Convexity of the support of the displacement interpolation: counterexamples
Filippo Santambrogio, Xu-Jia Wang

To cite this version:
Filippo Santambrogio, Xu-Jia Wang. Convexity of the support of the displacement interpolation: counterexamples. 2015. hal-01214200

HAL Id: hal-01214200
https://hal.archives-ouvertes.fr/hal-01214200
Preprint submitted on 10 Oct 2015

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

L’archive ouverte pluridisciplinaire HAL, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d’enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.
Convexity of the support of the displacement interpolation: counterexamples

Filippo Santambrogio, Xu-Jia Wang

October 12, 2015

Abstract

Given two smooth and positive densities $\rho_0, \rho_1$ on two compact convex sets $K_0, K_1$, respectively, we consider the question whether the support of the measure $\rho_t$ obtained as the geodesic interpolant of $\rho_0$ and $\rho_1$ in the Wasserstein space $W_2(\mathbb{R}^d)$ is necessarily convex or not. We prove that this is not the case, even when $\rho_0$ and $\rho_1$ are uniform measures.

1 Introduction

One of the main features of optimal transport theory (we refer to [16] and [15] for a general presentation) is the fact that it provides an original and efficient way to define interpolations between probability measures.

Given a domain $\Omega \subset \mathbb{R}^d$ (that we take compact for simplicity and convex for the sake of the interpolation), we define the space $W_2(\Omega)$ as the space of probabilities on $\Omega$ endowed with the distance $W_2^2$, defined through

$$W_2^2(\mu, \nu) = \min \left\{ \int_{\Omega \times \Omega} |x - y|^2 \, d\gamma : \gamma \in \Pi(\mu, \nu) \right\},$$

where $\Pi(\mu, \nu)$ is the set of the so-called transport plans, i.e.

$$\Pi(\mu, \nu) = \{ \gamma \in \mathcal{P}(\Omega \times \Omega) : (\pi_x)_\#\gamma = \mu, (\pi_y)_\#\gamma = \nu \},$$

where $\pi_x(x, y) := x$ and $\pi_y(x, y) = y$ are the standard projections on the two factors of $\Omega \times \Omega$.

It is possible to prove that the above minimization problem has a solution, which is unique and of the form $\gamma = (id, T)_\#\mu$ (i.e. it is concentrated on the graph of a map $T : \Omega \to \Omega$, called optimal transport map) in many situations (in particular if $\mu \ll \mathcal{L}^d$ and that $W_2$, i.e. the square root of the minimal value above, is indeed a distance on $\mathcal{P}(\Omega)$.

The space $W_2(\Omega)$ can be checked to be a geodesic space, where for $\mu, \nu \in \mathcal{P}(\Omega)$ the (unique) geodesic curve connecting them is obtained through

$$\rho_t := ((1-t) \text{id} + tT)_\#\mu,$$  \hspace{1cm} (1.1)
where \( T \) is the optimal transport map from \( \mu \) to \( \nu \). This provides a useful interpolation between \( \mu \) and \( \nu \) which is very different from the linear interpolation \( (1 - t)\mu + t\nu \). In particular, it is useful in many applications, for instance in image processing: this is the case when doing histograms interpolations (roughly speaking, the average between a distribution of pixel colors which are almost white and another were they are almost black should be a a distribution with intermediate grey pixels, and not one with half pixels which are white and half which are black), or when artificially creating intermediate images between two pictures representing a same object which has moved (and the goal is to find the same object in the middle, instead of two half objects at the starting and arrival spots). But this very interpolation has also a lot of mathematical applications, as it was first pointed out by McCann in [14]. Indeed, McCann found a class of functionals \( F : \mathcal{P}(\Omega) \to \mathbb{R} \) which are convex along these geodesic lines (but not necessarily convex in the usual sense, think for instance at \( \mu \mapsto \int \int |x - y|^2d\mu(x)d\mu(y) \)), thus making possible to obtain uniqueness results or sufficient optimality conditions (see [4], for instance) for variational problems involving them. Also, this notion of convexity, called displacement convexity (the above interpolation is also called displacement interpolation) is an important notion in the study of gradient flows of these functionals (see [2]). We also recall that the interpolation \( \rho_t \) can be found numerically via one of the most classical algorithm for optimal transport, the so-called Benamou-Brenier method [3]. By solving the kinetic energy minimization problem

\[
\min \left\{ \int_0^1 \int_{\Omega} |v_t|^2d\rho_t dt : \partial_t \rho_t + \nabla \cdot (v_t \rho_t) = 0 \right\},
\]

among curves of measures with given initial and final data, one recovers the above interpolation, and the optimal velocity field \( v_t \) allows to find \( T \) (as we have \( v_t(x) = (T - \text{id}) \circ (T_t)^{-1} \) with \( T_t := (1 - t)\text{id} + tT \)).

Finally, let us remark that the interpolated measures \( \rho_t \) can also be considered as weighted barycenters between \( \rho_0 \) and \( \rho_1 \), as they solve the minimization problem

\[
\min \left\{ (1 - t)W_2^2(\rho, \rho_0) + tW_2^2(\rho, \rho_1) : \rho \in \mathcal{P}(\Omega) \right\},
\]
a minimization problem which has \( \rho_t \) as unique solution (provided one of the measures \( \rho_0, \rho_1 \) is absolutely continuous). This problem has also been considered when more than two measures are given, in order to find the weighted barycenter of many of them, solving (see [1])

\[
\min \left\{ \sum_{i=1}^{n} \lambda_i W_2^2(\rho, \rho_i) : \rho \in \mathcal{P}(\Omega) \right\},
\]

(a minimization problem that can be also recast in the setting of multi-marginal problems, as in [11]). Note that when \( n \geq 3 \) this variational definition of the barycenter is the only possible one, differently from the case \( n = 2 \) where one can simply use the geodesic in \( W_2 \).

In optimal transport theory, and in particular for regularity issues, the convexity of the support of a measure is an important fact, and some results are only available under this assumption. In particular, the optimal transport map between two smooth densities \( \rho_0, \rho_1 \) is smooth itself, provided
their supports are convex (this theory was first developed by Caffarelli, [7, 9, 8], see [10] for a survey). Since it is possible to write optimality conditions for the minimizers of (1.2) in terms of the optimal maps $T_i$ sending the optimal $\rho$ onto the various $\rho_i$ (which actually amounts to a system of Monge-Ampère equations), one could imagine some bootstrap strategy to study the regularity of $\rho$ using that of the $T_i$. Yet, this requires as a preliminary fact to check the convexity of $\text{spt}(\rho)$! Note that this is not needed for $n = 2$, since formula (1.1) allows to deduce the regularity of $\rho_i$ from that of $T$, which is the optimal map from $\rho_0$ to $\rho_1$: in particular, only the convexity of the supports of these two measures is needed. Anyway, the question whether the support of the barycenter is guaranteed to be convex whenever the supports of the $\rho_i$ are so is a very natural question. The case $n = 2$ is the easiest to study and if the answer is no in such a case there is no hope to get yes in the case $n \geq 3$.

Moreover, independently of the regularity motivation, this question is very natural from a geometric point of view, as a question on convex bodies. One can for instance insist on the case where $\rho_0$ and $\rho_1$ are uniform measures on convex sets and look at $\text{spt}(\rho_i)$ as an interpolation between these sets. More generally, one could wonder if the interpolation between log-concave densities (of their supports are convex (this theory was first developed by Caffarelli, [7, 9, 8], see [10] for a survey).

Unfortunately, the present paper shows that the answer to these questions is no. More precisely, we will see (Theorem 1) that on any smooth convex domain on $\mathbb{R}^2$ we can find smooth and positive densities $\rho_0$, $\rho_1$ such that the support of $\rho_{1/2}$ is not convex, and that this can also be done with uniform measures on (different, of course) convex sets.

2 The case of two segments or two curves

Consider two Lipschitz curves $\omega_0, \omega_1 : [a, b] \to \mathbb{R}^d$, and two measures $\nu_0, \nu_1 \in \mathcal{P}([a, b])$. We need to discuss the optimal transport between the measures $\rho_0$ and $\rho_1$, defined through $\mu_i := (\omega_i)^\# \nu_i$ for $i = 0, 1$, in connection with the optimal transport between $\nu_0$ and $\nu_1$. Note that, if $d > 1$, these measures fall out of the usual assumptions for the existence of an optimal transport map between them, as it is not true that $\mu_0$ does not give mass to “small” (i.e. $(d - 1)$–dimensional, or smaller) sets (see [5, 13] and, more recently, [12], for the exact conditions for this existence result).

Yet, the fact that both measures are one-dimensional, allows to act differently (by the way, let us mention a recent preprint, [6], on the case where optimal transport between measures in different dimensions is considered). We will suppose that $\omega_0$ and $\omega_1$ are such that $\dot{\omega}_0(t) \cdot \dot{\omega}_1(s) > 0$ for a.e. $(t, s) \in [a, b]^2$. In this case what happens is that optimal transport plans $\gamma \in \Pi(\mu_0, \mu_1)$ for the quadratic cost $|x - y|^2$ are just obtained as the image of optimal transport plans in $\Pi(\nu_0, \nu_1)$ through the map $(t, s) \mapsto (\omega_0(t), \omega_1(s))$.

Indeed, it is well-known that transport plans $\gamma$ are optimal for the quadratic cost if and only if their support is monotone, in the sense that $(x, y), (x', y') \in \text{spt}(\gamma)$ implies $(x - x') \cdot (y - y') \geq 0$.

\footnote{This result has been proven by Young-Heon Kim via un-published computations, in the hope of generalizing to higher dimensions.}
Yet, for \((x, y) = (\omega_0(t), \omega_1(s))\) and \(x' = (\omega_0(t'), \omega_1(s'))\), we have
\[
(x - x') \cdot (y - y') = \left( \int_t^{t'} \omega_0(\tau) \mathrm{d}\tau \right) \cdot \left( \int_s^{s'} \omega_1(\sigma) \mathrm{d}\sigma \right) = \int_t^{t'} \int_s^{s'} \omega_0(\tau) \cdot \omega_1(\sigma) \mathrm{d}\sigma \mathrm{d}\tau,
\]
and, since the integrand in the last expression is positive, the sign of the integral is the same as the sign of \((t' - t)(s' - s)\). This proves that, if we write \(\gamma = (\omega_0 \times \omega_1)^\# \tilde{\gamma}\), the monotonicity of the support of \(\gamma\) is equivalent to the 1D monotonicity of that of \(\tilde{\gamma}\). In particular, when \(\nu_0 \ll \mathcal{L}^1\), the optimal transport plan in \(\Pi(\nu_0, \nu_1)\) is unique and easy to find, and it is induced by a map, the unique monotone increasing mapping \(T\) sending \(\nu_0\) onto \(\nu_1\); as a consequence, there is a unique optimal transport plan from \(\mu_0\) to \(\mu_1\), and it is induced by the map \(\omega_1 \circ T \circ \omega_0^{-1}\).

We just point out that the situation is even simpler in case \(\omega_0\) and \(\omega_1\) are affine functions, i.e. their images are two segments \(S_0\) and \(S_1\). Up to translations, we can suppose that \(S_0\) and \(S_1\) are of the form \(S_i = \{tv_i, t \in [0, 1]\}\), for some vectors \(v_i \in \mathbb{R}^d\) \((i = 0, 1)\). Translating the segments is not restrictive as we know that optimal transport and translations commute.

Then, we distinguish three cases: either \(v_0 \cdot v_1 = 0\), or \(v_0 \cdot v_1 > 0\), or \(v_0 \cdot v_1 < 0\). In the first case it is easy to check that, for every \(x \in S_0\) and \(y \in S_1\), we have \(|x - y|^2 = |x|^2 + |y|^2\). This implies \(\int |x - y|^2 \mathrm{d}\gamma = \int |x|^2 \mathrm{d}\mu_0 + \int |y|^2 \mathrm{d}\mu_1\) for every \(\gamma \in \Pi(\mu_0, \mu_1)\), and hence every transport plan between these two measures is optimal (in particular, there is no uniqueness). In the case \(v_0 \cdot v_1 > 0\), we have, for \(x = tv_0, y = sv_1\) and \(c = v_0 \cdot v_1\),
\[
|x - y|^2 = c|t - s|^2 + (|v_0|^2 - c)t^2 + (|v_1|^2 - c)s^2,
\]
where the first term in the right hand side is a positive multiple of the quadratic cost in the variables \((t, s)\) and the second term is separable. The case \(v_0 \cdot v_1 < 0\) can be reduced to the previous one up to replacing \(v_0\) with \(-v_0\) and translating again. This means that, whenever \(v_0\) and \(v_1\) are not orthogonal, the optimal transport between the two measures can be computed as if we were in dimension one, i.e. it is the monotone increasing map. The orientation of the segments (which is crucial to define the monotonicity of the map) has to be chosen so that the scalar product is positive.

With these considerations in mind, we consider the following example.

Let \(S_0 = \{(t, at) : t \in [-1, 1]\}\) and \(S_1 = \{(t, -at) : t \in [-1, 1]\}\) be two segments in \(\mathbb{R}^2\) with middle point at the origin, and connecting \((-1, -a)\) to \((1, a)\) and \((-1, a)\) to \((1, -a)\), respectively. We select \(a \in [0, 1]\) so that the correct orientation of the two segments is the one based on the abscissas. Then we consider a probability measure \(\mu_0\) on \(S_0\) with linear density, proportional to \(1 - t\), and \(\mu_1\) a probability measure on \(S_1\) with density proportional to \(t + 1\). A simple calculation provides the formula for the optimal transport map
\[
T(t, at) = (-1 + \sqrt{4 - (1 - t)^2}, a(1 - \sqrt{4 - (1 - t)^2}))
\]
and the shape of the interpolated measure is represented in the following figure. The fact that this interpolation is supported on a curve which is not a segment (it is actually part of the boundary of an ellipse) will be the key point for our counter-examples.
Figure 1: The support of $\mu_{1/2}$ when $\mu_0$ and $\mu_1$ have linear densities on the segments $S_0$ and $S_1$.

Figure 2: The supports of the interpolations $\mu_t$ for $t = k/6$.

3 Counterexamples with convex bodies

In this section we use the previous considerations to provide a wide class of cases where two measures $\rho_0$ and $\rho_1$ have smooth and positive densities on their supports, their supports are convex sets, but the support of the interpolant $\rho_{1/2}$ cannot be convex.

Geometric setting We consider now a convex set $\Omega \subset \mathbb{R}^2$ and two portions $\Gamma_0$ and $\Gamma_1$ of its boundary, and we suppose that, in suitable coordinates, we have $\Gamma_0 = \{(t, f(t)) : t \in I\}$ and $\Gamma_1 = \{(t, g(t)) : t \in J\}$ for some disjoint intervals $I = [a, b]$ and $J = [c, d]$ and $f : I \to \mathbb{R}, g : J \to \mathbb{R}$ Lipschitz functions, with $0 < f' < \lambda$ and $0 > g' > -\lambda$ for a suitable constant $\lambda \in (0, 1)$. We suppose $f(b) = g(c)$ and $f(a) = g(d)$. Let us call $x = (a, f(a))$ and $x' = (b, f(b))$ the endpoints of $\Gamma_0$, and $y = (c, g(c))$ and $y' = (d, g(d))$ those of $\Gamma_1$. We also take two points $x'' \in \Gamma_0$ and $y'' \in \Gamma_1$, with the same ordinate, smaller than $(f(a) + f(b))/2 = (g(c) + g(d))/2$. We consider two measures $\mu_0 \in \mathcal{P}(\Gamma_0)$ and $\mu_1 \in \mathcal{P}(\Gamma_1)$ such that the optimal transport plan between them exists, is unique, and is induced by a map $T$ with $T(x) = y$, $T(x') = y'$ and $T(x'') = y''$.

The situation is the one sketched in the picture below.

We now consider two sequence of densities $\rho_n^0$ and $\rho_n^1$ weakly converging to $\rho_0$ and $\rho_1$, respectively, and concentrated on $\Omega$ (we will see different examples later where their support can be, or not, the whole domain $\Omega$). Let us call $\rho_{1/2}^n$ the interpolated density between them at time $t = 1/2$. We will prove that, for $n$ large enough, it is not possible that the support of $\rho_{1/2}^n$ is convex.

Lemma 1. Let $\Omega$ be a convex domain and $\mu_0$ and $\mu_1$ two measures on the segments $S_0$ and $S_1$ as in the geometric setting above. Let $\rho_0^n$ and $\rho_1^n$ be smooth densities weakly converging to $\rho_0$ and $\rho_1$, respectively, and concentrated on $\Omega$. Let $\rho_{1/2}^n$ be the middle point of the geodesic in $\mathbb{W}_2(\Omega)$ between them. Then, for $n$ large enough, the support of $\rho_{1/2}^n$ is not convex.

Proof. Consider the optimal transport plan $\gamma^n \in \Pi(\rho_0^n, \rho_1^n)$ for the quadratic cost. By uniqueness
of the optimal plan \( \gamma \) from \( \mu_0 \) to \( \mu_1 \), it is clear that \( \gamma^n \rightharpoonup \gamma \). From the fact that the support of \( \gamma \) is included in the Hausdorff limit of the supports of \( \gamma^n \), we deduce the existence of points \((x_n, y_n), (x'_n, y'_n), (x''_n, y''_n)\) in \(\text{spt}(\gamma^n)\) converging to \((x, y), (x', y')\) and \((x'', y'')\), respectively. Since \(((x_n + y_n)/2, (x'_n + y'_n)/2)\) belong to the support of \(\rho^n_{1/2}\), if this support were convex, it should also contain \(p_n := (x_n + y_n + x'_n + y'_n)/4\).

For simplicity, and without loss of generality, we will suppose that \((x + y + x' + y')/4 = 0\) (this is possible up to a translation).

Now, suppose \(p_n \in \text{spt}(\rho^n_{1/2})\). This means that there exist \(z_n, w_n \in \Omega\) such that \((z_n + w_n)/2 = p_n\) and \((z_n, w_n) \in \text{spt}(\gamma^n)\). Note that the monotonicity of \(\text{spt}(\gamma^n)\) implies the inequality

\[
(w_n - y''_n) \cdot (z_n - x''_n) \geq 0.
\]

Up to subsequences, we can pass to the limit and obtain the existence of two points \(z, w \in \bar{\Omega}\) such that \(z + w = 0\) and \((w - y'') \cdot (z - x'') \geq 0\).

Using \(w = -z\) this translates into \(z \in \bar{\Omega} \cap (-\bar{\Omega})\) and

\[
\left| z - \frac{x'' - y''}{2} \right|^2 \leq \left| \frac{x'' + y''}{2} \right|^2.
\]

This means that \(z\) belongs to the intersection of \(\bar{\Omega} \cap (-\bar{\Omega})\) and a ball centered at \(q := (x'' - y'')/2\) with radius \(|(x'' + y'')/2|\). But it is possible to see that this intersection is empty, as one can observe from the next figures.

Indeed, the closest point of \(\bar{\Omega} \cap (-\bar{\Omega})\) to \(q\) is the point \(\tilde{q}\) on the intersection of horizontal line through 0 with \(\partial \Omega\), yet it does not belong to the required ball, as the segment \([x'', -y'']\) is a diameter of this ball, and \(\tilde{q}\) sees both segments \([x'', q]\) and \([q, -y'']\) under an angle smaller than 45° because of the derivative condition on \(f\) and \(g\).

We can now provide our main counter-examples.

**Theorem 1.** For every smooth convex body \(\Omega \subset \mathbb{R}^2\) there exist two smooth probability densities \(\rho_0\) and \(\rho_1\) in \(\mathcal{P}(...\Omega)\), strictly positive on \(\Omega\), such that the support of their interpolation \(\rho_{1/2}\) is not convex.
\[ 0 = x + y + x' + y' \]

\[ 4x'' + y'' = q \]

\[ \begin{aligned} &\Omega, -\Omega, \text{ and the reflected points.} \\ &\begin{figure}[h] \begin{center} \includegraphics[width=\textwidth]{figure4.png} \end{center} \end{figure} \end{aligned} \]

\[ \begin{aligned} &\begin{figure}[h] \begin{center} \includegraphics[width=\textwidth]{figure5.png} \end{center} \end{figure} \end{aligned} \]

**Proof.** It is enough to notice that it is always possible to find two portions of boundary \( \Gamma_0, \Gamma_1 \subset \partial \Omega \) such that the assumptions of Lemma 1 (i.e. those of the above geometrical setting) are satisfied (it is enough to choose \( \Gamma_0 \) and \( \Gamma_1 \) close enough to an extremal point of \( \Omega \), and orientate the coordinate system so that this point is the unique point of \( \Omega \) with maximal ordinate and to choose suitable non-uniform probabilities \( \mu_0 \) and \( \mu_1 \) on \( \Gamma_0 \) and \( \Gamma_1 \)). Then, we take any smooth and strictly positive approximations \( \rho_0^n \) and \( \rho_1^n \) and choose \( \rho_0 = \rho_0^n \) and \( \rho_1 = \rho_1^n \), for \( n \) large enough.

**Remark 1.** The same procedure can be applied to any convex domain \( \Omega \), provided that there exist portions of its boundary with tangent vectors forming angle strictly smaller than \( 90^\circ \), which allows to choose a suitable coordinate system and fit the geometrical setting. The only convex domains in \( \mathbb{R}^2 \) where this is not possible are the rectangles and the acute (or right) triangles.

**Theorem 2.** There exist convex sets \( A, B \subset \mathbb{R}^2 \) such that, setting \( \rho_0 = |A|^{-1} \mathcal{L}^{2} A \) and \( \rho_1 = |B|^{-1} \mathcal{L}^{2} B \), the support of the middle point interpolation \( \rho_{1/2} \) is not convex.

**Proof.** It is enough to choose \( a \in (0, 1) \) and use, in Lemma 1, \( \Omega = \{(t, s) \in \mathbb{R}^2 : s \leq -a|t|, s \geq -1, |t| \leq 1\} \), \( \Gamma_0 = [(-1, -a), (0, 0)] \), \( \Gamma_1 = [(1, -a), (0, 0)] \), and measures \( \mu_0 \) and \( \mu_1 \) supported on \( \Gamma_0 \) and \( \Gamma_1 \) and constructed as in the example at the end of the previous section. Then we take \( A_n \) to be the triangle with vertices \((0, 0), (-1, -a)\) and \((-1, -a - \frac{1}{n})\), \( B_n \) the triangle with vertices \((0, 0), (1, -a)\) and \((1, -a - \frac{1}{n})\), \( \rho_0^n = |A_n|^{-1} \mathcal{L}^{2} A_n \) and \( \rho_1^n = |B_n|^{-1} \mathcal{L}^{2} B_n \). These measures converge to \( \mu_0 \) and \( \mu_1 \), respectively. Hence, for \( n \) large enough we can take \( A = A_n \) and \( B = B_n \).

**Remark 2.** Note that the domains \( A_n \) and \( B_n \) in the above proof can be smoothened, and made strictly convex, with no extra difficulty.
Figure 6: Counter-example with uniform measures.

References


