LECTURE NOTES BY FELIX OTTO FOR THE PIMS SUMMER SCHOOL ON OPTIMAL TRANSPORT

Version of July 18th 2022.

These notes are based on the paper "Quantitative linearization results for the Monge-Ampère equation" with M. Goldman and M. Huesmann, in particular Section 3 therein. Compared to this paper, the notes come with more motivations, less PDE machinery, and a couple of simplifications. They should allow for an independent reading.

In the next weeks, Lukas Koch and I plan to sketch the remaining part of the proof, so that it is logically complete, and post the notes. I thank Francesco Mattesini and Christian Wagner for a careful reading. If you have questions, find typos – or more serious issues – please let me/us know.

1. Connection of optimal transportation and the Neumann problem for the Poisson equation

sec:orth

In this section, we motivate the connection between optimal transportation (OT) and the Neumann boundary value problem for the Poisson equation.

1.1. **Trajectories.** For the above connection, it is convenient to adopt a dynamical view upon OT, identifying a pair (x, y) of (matched) points with the (straight) trajectory

ao00 (1)
$$[0,1] \ni t \mapsto X(t) := ty + (1-t)x.$$

Given an optimal transfer plan π for λ, μ, we ask the question on how to choose a function ϕ in such a way that its gradient $\nabla\phi$ captures the velocity of the trajectories, meaning

ao65 (2)
$$\dot{X}(t) \approx \nabla\phi(X(t)) \quad \text{for } (x,y) \in \mathrm{supp}\pi.$$

As we shall see, the answer relates to the Poisson equation $-\triangle\phi = \mu - \lambda$.

We are interested in connecting to a boundary value problem for the Poisson equation on some domain, say a ball B_R of some radius R (to be optimized later) and center w. l. o. g. given by the origin. We are thus led to restrict ourselves to the set of trajectories that spend some time in the closure $\bar{B}_R$:

ao17 (3)
$$\Omega := \{\, (x,y) \,|\, \exists t \in [0,1]\ X(t) \in \bar{B}_R \,\}.$$

To every $(x,y) \in \Omega$, we associate the entering and exiting times $0 \leq \sigma \leq \tau \leq 1$ of the corresponding trajectory

ao66
$$\begin{aligned} \sigma &:= \min\{t \in [0,1] \,|\, X(t) \in \bar{B}_R\}, \\ \tau &:= \max\{t \in [0,1] \,|\, X(t) \in \bar{B}_R\}. \end{aligned} \tag{4}$$

(Note that some trajectories may both enter and exit.) Given a transfer plan π, we keep track of *where* the trajectories enter and exit B_R, which is captured by two (non-negative) measures f and g concentrated on ∂B_R, defined through

ao09
$$\int \zeta df = \int_{\Omega\cap\{X(\sigma)\in\partial B_R\}} \zeta(X(\sigma)) d\pi, \tag{5}$$

ao09bis
$$\int \zeta dg = \int_{\Omega\cap\{X(\tau)\in\partial B_R\}} \zeta(X(\tau)) d\pi \tag{6}$$

for all continuous and compactly supported functions ζ. Note that the set of trajectories $\Omega \cap \{X(\sigma) \in \partial B_R\} = \{\exists t \in [0,1] \;\; X(t) \in \partial B_R\}$ implicitly defines a Borel measurable subset of $\mathbb{R}^d \times \mathbb{R}^d$, namely the pre-image under the mapping (1), which is continuous from $\mathbb{R}^d \times \mathbb{R}^d$ into $C^0([0,1])$. Hence it is legitimate to integrate against π as in (5).

lem:orth

Lemma 1. *We have for any admissible π and any continuously differentiable function ϕ on $\bar{B}_R$*

ao08
$$\begin{aligned} &\int_\Omega \int_\sigma^\tau |\dot{X}(t) - \nabla\phi(X(t))|^2 dt d\pi \\ &= \int_\Omega \int_\sigma^\tau |\dot{X}(t)|^2 dt d\pi + \int_\Omega \int_\sigma^\tau |\nabla\phi(X(t))|^2 dt d\pi \\ &- 2\int_{B_R} \phi d(\mu - \lambda) - 2\int_{\partial B_R} \phi d(g - f). \end{aligned} \tag{7}$$

For later purpose, we record

ao44
$$\lambda(B_R) + f(\partial B_R) = \mu(B_R) + g(\partial B_R). \tag{8}$$

Proof of Lemma 1. For identity (7) we note that for the mixed term we have by the chain rule $\dot{X}(t) \cdot \nabla\phi(X(t)) = \frac{d}{dt}[\phi(X(t))]$ and thus by the fundamental theorem of calculus $\int_\sigma^\tau \dot{X}(t) \cdot \nabla\phi(X(t)) dt = \phi(X(\tau)) - \phi(X(\sigma))$. In view of definition (4) we either have $X(\sigma) \in \partial B_R$, or we have $X(\sigma) \in B_R$ and thus $\sigma = 0$ and $X(\sigma) = x$, so that we may ignore $(x,y) \in \Omega$ in this latter case. Hence $\int_\Omega \phi(X(\sigma)) d\pi$ $= \int_{\Omega\cap\{X(\sigma)\in\partial B_R\}} \phi(X(\sigma)) d\pi + \int_{\{x\in B_R\}} \phi(x) d\pi$. By definition (5), the first integral is $\int \phi df$. By admissibility of π, the second integral is $\int_{B_R} \phi d\lambda$. Likewise, one obtains $\int_\Omega \phi(X(\tau)) d\pi = \int \phi dg + \int_{B_R} \phi d\mu$.

Specifying to $\phi = 1$, and thus $\nabla\phi = 0$ so that the mixed term vanishes, we learn (8) from the above two identities.

ss:pert

1.2. **Perturbative regime.** We will focus on a "perturbative regime", which comes in form of two local smallness conditions. Smallness conditions have to be formulated in a non-dimensionalized way, which we implement by expressing this local smallness condition on a ball of non-dimensionalized radius, which for convenience is taken to be 3.

The first smallness condition involves the data (thus the letter D), that is, the two measures λ and μ. We monitor how close these measures are to the Lebesgue measure on B_3. It is natural to quantify this in terms of the Wasserstein distance. Since the mass $\lambda(B_3)$ in general is not equal to the Lebesgue volume $|B_3|$, we have to split this into two: We monitor how Wasserstein-close the restriction $\lambda \llcorner B_3$ is to the uniform measure $\kappa_\lambda dx \llcorner B_3$, where $\kappa_\lambda := \frac{\lambda(B_3)}{|B_3|}$, and we monitor how close this density κ_λ is to unity. It is convenient to do both on the squared level:

ao88

$$\begin{aligned} D :&= W^2(\lambda \llcorner B_3, \kappa_\lambda dx \llcorner B_3) + (\kappa_\lambda - 1)^2 \\ &\quad + \text{same expression with } \lambda \rightsquigarrow \mu. \end{aligned} \tag{9}$$

The second smallness condition involves the solution itself, i. e. π. It monitors the length of trajectories that start or end in B_3. It does so in a square-averaged sense, like the total cost function itself. In fact, it is a localization of the cost functional (or energy, thus the letter E):

ao45

$$E := \int_{(B_3\times\mathbb{R}^d)\cup(\mathbb{R}^d\times B_3)} |x-y|^2 d\pi. \tag{10}$$

We expect (and shall rigorously argue later) that in the perturbative regime $E + D \ll 1$ and provided $R \in [1,2]$ we have for the second r. h. s. term in (7)

ao14

$$\int_\Omega \int_\sigma^\tau |\nabla\phi(X(t))|^2 dt d\pi \approx \int_{B_R} |\nabla\phi|^2. \tag{11}$$

Indeed, for $E \ll 1$, trajectories do not move much so that

$$\int_\Omega \int_\sigma^\tau |\nabla\phi(X(t))|^2 dt d\pi \approx \int_{\{x\in B_R\}} |\nabla\phi(x)|^2 d\pi = \int_{B_R} |\nabla\phi|^2 d\lambda,$$

where the last identity follows from admissibility. Furthermore, for $D \ll 1$, λ is close to Lebesgue so that

$$\int_{B_R} |\nabla\phi|^2 d\lambda \approx \int_{B_R} |\nabla\phi|^2.$$

1.3. **Connection to the Neumann problem for the Poisson equation.** Hence in order to achieve (2), in view of (7) and (11), we are led to minimize

$$\int_{B_R} |\nabla\phi|^2 - 2\int_{B_R} \phi d(\mu-\lambda) - 2\int_{\partial B_R} \phi d(g-f) \tag{12}$$

in ϕ. A minimizer ϕ of (12), if it exists as a continuously differentiable function on $\bar{B}_R$, would be characterized by the Euler-Lagrange equation

$$\int_{B_R} \nabla\zeta\cdot\nabla\phi - \int_{B_R} \zeta d(\mu-\lambda) - \int_{\partial B_R} \zeta d(g-f) = 0 \tag{13}$$

for all continuously differentiable test functions ζ on B_R. If ϕ even exists as a twice continuously differentiable function on $\bar{B}_R$, we could appeal to the calculus identity $\nabla\zeta\cdot\nabla\phi = \nabla\cdot(\zeta\nabla\phi) - \zeta\triangle\phi$ and the divergence theorem in form of $\int_{B_R}\nabla\cdot(\zeta\nabla\phi) = \int_{\partial B_R}\zeta\nu\cdot\nabla\phi$, where $\nu(x) = \frac{x}{R}$ denotes the outer normal to ∂B_R in a point x, to obtain the integration by parts formula

$$\int_{B_R} \nabla\zeta\cdot\nabla\phi = \int_{B_R} \zeta(-\triangle\phi) + \int_{\partial B_R} \zeta\nu\cdot\nabla\phi. \tag{14}$$

Hence (13) can be reformulated and regrouped as

$$\int_{B_R} \zeta(-\triangle\phi - d(\mu-\lambda)) + \int_{\partial B_R} \zeta(\nu\cdot\nabla\phi - d(g-f)) = 0. \tag{15}$$

Considering first all test functions ζ's that vanish on ∂B_R, we learn from (15) that $-\triangle\phi = \mu - \lambda$ distributionally in B_R. Since $\mu-\lambda$ is a bounded measure, the first term in (15) thus vanishes also for test functions that do not vanish on ∂B_R. Hence the second term in (15) vanishes individually, which means $\nu\cdot\nabla\phi = g - f$ distributionally on ∂B_R. Hence we end up with what is called the Poisson equation with Neumann boundary conditions

$$-\triangle\phi = \mu-\lambda \text{ in } B_R, \quad \nu\cdot\nabla\phi = g-f \text{ on } \partial B_R. \tag{16}$$

This is a classical elliptic boundary value problem, which for sufficiently regular $\mu-\lambda$ and $g-f$ has a unique twice differentiable solution, provided (8) holds, and

$$\int_{B_R} \phi = 0 \tag{17}$$

is imposed. This motivates the connection between optimal transportation and the (short) Neumann-Poisson problem.

However, for rough (like sum of Diracs) measures λ,μ, and thus also rough measures f,g, the solution ϕ of (16), even if it exists for this linear problem, will be rough, too. In particular, (11) may not be true; even worse, both the l. h. s. and the r. h. s. might be infinite. Hence we shall approximate both $\mu-\lambda$ and $g-f$ by smooth functions (in fact, we shall approximate $\mu-\lambda$ by a constant function). The best way to organize the output of Lemma 1 is given by

cor:orth **Corollary 1.** *We have for any admissible π and any twice continuously differentiable function ϕ on $\bar B_R$*

$$\begin{aligned}
&\int_\Omega\int_\sigma^\tau |\dot X(t)-\nabla\phi(X(t))|^2 dt d\pi\\
&\le \int_\Omega |x-y|^2 d\pi - \int_{B_R}|\nabla\phi|^2\\
&+2\int_{B_R}\phi(-\triangle\phi-d(\mu-\lambda))+2\int_{\partial B_R}\phi(\nu\cdot\nabla\phi-d(g-f))\\
&+\int_\Omega\int_\sigma^\tau|\nabla\phi(X(t))|^2 dt d\pi-\int_{B_R}|\nabla\phi|^2. \qquad (18)
\end{aligned}$$

ao10

As we already argued, see (11), we expect the term in last line (18) to be negligible. The integrals in the second r. h. s. line can be made small by approximately solving (16) – so that there is a trade-off between making the last line and the second line small. However, the main open task is to argue, based on the optimality of π, that the term in the first r. h. s. line is small for an approximate solution of (16)

Proof of Corollary 1. The upgrade of identity (7) to inequality (18) relies on

ao46
$$\int_\Omega\int_\sigma^\tau |\dot X(t)|^2 dt d\pi \le \int_\Omega |x-y|^2 d\pi, \qquad (19)$$

ao48
$$\int_{B_R}|\nabla\phi|^2=\int_{B_R}\phi(-\triangle\phi)+\int_{\partial B_R}\phi\nu\cdot\phi. \qquad (20)$$

Inequality (19) follows from $\int_\sigma^\tau|\dot X(t)|^2dt\le\int_0^1|\dot X(t)|^2dt=|x-y|^2$. Identity (20) follows from (14) for $\zeta=\phi$.

1.4. **Localizing optimality.** As mentioned after Corollary 1, the main open task is to estimate the first r. h. s. line of (18). For this, we will (for the first time) use that π is optimal. In order to connect to the Neumann-Poisson problem on B_R, we need to leverage optimality in a localized way. Of course, it will in general not be true that the cost of π localized to $(B_R\times\mathbb{R}^d)\cup(\mathbb{R}^d\times B_R)$ is estimated by the transportation cost between the localized measures $\lambda\llcorner B_R$ and $\mu\llcorner B_R$. However, this is almost true if one adds the distribution of the entering points f, see (5), and exiting points g, see (6), respectively:

lem:opt **Lemma 2.** *For π optimal we have*

ao70
$$\Big(\int_\Omega |x-y|^2 d\pi\Big)^{\frac{1}{2}} \le W(\lambda\llcorner B_R + f, \mu\llcorner B_R + g) + \Big(2\int_{\{\exists t\in[0,1]\; X(t)\in\partial B_R\}} |x-y|^2 d\pi\Big)^{\frac{1}{2}}. \tag{21}$$

Lemma 2 controls the transportation cost coming from those trajectories that spend some time in $\bar B_R$, which amounts to the l. h. s. of (21) according to definition (3), by an OT problem localized to $\bar B_R$ as described by the first r. h. s. term. It does so up to the transportation cost coming from those (fewer) trajectories that cross (or touch) the boundary ∂B_R, see the second r. h. s. term. Eventually, one has to argue that this last term is negligible for a good choice of R.

As its form suggests, (21) has the structure of a triangle inequality. In fact, its proof has similarities with the proof of the triangle inequality for W, using a disintegration (or conditioning) argument.

Proof of Lemma 2. We now introduce the distribution of $x = X(0)$ under π conditioned on the event that the trajectory X enters at $z \in \partial B_R$. In less probabilistic and more measure-theoretic terms ("disintegration"), we introduce the (weakly continuous) family of probability measures $\{\lambda_z\}_{z\in\partial B_R}$ such that

ao79
$$\int_{\Omega\cap\{X(\sigma)\in\partial B_R\}} \zeta(x, X(\sigma))\pi(dxdy) = \int_{\partial B_R}\int \zeta(x,z)\lambda_z(dx)f(dz), \tag{22}$$

which is possible by (5). Here, ζ is an arbitrary test function on $\mathbb{R}^d\times\mathbb{R}^d$. Likewise, we introduce the probability distribution $\{\mu_w\}_{w\in\partial B_R}$ of the end points of trajectories that exit in w:

ao80
$$\int_{\Omega\cap\{X(\tau)\in\partial B_R\}} \zeta(X(\tau), y)\pi(dxdy) = \int_{\partial B_R}\int \zeta(w,y)\mu_w(dy)f(dw). \tag{23}$$

Let $\bar\pi$ denote an optimal plan for $W(\lambda\llcorner B_R + f, \mu\llcorner B_R + g)$. Equipped with these objects, we now define a competitor $\tilde\pi$ for π that mixes π with $\bar\pi$, in the sense that it takes the trajectories from π that stay outside of $\bar B_R$, the trajectories from $\bar\pi$ that stay inside (the open) B_R, and concatenates trajectories X from π that enter or exit $\bar B_R$ with trajectories of $\bar\pi$ that start or end in ∂B_R:

$$\begin{aligned}\int\zeta(x,y)\tilde\pi(dxdy) =& \int_{\Omega^c}\zeta(x,y)\pi(dxdy)\\ &+\int_{B_R\times B_R}\zeta(x,y)\bar\pi(dxdy)\\ &+\int_{\partial B_R\times B_R}\int\zeta(x,y)\lambda_z(dx)\bar\pi(dzdy)\\ &+\int_{B_R\times\partial B_R}\int\zeta(x,y)\mu_w(dy)\bar\pi(dxdw)\\ &+\int_{\partial B_R\times\partial B_R}\int\int\zeta(x,y)\mu_w(dy)\lambda_z(dx)\bar\pi(dzdw).\end{aligned}\tag{24}$$

It is straightforward to see that $\tilde\pi$ has marginals λ and μ; let us check the first by using (24) for a function $\zeta=\zeta(x)$: Since μ_w is a probability measure to the effect of $\int\zeta(x)\mu_w(dy)=\zeta(x)$, the second and the forth r. h. s. term of (24) combine to $\int_{B_R\times\mathbb{R}^d}\zeta(x)\bar\pi(dxdy)$. Here, we used that $\bar\pi$ is supported on $\bar B_R\times\bar B_R$. Likewise, the third and the fifth term combine to $\int_{\partial B_R\times\mathbb{R}^d}\int\zeta(x)\lambda_z(dx)\bar\pi(dzdy)$. By admissibility of $\bar\pi$, the combination of the second and forth term gives $\int_{B_R}\zeta(x)\mu(dx)$, which as in the proof of Lemma 1 (by admissibility of π) can be seen to agree with $\int_{\Omega\cap\{X(\sigma)\in B_R\}}\zeta(x)\pi(dxdy)$. Since $\int\zeta(x)\lambda_z(dx)$ does not depend on y, for the same reason, the combination of the third and fifth term renders $\int_{\partial B_R}\zeta(z)f(dz)$, which by definition (5) is equal to $\int_{\Omega\cap\{X(\sigma)\in\partial B_R\}}\zeta(x)\pi(dxdy)$. Hence these four terms combine to $\int_\Omega\zeta(x)\pi(dxdy)$. Therefore, the r. h. s. of (24) collapses as desired to $\int\zeta(x)\pi(dxdy)$, which coincides with $\int\zeta(x)\lambda(dx)$ by admissibility of π.

By optimality of π, we have $\int|x-y|^2d\pi\le\int|x-y|^2d\tilde\pi$; rewriting this as $\int_\Omega|x-y|^2d\pi+\int_{\Omega^c}|x-y|^2d\pi\le\int|x-y|^2d\tilde\pi$, and using (24) for $\zeta(x,y)=|x-y|^2$, we gather

$$\Big(\int_\Omega|x-y|^2d\pi\Big)^{\frac12}\le\|(f_2,f_3,f_4,f_5)\|,\tag{25}$$

where the four functions $f_2,\cdots,f_5\ge0$ are defined by

$$\begin{aligned}f_2(x,y):=|x-y|,\quad f_5^2(z,w):=\int\int|x-y|^2\mu_w(dy)\lambda_z(dx),\\ f_3^2(z,y):=\int|x-y|^2\lambda_z(dx),\quad f_4^2(x,w):=\int|x-y|^2\mu_w(dy),\end{aligned}$$

and the norm $\|\cdot\|$ is defined through

ao75

$$\begin{aligned}
&\|(f_2,f_3,f_4,f_5)\|^2\\
&=\int_{B_R\times B_R} f_2^2(x,y)\bar\pi(dxdy)+\int_{\partial B_R\times B_R} f_3^2(z,y)\bar\pi(dzdy)\\
&+\int_{B_R\times\partial B_R} f_4^2(x,w)\bar\pi(dxdw)+\int_{\partial B_R\times\partial B_R} f_5^2(z,w)\bar\pi(dzdw).
\end{aligned}\tag{26}$$

By the triangle inequality w. r. t. $L^2(\lambda_z)$ and $L^2(\mu_w)$, and using that λ_z, μ_w are probability measures, we obtain

ao82

$$f_3\le|z-y|+\tilde f_3,\quad f_4\le|x-w|+\tilde f_4,\quad f_5\le|z-w|+\sqrt2\tilde f_5,\tag{27}$$

where the three functions $\tilde f_3,\tilde f_4,\tilde f_5\ge0$ are defined by

ao78

$$\tilde f_3^2(z,y):=\int|x-z|^2\lambda_z(dx),\quad \tilde f_4^2(x,w):=\int|w-y|^2\mu_w(dy),\tag{28}$$

ao77

$$\tilde f_5^2(z,w):=\tilde f_3^2(z,y)+\tilde f_4^2(x,w).\tag{29}$$

The factor of $\sqrt2$ in (27) arises because of $\tilde f_3+\tilde f_4\le\sqrt2\tilde f_5$.

From (27) we obtain by the triangle inequality for $\|\cdot\|$

ao76

$$\begin{aligned}
&\Big(\int_\Omega|x-y|^2d\pi\Big)^{\frac12}\overset{(25)}{\le}\|(f_2,f_3,f_4,f_5)\|\\
&\le\|(|x-y|,|z-y|,|x-w|,|z-w|)\|+\sqrt2\|(0,\tilde f_3,\tilde f_4,\tilde f_5)\|,
\end{aligned}\tag{30}$$

where we gave up a factor of $\sqrt2$ on $\tilde f_3,\tilde f_4$. By definition (26) of $\|\cdot\|$, the first r. h. s. term in (30) coincides with the square root of $\int_{\bar B_R\times\bar B_R}|x-y|^2\bar\pi$, which by optimality of $\bar\pi$ is $W(\lambda\llcorner B_R+f,\,\mu\llcorner B_R+g)$, as desired. By definitions (26), (28), and (29), the second r. h. s. term in (30) is equal to 2 times $\int_{\partial B_R\times\mathbb R^d}\int|x-z|^2\lambda_z(dx)\bar\pi(dzdy)+\int_{\mathbb R^d\times\partial B_R}\int|w-y|^2\mu_y(dy)\bar\pi(dxdw)$. By admissibility of $\bar\pi$, this sum is equal to $\int_{\partial B_R}\int|x-z|^2\lambda_z(dx)f(dz)+\int_{\partial B_R}\int|w-y|^2\mu_y(dy)g(dw)$. By definitions (22) and (23), this coincides with $\int_{\Omega\cap\{X(\sigma)\in\partial B_R\}}|x-X(\sigma)|^2d\pi+\int_{\Omega\cap\{X(\tau)\in\partial B_R\}}|X(\tau)-y|^2d\pi$. Since on the intersection $\Omega\cap\{X(\sigma)\in\partial B_R\}\cap\{X(\tau)\in\partial B_R\}$ we have $|x-X(\sigma)|^2+|X(\tau)-y|^2\le|x-y|^2$, this sum is estimated by $\int_{\Omega\cap(\{X(\sigma)\in\partial B_R\}\cup\{X(\tau)\in\partial B_R\})}|x-y|^2d\pi$. Note that this set of integration coincides with $\{\exists t\in[0,1]\ X(t)\in\partial B_R\}$, as desired.

1.5. **Constructing a competitor based on the Neumann-Poisson problem.** As mentioned after Corollary 1, the remaining task is to estimate the first r. h. s. line of (18). For this, we will use Lemma 2 and construct a competitor for $W(\lambda\llcorner B_R+f,\mu\llcorner B_R+g)$ based on ϕ, the solution of the Neumann-Poisson problem (16), where we think of the measures λ,μ as having continuous densities with respect to the Lebesgue measure.

lem:comp2 **Lemma 3.**

ao20
$$W^2(\lambda\llcorner B_R+f,\mu\llcorner B_R+g)\le\frac{1}{\min\{\min_{\bar B_R}\lambda,\min_{\bar B_R}\mu\}}\int_{B_R}|\nabla\phi|^2. \tag{31}$$

Lemma 3 makes a second dilemma apparent: The intention was to use it in conjunction with Lemma 2 to obtain an estimate on the first r. h. s. line in (18). This however would require that $\min_{\bar B_R}\lambda,\min_{\bar B_R}\mu\gtrapprox 1$, so a (one-sided) closeness of μ and λ to the Lebesgue measure in a strong topology, as opposed to the closeness in a weak topology as expressed by (9). Hence this provides another reason for approximating λ and μ by more regular versions.

Proof Lemma 3. The proof is short if one uses the Benamou-Brenier formulation in its distributional version, as we shall do. For every $t\in[0,1]$ we introduce the (singular non-negative) measure

ao84
$$\rho_t:=t(\mu\llcorner B_R+g)+(1-t)(\lambda\llcorner B_R+f) \tag{32}$$

and the (t-independent) vector-valued measure

ao85
$$j_t:=\nabla\phi dx\llcorner B_R. \tag{33}$$

We note that (16) in its distributional form of (13) can be re-expressed as

ao83
$$\frac{d}{dt}\int\zeta d\rho_t=\int\nabla\zeta\cdot dj_t \tag{34}$$

for all test functions ζ. In the jargon of the Benamou-Brenier formulation, which is inspired from continuum mechanics, ρ_t is a (mass) density, j_t is a flux, and (34) is the distributional version of the continuity equation $\partial_t\rho_t+\nabla\cdot j_t=0$, expressing conservation of mass.

Following Benamou-Brenier one takes the Radon-Nikodym derivative $\frac{dj_t}{d\rho_t}$ of the (vectorial) measure j_t w. r. t. ρ_t (it plays the role of an Eulerian velocity field), and considers the expression that corresponds to the total kinetic energy:

ao86
$$\frac12\int\left|\frac{dj_t}{d\rho_t}\right|^2d\rho_t:=\sup\left\{\int\xi\cdot dj_t-\int\frac12|\xi|^2d\rho_t\right\}\in[0,\infty], \tag{35}$$

where the supremum is taken over all continuous vector fields ξ with compact support. Benamou-Brenier gives

ao87
$$W^2(\rho_0,\rho_1)\le\int_0^1\int\left|\frac{dj_t}{d\rho_t}\right|^2d\rho_tdt. \tag{36}$$

Since in our case, j_t is supported in (the open) B_R, see (33), in the r. h. s. of (35) we may restrict ourselves to ξ supported in B_R. In this case we have by definition (32) that $\int\xi\cdot dj_t-\int\frac12|\xi|^2d\rho_t=\int_{B_R}\big(\xi\cdot\nabla\phi$ $-\frac12|\xi|^2(t\mu+(1-t)\lambda)\big)$. Since we have that $\mu,\lambda>0$ a. e. , by Young's inequality in form of $\xi\cdot\nabla\phi \le \frac{1}{2}(t\mu+(1-t)\lambda)|\xi|^2+\frac{1}{2(t\mu+(1-t)\lambda)}|\nabla\phi|^2$ we thus obtain for the r. h. s. of (36)

$$\int\left|\frac{dj_t}{d\rho_t}\right|^2 d\rho_t \le \int_{B_R}\frac{|\nabla\phi|^2}{t\mu+(1-t)\lambda} \le \frac{1}{\min\{\min_{\bar B_R}\lambda,\min_{\bar B_R}\mu\}}\int_{B_R}|\nabla\phi|^2.$$

Since by definition (32), the l. h. s. of (36) coincides with the l. h. s. of (31), we are done.

2. Harmonic approximation

The purpose of this section is to establish that the displacement in an optimal plan π can locally be approximated by a harmonic gradient $\nabla\phi$ (by which we mean that for each Cartesian direction $i=1,\cdots,d$, the component $\partial_i\phi$ is harmonic, as a consequence of $-\triangle\phi=const$). This holds provided we are in the perturbative regime, see Subsection 1.2, where E and D are defined. More precisely, given any fraction $0<\theta\ll 1$, there exists a threshold $\epsilon>0$ for $E+D$ so that below that threshold, the l. h. s. of (37) is only a fraction θ of E, plus a possibly large multiple of D.

Proposition 1. *For every $\theta>0$, there exist $\epsilon(d,\theta)>0$ and $C(d,\theta)<\infty$ such that the following holds. Let π be optimal for λ,μ; provided $E+D\le\epsilon$, there exists a harmonic $\nabla\phi$ on B_1 such that*

$$\int_{(B_1\times\mathbb{R}^d)\cup(\mathbb{R}^d\times B_1)}|(y-x)-\nabla\phi(x)|^2 d\pi \le \theta E + CD, \tag{37}$$

$$\int_{B_1}|\nabla\phi|^2 \le C(E+D). \tag{38}$$

(The proof actually reveals an explicit dependence of ϵ and C on θ.) We will obtain $\nabla\phi$ by solving the Neumann-Poisson problem

$$-\triangle\phi = \frac{\mu(B_R)}{|B_R|}-\frac{\lambda(B_R)}{|B_R|} \text{ in } B_R \quad\text{and}\quad \nu\cdot\nabla\phi=\bar g-\bar f \text{ on } \partial B_R, \tag{39}$$

where $\bar f,\bar g$ are suitable regular approximations of f,g, and where the radius $R\in[1,2]$ is well-chosen. Hence by an application of Lemma 3 to this setting, we have

$$\begin{aligned}&W^2\Big(\frac{\lambda(B_R)}{|B_R|}dx\llcorner B_R+\bar f,\frac{\mu(B_R)}{|B_R|}dx\llcorner B_R+\bar g\Big)\\ &\qquad\le \frac{1}{\min\{\frac{\lambda(B_R)}{|B_R|},\frac{\mu(B_R)}{|B_R|}\}}\int_{B_R}|\nabla\phi|^2.\end{aligned}\tag{40}$$

Working with (39) creates the additional task of estimating the first r. h. s. term (21) of Lemma 2 by the l. h. s. of (40), which is conveniently done with help of the triangle inequality:

ao91

$$\begin{aligned}W(\lambda\llcorner B_R+f,\mu\llcorner B_R+g)\le\ &W(\frac{\lambda(B_R)}{|B_R|}dx\llcorner B_R+\bar f,\frac{\mu(B_R)}{|B_R|}dx\llcorner B_R+\bar g)\\ &+W(\lambda\llcorner B_R,\frac{\lambda(B_R)}{|B_R|}dx\llcorner B_R)+\text{same term with }\lambda\rightsquigarrow\mu\\ &+W(f,\bar f)+\text{same term with }f\rightsquigarrow g.\end{aligned}\tag{41}$$

Combining (21), (40) and (41) we see that the first r. h. s. line $\int_\Omega|x-y|^2d\pi-\int_{B_R}|\nabla\phi|^2$ in Corollary 1 is less than or equal to

ao94

$$\Big(1-\frac{1}{\min\{\frac{\lambda(B_R)}{|B_R|},\frac{\mu(B_R)}{|B_R|}\}}\Big)\int_{B_R}|\nabla\phi|^2\tag{42}$$

plus the square of

ao95

$$\begin{aligned}&\Big(2\int_{\{\exists t\in[0,1]\ X(t)\in\partial B_R\}}|x-y|^2d\pi\Big)^{\frac12}\\ &+W(\lambda\llcorner B_R,\frac{\lambda(B_R)}{|B_R|}dx\llcorner B_R)+\text{same term with }\lambda\rightsquigarrow\mu\\ &+W(f,\bar f)+\text{same term with }f\rightsquigarrow g.\end{aligned}\tag{43}$$

We expect (and can show for a good radius $R\in[1,2]$) that in the regime $D\ll1$, the prefactor on the r. h. s. of (42) is $\lesssim\sqrt{D}$, and that the second line in (43) is $\lesssim D$, as consistent with (37).

The main remaining work is to identify a good radius $R\in[1,2]$ and to construct $\bar f$ and $\bar g$. Again, there is a trade-off/conflict of interest:

- On the one hand, $\bar g-\bar f$ has to be sufficiently regular so that via (39), ϕ is sufficiently regular. In particular, we need (38) (with B_1 replaced by the larger B_R) to obtain that the error (42) is $o(E+D)$. This is ensured by (46) in the upcoming Lemma 4. In fact, it even yields uniform integrability of $|\nabla\phi|^2$ on B_R, which is crucial to show that also the last line in (18) is $o(E+D)$.
- On the other hand, $(\bar f,\bar g)$ has to be sufficiently close to (f,g). In particular, in view of the last term in (43) we need $W^2(f,\bar f)$ $+W^2(g,\bar g)=o(E)+O(D)$. This is ensured by (46) in the upcoming Lemma 4. Here, we have to rely on $M\ll1$ in our regime $E+D\ll1$, which is a consequence of the monotonicity of $\mathrm{supp}\pi$ by an argument not (yet) contained in these notes. $M\ll1$ will also be needed to show that the previous to last line in (18) is $o(E)+O(D)$ for a good R.

In the upcoming approximation lemma we restrict to g for brevity.

lem:approx

Lemma 4. *Consider the maximal length of trajectories that spend some time in $\bar B_2$*

ho03

$$M:=\sup\{\,|x-y|\,|\,(x,y)\in\mathrm{supp}\pi\cap\{\exists t\in[0,1]\,|\,X(t)\in\bar B_2\}\,\}.\tag{44}$$

Provided $M<1$, for every $R\in[1,2]$ exists a function $\bar g_R$ on ∂B_R such that

ao97 (45)
$$\int_1^2 W^2(g_R,\bar g_R)\,dR \le 8M(E+D),$$

ao96 (46)
$$\int_1^2\int_{\partial B_R}\bar g_R^2\,dR \le 3^d\kappa_\mu(E+D).$$

Note that we put an index R on g because the definition (6) obviously depends on R. Adding (45), divided by M, and (46), we learn that there exists $R\in[1,2]$ such that $\frac{1}{M}W^2(g_R,\bar g_R)+\int_{\partial B_R}\bar g_R^2 \le (8+3^d\kappa_\mu)$ $(E+D)$. We use the same addition argument to show that we can find a common R that suits both f and g.

Proof of Lemma 4. We fix an $R\in[1,2]$ and start with the construction of $\bar g_R$, momentarily returning to our short-hand notation $\bar g$. Let $\bar\pi$ be optimal for $W^2(\mu\llcorner B_3,\kappa_\mu dz\llcorner B_3)$; note that $\bar\pi$ is supported on $B_3\times B_3$. We extend it (trivially) by the identity to $\mathbb{R}^d\times\mathbb{R}^d$; the extension (which we still call) $\bar\pi$ is admissible for $W^2(\mu,\kappa_\mu dz\llcorner B_3+\mu\llcorner B_3^c)$. We retain

ho05 (47)
$$\int |y-z|^2 d\bar\pi = W^2(\mu\llcorner B_3,\kappa_\mu dz\llcorner B_3)\overset{(9)}{\le} D.$$

Like in the proof of the triangle inequality for the Wasserstein metric, we disintegrate $\bar\pi$ according to

ho09 (48)
$$\int \zeta(y,z)\bar\pi(dz|y)\mu(dy)=\int\zeta(y,z)\bar\pi(dydz),$$

since this family of (conditional) probability measures $\{\bar\pi(\cdot|y)\}_{y\in\mathbb{R}^d}$ allows us to define the measure $\tilde\pi$ on $\mathbb{R}^d\times\mathbb{R}^d\times\mathbb{R}^d$ via

$$\int\zeta(x,y,z)\tilde\pi(dxdydz)=\int\int\zeta(x,y,z)\bar\pi(dz|y)\pi(dxdy),$$

which has desired property

ao99 (49)
$$\begin{aligned}\text{marginal of }\tilde\pi\text{ w. r. t. }(x,y)\quad&=\quad\pi,\\ \text{marginal of }\tilde\pi\text{ w. r. t. }(y,z)\quad&=\quad\bar\pi.\end{aligned}$$

We still associate the trajectory X to a pair (x,y) as in (1). We extend this definition by associating to a triplet (x,y,z) the continuous piecewise affine trajectory defined by (1) followed by

$$X(t)=(t-1)z+(2-t)y\quad\text{for } t\in[1,2].$$

We now are interested in the distribution κ of the endpoint $z=X(2)$ of those trajectories that exit $\bar B_R$ during the first time interval $[0,1]$,

i. e. those trajectories $X \in \Omega$ with $X(\tau) \in \partial B_R$. This distribution is defined by

$$\int \zeta d\kappa = \int_{\Omega \cap \{X(\tau) \in \partial B_R\}} \zeta(z) \tilde{\pi}(dxdydz). \tag{50}$$

Starting from the triangle inequality in form of $|y| \le |y - X(\tau)| + R \le |x - y| + R$, we see that our assumption $M < 1$, cf. (49), in conjunction with $R \le 2$ leads to $y = X(1) \in B_3$ for $(x, y) \in \mathrm{supp}\pi$. By the first item in (49), this extends to $(x, y, z) \in \mathrm{supp}\tilde{\pi}$. By the second item in (49), this yields $z = X(2) \in B_3$ by $\bar{\pi}(B_3 \times B_3^c) = 0$ according to our extension of $\bar{\pi}$. Hence we have

$$\kappa \text{ is supported in } B_3. \tag{51}$$

Therefore by the second item in (49) and the admissibility of $\bar{\pi}$ for $W^2(\mu, \kappa_\mu dz \llcorner B_3 + \mu \llcorner B_3^c)$, we obtain

$$\int \zeta d\kappa \le \int_{\{z \in B_3\}} \zeta(z) \bar{\pi}(dydz) = \kappa_\mu \int_{B_3} \zeta \quad \text{provided } \zeta \ge 0.$$

Hence κ has a Lebesgue density, we still denote by κ, that satisfies

$$\kappa \le \kappa_\mu \quad \text{Lebesgue almost surely.} \tag{52}$$

Finally, we radially project κ onto ∂B_R:

$$\int \zeta d\bar{g} = \int \zeta(R\frac{z}{|z|}) \kappa(dz). \tag{53}$$

This concludes the construction of $\bar{g}$, we now turn to its estimate.

We start with (45) and note that an admissible plan for $W^2(g, \bar{g})$ is given by

$$\int_{\Omega \cap \{X(\tau) \in \partial B_R\}} \zeta(X(\tau), R\frac{z}{|z|}) d\tilde{\pi}.$$

Indeed, on the one hand, for ζ only depending on the first variable, $\tilde{\pi}$ may be replaced by π according to the first item in (49) so that we obtain $\int \zeta dg$ by its definition (6). On the other hand, for ζ only depending on the second variable, we obtain $\int \zeta d\bar{g}$ by combining (50) and (53). Hence we have

$$W^2(g, \bar{g}) \le \int_{\Omega \cap \{X(\tau) \in \partial B_R\}} |X(\tau) - R\frac{z}{|z|}|^2 d\tilde{\pi}. \tag{54}$$

Since $X(\tau) \in \partial B_R$, an elementary geometric argument on the radial projection yields for the integrand $|X(\tau) - R\frac{z}{|z|}| \le 2|X(\tau) - z|$, which in turn is $\le 2(|x - y| + |y - z|)$. We also note that for X in the domain of integration $\Omega \cap \{X(\tau) \in \partial B_R\}$ we have $\min_{[0,1]} |X| \le R \le \max_{[0,1]} |X|$; by $R \le 2$ and definition (44), this implies $\min_{[0,1]} |X| \le$

$R \le \min_{[0,1]}|X| + M$ on the support of $\tilde\pi$. We recall that by $M \le 1$ and $R \le 2$ we also have $y \in B_3$ there. Summing up, we infer from (54)

$$\begin{aligned}
&W^2(g,\bar g)\\
&\le 4\int_{\{y\in B_3\}\cap\{\min_{[0,1]}|X|\le R\le \min_{[0,1]}|X|+M\}} (|x-y|+|y-z|)^2 d\tilde\pi.
\end{aligned}$$

We now (re)introduce the index R and integrate over $R \in [1,2]$; expressing the domain of integration $\{\min_{[0,1]}|X| \le R \le \min_{[0,1]}|X|+M\}$ by a characteristic function and exchanging the order of integration we obtain

$$\int_1^2 W^2(g_R,\bar g_R)dR \le 4M\int_{\{y\in B_3\}} (|x-y|+|y-z|)^2 d\tilde\pi.$$

The use of $(|x-y|+|y-z|)^2 \le 2(|x-y|^2+|y-z|^2)$ allows us to appeal to the compatibility (49):

$$\int_1^2 W^2(g_R,\bar g_R) \le 8M\Big(\int_{\mathbb{R}^d\times B_3}|x-y|^2 d\pi + \int |y-z|^2 d\bar\pi\Big).$$

By definition (10) and by (47) this turns into (45).

In preparation for establishing (46), we first provide an estimate of the measure κ defined in (50), which shows that it is concentrated near ∂B_R, see (55). By definition (50) we have

$$\int ||z|-R|d\kappa = \int_{\Omega\cap\{X(\tau)\in\partial B_R\}} ||z|-R|\tilde\pi(dxdydz).$$

Since $|X(\tau)| = R$, we may write $||z|-R| = ||z|-|X(\tau)|| \le |x-y| + |y-z|$. By the same argument on the domain of integration as after (54), we obtain

$$\begin{aligned}
&\int ||z|-R|d\kappa\\
&\le \int_{\{y\in B_3\}\cap\{\min_{[0,1]}|X|\le R\le \max_{[0,1]}|X|\}} (|x-y|+|y-z|)\tilde\pi(dxdydz).
\end{aligned}$$

Making the index R appear again and integrating in R, this gives

$$\begin{aligned}
&\int_1^2\int ||z|-R|d\kappa_R\,dR\\
&= \int_{\{y\in B_3\}} (\max_{[0,1]}|X| - \min_{[0,1]}|X|)(|x-y|+|y-z|)\tilde\pi(dxdydz).
\end{aligned}$$

Using $\max_{[0,1]}|X| - \min_{[0,1]}|X| \le |x-y|$ and then Young's inequality in form of $|x-y|(|x-y|+|y-z|) \le \frac{3}{2}|x-y|^2 + \frac{1}{2}|y-z|^2$ we thus obtain from (49)

$$\int_1^2 \int ||z| - R| d\kappa_R \, dR$$
$$= \frac{3}{2} \int_{\{y \in B_3\}} |x-y|^2 \pi(dxdy) + \frac{1}{2} \int |y-z|^2 \bar{\pi}(dydz).$$

By definition (10) and by (47) this turns into

ho07 (55)
$$\int_1^2 \int ||z| - R| d\kappa_R \, dR \le \frac{3}{2} E + \frac{1}{2} D.$$

The passage from (55) to (46) relies on

ho08 (56)
$$\int_{\partial B_R} \frac{1}{2} \bar{g}^2 \le 3^{d-1} \kappa_\mu \int ||z| - R| d\kappa.$$

Indeed, by (52), (56) reduces to

ho12 (57)
$$\int_{\partial B_R} \frac{1}{2} \bar{g}^2 \le 3^{d-1} (\mathrm{esssup} \kappa) \int ||z| - R| \kappa dz.$$

Introducing polar coordinates $z = r\hat{z}$ with $r \in (0, \infty)$ and $\hat{z} \in \partial B_1$, which are natural to re-express (53), (57) reduces to the single-variable statement

$$\frac{1}{2} \Big(\int \kappa r^{d-1} dr \Big)^2 \le 3^{d-1} (\mathrm{esssup} \kappa) \int |r - R| \kappa r^{d-1} dr.$$

It is convenient to rephrase this in terms of $\bar{\kappa} = \kappa r^{d-1}$; since because of (51) we have $\mathrm{esssup} \bar{\kappa} \le 3^{d-1} \mathrm{esssup} \kappa$, it suffices to show for an arbitrary function $\bar{\kappa} \ge 0$ of $r \in (-\infty, \infty)$ that

ho13 (58)
$$\frac{1}{2} \Big(\int \bar{\kappa} dr \Big)^2 \le (\mathrm{esssup} \bar{\kappa}) \int |r - R| \bar{\kappa} dr.$$

The argument for (58) is elementary: By translation in r, we may assume $R = 0$; by homogeneity in $\bar{\kappa}$, we may assume $\mathrm{esssup} \bar{\kappa} = 1$, that is, $\bar{\kappa} \in [0, 1]$. We now change perspective and seek to minimize the r. h. s. $\int |r| \bar{\kappa} dr$ under constraining the l. h. s. through prescribing $m = \int \bar{\kappa} dr$. By linearity of $\int |r| \bar{\kappa} dr$ in $\bar{\kappa}$, this functional assumes its minimum on extremal points w. r. t. the constraints $\bar{\kappa} \in [0, 1]$ and $\int \bar{\kappa} dr = m$. Those are characteristic functions of sets of Lebesgue measure m. Clearly, the set I with $|I| = m$ that minimizes $\int_I |r| dr$ is given by $I = [-\frac{m}{2}, \frac{m}{2}]$; the minimum is $\frac{m^2}{2}$, as desired.