heuristic justification of $\varepsilon\mathcal{T}^\varepsilon(\mu,\nu) \to \frac{1}{2}W_2^2(\mu,\nu)$.

Idea of the proof 1/3

Heuristic justification of $\varepsilon \mathcal{T}^{\varepsilon}(\mu,\nu) \to \frac{1}{2} W_2^2(\mu,\nu)$.
The law R^{ε} is explicitly given by

$$R^{\varepsilon} = \mathrm{Law}\left(X, Xe^{-\varepsilon/2} + \sqrt{1-e^{-\varepsilon}}\,Y\right),$$

with X, Y two independent standard Gaussian random vectors on $\mathbf{R}^d$.

Idea of the proof 1/3

Heuristic justification of $\varepsilon \mathcal{T}^\varepsilon(\mu,\nu) \to \frac{1}{2} W_2^2(\mu,\nu)$.
The law R^ε is explicitly given by

$$R^\varepsilon = \mathrm{Law}\left(X, Xe^{-\varepsilon/2} + \sqrt{1-e^{-\varepsilon}}\,Y\right),$$

with X, Y two independent standard Gaussian random vectors on $\mathbf{R}^d$.
In other words,

$$R^\varepsilon(dxdy) = \frac{1}{Z_\varepsilon} \exp\left(-\frac{|x|^2}{2} - \frac{|y - xe^{-\varepsilon/2}|^2}{2(1-e^{-\varepsilon})}\right) dxdy.$$

Idea of the proof 1/3

Heuristic justification of $\varepsilon\mathcal{T}^{\varepsilon}(\mu,\nu) \to \frac{1}{2}W_2^2(\mu,\nu)$.
The law R^{ε} is explicitly given by

$$R^{\varepsilon} = \mathrm{Law}\left(X, Xe^{-\varepsilon/2} + \sqrt{1-e^{-\varepsilon}}\,Y\right),$$

with X, Y two independent standard Gaussian random vectors on $\mathbf{R}^d$.
In other words,

$$R^{\varepsilon}(dxdy) = \frac{1}{Z_{\varepsilon}}\exp\left(-\frac{|x|^2}{2} - \frac{|y-xe^{-\varepsilon/2}|^2}{2(1-e^{-\varepsilon})}\right)dxdy.$$

Therefore,

$$\begin{aligned}
\varepsilon H(\pi|R^{\varepsilon}) &= \varepsilon\int \log\left(\frac{d\pi}{dx}\right)d\pi - \varepsilon\int\log\left(\frac{dR^{\varepsilon}}{dx}\right)d\pi\\
&= \varepsilon\int\log\left(\frac{d\pi}{dx}\right)d\pi + \frac{\varepsilon}{2(1-e^{-\varepsilon})}\int |y-e^{-\varepsilon/2}x|^2\,\pi(dxdy)\\
&+ \frac{\varepsilon}{2}\int |x|^2\,\mu(dx) + c(\varepsilon),
\end{aligned}$$

where $c(\varepsilon)\to 0$ (and is independent of μ,ν,π).

Heuristic justification of $\varepsilon \mathcal{T}^\varepsilon(\mu,\nu) \to \frac{1}{2} W_2^2(\mu,\nu)$.
The law R^ε is explicitly given by

$$R^\varepsilon = \mathrm{Law}\left(X, Xe^{-\varepsilon/2} + \sqrt{1-e^{-\varepsilon}}Y\right),$$

with X, Y two independent standard Gaussian random vectors on $\mathbf{R}^d$.
In other words,

$$R^\varepsilon(dxdy) = \frac{1}{Z_\varepsilon} \exp\left(-\frac{|x|^2}{2} - \frac{|y - xe^{-\varepsilon/2}|^2}{2(1-e^{-\varepsilon})}\right) dxdy.$$

Therefore, as $\varepsilon \to 0$

$$\varepsilon H(\pi|R^\varepsilon) \sim \frac{1}{2}\int |x-y|^2 \, \pi(dxdy)$$

Idea of the proof 1/3

Heuristic justification of $\varepsilon \mathcal{T}^\varepsilon(\mu,\nu) \to \frac{1}{2} W_2^2(\mu,\nu)$.
The law R^ε is explicitly given by

$$R^\varepsilon = \mathrm{Law}\left(X, Xe^{-\varepsilon/2} + \sqrt{1-e^{-\varepsilon}}\,Y\right),$$

with X, Y two independent standard Gaussian random vectors on $\mathbf{R}^d$.
In other words,

$$R^\varepsilon(dxdy) = \frac{1}{Z_\varepsilon} \exp\left(-\frac{|x|^2}{2} - \frac{|y - xe^{-\varepsilon/2}|^2}{2(1-e^{-\varepsilon})}\right) dxdy.$$

Therefore, as $\varepsilon \to 0$

$$\varepsilon H(\pi|R^\varepsilon) \sim \frac{1}{2} \int |x-y|^2\, \pi(dxdy)$$

So for small ε, minimizing $\pi \mapsto H(\pi|R^\varepsilon)$ amounts to minimizing
$\pi \mapsto \frac{1}{2} \int |x-y|^2\, \pi(dxdy)$.

Idea of the proof 1/3

Heuristic justification of $\varepsilon\mathcal{T}^\varepsilon(\mu,\nu) \to \frac{1}{2}W_2^2(\mu,\nu)$.
The law R^ε is explicitly given by

$$R^\varepsilon = \mathrm{Law}\left(X, Xe^{-\varepsilon/2} + \sqrt{1-e^{-\varepsilon}}\,Y\right),$$

with X, Y two independent standard Gaussian random vectors on $\mathbf{R}^d$.
In other words,

$$R^\varepsilon(dxdy) = \frac{1}{Z_\varepsilon}\exp\left(-\frac{|x|^2}{2} - \frac{|y-xe^{-\varepsilon/2}|^2}{2(1-e^{-\varepsilon})}\right)dxdy.$$

Therefore, as $\varepsilon \to 0$

$$\varepsilon H(\pi|R^\varepsilon) \sim \frac{1}{2}\int |x-y|^2\,\pi(dxdy)$$

So for small ε, minimizing $\pi \mapsto H(\pi|R^\varepsilon)$ amounts to minimizing $\pi \mapsto \frac{1}{2}\int |x-y|^2\,\pi(dxdy)$.
Rigorous proof can be found in Mikami ('04), Léonard ('12), Carlier-Duval-Peyré-Schmitzer ('17)

Idea of the proof 2/3

Key point : The optimal coupling π^* for $\mathcal{T}^\varepsilon(\gamma,\nu)$ is of the form

$$d\pi^*(x,y) = f(x)g(y)\, dR^\varepsilon(x,y),$$

with f *log-convex* and g *log-concave*.

Classical fact : A coupling π is optimal for $\mathcal{T}^\varepsilon(\gamma,\nu)$ if and only if it is of the form

$$\pi(dxdy) = f(x)g(y)dR^\varepsilon(x,y).$$

Idea of the proof 3/3

A p.m. $\pi(dxdy) = f(x)g(y)dR^{\varepsilon}(x,y)$ is a coupling between γ and $\nu = e^{-W}\,d\gamma$ if and only if f and g solve the following non-linear system of equations

$$\begin{cases} 1 = f(x)P_{\varepsilon}g(x) \\ e^{-W(y)} = P_{\varepsilon}(f)(y)g(y) \end{cases}$$

Idea of the proof 3/3

A p.m. $\pi(dxdy) = f(x)g(y)dR^\varepsilon(x,y)$ is a coupling between γ and $\nu = e^{-W}\,d\gamma$ if and only if f and g solve the following non-linear system of equations

$$\begin{cases} 1 = f(x)P_\varepsilon g(x) \\ e^{-W(y)} = P_\varepsilon(f)(y)g(y) \end{cases}$$

which is equivalent to a fixed point equation

$$g(y) = \frac{e^{-W(y)}}{P_\varepsilon f(y)} \quad \text{and} \quad f = \Phi_\varepsilon(f) \quad \text{where} \quad \Phi_\varepsilon(h) = \frac{1}{P_\varepsilon\left(e^{-W}\frac{1}{P_\varepsilon(h)}\right)}$$

Idea of the proof 3/3

A p.m. $\pi(dxdy) = f(x)g(y)dR^{\varepsilon}(x,y)$ is a coupling between γ and $\nu = e^{-W}\,d\gamma$ if and only if f and g solve the following non-linear system of equations

$$\begin{cases} 1 = f(x)P_\varepsilon g(x) \\ e^{-W(y)} = P_\varepsilon(f)(y)g(y) \end{cases}$$

which is equivalent to a fixed point equation

$$g(y) = \frac{e^{-W(y)}}{P_\varepsilon f(y)} \quad \text{and} \quad f = \Phi_\varepsilon(f) \quad \text{where} \quad \Phi_\varepsilon(h) = \frac{1}{P_\varepsilon\left(e^{-W}\frac{1}{P_\varepsilon(h)}\right)}$$

$\rightsquigarrow$ We want to show that this system admits a solution (f,g) with f log-convex and g log-concave.

Idea of the proof 3/3

A p.m. $\pi(dxdy) = f(x)g(y)dR^\varepsilon(x,y)$ is a coupling between γ and $\nu = e^{-W}\,d\gamma$ if and only if f and g solve the following non-linear system of equations

$$\begin{cases} 1 = f(x)P_\varepsilon g(x) \\ e^{-W(y)} = P_\varepsilon(f)(y)g(y) \end{cases}$$

which is equivalent to a fixed point equation

$$g(y) = \frac{e^{-W(y)}}{P_\varepsilon f(y)} \quad \text{and} \quad f = \Phi_\varepsilon(f) \quad \text{where} \quad \Phi_\varepsilon(h) = \frac{1}{P_\varepsilon\left(e^{-W}\frac{1}{P_\varepsilon(h)}\right)}$$

$\rightsquigarrow$ We want to show that this system admits a solution (f,g) with f log-convex and g log-concave.

Crucial property : Φ leaves stable the class of log-convex functions.

Idea of the proof 3/3

A p.m. $\pi(dxdy) = f(x)g(y)dR^\varepsilon(x,y)$ is a coupling between γ and $\nu = e^{-W}\,d\gamma$ if and only if f and g solve the following non-linear system of equations

$$\begin{cases} 1 = f(x)P_\varepsilon g(x) \\ e^{-W(y)} = P_\varepsilon(f)(y)g(y) \end{cases}$$

which is equivalent to a fixed point equation

$$g(y) = \frac{e^{-W(y)}}{P_\varepsilon f(y)} \quad \text{and} \quad f = \Phi_\varepsilon(f) \quad \text{where} \quad \Phi_\varepsilon(h) = \frac{1}{P_\varepsilon\left(e^{-W}\frac{1}{P_\varepsilon(h)}\right)}$$

$\rightsquigarrow$ We want to show that this system admits a solution (f,g) with f log-convex and g log-concave.

Crucial property : Φ leaves stable the class of log-convex functions.

Imagine Φ_ε is contractant for a good metric, then starting with f_0 log-convex, the sequence f_n defined by $f_{n+1} = \Phi_\varepsilon(f_n)$ will converge to a fixed point f_∞ which will be log-convex. And $g_\infty = e^{-W}/P_\varepsilon(f_\infty)$ will be log-concave.

Idea of the proof 3/3

A p.m. $\pi(dxdy) = f(x)g(y)dR^\varepsilon(x,y)$ is a coupling between γ and $\nu = e^{-W}\, d\gamma$ if and only if f and g solve the following non-linear system of equations

$$\begin{cases} 1 = f(x)P_\varepsilon g(x) \\ e^{-W(y)} = P_\varepsilon(f)(y)g(y) \end{cases}$$

which is equivalent to a fixed point equation

$$g(y) = \frac{e^{-W(y)}}{P_\varepsilon f(y)} \quad \text{and} \quad f = \Phi_\varepsilon(f) \quad \text{where} \quad \Phi_\varepsilon(h) = \frac{1}{P_\varepsilon\left(e^{-W}\frac{1}{P_\varepsilon(h)}\right)}$$

$\rightsquigarrow$ We want to show that this system admits a solution (f, g) with f log-convex and g log-concave.

Crucial property : Φ leaves stable the class of log-convex functions.

Imagine Φ_ε is contractant for a good metric, then starting with f_0 log-convex, the sequence f_n defined by $f_{n+1} = \Phi_\varepsilon(f_n)$ will converge to a fixed point f_∞ which will be log-convex. And $g_\infty = e^{-W}/P_\varepsilon(f_\infty)$ will be log-concave.

Φ_ε is not contractant in general. Fortunately it is possible to adapt the arguments of a paper of Fortet (1940) to show the convergence of f_n to a fixed point of Φ_ε.

Thank you for your attention !