Some hints and discussions about the regularity problem for the case $c(x, y) = |x - y|$ and its approximation

Filippo Santambrogio

Laboratoire de Mathématiques d'Orsay, Université Paris-Sud
http://www.math.u-psud.fr/∼santambr/

Pisa, November 6th, 2012

1. Costs and regularity in optimal transport
   - The quadratic cost
   - The linear cost and its properties
   - Costs satisfying the MTW condition
2. Approximation of the optimal transport for $c(x, y) = |x - y|$
3. Uniform estimates
4. Difficulties and hints

Costs and regularity in optimal transport

Linear, quadratic, and more exotic costs

The two most studied costs

We dare say that the two main cost functions used for optimal transport in the Euclidean space, due to the interesting features of the related minimization problems, are

$$c_1(x, y) = |x - y|$$

(the original one proposed by Monge) and

$$c_2(x, y) = |x - y|^2,$$

which is equivalent to using $c(x, y) = -x \cdot y$, and has many applications.

The quadratic case

For the cost c_2 at least two important properties found by Y. Brenier are crucial : the optimal transport is always the gradient of a convex function, $T = \nabla\psi$, and there is a dynamical formulation where the kinetic energy is minimized,

$$\min\left\{\int |x - T(x)|^2 f(x)dx, \;:\; T_\# f = g\right\}$$
$$= \min\left\{\int_0^1 \int \rho_t |v_t|^2 \, dx\, dt, \;:\; \partial_t \rho_t + \nabla\cdot(\rho_t v_t) = 0, \quad \rho_0 = f,\, \rho_1 = g\right\}.$$

The optimal transport can be retrieved via the Monge-Ampère equation on ψ :

$$\det(D^2\psi) = \frac{f}{g\circ\nabla\psi}, \quad (\nabla\psi)(\Omega) = \Omega'$$

where $\Omega = \operatorname{spt} f$ and $\Omega' = \operatorname{spt} g$. This is a fully nonlinear elliptic equation in the class of convex functions.

The quadratic case

For the cost c_2 at least two important properties found by Y. Brenier are crucial : the optimal transport is always the gradient of a convex function, $T = \nabla\psi$, and there is a dynamical formulation where the kinetic energy is minimized,

$$\min\left\{\int |x - T(x)|^2 f(x)dx, \ : \ T_\# f = g\right\}$$

$$= \min\left\{\int_0^1 \int \rho_t |v_t|^2 \, dx \, dt, \ : \ \partial_t \rho_t + \nabla \cdot (\rho_t v_t) = 0, \quad \rho_0 = f, \, \rho_1 = g\right\}.$$

The optimal transport can be retrieved via the Monge-Ampère equation on ψ :

$$\det(D^2\psi) = \frac{f}{g \circ \nabla\psi}, \quad (\nabla\psi)(\Omega) = \Omega'$$

where $\Omega = \operatorname{spt} f$ and $\Omega' = \operatorname{spt} g$. This is a fully nonlinear elliptic equation in the class of convex functions.

Regularity for the quadratic case

Many regularity results have been established for the optimal transport T for c_2 thanks to a refined analysis of the Monge-Ampère equation made by Caffarelli, Urbas, and, more recently, by DePhilippis and Figalli.
In particular we have

Theorem

Suppose that $\operatorname{spt} f = \Omega$ *and* $\operatorname{spt} g = \Omega'$ *are convex sets, and that* f *and* g *are bounded from above and below on* Ω *and* Ω'. *Then* $\psi \in C^{1,\alpha} \cap W^{2,1}$ *and* $T \in C^{0,\alpha} \cap W^{1,1}$. *Moreover,*
$f, g \in C^{k,\alpha} \Rightarrow \psi \in C^{k+2,\alpha} \Rightarrow T \in C^{k+1,\alpha}$. *This regularity arrives up to the boundary in case of strict convexity of the domains.*

The linear case

The linear case has some different interesting features. Set

$$W_1(f,g) := \min\left\{ \int |x - T(x)| f(x) dx : T_\# f = g \right\} = \min\left\{ \int c_1 d\gamma : \begin{array}{ll} (\pi_x)_\# \gamma & = f \\ (\pi_y)_\# \gamma & = g \end{array} \right\}$$

Duality It is expressed in terms of Lipschitz functions

$$W_1(f,g) = \max\left\{ \int u\, d(f-g) \;:\; u \in \mathrm{Lip}_1 \right\}.$$

The optimal u (*Kantorovich potential*) gives a partition of Ω into segments (*transport rays*) where u is affine with slope 1, which are preserved by the transport, since $|x-y| = u(x) - u(y)$ γ−a.e.

Non-uniqueness of the optimizer Any plan γ "respecting" the transport rays is optimal. This reduces to a family of one-dimensional problems, which have many solutions. One of them plays a special role : the optimal map T_{mon} which is monotone increasing on each transport ray. It is also the solution of the secondary variational problem

$$\min\left\{ \int c_2 d\gamma \;:\; \gamma \text{ optimal from } f \text{ to } g \text{ for the cost } c_1 \right\}$$

The linear case

The linear case has some different interesting features. Set

$$W_1(f,g) := \min\left\{\int |x - T(x)| f(x)dx : T_\# f = g\right\} = \min\left\{\int c_1 d\gamma : \begin{array}{ll}(\pi_x)_\#\gamma & = f\\ (\pi_y)_\#\gamma & = g\end{array}\right\}$$

Duality It is expressed in terms of Lipschitz functions

$$W_1(f,g) = \max\left\{\int u\, d(f-g) \ : \ u \in \mathrm{Lip}_1\right\}.$$

The optimal u (*Kantorovich potential*) gives a partition of Ω into segments (*transport rays*) where u is affine with slope 1, which are preserved by the transport, since $|x-y| = u(x) - u(y)$ $\gamma-$a.e.

Non-uniqueness of the optimizer Any plan γ "respecting" the transport rays is optimal. This reduces to a family of one-dimensional problems, which have many solutions. One of them plays a special role : the optimal map T_{mon} which is monotone increasing on each transport ray. It is also the solution of the secondary variational problem

$$\min\left\{\int c_2 d\gamma \ : \ \gamma \text{ optimal from } f \text{ to } g \text{ for the cost } c_1\right\}$$

The linear case

The linear case has some different interesting features. Set

$$W_1(f,g):=\min\left\{\int|x-T(x)|f(x)dx : T_\# f=g\right\}=\min\left\{\int c_1 d\gamma : \begin{array}{ll}(\pi_x)_\#\gamma & =f\\ (\pi_y)_\#\gamma & =g\end{array}\right\}$$

Duality It is expressed in terms of Lipschitz functions

$$W_1(f,g)=\max\left\{\int u\, d(f-g)\ :\ u\in \mathrm{Lip}_1\right\}.$$

The optimal u (*Kantorovich potential*) gives a partition of Ω into segments (*transport rays*) where u is affine with slope 1, which are preserved by the transport, since $|x-y|=u(x)-u(y)$ γ−a.e.

Non-uniqueness of the optimizer Any plan γ "respecting" the transport rays is optimal. This reduces to a family of one-dimensional problems, which have many solutions. One of them plays a special role : the optimal map T_{mon} which is monotone increasing on each transport ray. It is also the solution of the secondary variational problem

$$\min\left\{\int c_2 d\gamma\ :\ \gamma \text{ optimal from } f \text{ to } g \text{ for the cost } c_1\right\}$$

The linear case and the transport density

Prescribed divergence minimization There's equivalence with the Beckmann's problem

$$W_1(f,g) = \min\left\{\int |v(x)|dx \,:\, \nabla\cdot v = f-g\right\}.$$

The optimal v is of the form $\sigma\nabla u$, and σ, given by

$$\sigma = \int_0^1 (\pi_t)_\#(c_1\cdot\gamma)\,dt, \quad \text{where } \pi_t(x,y) = (1-t)x + ty,$$

is called *transport density*. It does not really depend on the choice of the optimal γ and only depends on $f-g$.

Mild regularity results exist, for instance

$$f,g\in L^p \Rightarrow \sigma, v\in L^p \qquad (1\leq p\leq\infty)$$

(DePascale, Pratelli, Evans, S. . .)

Yet, higher regularity, such as $f,g\in C^0\Rightarrow \sigma,v\in C^0$ are not known.

The linear case and the transport density

Prescribed divergence minimization There's equivalence with the Beckmann's problem

$$W_1(f,g) = \min\left\{\int |v(x)|dx \,:\, \nabla\cdot v = f-g\right\}.$$

The optimal v is of the form $\sigma\nabla u$, and σ, given by

$$\sigma = \int_0^1 (\pi_t)_\#(c_1\cdot\gamma)\,dt, \quad \text{where } \pi_t(x,y) = (1-t)x+ty,$$

is called *transport density*. It does not really depend on the choice of the optimal γ and only depends on $f-g$.
Mild regularity results exist, for instance

$$f,g\in L^p \Rightarrow \sigma, v\in L^p \qquad (1\le p\le\infty)$$

(DePascale, Pratelli, Evans, S. . .)
Yet, higher regularity, such as $f,g\in C^0\Rightarrow \sigma, v\in C^0$ are not known.

Examples of regularity and non-regularity in the linear case

Consider $\Omega = \Omega' = [0,1] \times [0,1]$ and write f_x and g_x for $f(x,\cdot)$ and $g(x,\cdot)$ and F_x and G_x for their primitives. Suppose f, g bounded from below and above. Suppose $F_x(0) = G_x(0) = 0$, $F_x(1) = G_x(1)$ and $F_x(y) \geq G_x(y)$ for all x, y.

Then the optimal potential u is given by $u(x,y) = -y$ and the monotone optimal transport T by

$$T_{\mathrm{mon}}(x,y) = G_x^{-1}(F_x(y)).$$

Its regularity w.r.t. the y variable is better than that of f and g (because of the primitives) but w.r.t. x it is the same as that of F_x and G_x, i.e. that of f and g.

Regularity : the regularity of T_{mon} is non-trivial and has not been investigated that much so far, apart from one paper by Fragalà, Gelli, Pratelli in the 2D case when f and g have disjoint and convex supports.

Notice that if we replace the square with a non-convex domain the regularity of (F_x, G_x) may be worse than that of (f_x, g_x) and the same example would not work.

Examples of regularity and non-regularity in the linear case

Consider $\Omega = \Omega' = [0,1] \times [0,1]$ and write f_x and g_x for $f(x,\cdot)$ and $g(x,\cdot)$ and F_x and G_x for their primitives. Suppose f,g bounded from below and above. Suppose $F_x(0) = G_x(0) = 0$, $F_x(1) = G_x(1)$ and $F_x(y) \geq G_x(y)$ for all x,y.

Then the optimal potential u is given by $u(x,y) = -y$ and the monotone optimal transport T by

$$T_{\mathrm{mon}}(x,y) = G_x^{-1}(F_x(y)).$$

Its regularity w.r.t. the y variable is better than that of f and g (because of the primitives) but w.r.t. x it is the same as that of F_x and G_x, i.e. that of f and g.

Regularity : the regularity of T_{mon} is non-trivial and has not been investigated that much so far, apart from one paper by Fragalà, Gelli, Pratelli in the 2D case when f and g have disjoint and convex supports.

Notice that if we replace the square with a non-convex domain the regularity of (F_x, G_x) may be worse than that of (f_x, g_x) and the same example would not work.

Examples of regularity and non-regularity in the linear case

Consider $\Omega = \Omega' = [0,1] \times [0,1]$ and write f_x and g_x for $f(x,\cdot)$ and $g(x,\cdot)$ and F_x and G_x for their primitives. Suppose f, g bounded from below and above. Suppose $F_x(0) = G_x(0) = 0$, $F_x(1) = G_x(1)$ and $F_x(y) \geq G_x(y)$ for all x, y.

Then the optimal potential u is given by $u(x,y) = -y$ and the monotone optimal transport T by

$$T_{\mathrm{mon}}(x,y) = G_x^{-1}(F_x(y)).$$

Its regularity w.r.t. the y variable is better than that of f and g (because of the primitives) but w.r.t. x it is the same as that of F_x and G_x, i.e. that of f and g.

Regularity : the regularity of T_{mon} is non-trivial and has not been investigated that much so far, apart from one paper by Fragalà, Gelli, Pratelli in the 2D case when f and g have disjoint and convex supports.

Notice that if we replace the square with a non-convex domain the regularity of (F_x, G_x) may be worse than that of (f_x, g_x) and the same example would not work.

Ma-Trudinger-Wang regularity for more exotic costs

Consider a cost c satisfying the *twist condition* :

$$\forall x \quad y \mapsto \nabla_x c(x,y) \text{ is injective.}$$

Then, the optimal transport exists, is unique and may be computed from the Kantorovich potential φ : $T(x) = \big(\nabla_x c(x,\cdot)\big)^{-1}(\nabla\varphi(x))$. If c is smooth and $D_{xy}c$ is non-singular, then T has the same regularity as $\nabla\varphi$. In 2005 Ma, Trudinger and Wang found the condition which is the corner-stone for regularity results for more general costs. For $c \in C^4$, they proved some estimates on φ under the horrible assumption that there exists $c_0 > 0$ such that

$$\sum_{i,j,k,l,r,s} \left(\frac{\partial^3 c}{\partial x_i \partial x_j \partial y_r} (D_{xy}c)^{-1}_{r,s} \frac{\partial^3 c}{\partial x_s \partial y_k \partial y_l} - \frac{\partial^4 c}{\partial x_i \partial x_j \partial y_k \partial y_l} \right) \xi_i \xi_j \eta_k \eta_l \geq c_0 |\xi|^2 |\eta|^2$$

for every $\xi \perp \eta \in \mathbb{R}^n$.

Ma-Trudinger-Wang regularity for more exotic costs

Consider a cost c satisfying the *twist condition* :

$$\forall x \quad y \mapsto \nabla_x c(x,y) \text{ is injective.}$$

Then, the optimal transport exists, is unique and may be computed from the Kantorovich potential φ : $T(x) = \big(\nabla_x c(x,\cdot)\big)^{-1}(\nabla\varphi(x))$. If c is smooth and $D_{xy}c$ is non-singular, then T has the same regularity as $\nabla\varphi$.In 2005 Ma, Trudinger and Wang found the condition which is the corner-stone for regularity results for more general costs. For $c \in C^4$, they proved some estimates on φ under the horrible assumption that there exists $c_0 > 0$ such that

$$\sum_{i,j,k,l,r,s}\left(\frac{\partial^3 c}{\partial x_i \partial x_j \partial y_r}\big(D_{xy}c\big)^{-1}_{r,s}\frac{\partial^3 c}{\partial x_s \partial y_k \partial y_l} - \frac{\partial^4 c}{\partial x_i \partial x_j \partial y_k \partial y_l}\right)\xi_i\xi_j\eta_k\eta_l \geq c_0|\xi|^2|\eta|^2$$

for every $\xi \perp \eta \in \mathbb{R}^n$.

Ma-Trudinger-Wang regularity for more exotic costs II

The assumption on c has to be coupled with an assumption on the domains Ω and Ω', which should be $c-$convex :

$$\forall x \in \Omega \quad \nabla_y c(x,\cdot)(\Omega') \text{ and } \quad \forall y \in \Omega' \quad \nabla_x c(\cdot,y)(\Omega) \text{ are convex.}$$

Thanks to MTW's results and other subsequent works (Loeper, Liu...) it is possible to prove

Theorem

Suppose that $\operatorname{spt} f = \Omega$ *and* $\operatorname{spt} g = \Omega'$ *are a $c-$convex pair, that c satsifies MTW condition, and that f and g are bounded from above and below on Ω and Ω'. Then $\psi \in C^{1,\alpha}$ and $T \in C^{0,\alpha}$. Moreover, $f, g \in C^{k,\alpha}$ implies $\psi \in C^{k+2,\alpha}$ and $T \in C^{k+1,\alpha}$. This regularity arrives up to the boundary in case of strict $c-$convexity and smoothness of the domains.*

The results have also be extended to the case $c_0 = 0$.

Approximation of the cost $|x - y|$

paying attention to the MTW condition

MTW again : the case we are interested in

It happens that, for every $\varepsilon > 0$

$$c_{(\varepsilon)}(x, y) := \sqrt{\varepsilon^2 + |x - y|^2}$$

satisfies MTW's assumptions and that if we take $\Omega = \Omega'$ a strictly convex smooth domain (say, a ball), then Ω and Ω' form a $c-$strictly convex pair.
And we easily see that $\lim_{\varepsilon \to 0} \sqrt{\varepsilon^2 + |x - y|^2} = |x - y|$.
Selection of the limit : more precisely, if γ_ε is the optimal transport plan for $c_{(\varepsilon)}$, and $\gamma_\varepsilon \rightharpoonup \gamma$ as $\varepsilon \to 0$, what is the limit transport plan γ ?
By using

$$\sqrt{\varepsilon^2 + |x - y|^2} = |x - y| + \frac{\varepsilon^2}{2|x - y|} + o(\varepsilon^2)$$

it is possible to see that γ must be optimal for c_1 and, among the optimizers, must also minimize $\gamma \mapsto \int |x - y|^{-1} d\gamma$. This is again a one dimensional problem formally solved by T_{mon}.
Difficulty : if $\operatorname{spt} f \cap \operatorname{spt} g \neq \emptyset$, this secondary variational problem could give constantly the value $+\infty$.

MTW again : the case we are interested in

It happens that, for every $\varepsilon > 0$

$$c_{(\varepsilon)}(x, y) := \sqrt{\varepsilon^2 + |x - y|^2}$$

satisfies MTW's assumptions and that if we take $\Omega = \Omega'$ a strictly convex smooth domain (say, a ball), then Ω and Ω' form a $c-$strictly convex pair. And we easily see that $\lim_{\varepsilon \to 0} \sqrt{\varepsilon^2 + |x - y|^2} = |x - y|$.

Selection of the limit : more precisely, if γ_ε is the optimal transport plan for $c_{(\varepsilon)}$, and $\gamma_\varepsilon \rightharpoonup \gamma$ as $\varepsilon \to 0$, what is the limit transport plan γ ?

By using

$$\sqrt{\varepsilon^2 + |x - y|^2} = |x - y| + \frac{\varepsilon^2}{2|x - y|} + o(\varepsilon^2)$$

it is possible to see that γ must be optimal for c_1 and, among the optimizers, must also minimize $\gamma \mapsto \int |x - y|^{-1} d\gamma$. This is again a one dimensional problem formally solved by T_{mon}.

Difficulty : if $\operatorname{spt} f \cap \operatorname{spt} g \neq \emptyset$, this secondary variational problem could give constantly the value $+\infty$.

MTW again : the case we are interested in

It happens that, for every $\varepsilon > 0$

$$c_{(\varepsilon)}(x, y) := \sqrt{\varepsilon^2 + |x - y|^2}$$

satisfies MTW's assumptions and that if we take $\Omega = \Omega'$ a strictly convex smooth domain (say, a ball), then Ω and Ω' form a $c-$strictly convex pair. And we easily see that $\lim_{\varepsilon\to 0} \sqrt{\varepsilon^2 + |x - y|^2} = |x - y|$.

Selection of the limit : more precisely, if γ_ε is the optimal transport plan for $c_{(\varepsilon)}$, and $\gamma_\varepsilon \rightharpoonup \gamma$ as $\varepsilon \to 0$, what is the limit transport plan γ ?

By using

$$\sqrt{\varepsilon^2 + |x - y|^2} = |x - y| + \frac{\varepsilon^2}{2|x - y|} + o(\varepsilon^2)$$

it is possible to see that γ must be optimal for c_1 and, among the optimizers, must also minimize $\gamma \mapsto \int |x - y|^{-1} d\gamma$. This is again a one dimensional problem formally solved by T_{mon}.

Difficulty : if $\operatorname{spt} f \cap \operatorname{spt} g \neq \emptyset$, this secondary variational problem could give constantly the value $+\infty$.

MTW again : the case we are interested in

Help !!

(A question which could be useful to prove convergence to T_{mon} if the supports of f and g are not disjoint)

Given a sequence of absolutely continuous measures $f_n \rightharpoonup f$ and $g_n \rightharpoonup g$ (f ad g also absolutely continuous), let γ_n be the optimal transport plan which is monotone on transport rays between them. Is it true that it weakly converges to the corresponding plan between f and g ?

Application : take the optimal plan γ_ε for $c_{(\varepsilon)}$, fix $\delta > 0$, and define $\gamma_{\varepsilon,\delta}$ to be its restriction to the set $\{(x,y) : |x-y| > \delta\}$. It is an optimal plan between some measures $f_{\varepsilon,\delta}$ and $g_{\varepsilon,\delta}$. Look at its limit as $\varepsilon \to 0\dots$

Regularity as $\varepsilon \to 0$

Can the MTW estimates on the optimal transport T_ε for $c_{(\varepsilon)}$, pass to the limit as $\varepsilon \to 0$ and provide regularity for T_{mon} ?

- Of course, since $c_0 = 0$ is allowed any limit of costs satisfying MTW will also be OK **(but it does not work)**
- Obviously not, just look at the examples with T_{mon} : $f, g \in C^{0,\alpha}$ is not enough for $T_{\mathrm{mon}} \in C^{1,\alpha}$.

This suggests that

- The regularity assumptions on the cost itself and the twist condition are crucial for MTW
- Maybe something weaker can pass to the limit, and in particular estimates on the same order of regularity of f and g, not more.

We just need to check if we have some uniform estimates as $\varepsilon \to 0$.

Regularity as $\varepsilon \to 0$

Can the MTW estimates on the optimal transport T_ε for $c_{(\varepsilon)}$, pass to the limit as $\varepsilon \to 0$ and provide regularity for T_{mon} ?

- Of course, since $c_0 = 0$ is allowed any limit of costs satisfying MTW will also be OK **(but it does not work)**
- Obviously not, just look at the examples with T_{mon} : $f, g \in C^{0,\alpha}$ is not enough for $T_{\mathrm{mon}} \in C^{1,\alpha}$.

This suggests that

- The regularity assumptions on the cost itself and the twist condition are crucial for MTW
- Maybe something weaker can pass to the limit, and in particular estimates on the same order of regularity of f and g, not more.

We just need to check if we have some uniform estimates as $\varepsilon \to 0$.

Regularity as $\varepsilon \to 0$

Can the MTW estimates on the optimal transport T_ε for $c_{(\varepsilon)}$, pass to the limit as $\varepsilon \to 0$ and provide regularity for T_{mon} ?

- Of course, since $c_0 = 0$ is allowed any limit of costs satisfying MTW will also be OK **(but it does not work)**
- Obviously not, just look at the examples with T_{mon} : $f, g \in C^{0,\alpha}$ is not enough for $T_{\mathrm{mon}} \in C^{1,\alpha}$.

This suggests that

- The regularity assumptions on the cost itself and the twist condition are crucial for MTW
- Maybe something weaker can pass to the limit, and in particular estimates on the same order of regularity of f and g, not more.

We just need to check if we have some uniform estimates as $\varepsilon \to 0$.

Regularity as $\varepsilon \to 0$

Can the MTW estimates on the optimal transport T_ε for $c_{(\varepsilon)}$, pass to the limit as $\varepsilon \to 0$ and provide regularity for T_{mon} ?

- Of course, since $c_0 = 0$ is allowed any limit of costs satisfying MTW will also be OK **(but it does not work)**
- Obviously not, just look at the examples with T_{mon} : $f, g \in C^{0,\alpha}$ is not enough for $T_{\mathrm{mon}} \in C^{1,\alpha}$.

This suggests that

- The regularity assumptions on the cost itself and the twist condition are crucial for MTW
- Maybe something weaker can pass to the limit, and in particular estimates on the same order of regularity of f and g, not more.

We just need to check if we have some uniform estimates as $\varepsilon \to 0$.

Regularity as $\varepsilon \to 0$

Can the MTW estimates on the optimal transport T_ε for $c_{(\varepsilon)}$, pass to the limit as $\varepsilon \to 0$ and provide regularity for $T_{\rm mon}$?

- Of course, since $c_0 = 0$ is allowed any limit of costs satisfying MTW will also be OK **(but it does not work)**
- Obviously not, just look at the examples with $T_{\rm mon}$: $f, g \in C^{0,\alpha}$ is not enough for $T_{\rm mon} \in C^{1,\alpha}$.

This suggests that

- The regularity assumptions on the cost itself and the twist condition are crucial for MTW
- Maybe something weaker can pass to the limit, and in particular estimates on the same order of regularity of f and g, not more.

We just need to check if we have some uniform estimates as $\varepsilon \to 0$.

Regularity as $\varepsilon \to 0$

FAQ

Should we care about a limit notion of $c-$convexity as $\varepsilon \to 0$ for the domain ?

No, MTW already cared about it, we only need to get uniform bounds.

Should we prove, besides uniform a priori estimates, that we can actually approximate our transport map with smooth ones ?

No, the maps T_ε exactly do the job.

Should we worry about this fact : T_{mon} will never be smooth if Ω is not convex, but local interior estimates will never see this fact ?

No, because to apply a priori estimates to T_ε we already need to assume convexity.

Regularity as $\varepsilon \to 0$

FAQ

Should we care about a limit notion of $c-$convexity as $\varepsilon \to 0$ for the domain ?

No, MTW already cared about it, we only need to get uniform bounds.

Should we prove, besides uniform a priori estimates, that we can actually approximate our transport map with smooth ones ?

No, the maps T_ε exactly do the job.

Should we worry about this fact : T_{mon} will never be smooth if Ω is not convex, but local interior estimates will never see this fact ?

No, because to apply a priori estimates to T_ε we already need to assume convexity.

Regularity as $\varepsilon \to 0$

FAQ

Should we care about a limit notion of $c-$convexity as $\varepsilon \to 0$ for the domain ?

No, MTW already cared about it, we only need to get uniform bounds.

Should we prove, besides uniform a priori estimates, that we can actually approximate our transport map with smooth ones ?

No, the maps T_ε exactly do the job.

Should we worry about this fact : T_{mon} will never be smooth if Ω is not convex, but local interior estimates will never see this fact ?

No, because to apply a priori estimates to T_ε we already need to assume convexity.

Regularity as $\varepsilon \to 0$

FAQ

Should we care about a limit notion of $c-$convexity as $\varepsilon \to 0$ for the domain ?

No, MTW already cared about it, we only need to get uniform bounds.

Should we prove, besides uniform a priori estimates, that we can actually approximate our transport map with smooth ones ?

No, the maps T_ε exactly do the job.

Should we worry about this fact : T_{mon} will never be smooth if Ω is not convex, but local interior estimates will never see this fact ?

No, because to apply a priori estimates to T_ε we already need to assume convexity.

Regularity as $\varepsilon \to 0$

FAQ

Should we care about a limit notion of c−convexity as $\varepsilon \to 0$ for the domain ?

No, MTW already cared about it, we only need to get uniform bounds.

Should we prove, besides uniform a priori estimates, that we can actually approximate our transport map with smooth ones ?

No, the maps T_ε exactly do the job.

Should we worry about this fact : T_{mon} will never be smooth if Ω is not convex, but local interior estimates will never see this fact ?

No, because to apply a priori estimates to T_ε we already need to assume convexity.

Regularity as $\varepsilon \to 0$

FAQ

Should we care about a limit notion of $c-$*convexity as* $\varepsilon \to 0$ *for the domain ?*

No, MTW already cared about it, we only need to get uniform bounds.

Should we prove, besides uniform a priori estimates, that we can actually approximate our transport map with smooth ones ?

No, the maps T_ε exactly do the job.

Should we worry about this fact : T_{mon} *will never be smooth if* Ω *is not convex, but local interior estimates will never see this fact ?*

No, because to apply a priori estimates to T_ε we already need to assume convexity.

A more refined analysis of the cost $c_{(\varepsilon)}$

Since $c_{(\varepsilon)}$ satisfies the twist condition (even better : it is of the form $h(y-x)$ for $h(z)=\sqrt{\varepsilon^2+|z|^2}$, which is strictly convex), we can infer the optimal T from the potential φ : we know that, for $(x_0, y_0) \in \mathrm{spt}(\gamma)$ we have

$$x_0 \in \mathrm{argmin}_x \, h(x-y_0)-\varphi(x)$$

and hence, at $x=x_0$

$$\nabla\varphi(x)=\nabla h(x-T(x)) \Rightarrow T(x)=x-\frac{\varepsilon}{\sqrt{1-|\nabla\varphi(x)|^2}}\nabla\varphi(x).$$

The function h has Lipschitz constant strictly less than 1 on bounded domains, so the same is true for φ and $1-|\nabla\varphi|^2>0$.

Set $\alpha := \sqrt{1-|\nabla\varphi|^2}$ and $d:=\varepsilon/\alpha$. Notice that $d=\sqrt{\varepsilon^2+|T(x)-x|^2}$ is a (strictly positive and smoother) approximation of $|T(x)-x|$, and we have $T(x)=x-d\nabla\varphi$.

Main object : The function d is the key point of our analysis : its regularity implies that of T.

A more refined analysis of the cost $c_{(\varepsilon)}$

Since $c_{(\varepsilon)}$ satisfies the twist condition (even better : it is of the form $h(y-x)$ for $h(z)=\sqrt{\varepsilon^2+|z|^2}$, which is strictly convex), we can infer the optimal T from the potential φ : we know that, for $(x_0,y_0)\in\operatorname{spt}(\gamma)$ we have

$$x_0 \in \operatorname{argmin}_x h(x-y_0)-\varphi(x)$$

and hence, at $x=x_0$

$$\nabla\varphi(x)=\nabla h(x-T(x)) \Rightarrow T(x)=x-\frac{\varepsilon}{\sqrt{1-|\nabla\varphi(x)|^2}}\nabla\varphi(x).$$

The function h has Lipschitz constant strictly less than 1 on bounded domains, so the same is true for φ and $1-|\nabla\varphi|^2>0$.

Set $\alpha:=\sqrt{1-|\nabla\varphi|^2}$ and $d:=\varepsilon/\alpha$. Notice that $d=\sqrt{\varepsilon^2+|T(x)-x|^2}$ is a (strictly positive and smoother) approximation of $|T(x)-x|$, and we have $T(x)=x-d\nabla\varphi$.

Main object : The function d is the key point of our analysis : its regularity implies that of T.

Going on with the computations of the derivatives

The condition $x_0 \in \mathrm{argmin}_x\, h(x - T(x_0)) - \varphi(x)$ also implies (with $z_0 = T(x_0) - x_0 = -d\nabla\varphi$)

$$D^2\varphi(x_0) \le D^2 h(z_0) = \frac{\mathrm{I}}{\sqrt{\varepsilon^2 + |z_0|^2}} - \frac{z_0 \otimes z_0}{(\varepsilon^2 + |z_0|^2)^{3/2}} = \frac{1}{d}\left(\mathrm{I} - \nabla\varphi \otimes \nabla\varphi\right).$$

Hence, we can introduce a point-dependent, symmetric, positive definite matrix B such that

$$dD^2\varphi = \mathrm{I} - \nabla\varphi \otimes \nabla\varphi - B.$$

We can also compute

$$\nabla d = -\frac{1}{2}\frac{\varepsilon}{\alpha^3}(-2)D^2\varphi\nabla\varphi = \frac{dD^2\varphi\nabla\varphi}{\alpha^2} = \frac{\nabla\varphi - |\nabla\varphi|^2\nabla\varphi - B\nabla\varphi}{\alpha^2} = \nabla\varphi - \frac{B\nabla\varphi}{\alpha^2}.$$

Warning : Here the dangerous guy is α, which is of order $\varepsilon \to 0$.

Going on with the computations of the derivatives

The condition $x_0 \in \mathrm{argmin}_x\, h(x - T(x_0)) - \varphi(x)$ also implies (with $z_0 = T(x_0) - x_0 = -d\nabla\varphi$)

$$D^2\varphi(x_0) \leq D^2 h(z_0) = \frac{\mathrm{I}}{\sqrt{\varepsilon^2 + |z_0|^2}} - \frac{z_0 \otimes z_0}{(\varepsilon^2 + |z_0|^2)^{3/2}} = \frac{1}{d}\left(\mathrm{I} - \nabla\varphi \otimes \nabla\varphi\right).$$

Hence, we can introduce a point-dependent, symmetric, positive definite matrix B such that

$$dD^2\varphi = \mathrm{I} - \nabla\varphi \otimes \nabla\varphi - B.$$

We can also compute

$$\nabla d = -\frac{1}{2}\frac{\varepsilon}{\alpha^3}(-2)D^2\varphi\nabla\varphi = \frac{dD^2\varphi\nabla\varphi}{\alpha^2} = \frac{\nabla\varphi - |\nabla\varphi|^2\nabla\varphi - B\nabla\varphi}{\alpha^2} = \nabla\varphi - \frac{B\nabla\varphi}{\alpha^2}.$$

Warning : Here the dangerous guy is α, which is of order $\varepsilon \to 0$.

Determinant conditions

If we compute DT we get

$$DT = \mathrm{I} - dD^2\varphi - \nabla\varphi\otimes\nabla d = B + \nabla\varphi\otimes\frac{B\nabla\varphi}{\alpha^2},$$

and hence

$$\det(DT) = \frac{\det B}{\alpha^2}.$$

Suppose for simplicity $g = 1$ (all the computations can also be done for more general g, they are only more complicated) ; the transport condition implies

$$\det(DT) = f \quad \Rightarrow \quad \det B = \alpha^2 f.$$

Idea : the degeneracy of B should compensate the α^{-2} in ∇d.

Denote the inverse matrix B^{-1} by C. Taking the logarithm and differentiating the above equation, using $(\ln(\det B))' = \mathrm{trace}(CB')$ we get

$$C^{ij}B^{ij}_k = (2\ln\alpha)_k + \tilde{f}_k,$$

where $\tilde{f} = \ln f$ and $\nabla(\ln\alpha)$ may be computed thanks to

$$\ln\alpha = -\ln d + \ln\varepsilon \quad \Rightarrow \quad \nabla(\ln\alpha) = -\frac{\nabla d}{d}.$$

Determinant conditions

If we compute DT we get

$$DT = \mathrm{I} - dD^2\varphi - \nabla\varphi \otimes \nabla d = B + \nabla\varphi \otimes \frac{B\nabla\varphi}{\alpha^2},$$

and hence

$$\det(DT) = \frac{\det B}{\alpha^2}.$$

Suppose for simplicity $g = 1$ (all the computations can also be done for more general g, they are only more complicated) ; the transport condition implies

$$\det(DT) = f \quad \Rightarrow \quad \det B = \alpha^2 f.$$

Idea : the degeneracy of B should compensate the α^{-2} in ∇d.
Denote the inverse matrix B^{-1} by C. Taking the logarithm and differentiating the above equation, using $(\ln(\det B))' = \mathrm{trace}(CB')$ we get

$$C^{ij}B^{ij}_k = (2\ln\alpha)_k + \tilde{f}_k,$$

where $\tilde{f} = \ln f$ and $\nabla(\ln\alpha)$ may be computed thanks to

$$\ln\alpha = -\ln d + \ln\varepsilon \quad \Rightarrow \quad \nabla(\ln\alpha) = -\frac{\nabla d}{d}.$$

Uniform estimates

on B, etc. . .

Pogorelov-MTW method

At a first step, we need at least to investigate bounds on the matrix B. Consider $\overline{B} := \mathrm{trace}(B)$.

Following the Pogorelov-type method used by MTW, we can we get some bounds on some quantities $\overline{B} \geq 0$ in the following way :

- fix a cutoff function η with $\eta = 0$ on $\partial\Omega$ and look for $\max \overline{B}\eta^2$
- the maximum is attained in the interior, and we can write $\frac{\overline{B}_i}{\overline{B}} + 2\frac{\eta_i}{\eta} = 0$; $\left(\frac{\overline{B}_{ij}}{\overline{B}} + 2\frac{\eta_{ij}}{\eta} - 6\frac{\eta_i\eta_j}{\eta^2}\right)$ is a negative definite matrix
- multiply the last matrix times C, take the trace, and get $C^{ij}\overline{B}_{ij} \leq \dots$
- use $B_{ij}^{kk} = B_{kk}^{ij}$+other terms, differentiate twice the determinant condition to control $C^{ij}B_{kk}^{ij}$ (supposing $f \in C^{1,1}$)
- use these computations to show something like
$$C^{ij}\overline{B}_{ij} \geq M\frac{\overline{B}^2}{\alpha^2} + O\left(\frac{\overline{B}^2}{\alpha^2}\right),$$
where $M := dC^{ij}\varphi_{ij}$ is such that $\lim_{\overline{B}\to\infty} M = +\infty$
- conclude that $\overline{B}$ is actually uniformly bounded on $\{\eta = 1\}$.

Pogorelov-MTW method

At a first step, we need at least to investigate bounds on the matrix B. Consider $\overline{B} := \mathrm{trace}(B)$.
Following the Pogorelov-type method used by MTW, we can we get some bounds on some quantities $\overline{B} \geq 0$ in the following way :

- fix a cutoff function η with $\eta = 0$ on $\partial\Omega$ and look for $\max \overline{B}\eta^2$
- the maximum is attained in the interior, and we can write
$$\frac{\overline{B}_i}{\overline{B}} + 2\frac{\eta_i}{\eta} = 0 \; ; \quad \left(\frac{\overline{B}_{ij}}{\overline{B}} + 2\frac{\eta_{ij}}{\eta} - 6\frac{\eta_i \eta_j}{\eta^2} \right) \text{ is a negative definite matrix}$$
- multiply the last matrix times C, take the trace, and get $C^{ij}\overline{B}_{ij} \leq \dots$
- use $B_{ij}^{kk} = B_{kk}^{ij}$+other terms, differentiate twice the determinant condition to control $C^{ij} B_{kk}^{ij}$ (supposing $f \in C^{1,1}$)
- use these computations to show something like
$$C^{ij}\overline{B}_{ij} \geq M\frac{\overline{B}^2}{\alpha^2} + O\left(\frac{\overline{B}^2}{\alpha^2}\right),$$
where $M := dC^{ij}\varphi_{ij}$ is such that $\lim_{\overline{B}\to\infty} M = +\infty$
- conclude that $\overline{B}$ is actually uniformly bounded on $\{\eta = 1\}$.

Pogorelov-MTW method

At a first step, we need at least to investigate bounds on the matrix B. Consider $\overline{B} := \mathrm{trace}(B)$.
Following the Pogorelov-type method used by MTW, we can we get some bounds on some quantities $\overline{B} \geq 0$ in the following way :

- fix a cutoff function η with $\eta = 0$ on $\partial\Omega$ and look for $\max \overline{B}\eta^2$
- the maximum is attained in the interior, and we can write $\frac{\overline{B}_i}{\overline{B}} + 2\frac{\eta_i}{\eta} = 0$; $\left(\frac{\overline{B}_{ij}}{\overline{B}} + 2\frac{\eta_{ij}}{\eta} - 6\frac{\eta_i\eta_j}{\eta^2}\right)$ is a negative definite matrix
- multiply the last matrix times C, take the trace, and get $C^{ij}\overline{B}_{ij} \leq \dots$
- use $B^{kk}_{ij} = B^{ij}_{kk}$+other terms, differentiate twice the determinant condition to control $C^{ij}B^{ij}_{kk}$ (supposing $f \in C^{1,1}$)
- use these computations to show something like

$$C^{ij}\overline{B}_{ij} \geq M\frac{\overline{B}^2}{\alpha^2} + O(\frac{\overline{B}^2}{\alpha^2}),$$

where $M := dC^{ij}\varphi_{ij}$ is such that $\lim_{\overline{B}\to\infty} M = +\infty$
- conclude that $\overline{B}$ is actually uniformly bounded on $\{\eta = 1\}$

Pogorelov-MTW method

At a first step, we need at least to investigate bounds on the matrix B. Consider $\overline{B} := \mathrm{trace}(B)$.
Following the Pogorelov-type method used by MTW, we can we get some bounds on some quantities $\overline{B} \geq 0$ in the following way :

- fix a cutoff function η with $\eta = 0$ on $\partial\Omega$ and look for $\max \overline{B}\eta^2$
- the maximum is attained in the interior, and we can write $\frac{\overline{B}_i}{\overline{B}} + 2\frac{\eta_i}{\eta} = 0$; $\left(\frac{\overline{B}_{ij}}{\overline{B}} + 2\frac{\eta_{ij}}{\eta} - 6\frac{\eta_i\eta_j}{\eta^2}\right)$ is a negative definite matrix
- multiply the last matrix times C, take the trace, and get $C^{ij}\overline{B}_{ij} \leq \dots$
- use $B^{kk}_{ij} = B^{ij}_{kk}$+other terms, differentiate twice the determinant condition to control $C^{ij}B^{ij}_{kk}$ (supposing $f \in C^{1,1}$)
- use these computations to show something like
$$C^{ij}\overline{B}_{ij} \geq M\frac{\overline{B}^2}{\alpha^2} + O\left(\frac{\overline{B}^2}{\alpha^2}\right),$$
where $M := dC^{ij}\varphi_{ij}$ is such that $\lim_{\overline{B}\to\infty} M = +\infty$
- conclude that $\overline{B}$ is actually uniformly bounded on $\{\eta = 1\}$

Pogorelov-MTW method

At a first step, we need at least to investigate bounds on the matrix B. Consider $\overline{B} := \mathrm{trace}(B)$.
Following the Pogorelov-type method used by MTW, we can we get some bounds on some quantities $\overline{B} \geq 0$ in the following way :

- fix a cutoff function η with $\eta = 0$ on $\partial\Omega$ and look for $\max \overline{B}\eta^2$
- the maximum is attained in the interior, and we can write $\frac{\overline{B}_i}{\overline{B}}+2\frac{\eta_i}{\eta} = 0$; $\left(\frac{\overline{B}_{ij}}{\overline{B}} + 2\frac{\eta_{ij}}{\eta} - 6\frac{\eta_i\eta_j}{\eta^2}\right)$ is a negative definite matrix
- multiply the last matrix times C, take the trace, and get $C^{ij}\overline{B}_{ij} \leq \dots$
- use $B^{kk}_{ij} = B^{ij}_{kk}$+other terms, differentiate twice the determinant condition to control $C^{ij}B^{ij}_{kk}$ (supposing $f \in C^{1,1}$)
- use these computations to show something like

$$C^{ij}\overline{B}_{ij} \geq M\frac{\overline{B}^2}{\alpha^2} + O(\frac{\overline{B}^2}{\alpha^2}),$$

where $M := dC^{ij}\varphi_{ij}$ is such that $\lim_{\overline{B}\to\infty} M = +\infty$
- conclude that $\overline{B}$ is actually uniformly bounded on $\{\eta = 1\}$

Pogorelov-MTW method

At a first step, we need at least to investigate bounds on the matrix B. Consider $\overline{B} := \mathrm{trace}(B)$.

Following the Pogorelov-type method used by MTW, we can we get some bounds on some quantities $\overline{B} \geq 0$ in the following way :

- fix a cutoff function η with $\eta = 0$ on $\partial\Omega$ and look for $\max \overline{B}\eta^2$
- the maximum is attained in the interior, and we can write $\frac{\overline{B}_i}{\overline{B}} + 2\frac{\eta_i}{\eta} = 0$; $\left(\frac{\overline{B}_{ij}}{\overline{B}} + 2\frac{\eta_{ij}}{\eta} - 6\frac{\eta_i\eta_j}{\eta^2}\right)$ is a negative definite matrix
- multiply the last matrix times C, take the trace, and get $C^{ij}\overline{B}_{ij} \leq \dots$
- use $B_{ij}^{kk} = B_{kk}^{ij}$+other terms, differentiate twice the determinant condition to control $C^{ij}B_{kk}^{ij}$ (supposing $f \in C^{1,1}$)
- use these computations to show something like

$$C^{ij}\overline{B}_{ij} \geq M\frac{\overline{B}^2}{\alpha^2} + O\left(\frac{\overline{B}^2}{\alpha^2}\right),$$

  where $M := dC^{ij}\varphi_{ij}$ is such that $\lim_{\overline{B}\to\infty} M = +\infty$
- conclude that $\overline{B}$ is actually uniformly bounded on $\{\eta = 1\}$

Pogorelov-MTW method

At a first step, we need at least to investigate bounds on the matrix B. Consider $\overline{B} := \mathrm{trace}(B)$.
Following the Pogorelov-type method used by MTW, we can we get some bounds on some quantities $\overline{B} \geq 0$ in the following way :

- fix a cutoff function η with $\eta = 0$ on $\partial\Omega$ and look for $\max \overline{B}\eta^2$
- the maximum is attained in the interior, and we can write $\frac{\overline{B}_i}{\overline{B}} + 2\frac{\eta_i}{\eta} = 0$; $\left(\frac{\overline{B}_{ij}}{\overline{B}} + 2\frac{\eta_{ij}}{\eta} - 6\frac{\eta_i\eta_j}{\eta^2}\right)$ is a negative definite matrix
- multiply the last matrix times C, take the trace, and get $C^{ij}\overline{B}_{ij} \leq \dots$
- use $B^{kk}_{ij} = B^{ij}_{kk}$+other terms, differentiate twice the determinant condition to control $C^{ij}B^{ij}_{kk}$ (supposing $f \in C^{1,1}$)
- use these computations to show something like
$$C^{ij}\overline{B}_{ij} \geq M\frac{\overline{B}^2}{\alpha^2} + O(\frac{\overline{B}^2}{\alpha^2}),$$
where $M := dC^{ij}\varphi_{ij}$ is such that $\lim_{\overline{B}\to\infty} M = +\infty$
- conclude that $\overline{B}$ is actually uniformly bounded on $\{\eta = 1\}$

After $\overline{B}$, a new estimate

We have got uniform bounds on $\overline{B}$, and hence on B and $dD^2\varphi$. We know $\det B \approx \alpha^2$; to check how the eigenvalues of B behave, we consider now

$$G := \frac{\nabla\varphi \cdot B\nabla\varphi}{\alpha^2} \geq 0$$

and we try to apply the same method to G.

Up to the trick of considering instead $\overline{B} + G$, it works, and the same kind of computations show a uniform local bound on G.

Structure of $\overline{B}$ and $\nabla\varphi$: This implies that B has one eigenvalue of the order of α^2 and all the others are of the order of 1 and that $\nabla\varphi$ is "almost" directed as the eigenvector e_1 corresponding to the first one :

$$\nabla\varphi = \sum_i \varphi_i e_i \quad \text{with} \quad |1-\varphi_1| \leq C\alpha^2, \; |\varphi_i| \leq C\alpha \;\; \text{for } i > 1$$

After $\overline{B}$, a new estimate

We have got uniform bounds on $\overline{B}$, and hence on B and $dD^2\varphi$. We know $\det B \approx \alpha^2$; to check how the eigenvalues of B behave, we consider now

$$G := \frac{\nabla\varphi \cdot B\nabla\varphi}{\alpha^2} \geq 0$$

and we try to apply the same method to G.
Up to the trick of considering instead $\overline{B} + G$, it works, and the same kind of computations show a uniform local bound on G.

Structure of $\overline{B}$ and $\nabla\varphi$: This implies that B has one eigenvalue of the order of α^2 and all the others are of the order of 1 and that $\nabla\varphi$ is "almost" directed as the eigenvector e_1 corresponding to the first one :

$$\nabla\varphi = \sum_i \varphi_i e_i \quad \text{with} \quad |1-\varphi_1| \leq C\alpha^2, \ |\varphi_i| \leq C\alpha \ \text{ for } i > 1$$

After $\overline{B}$, a new estimate

We have got uniform bounds on $\overline{B}$, and hence on B and $dD^2\varphi$. We know $\det B \approx \alpha^2$; to check how the eigenvalues of B behave, we consider now

$$G := \frac{\nabla\varphi \cdot B\nabla\varphi}{\alpha^2} \geq 0$$

and we try to apply the same method to G.
Up to the trick of considering instead $\overline{B} + G$, it works, and the same kind of computations show a uniform local bound on G.
Structure of $\overline{B}$ and $\nabla\varphi$: This implies that B has one eigenvalue of the order of α^2 and all the others are of the order of 1 and that $\nabla\varphi$ is "almost" directed as the eigenvector e_1 corresponding to the first one :

$$\nabla\varphi = \sum_i \varphi_i e_i \quad \text{with} \quad |1-\varphi_1| \leq C\alpha^2,\ |\varphi_i| \leq C\alpha \ \text{ for } i > 1$$

Useless and informal applications of this bound

Bounds on $\overline{B}$: we infer that $dD^2\varphi$ is uniformly bounded. At the limit, this means that the Kantorovich potential is $C^{1,1}$ on points x such that $|T(x)-x| \geq \delta > 0$, i.e. on particular points in the interior of the transport rays.

Definition and bounds on G : compute $\nabla d \cdot \nabla\varphi$: we have

$$\nabla d \cdot \nabla\varphi = |\nabla\varphi|^2 - G \in [-\max G, 1].$$

The upper bound stands for monotonicity of the transport along transport rays (at the limit, the transport follows the direction of $-\nabla\varphi$ and $T' = d' + 1$). The lower bound stands for Lipschitz behavior of T along the same rays.

Notice that this is not completely trivial, since we do not transport f onto g (which are supposed to be bounded above and below) but fJ onto gJ, where J is a Jacobian coefficient depending on transport rays, which could vanish.

Consequences on ∇d : the information above slightly improve the expected rate of explosion of ∇d, which is actually bounded by α^{-1}

Useless and informal applications of this bound

Bounds on $\overline{B}$: we infer that $dD^2\varphi$ is uniformly bounded. At the limit, this means that the Kantorovich potential is $C^{1,1}$ on points x such that $|T(x)-x| \geq \delta > 0$, i.e. on particular points in the interior of the transport rays.

Definition and bounds on G : compute $\nabla d \cdot \nabla\varphi$: we have

$$\nabla d \cdot \nabla\varphi = |\nabla\varphi|^2 - G \in [-\max G, 1].$$

The upper bound stands for monotonicity of the transport along transport rays (at the limit, the transport follows the direction of $-\nabla\varphi$ and $T' = d' + 1$). The lower bound stands for Lipschitz behavior of T along the same rays.

Notice that this is not completely trivial, since we do not transport f onto g (which are supposed to be bounded above and below) but fJ onto gJ, where J is a Jacobian coefficient depending on transport rays, which could vanish.

Consequences on ∇d : the information above slightly improve the expected rate of explosion of ∇d, which is actually bounded by α^{-1}.

Useless and informal applications of this bound

Bounds on $\overline{B}$: we infer that $dD^2\varphi$ is uniformly bounded. At the limit, this means that the Kantorovich potential is $C^{1,1}$ on points x such that $|T(x)-x| \geq \delta > 0$, i.e. on particular points in the interior of the transport rays.

Definition and bounds on G : compute $\nabla d \cdot \nabla\varphi$: we have

$$\nabla d \cdot \nabla\varphi = |\nabla\varphi|^2 - G \in [-\max G, 1].$$

The upper bound stands for monotonicity of the transport along transport rays (at the limit, the transport follows the direction of $-\nabla\varphi$ and $T' = d' + 1$). The lower bound stands for Lipschitz behavior of T along the same rays.

Notice that this is not completely trivial, since we do not transport f onto g (which are supposed to be bounded above and below) but fJ onto gJ, where J is a Jacobian coefficient depending on transport rays, which could vanish.

Consequences on ∇d : the information above slightly improve the expected rate of explosion of ∇d, which is actually bounded by α^{-1}.

Useless and informal applications of this bound

Bounds on $\overline{B}$: we infer that $dD^2\varphi$ is uniformly bounded. At the limit, this means that the Kantorovich potential is $C^{1,1}$ on points x such that $|T(x)-x| \geq \delta > 0$, i.e. on particular points in the interior of the transport rays.

Definition and bounds on G : compute $\nabla d \cdot \nabla \varphi$: we have

$$\nabla d \cdot \nabla \varphi = |\nabla \varphi|^2 - G \in [-\max G, 1].$$

The upper bound stands for monotonicity of the transport along transport rays (at the limit, the transport follows the direction of $-\nabla\varphi$ and $T' = d' + 1$). The lower bound stands for Lipschitz behavior of T along the same rays.

Notice that this is not completely trivial, since we do not transport f onto g (which are supposed to be bounded above and below) but fJ onto gJ, where J is a Jacobian coefficient depending on transport rays, which could vanish.

Consequences on ∇d : the information above slightly improve the expected rate of explosion of ∇d, which is actually bounded by α^{-1}.

Useless and informal applications of this bound

Question

How to pass to the limit the information on $\nabla d \cdot \nabla \varphi$?

The easiest way to answer would pass through a modification of the classical *div-curl lemma* to include the case we are facing :
$v_\varepsilon := \nabla d$ is bounded in $W^{-1,\infty}$,
$w_\varepsilon := \nabla \varphi$ is almost bounded in $W^{1,\infty}$ (up to the multiplication times d),
$\nabla \times v_\varepsilon = 0$,
$\nabla \cdot w_\varepsilon$ is almost bounded in L^∞ (again, up to the multiplication times d)... (discussion in progress with F. Murat)

Otherwise, one should finely estimate how much the integral curves of $-\nabla \varphi$ deviate w.r.t. the transport rays and compare this rate with the bounds on ∇d.

Another kind of quantified estimate

For fixed $\varepsilon > 0$, the transport $T = x - d\nabla\varphi$ is a diffeomorphism of Ω (if Ω is strictly convex and smooth). In particular, it sends the boundary onto the boundary.

Let us consider the case $\Omega = B(0,1)$. We have

$$\forall x_0 \; : \; |x_0| = 1, \quad |x_0 - d\nabla\varphi|^2 = 1 \;\Rightarrow\; d\nabla\varphi \cdot x_0 = \frac{1}{2}d^2|\nabla\varphi|^2$$

and

$$\forall x_0 \; : \; |x_0| = 1; \quad x \mapsto |T(x)|^2 \; \text{ is maximal at } x_0,$$

which implies

$$T(x_0) \cdot DT(x_0) = \lambda x_0 \; \text{ for some } \lambda \geq 0,$$

and in turn gives

$$B\left(x_0 - d\nabla\varphi\frac{1+\alpha^2}{2\alpha^2}\right) = \lambda x_0$$

and finally

$$\alpha \geq \sqrt[4]{\frac{\varepsilon^2}{4+\varepsilon^2}} \approx \frac{\sqrt{\varepsilon}}{2}, \quad d = \frac{\varepsilon}{\alpha} \leq \sqrt[4]{\varepsilon^2(4+\varepsilon^2)} \approx 2\sqrt{\varepsilon},$$

which gives a quantified bound on $d \to 0$ on $\partial\Omega$.

Another kind of quantified estimate

For fixed $\varepsilon > 0$, the transport $T = x - d\nabla\varphi$ is a diffeomorphism of Ω (if Ω is strictly convex and smooth). In particular, it sends the boundary onto the boundary.

Let us consider the case $\Omega = B(0,1)$. We have

$$\forall x_0 \,:\, |x_0| = 1, \quad |x_0 - d\nabla\varphi|^2 = 1 \;\Rightarrow\; d\nabla\varphi \cdot x_0 = \frac{1}{2} d^2 |\nabla\varphi|^2$$

and

$$\forall x_0 \,:\, |x_0| = 1; \quad x \mapsto |T(x)|^2 \text{ is maximal at } x_0,$$

which implies

$$T(x_0) \cdot DT(x_0) = \lambda x_0 \text{ for some } \lambda \geq 0,$$

and in turn gives

$$B\left(x_0 - d\nabla\varphi \frac{1+\alpha^2}{2\alpha^2}\right) = \lambda x_0$$

and finally

$$\alpha \geq \sqrt[4]{\frac{\varepsilon^2}{4+\varepsilon^2}} \approx \frac{\sqrt{\varepsilon}}{2}, \quad d = \frac{\varepsilon}{\alpha} \leq \sqrt[4]{\varepsilon^2(4+\varepsilon^2)} \approx 2\sqrt{\varepsilon},$$

which gives a quantified bound on $d \to 0$ on $\partial\Omega$.

Another kind of quantified estimate

For fixed $\varepsilon > 0$, the transport $T = x - d\nabla\varphi$ is a diffeomorphism of Ω (if Ω is strictly convex and smooth). In particular, it sends the boundary onto the boundary.

Let us consider the case $\Omega = B(0,1)$. We have

$$\forall x_0 \ : \ |x_0| = 1, \quad |x_0 - d\nabla\varphi|^2 = 1 \ \Rightarrow \ d\nabla\varphi \cdot x_0 = \frac{1}{2} d^2 |\nabla\varphi|^2$$

and

$$\forall x_0 \ : \ |x_0| = 1; \quad x \mapsto |T(x)|^2 \text{ is maximal at } x_0,$$

which implies

$$T(x_0) \cdot DT(x_0) = \lambda x_0 \text{ for some } \lambda \geq 0,$$

and in turn gives

$$B\left(x_0 - d\nabla\varphi \frac{1+\alpha^2}{2\alpha^2}\right) = \lambda x_0$$

and finally

$$\alpha \geq \sqrt[4]{\frac{\varepsilon^2}{4+\varepsilon^2}} \approx \frac{\sqrt{\varepsilon}}{2}, \quad d = \frac{\varepsilon}{\alpha} \leq \sqrt[4]{\varepsilon^2(4+\varepsilon^2)} \approx 2\sqrt{\varepsilon},$$

which gives a quantified bound on $d \to 0$ on $\partial\Omega$.

How to go on

if we want at least some continuity for d

Local Lipschitz or Hölder continuity of d

Lipschitz bounds on d would be natural under the assumption $f \in C^{0,1}$.

Yet, there is no reason to think that $\nabla\varphi$ and $B\nabla\varphi$ could not be orthogonal or almost orthogonal, so that the bound on G is not enough to bound $|\nabla d|$.

A Pogorelov estimate on $|\nabla d|^2$ would imply terms with B^2, which are difficult to handle with C^{ij} only.

Hence, here are some alternative ideas, based on some elliptic PDEs to get weaker continuity bounds on d.

Notice that for the moment even $f, g \in C^\infty \Rightarrow T \in C^0$ is not known at all !

Two elliptic PDEs

First idea : compute Δd (or other divergence-form operators on d) and hope to use some known results to get $C^{0,\alpha}$ estimates. This is morally third-order in φ.

$$d\Delta d = 3|\nabla d|^2 + \frac{-d\nabla\varphi\cdot\nabla\overline{B} + G(n-1-\overline{B}) - \overline{B} + \mathrm{trace}(B^2)}{\alpha^2} + \text{bdd terms.}$$

Second idea : write an equation with an operator $\nabla\cdot F(\nabla\varphi)$ and look for regularity of $F(\nabla\varphi)$ itself. This is only second-order in φ.

$$\nabla\cdot\left(\frac{\varepsilon\nabla\varphi}{\sqrt{1-|\nabla\varphi|^2}}\right) = \nabla\cdot(d\nabla\varphi) = d\Delta\varphi + \nabla d\cdot\nabla\varphi = n - \overline{B} - G \in L^\infty_{loc}.$$

Two elliptic PDEs

First idea : compute Δd (or other divergence-form operators on d) and hope to use some known results to get $C^{0,\alpha}$ estimates. This is morally third-order in φ.

$$d\Delta d = 3|\nabla d|^2 + \frac{-d\nabla\varphi\cdot\nabla\overline{B} + G(n-1-\overline{B}) - \overline{B} + \mathrm{trace}(B^2)}{\alpha^2} + \text{bdd terms.}$$

Second idea : write an equation with an operator $\nabla\cdot F(\nabla\varphi)$ and look for regularity of $F(\nabla\varphi)$ itself. This is only second-order in φ.

$$\nabla\cdot\left(\frac{\varepsilon\nabla\varphi}{\sqrt{1-|\nabla\varphi|^2}}\right) = \nabla\cdot(d\nabla\varphi) = d\Delta\varphi + \nabla d\cdot\nabla\varphi = n - \overline{B} - G \in L^\infty_{loc}.$$

Two elliptic PDEs

First idea

The general $C^{0,\alpha}$ theory for non-linear elliptic equations allows terms of the order of the square of the gradient at the right hand side.

Questions :

- how to get rid of the factor d at the left-hand side ? reasonably one should get estimates on $(d-\delta)_+$, which would be enough to get $d \in C^0$, but this seems delicate
- how the handle the term with α^2 at the denominator
  - the term with $\nabla\varphi \cdot \nabla\overline{B}$ should be dealt with separately, since probably $\overline{B}$ can have some extra continuity properties *in the direction of* $\nabla\varphi$
  - the other terms are bounded, thus getting something of the form $\frac{c}{\alpha^2}$: notice that this is the same estimated order of $|\nabla d|^2$ *in the worst possible scenario*. Is it possible to get continuity bounds from it ?

Two elliptic PDEs

Second idea

Many recent results have been done on the regularity of $F(\nabla\varphi)$ for solutions of degenerate PDEs like $\nabla\cdot F(\nabla\varphi)=h$, and in particular C^0 results (Brasco, Carlier, S., Vespri, Colombo, Figalli...).

Yet, F is degenerate here, but has no uniform continuity bounds w.r.t. $\nabla\varphi$. Hence, the usual strategy : proving that φ is smooth out of the degeneracy region and deducing continuity results for $F(\nabla\varphi)$ does not work.

Duality : if $F=\nabla H$ then $v=F(\nabla\varphi)$ is the solution of $\min\int H^*(v)\ :\ \nabla\cdot v=h$, and here $H^*(z)=\sqrt{\varepsilon^2+|z|^2}$. As $\varepsilon\to 0$, we tend to the minimization of the L^1 norm with prescribed divergence, as for the transport density. And it is well known that $h\in L^\infty\Rightarrow v\in L^\infty$ but not $v\in C^0$.

But maybe h is better than L^∞, and we could consider studying this operator under stronger assumptions on h.

Simplification in the applications to the transport density

Let us stress that if the goal is studying the continuity of the optimal transport in order to apply it to the regularity of the transport density, we can choose the very simplest case :

- σ only depends on $f - g$, so we can add a common density to both measures (and rescale their masses to 1) and guarantee
  - uniform lower bounds on both densities
  - $g = 1$
  - f as close to 1 as we want
- if f, g are smooth and compactly supported in $\mathbb{R}^n$, Ω can be chosen to be any domain containing their supports ; in particular we can take it convex, and even $\Omega = B(0, R)$
- we do not need to care about the optimal transport which is selected as $\varepsilon \to 0$ (whether it is T_{mon} or not), since anyway σ does not depend on the choice of the optimizer.

The end of the talk

Thanks for your attention

and sorry for so many questions with so few answers