Optimal transport between mutually singular measures

Giuseppe Buttazzo
Dipartimento di Matematica
Università di Pisa
buttazzo@dm.unipi.it
http://cvgmt.sns.it

"VIII Partial differential equations, optimal design and numerics"
Benasque (Spain), August 18–30, 2019

Research jointly made with

Guillaume Carlier
CEREMADE
Université Paris Dauphine, Paris

Maxime Laborde
Department of Mathematics and Statistics
McGill University, Montreal

G. Buttazzo, G. Carlier, M. Laborde: *On the Wasserstein distance between mutually singular measures.* Adv. Calc. Var., (2019).

In a metric space (X, d) for a given $x_0 \in X$ and $\mathcal{A} \subset X$ we denote

$$\text{dist}(x_0, \mathcal{A}) = \inf \Big\{ d(x_0, y) \ : \ y \in \mathcal{A} \Big\}.$$

We consider:

- X the space of Borel probabilities μ on $\mathbb{R}^d$ with a finite p-moment ($1 \le p < +\infty$ is fixed), metrized by a Wasserstein distance;

- $x_0 = \mu$ an element of X;

- $\mathcal{A}$ a subset of X consisting of measures which are singular with respect to μ.

In other words, we consider the problem

$$W(\mu, \mathcal{A}) = \inf \Big\{ W(\mu, \nu) \ : \ \nu \in \mathcal{A} \Big\}$$

where W denotes the p-Wasserstein distance and $\mathcal{A}$ is a suitable class of Borel probabilities ν that are singular with respect to μ, that is μ is concentrated on E and ν is concentrated on $\mathbb{R}^d \setminus E$ for a suitable set E. We will discuss several possible choices for the admissible class $\mathcal{A}$.

$$W(\mu, \nu) = \min \left(\int_{\mathbb{R}^d \times \mathbb{R}^d} |x - y|^p d\gamma \ : \ \gamma \in P(\mu, \nu) \right)^{1/p}$$

$P(\mu, \nu) =$ transport plans between μ and ν.

Problems of this kind arise in some models of biological bilayer membranes, for which we refer for instance to:

M.A. Peletier, M. Röger: Arch. Rational Mech. Anal. (2009)

L. Lussardi, M.A. Peletier, M. Röger: J. Fixed Point Theory Appl. (2014)

and references therein. Here we consider only the mathematical issues, that appear to be very rich, together with some numerical simulations.

About the terminology:

• μ is concentrated on A if $\mu(E) = \mu(A \cap E)$ for every E.

• μ and ν are singular if there exists a set A such that $\mu(A) = \nu(A^c) = 0$; notation $\mu \perp \nu$.

• μ is absolutely continuous with respect to ν if $\nu(E) = 0$ implies $\mu(E) = 0$; notation $\mu \ll \nu$. By Radon-Nikodym theorem this is equivalent to say that $\mu \in L^1(\nu)$.

The first remark is that, if we consider the entire class of all measures which are singular with respect to μ

$$\mathcal{A} = \mu^{\perp} = \left\{\nu \in \mathcal{P} \ : \ \nu \perp \mu\right\}$$

the optimization problem above is indeed trivial, as the following proposition shows.

Proposition. *For every $\mu \in \mathcal{P}$ we have*

$$W(\mu, \mu^{\perp}) = \inf \left\{W(\mu, \nu) \ : \ \nu \perp \mu\right\} = 0.$$

Proof. This is easy when $\mu \in L^1(\mathbb{R}^d)$; it is enough to take a sequence of probabilities

$$\nu_n = \sum_{k \in \mathbf{N}} a_{n,k} \delta_{x_k}$$

with x_k suitably chosen in a way that $\nu_n \xrightarrow{*} \mu$. For a general μ the x_k have to be taken in $\mathbb{R}^d \backslash N$ where N is the Lebesgue null set where the singular part μ^s of μ is concentrated.

In the next step, we consider the class

$$\mathcal{A} = \mu^\perp \cap L^1 = \left\{ \nu \perp \mu, \ \nu \ll \mathcal{L}^d \right\}$$

where $\mathcal{L}^d$ is the Lebesgue measure.

Proposition. *For every* $\mu \perp \mathcal{L}^d$ *we have* $W(\mu, \mu^\perp \cap L^1) = 0$.

Proof. Let N be a $\mathcal{L}^d$-negligible set where μ is concentrated. We can find

$$\nu_n = \sum_{k \in \mathbf{N}} a_{n,k} \delta_{x_k}$$

with $x_k \in \mathbb{R}^d \setminus N$ suitably chosen in a way that $\nu_n \overset{*}{\rightharpoonup} \mu$. It is now enough to take $\rho_n \in L^1$ such that $W(\nu_n, \rho_n) \leq 1/n$ and we have $\rho_n \perp \mu$ with

$$W(\mu, \rho_n) \leq W(\mu, \nu_n) + W(\nu_n, \rho_n) \to 0.$$

The situation becomes more interesting when μ is not singular with respect to $\mathcal{L}^d$, that is the absolutely continuous part μ_a of μ with respect to the Lebesgue measure does not vanish.

Proposition. *Assume there exists* $\delta > 0$ *such that the set* $\{\mu_a \geq \delta\}$ *contains an open set. Then* $W(\mu, \mu^\perp \cap L^1) > 0$.

Proof. Let A be an open set contained in $\{\mu_a \geq \delta\}$ and assume by contradiction that there exists $\rho_n \in \mu^\perp \cap L^1$ with $\rho_n \stackrel{*}{\rightharpoonup} \mu$. Then we have

$$\int_A \rho_n \, dx \leq \frac{1}{\delta} \int_{\{\mu_a \geq \delta\}} \rho_n \mu_a \, dx \leq \frac{1}{\delta} \int \rho_n \mu_a \, dx = 0$$

which is impossible, since

$$\delta \mathcal{L}^d(A) \leq \mu_a(A) \leq \mu(A) \leq \liminf_n \int_A \rho_n \, dx.$$

Corollary. *If μ_a is (or coincides a.e. with) a lower semicontinuous function and $\mu_a \neq 0$, then $W(\mu, \mu^\perp \cap L^1) > 0$.*

The quantity $W(\mu, \mu^{\perp} \cap L^1)$ can be better characterized in some cases.

Theorem. *For every $\mu \in \mathcal{P}$ there exists $\nu \in \mathcal{P}$ such that*

$$W(\mu, \mu^{\perp} \cap L^1) = W(\mu, \nu).$$

If $S = \operatorname{spt} \mu$ is Lipschitz the measure ν is concentrated on ∂S. Moreover, if $\mu \in L^1$ we have that ν is unique and given by

$$\nu = T^{\#}\mu \quad \textit{where} \quad T = id - d_{\partial S} \nabla d_{\partial S}$$

being $d_{\partial S}$ is the distance function to ∂S.

Proof. (sketch). Let $\rho_n \in \mu^\perp \cap L^1$ be such that

$$W(\mu, \mu^\perp \cap L^1) = \lim_n W(\mu, \rho_n);$$

we may assume that (up to a subsequence) $\rho_n \to \nu$ weakly*, so that

$$W(\mu, \mu^\perp \cap L^1) = W(\mu, \nu).$$

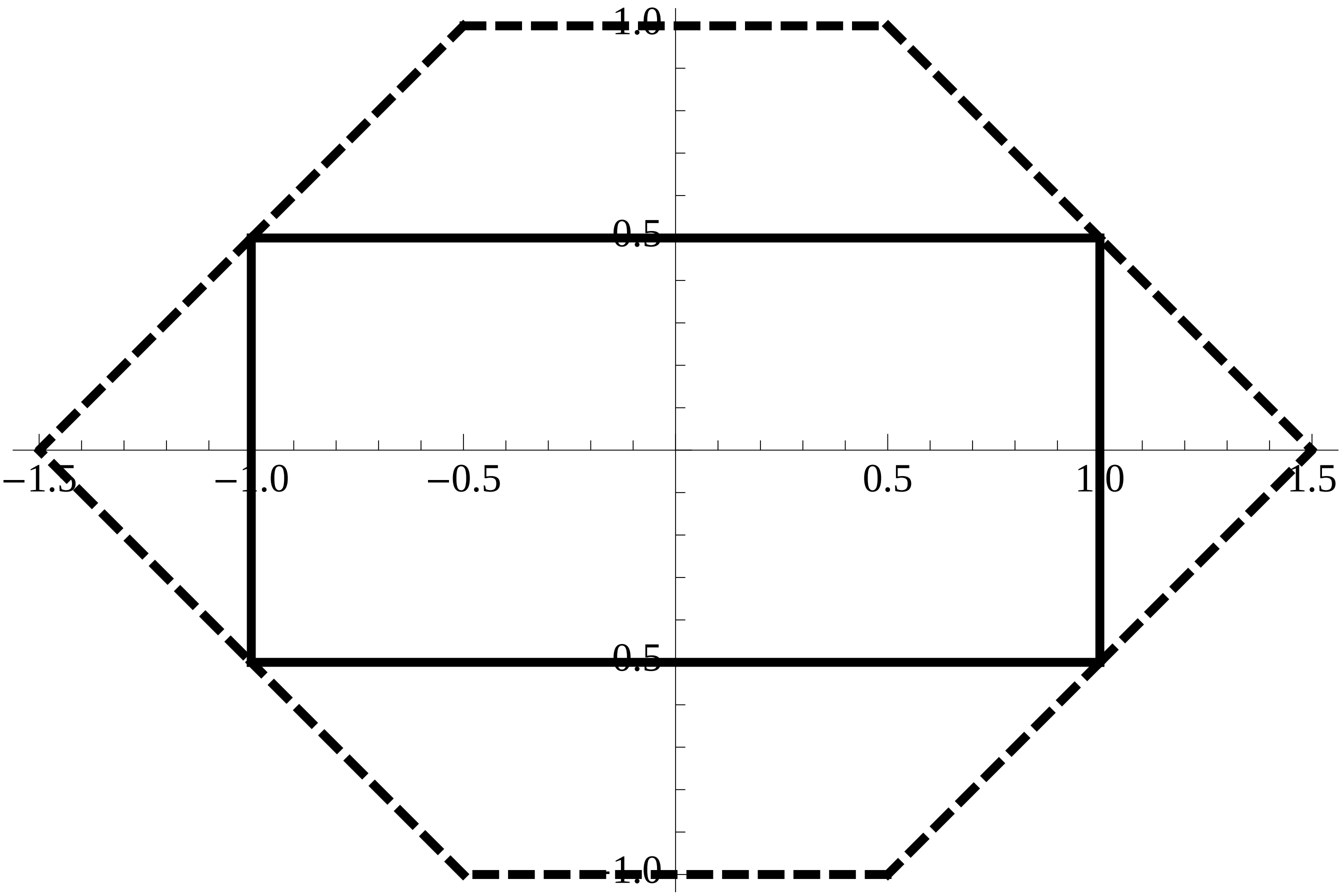
Heuristically, in order to achieve the minimal Wasserstein distance, every point on S has to be transported out of S in the shortest way. Hence ν has to be concentrated on ∂S.

In addition, if $\mu \in L^1$ the transport is achieved by a transport map T which maps a.e. point x into the closest point of ∂S following the direction of $\nabla d_{\partial S}$, hence

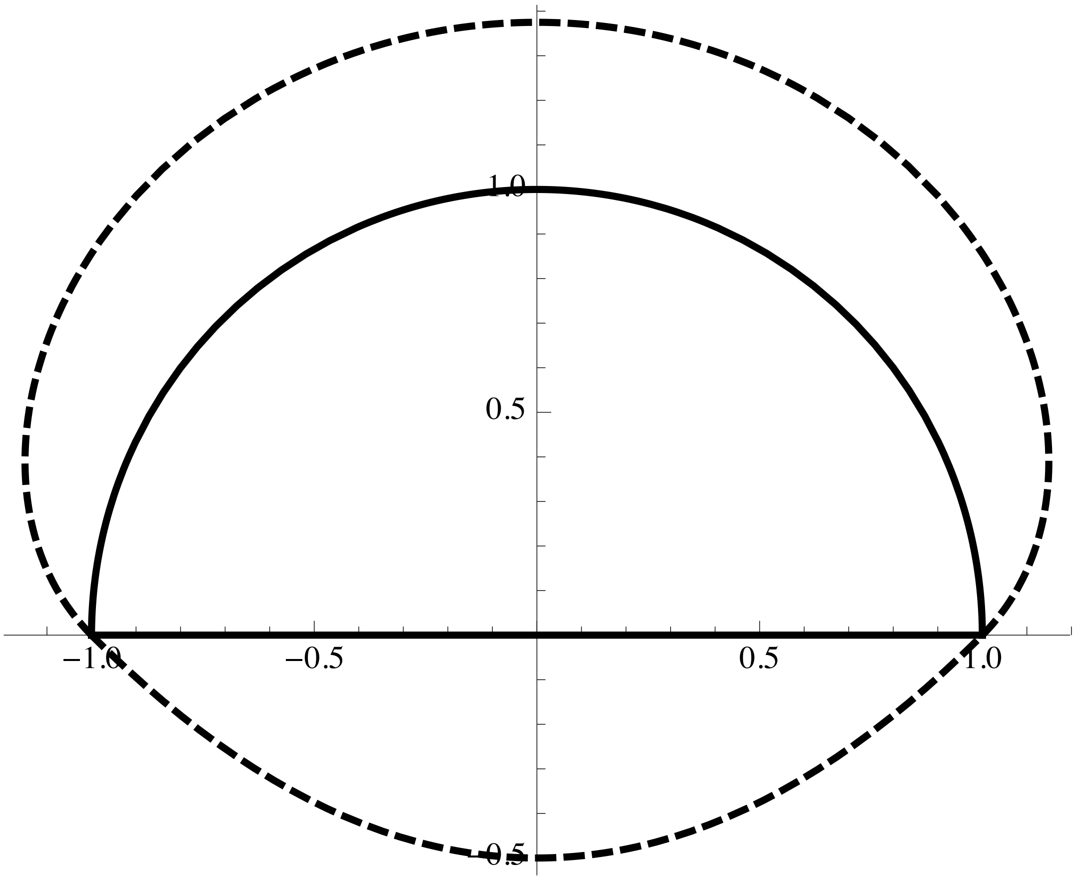
$$T(x) = x - d_{\partial S}(x)\nabla d_{\partial S}(x).$$

If μ is singular no uniqueness, for instance if $\mu = \mathcal{L}^d \lfloor B + \delta_0$ any $\nu = \frac{1}{d}\mathcal{H}^{d-1}\lfloor \partial B + \eta$ is good, for every probability η on ∂B.

Here are two pictures to illustrate the behavior of the optimal measure ν.



The measure ν when $\mu = 1_Q$ with Q a rectangle



The measure ν when $\mu = 1_Q$ with Q a half circle

We now fix a nonnegative integrable function ϕ with $\int \phi \, dx > 1$ and consider the class

$$\mathcal{A}_\phi = \mu^\perp \cap L^1_\phi = \left\{\rho \in \mathcal{P} \cap L^1,\ \rho \perp \mu,\ \rho \le \phi\right\}.$$

Theorem. *For every $\mu \in \mathcal{P}$ there exists a set A with $\mu(A) = 0$ and such that*

$$W(\mu, \mathcal{A}_\phi) = W(\mu, \phi 1_A).$$

The set A is unique when $\mu \in L^1$. In other words the optimization problem

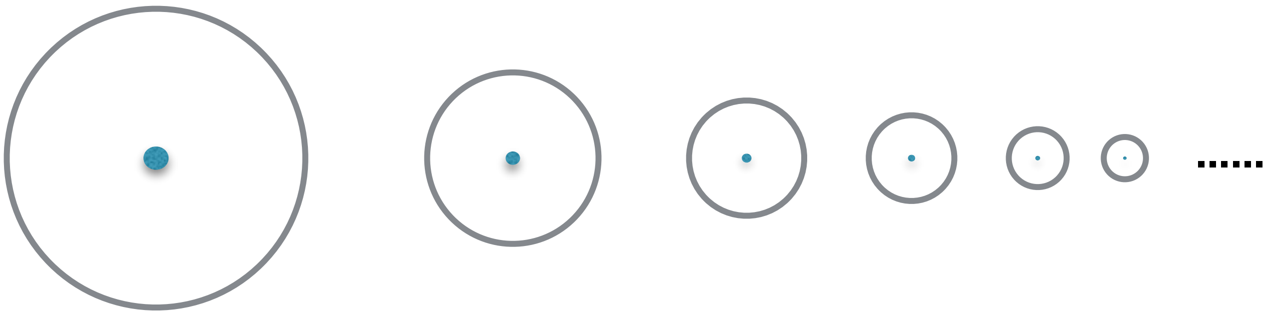
$$\min \left\{ W(\mu, \nu) \ :\ \nu \in \mu^\perp \cap L^1_\phi \right\}$$

has a solution of the form $\nu = \phi 1_A$.

In particular, for every fixed B, the problem

$$\min\left\{W(A,B)\ :\ |A|=|B|,\ |A\cap B|=0\right\}$$

has a unique solution. In general the set A can be very irregular (no finite perimeter).



$\mu=\sum_n n^{-2}\delta_{x_n}$ or small disks instead of Dirac masses

$\mathrm{Per}(A)=2\sqrt{\pi}\sum_n n^{-1}=+\infty.$

However, if μ is "good" the set A is more regular.

Theorem. *Let* $p = 2$ *and* $\phi = 1$*. Assume that* $\mu \in \mathcal{P} \cap BV$ *and that the set* $S_\mu = \{\mu(x) > 0\}$ *has a finite perimeter. Then the set* A *above has a finite perimeter and*

$$\mathrm{Per}(A) \leq \int |\nabla \mu| + 2\,\mathrm{Per}(S_\mu).$$

In particular, when $\mu = 1_B\,dx$, with B of finite perimeter, the optimal set A satisfies the inequality

$$\mathrm{Per}(A) \leq 3\,\mathrm{Per}(B).$$

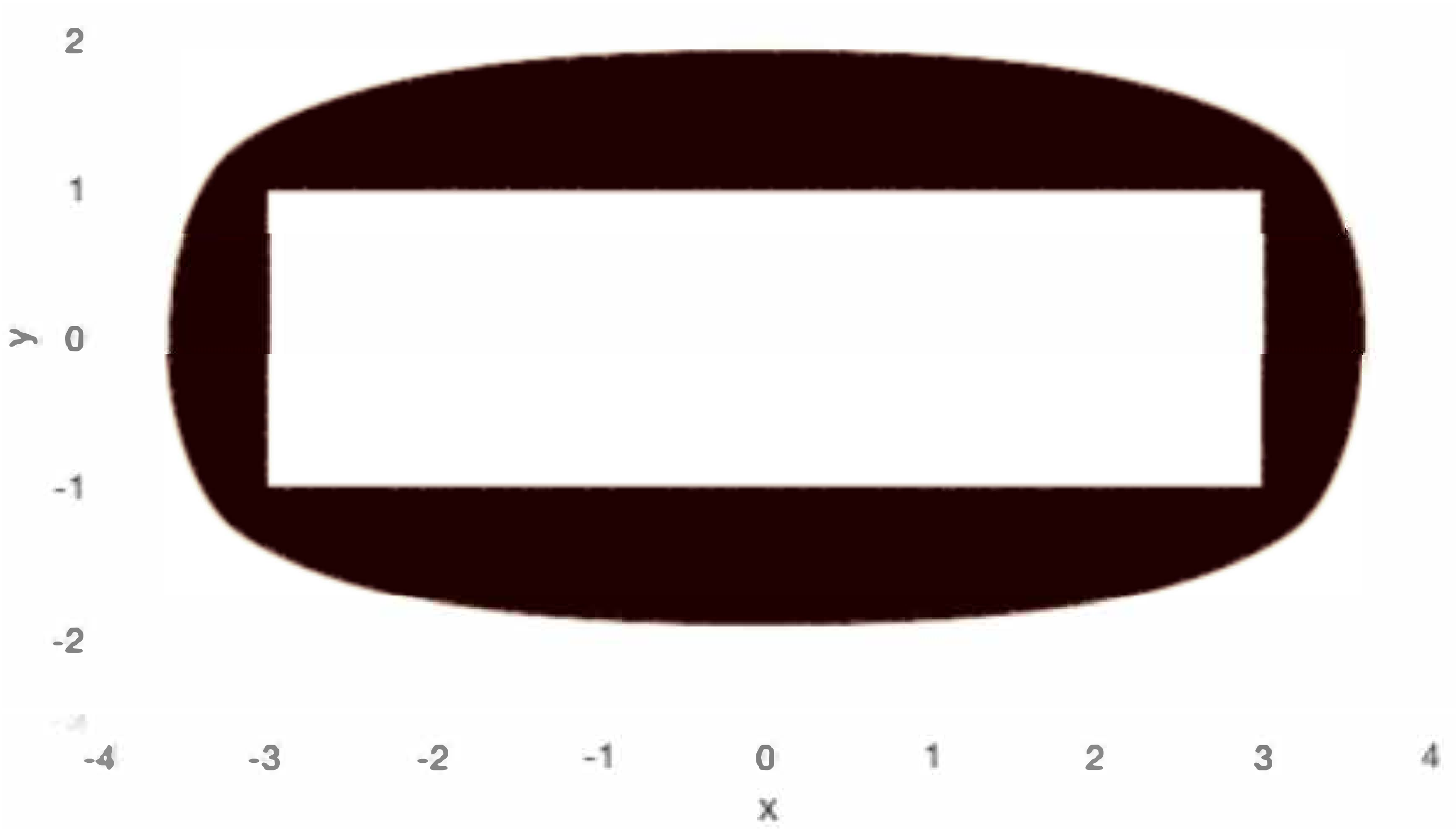
An interesting question then rises: for a smooth (finite perimeter is enough) set B let A_B be the corresponding optimal set A given by the theorem above. Find the best constant C such that

$$\frac{\mathrm{Per}(A_B)}{\mathrm{Per}(B)} \leq C.$$

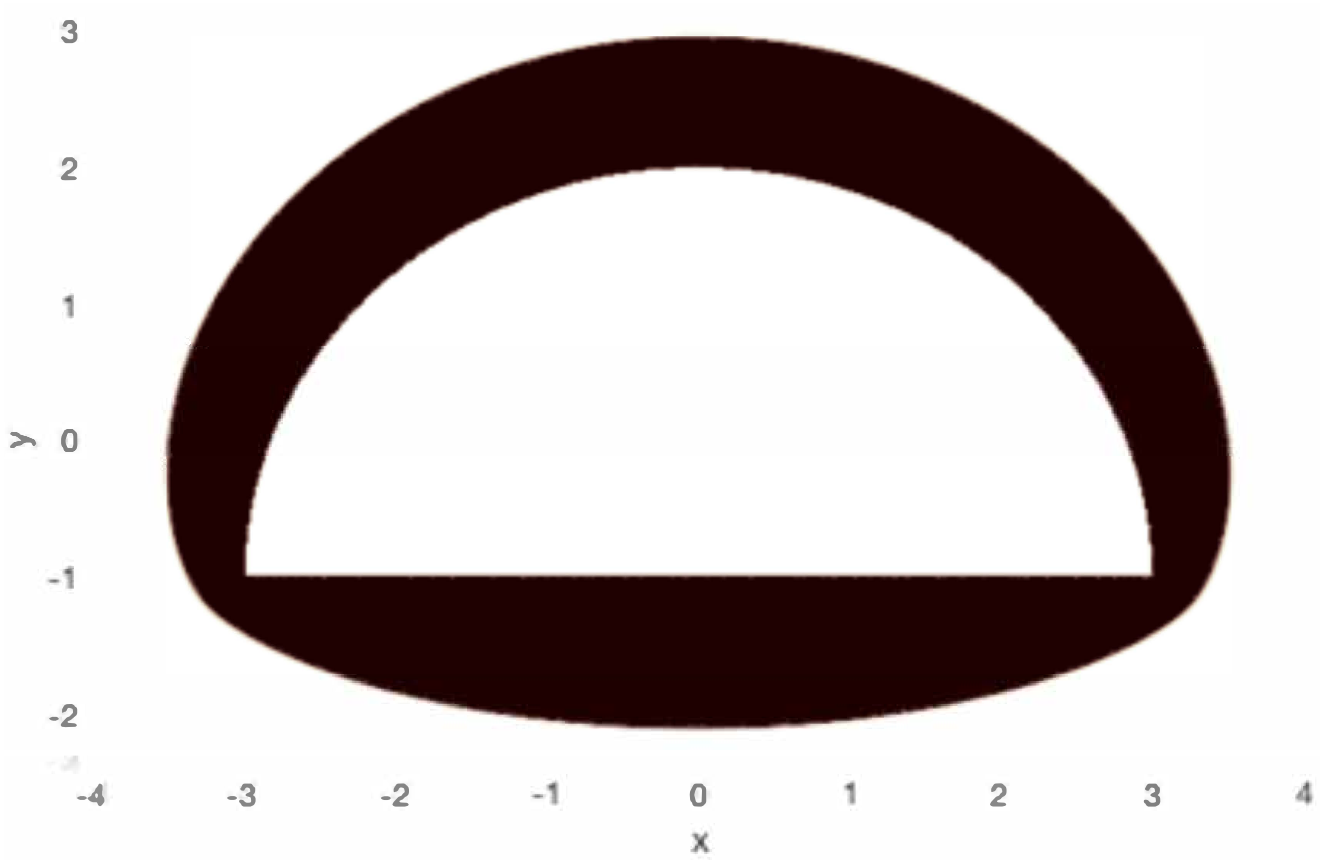
Is the best constant achieved when B is a ball? In that case we would have

$$C = 1 + 2^{(d-1)/d},$$

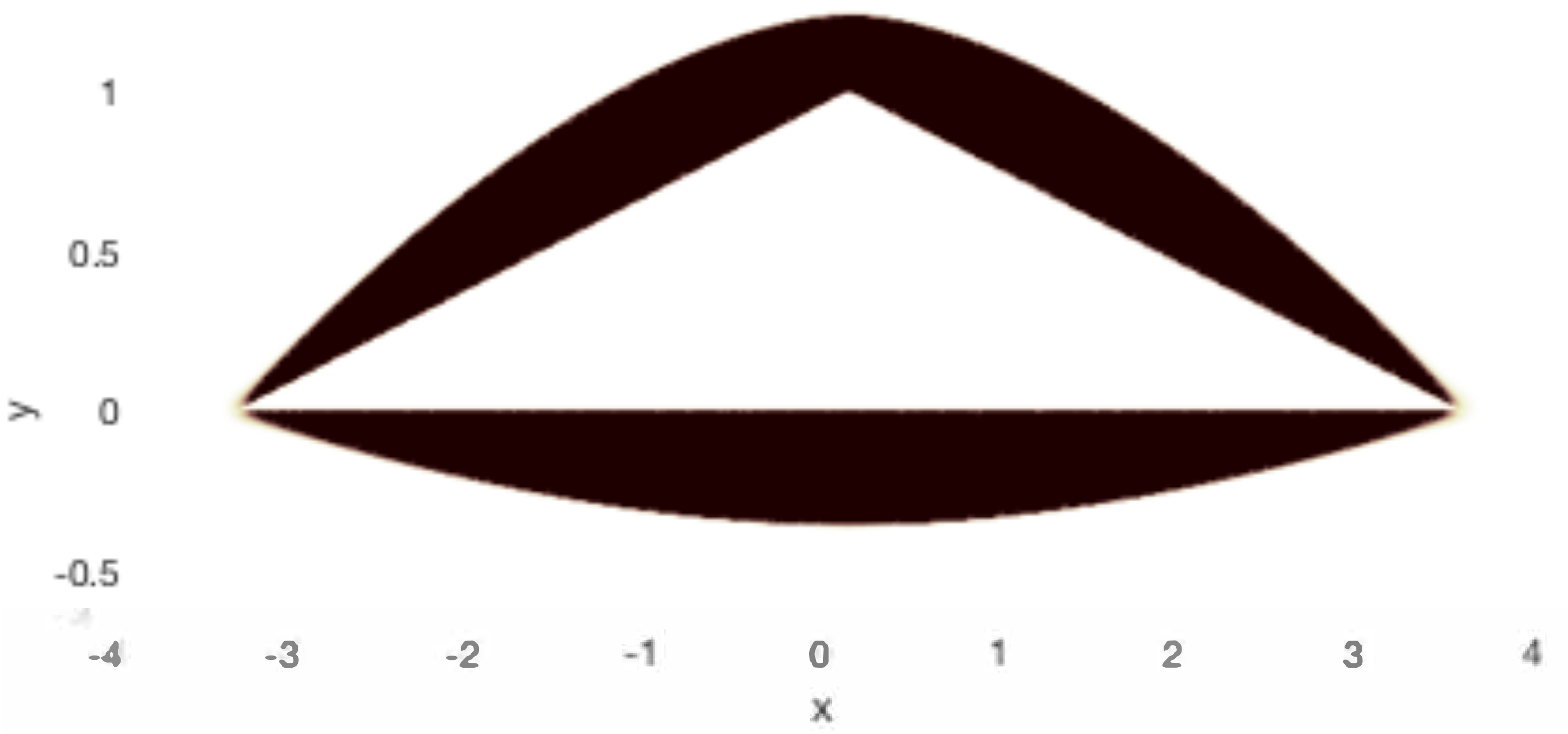
in agreement with the (non-sharp) evaluation $C \leq 3$ of the theorem.
Below some numerical outputs for the set A.



The set A when $\mu = 1_Q$ with Q a rectangle

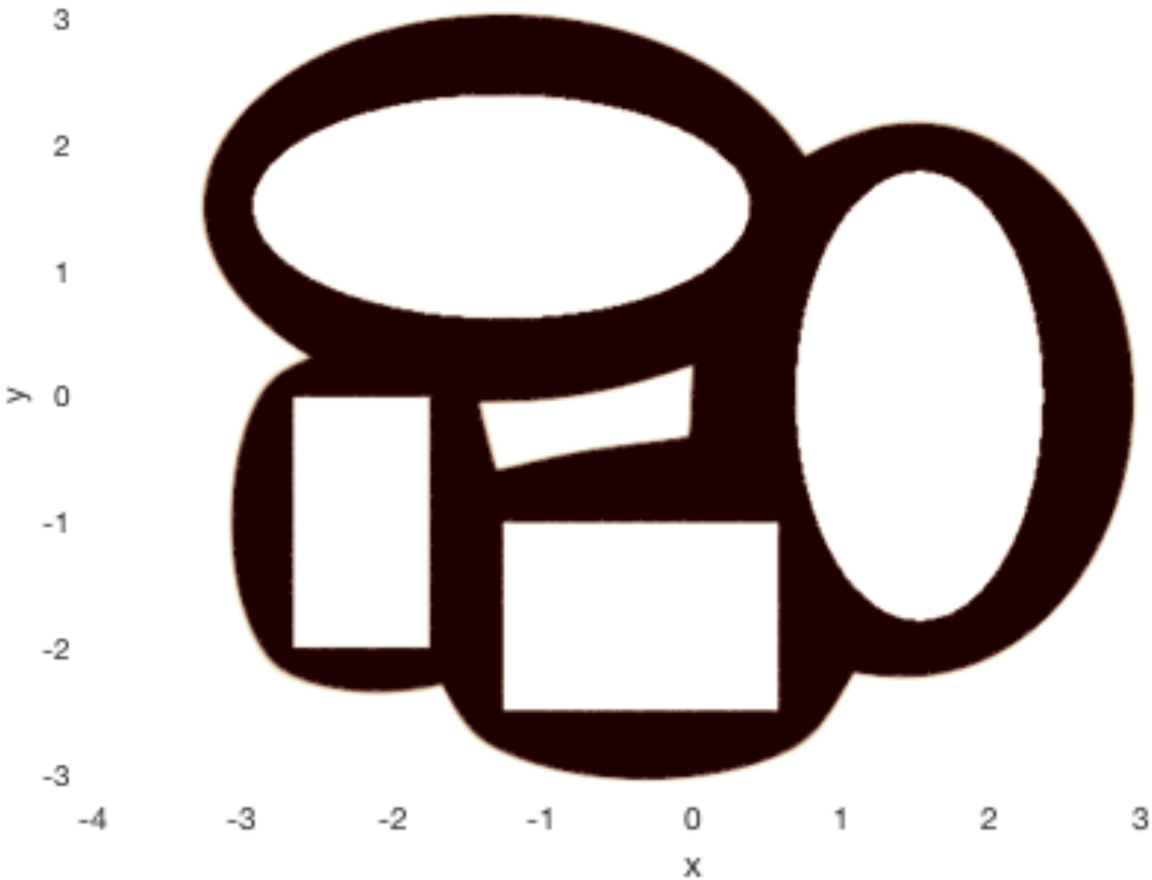


The set A when $\mu = 1_Q$ with Q a half-circle



The set A when $\mu = 1_Q$ with Q a triangle

The set A when $\mu = 1_Q$ with Q a pacman-like set



The set A when $\mu = 1_Q$ with Q a disconnected set

We now consider the minimum problem

$$\min\Big\{\operatorname{Per}(B)+kW(B,A)\ :\ |A\cap B|=0,\ |A|=|B|=1\Big\}$$

where we minimize both with respect to A and to B. For simplicity we assume that all the competing domains A, B are contained in a fixed bounded set D.

Theorem. *There exists an optimal pair A, B. In the two-dimensional case $d = 2$, we can take $D = \mathbb{R}^2$.*

Let (B_n, A_n) be a minimizing sequence. Since $P(B_n)$ are bounded and $B_n \subset D$, possibly passing to subsequences we may assume that $B_n \to B^*$ strongly in L^1. Analogously, we may suppose to have $1_{A_n} \rightharpoonup \bar{\theta}$ weakly* in L^∞ for a suitable $\bar{\theta}$ with $0 \leq \bar{\theta} \leq 1$, and

$$\int_{B^*} \bar{\theta}\, dx = \lim_n \int_{B_n} 1_{A_n}\, dx = 0.$$

By the theorem above with $\mu = 1_{B^*}$ and $\phi = 1$, the minimization problem

$$\min \left\{ W(B^*, \theta) \;:\; \int \theta\, dx = 1,\ 0 \leq \theta \leq 1,\ \int_{B^*} \theta\, dx = 0 \right\}$$

admits a solution $\bar{\theta}$ which is the characteristic function of a set A^*, that is $\bar{\theta} = 1_{A^*}$. By the minimality of A^* we have

$$W(B^*, A^*) \leq W(B^*, \bar{\theta}) = \lim_n W(B_n, A_n).$$

We may now conclude by the lower semicontinuity of the perimeter with respect to the strong L^1-convergence, so that

$$P(B^*) + kW(B^*, A^*) \leq \liminf_n \Big(P(B_n) + kW(B_n, A_n) \Big),$$

which concludes the proof.

Even if we expect that a result similar to the one above holds for every dimension, our proof uses the fact that for a connected set its diameter is bounded by its perimeter, which only holds in dimension two.

It would be interesting to find an alternative proof valid for every dimension d.

Circles are stationary. Are the solutions circles when k is small? Solutions are disconnected when k becomes large.

More generally, we may consider the shape optimization problem

$$\min\Big\{\mathrm{Per}(B)+kF_\alpha(B)\ :\ B\subset D,\ |B|=1\Big\}$$

where, for $\alpha>0$

$$F_\alpha(B)=\min\Big\{W^\alpha(A,B)\ :\ |A\cap B|=0,\ |A|=1\Big\}.$$

Thanks to the theory of quasi-minimizers of the perimeter the optimal domains B are such that, denoting by $\partial^* B$ the reduced boundary, $\partial^* B\cap D$ is a $C^{1,1/2}$ hypersurface and the Hausdorff dimension of $(\partial B\setminus\partial^* B)\cap D$ is at most $d-8$.