Asia-Pacific Analysis and PDE Seminar

A variational approach to the regularity theory for optimal transportation

Felix Otto, Max Planck Institute for Mathematics in the Sciences,

Leipzig, Germany

joint work with Michael Goldman (arXiv '17, Ann. ENS),
with MG & Martin Huesmann (arXiv '18, '19, CPAM)
with Tatsuya Miura (arXiv '20)
with Maxime Prod'homme & Tobias Ried (arXiv '20)

Follow de Giorgi's approach for Minimal Surfaces

Harmonic approximation $\overset{\text{elliptic regularity, change of variables}}{\Longrightarrow}$ One-Step improvement $\overset{\text{Campanato iteration}}{\Longrightarrow}$ ϵ-regularity $\overset{\text{measure theory}}{\Longrightarrow}$ partial regularity.

Here $C^{2,\alpha}$-regularity for Kantorowicz potential ψ

i. e. $C^{1,\alpha}$-regularity for Brenier map $T = \nabla\psi$.

Purely variational/ no maximum principle/ avoids Caffarelli's theory

Applications (old and new):

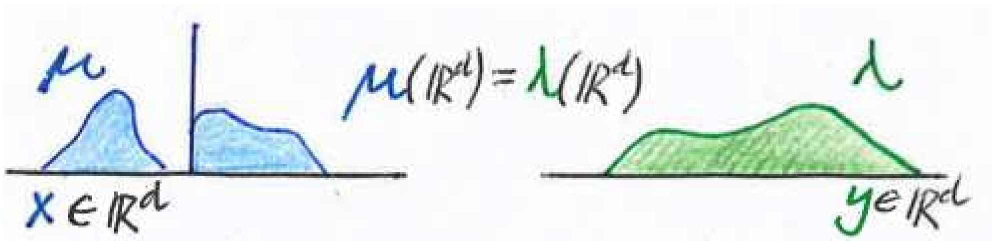
Partial regularity [Figalli-Kim, DePhilippis-Figalli],

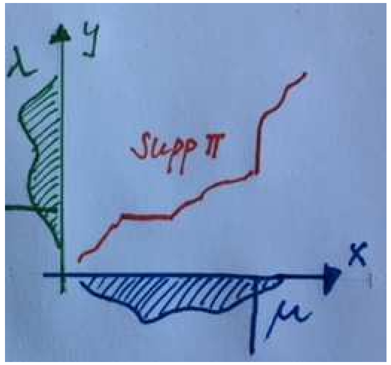
ϵ-regularity at boundary for $\partial\Omega \in C^{1,\alpha}$ [Chen-Figalli],

"Matching" Poisson to Lebesgue [Parisi et. al, Ambrosio et. al]

Recall Optimal Transportation

Given two measures μ and λ





seek transfer plan π, i. e. $\pi(U \times \mathbb{R}^d) = \mu(U)$, $\pi(\mathbb{R}^d \times V) = \lambda(V)$
that minimizes Euclidean transport cost $\int_{\mathbb{R}^d \times \mathbb{R}^d} |y - x|^2 \pi(dxdy)$.

Minimum $=: W^2(\mu, \lambda)$ (squared) Wasserstein distance.

From optimal transportation to Monge-Ampère

Minimize $\int_{\mathbb{R}^d \times \mathbb{R}^d} |y-x|^2 \pi(dxdy)$
among all $\pi(dxdy)$ with marginals $\mu(dx)$ and dy.

Support of optimal π is cyclically monotone; hence
$\exists$ convex ψ supp$\pi \subset \{\,(x,y)\,|\,y \in$ subgradient $\partial\psi(x)\,\}$.

$\forall$ test functions ζ $\int \zeta(\nabla\psi(x))\mu(dx) = \int \zeta(y)dy$.

In smooth case, this amounts to $\det D^2\psi = \mu$,
an instance of the Monge-Ampère equation.

Nature of the Monge-Ampère equation

Recall Monge-Ampère: $\det D^2\psi = 1$.
Fully non-linear with $F(A) := \det A - 1$.

However elliptic: $F(A) > F(A')$ for $A > A' \geq 0$;
satisfies comparison principle.

However degenerate: $\leftrightarrow$ affine invariant (non-compact SL(n)).
Cf. Laplacian $F(A) = \mathrm{tr} A$: rotation invariant (compact SO(n-1)).

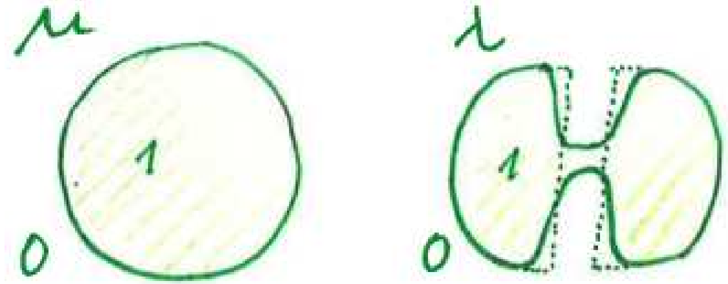
Caffarelli's '90 breakthrough:
comparison principle, affine invariance, compactness.

Monge-Ampère equation at crossroads
of fully nonlinear and variational.

Singularities in optimal transportation are generic

Caffarelli's example:

smooth data μ, λ
do not yield smooth $T = \nabla\psi$.



Riemannian setup: Loeper's example, Ma-Trudinger-Wang theory.

Thus ϵ-regularity is of interest:

$\int |y\text{-}x|^2 \pi(dxdy) \leq \epsilon$ locally, μ, λ smooth locally

$\Longrightarrow$ Kantorowicz potential ψ smooth locally.

Relies on harmonic approximation in our approach:

$\int |y\text{-}x|^2 \pi(dxdy) \leq \epsilon$ locally, $\mu, \lambda \approx 1$ locally

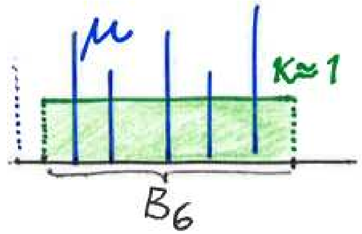
$\Longrightarrow$ displacement $(y\text{-}x)\pi(dxdy) \approx \nabla$harmonic locally.

At the core: harmonic approximation

Local energy $\quad E := \int_{(B_6\times\mathbb{R}^d)\cup(\mathbb{R}^d\times B_6)} |y-x|^2\pi(dxdy),$

Local data size^2 $\quad D := W_{B_6}^2(\mu,\kappa) + (\kappa-1)^2$

$+$ same for λ



Proposition 1 (Goldman&Huesmann&O.)

$\forall\ \tau>0 \quad \exists\ \epsilon(\tau,d)>0,\ C(\tau,d)<\infty \quad$ s. t. $\quad E+D\le\epsilon \Longrightarrow$

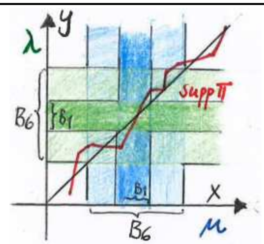
$\exists\ \nabla\phi$ harmonic, $\quad \int_{B_3}|\nabla\phi|^2 \le C(E+D),$

$$\int_{(B_1\times\mathbb{R}^d)\cup(\mathbb{R}^d\times B_1)} \left|y-x-\nabla\phi(\tfrac{1}{2}x+\tfrac{1}{2}y)\right|^2\pi(dxdy) \le \tau E + CD.$$

Amounts to:

Displacement $\ y-x$

$\approx$ harmonic gradient $\ \nabla\phi$



Harmonic approximation: correct homogeneities ...

$E := \int_{(B_6\times\mathbb{R}^d)\cup(\mathbb{R}^d\times B_6)} |y-x|^2\pi(dxdy)$, quadratic in solution,

$D := W^2_{B_6}(\mu,\kappa) + (\kappa-1)^2 +$ same for λ, quadratic in data.

Proposition 1

$\forall\ \tau>0 \quad \exists\ \epsilon(\tau,d)>0,\ C(\tau,d)<\infty \quad$ s. t. $\quad E+D\le\epsilon \implies$

$\exists\ \nabla\phi$ harmonic, $\quad \int_{B_2}|\nabla\phi|^2 \le C(E+D)$,

$$\int_{(B_1\times\mathbb{R}^d)\cup(\mathbb{R}^d\times B_1)} \left|y-x-\nabla\phi(\tfrac{1}{2}x+\tfrac{1}{2}y)\right|^2\pi(dxdy) \le \tau E + CD.$$

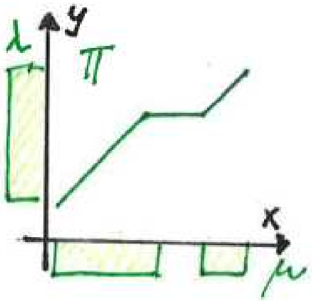
Compare to $\int_{B_1} L(\nabla u-\nabla\phi) \le \tau\int_{B_6} L(\nabla u) + C\int_{B_6}|f|^2$

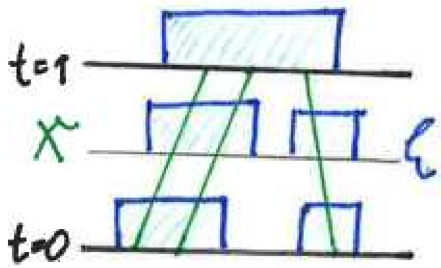
for $\ -\nabla\cdot DL(\nabla u) = \nabla\cdot f\ $ with uniformly convex L .

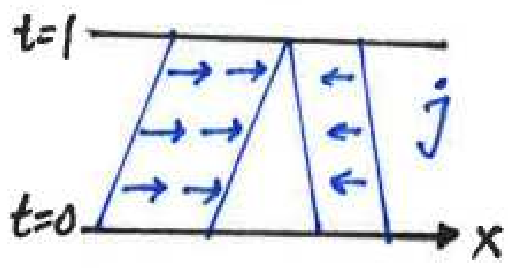
... and correct metric

From Lagrangian to Eulerian (Benamou-Brenier)

Transport plan π, trajectories $X(t)$, density/flux (ρ, j)







Continuity eqn. $\partial_t \rho + \nabla \cdot j = 0$, kinetic energy $\int_{\mathbb{R}^d \times (0,1)} \frac{1}{\rho} |j|^2$.

$$W^2(\mu, \lambda) = \inf\{ \int \frac{1}{\rho} |j|^2 \mid \partial_t \rho + \nabla \cdot j = 0,\ \rho(t{=}0){=}\mu,\ \rho(t{=}1){=}\lambda \}$$

Kinetic energy density $(\rho, j) \mapsto \frac{1}{\rho} |j|^2$ is strictly convex.

Linearization $\frac{1}{\rho} |j|^2 \rightsquigarrow |j|^2$ amounts to $W^2(\mu, \lambda) \rightsquigarrow \|\mu - \lambda\|^2_{\dot{H}^{-1}}$.

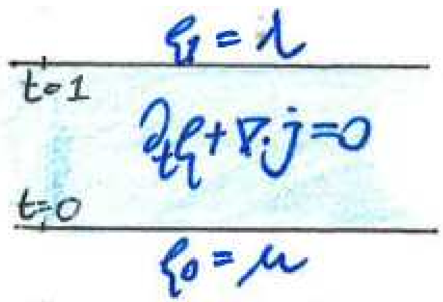
Analogy for minimal surfaces: varifolds vs. currents.

Eulerian version of harmonic approximation

Density/flux $(\rho, j) = (\rho_t dt, j_t dt)$ where

$\int \zeta d\rho_t = \int \zeta(ty+(1\text{-}t)x)\pi(dxdy), \ \int \xi \cdot dj_t = \int \xi(ty+(1\text{-}t)x)\cdot(y\text{-}x)\pi(dxdy)$

continuity equation



$E \rightsquigarrow \int_{B_5\times(0,1)} \frac{1}{\rho}|j|^2$

reveals strict convexity of variational problem

Proposition 1'

$\forall \ \tau > 0 \quad \exists \ \epsilon(\tau, d) > 0, \ C(\tau, d) < \infty \quad$ s. t. $\quad E + D \leq \epsilon \Longrightarrow$

$\exists \ \nabla\phi$ harmonic, $\quad \int_{B_2} |\nabla\phi|^2 \leq C(E + D),$

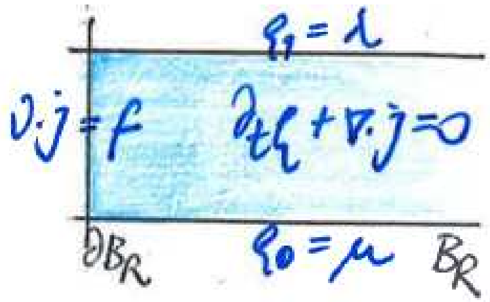
$\int_{B_2\times(0,1)} \frac{1}{\rho}\left|j - \rho\nabla\phi\right|^2 \leq \tau E + CD.$

Amounts to: Eulerian velocity $\frac{j}{\rho} \approx \nabla\phi$ harmonic gradient

Construct $\nabla\phi$ via flux (Neumann) data

Normal flux $f := \nu \cdot j$ across bdry ∂B_R,

its time integral $\bar{f} := \int_0^1 f\,dt$.



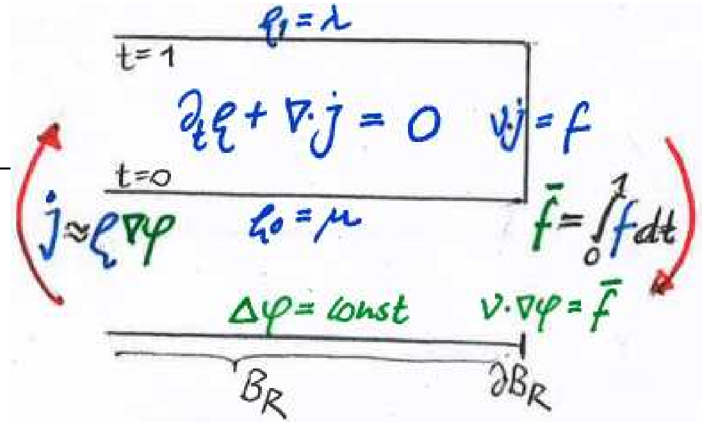
Proposition 1” $\forall\, \tau > 0 \;\exists\, \epsilon > 0, C < \infty :\; E + D \leq \epsilon \implies$
$\exists\, R \in (3,4)$ s. t. $\triangle\phi = \text{const}$ in B_R, $\nu \cdot \nabla\phi = \bar{f}$ on ∂B_R

satisfies $\displaystyle\int_{B_2} |\nabla\phi|^2 \leq C(E+D)$,

$$\int_{B_2\times(0,1)} \frac{1}{\rho}\left|j - \rho\nabla\phi\right|^2 \leq \tau E + CD.$$

cf. Dacorogna-Moser.

Choice of “good” radius R.



Old application: ϵ-regularity as Schauder theory

Recall Hölder semi-norms $[f]_{\alpha,B} := \sup_{x\neq x'\in B}\frac{|f(x)-f(x')|}{|x-x'|^\alpha}$, $\alpha\in(0,1)$.

Suppose $\mu = f\,dx,\ \lambda = g\,dy$

with Hölder continuous f, g and $f(0) = g(0) = 1$.

Monitor (the dimensionless) $E := \frac{1}{R^2|B_R|}\int_{\Omega\cap B_{2R}}|T-x|^2dx$

with $T = \nabla\psi$ Brenier map.

Monitor (the dimensionless) $D := R^{2\alpha}[f]^2_{\alpha,B_{2R}} + R^{2\alpha}[g]^2_{\alpha,B_{2R}}$.

Theorem 1 (Goldman&O., à la DePhilippis&Figalli)
If $E + D \ll 1$ then $R^{2\alpha}[\nabla T]^2_{\alpha,B_R} \lesssim E + D$.

Amounts to $C^{2,\alpha}$-regularity for Monge-Ampère $\det D^2\psi = f$ ($g\equiv 1$).

Comparison: DePhilippis&Figalli vs. Goldman&O.

Theorem 1 (Goldman&O., à la DePhilippis&Figalli)
$E+[f]^2_{\alpha,B_2}+[g]^2_{\alpha,B_2}\ll 1 \implies [D^2\psi]^2_{\alpha,B_1}\lesssim E+[f]^2_{\alpha,B_2}+[g]^2_{\alpha,B_2}.$

Perturbation around linear $\triangle\phi = const$,
as opposed to nonlinear $\det D^2\psi = 1$
(following Caffarelli '92 approach to $W^{2,p}/C^{2,\alpha}$, but using ϵ-regularity Figalli&Kim).

Get (immediately) linear homogeneities $[D^2\psi]_{\alpha,B_1}\lesssim [f]_{\alpha,B_2}+[g]_{\alpha,B_2}$.

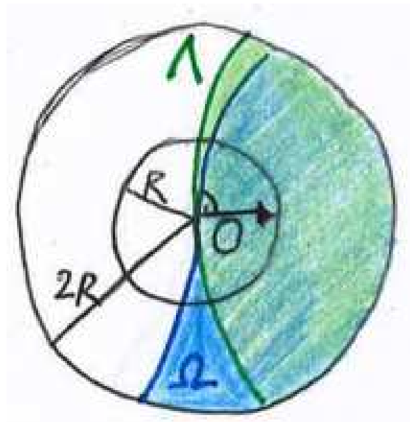
Get $\psi\in C^{2,\alpha}$ in one step (also for general cost functions),
as opposed to three-step bootstrap $C^{1,\alpha}$, $C^{1,1}$, $C^{2,\alpha}$.

Competitors in strictly convex variational problem,
instead of comparison principle.

Refined application: Boundary ϵ-regularity

Suppose $\mu = dx_{|\Omega}$, $\lambda = dy_{|\Lambda}$ for two open sets $\Omega, \Lambda \subset \mathbb{R}^d$.

Consider $E := \frac{1}{R^2|B_R|}\int_{\Omega\cap B_{2R}} |T-x|^2 dx$ where T Brenier map



Topological condition:

$T(\Omega \cap B_R) \subset \Lambda \cap B_{2R}, \quad T^{-1}(\Lambda \cap B_R) \subset \Omega \cap B_{2R}.$

Geometric condition:

$\Omega \cap B_{2R} = \{x_1 > g(x')\} \cap B_{2R}$, $\Lambda \cap B_{2R} = \{x_1 > h(x')\} \cap B_{2R}$

with $g(0) = h(0) = 0$, $\nabla' g(0) = \nabla' h(0) = 0$.

Consider $D := R^{2\alpha}[\nabla' g]^2_{\alpha, B_{2R}} + R^{2\alpha}[\nabla' h]^2_{\alpha, B_{2R}}$.

Theorem 2 (Miura&O., à la Chen-Figalli)

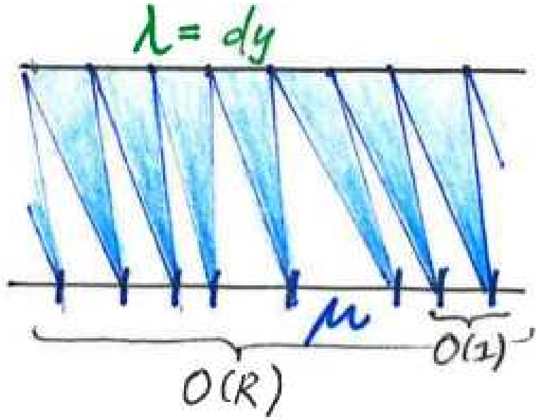
If $E + D \ll 1$ then $R^{2\alpha}[\nabla T]^2_{\alpha, B_R} \lesssim E + D$

Optimal by Jhaveri '19

New application: Quantitative linearization ...

Now general μ but $\lambda = dy$ Lebesgue. Expect:

$\mu \approx dx \implies y - x \approx \nabla u$ under π, where $\triangle u = \mu - 1$.



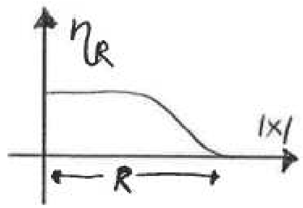
With an eye on
matching Poisson (scale 1)
to Lebesgue (on scale $R \gg 1$),

want: $\mu \approx dx$ *on large-scale average*,

eg. control of $\frac{1}{|B_R|} W^2_{B_R}(\mu, \kappa) + (\kappa - 1)^2$

$\implies$ $y - x \approx \nabla u$ *on large-scale average*,

eg. control of $\int \eta_R (y - x)\pi(dxdy) - \int \eta_R \nabla u$.

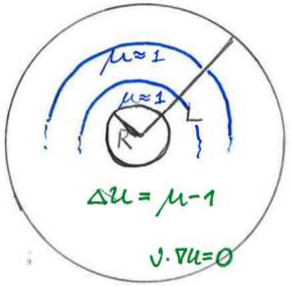


... on large scales for matching problems

Quantitative linearization on large scales

Admissible rate function:

$$\frac{1}{C}\sum_{k=0}^{\infty}\frac{\beta(2^kR)}{2^kR}\le\frac{\beta(R)}{R}\le C\quad\text{for}\quad R\ge 1.$$



Poisson: $\beta(R)\sim\ln R$ for $d=2$,

$\beta(R)\sim 1$ for $d>2$.

Theorem 3 (Goldman&Huesmann&O.)

Suppose $\operatorname{supp}\mu\subset B_L$, $\mu(\mathbb{R}^d)=|B_L|$

and $\frac{1}{|B_R|}W^2_{B_R}(\mu,\kappa)+(\kappa\text{-}1)^2\le\beta(R)$ for all $1\le R\le L$.

Then $$\left|\frac{\int\eta_{R(x)}(y-x)\pi(dxdy)}{\int\eta_{R(x)}\pi(dxdy)}-\frac{\int\eta_R\nabla u}{\int\eta_R}\right|\lesssim\frac{\beta(R)}{R},\ 1\ll R\le L,$$

where $\triangle u=\mu-1$ in B_L, $\nu\cdot\nabla u=0$ on ∂B_L.

Harmonic approximation even interesting for μ, λ (locally) constant

Suppose $\mu = dx$ and $\lambda = dy$ on B_6 i.e. $D = 0$

Recall: $f := \nu \cdot j$, $\bar{f} := \int_0^1 f\,dt$, $E := \int_{B_5 \times (0,1)} \frac{1}{\rho}|j|^2$.

Lemma 1 (Goldman&O.)

$\exists\ R \in (3,4)$ s. t. $\left\{ \begin{array}{rcl} \triangle\phi & = & 0 \text{ in } B_R, \\ \nu \cdot \nabla\phi & = & \bar{f} \text{ on } \partial B_R \end{array} \right\}$ satisfies

$$\int_{B_R} |\nabla\phi|^2 \le C(d)E,$$

$$\int_{B_R \times (0,1)} \frac{1}{\rho}\left|j - \rho\nabla\phi\right|^2 \le C(d)E^{\frac{d+2}{d+1}}.$$

transport velocity $\frac{j}{\rho} \approx \nabla\phi$ gradient of harmonic function

Reflects that Wasserstein $W_2 \approx \dot{H}^{-1}$ for densities ≈ 1.

Ingredients for Lemma 1

Lemma 1: $\exists\,\phi$ harmonic such that

$$\int_0^1\int_{B_1}\frac{1}{\rho}|j\text{-}\rho\nabla\phi|^2 \lesssim \Big(\int_0^1\int_{B_2}\frac{1}{\rho}|j|^2\Big)^{\frac{d+2}{d+1}}.$$

1) Construction of ϕ: $\triangle\phi = 0$ in B_R, $\nu\cdot\nabla\phi = \bar f$ on ∂B_R,

where $\bar f := \int_0^1 f$, $f := \nu\cdot j$

and $R\in[1,2]$ such that $\int_0^1\int_{\partial B_R}\frac{1}{\rho}|j|^2 \lesssim \int_0^1\int_{B_2}\frac{1}{\rho}|j|^2$ (good radius)

2) Orthogonality:

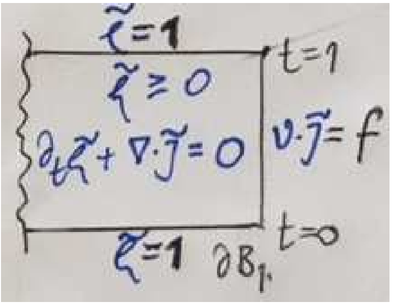
$$\int_0^1\int_{B_R}\frac{1}{\rho}|j\text{-}\rho\nabla\phi|^2 \le \int_0^1\int_{B_R}\frac{1}{\rho}|j|^2 - \int_{B_R}|\nabla\phi|^2.$$

Follows from McCann's *displacement convexity* in form of $\rho\le 1$.

3) Construction of competitor: $(\tilde\rho,\tilde j)$ admissible with bc of (ρ,j) and

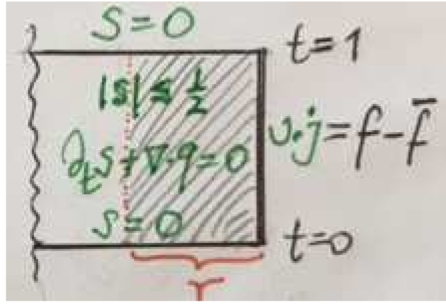
$$\int_0^1\int_{B_R}\frac{1}{\tilde\rho}|\tilde j|^2 - \int_{B_R}|\nabla\phi|^2 \lesssim \Big(\int_0^1\int_{\partial B_R} f^2\Big)^{\frac{d+2}{d+1}}.$$

Ingredients for Lemma 1: boundary layer construction



Ansatz $(\tilde{\rho}, \tilde{j}) = (1, \nabla\phi) + (s, q)$

$\mathrm{supp}(s, q) \subset [0,1] \times (\bar{B}_R - B_{R-r})$



Enough to construct (s, q) such that $|s| \le \frac{1}{2}$ and

$$\int_0^1 \int_{B_R} |q|^2 \lesssim r \int_0^1 \int_{\partial B_R} (f\text{-}\bar{f})^2 \quad \text{for} \quad r \gg \Big(\int_0^1 \int_{\partial B_R} (f\text{-}\bar{f})^2\Big)^{\frac{1}{d+1}}.$$

Reduce to *isoperimetric estimate* $\Big(\int_0^1 \int_{\partial B_R} (\zeta - \bar{\zeta})^2\Big)^{\frac{1}{2}}$

$$\lesssim r^{\frac{1}{2}} \Big(\int_0^1 \int_{B_R - B_{R-r}} |\nabla\zeta|^2\Big)^{\frac{1}{2}} + r^{-\frac{d+1}{2}} \int_0^1 \int_{B_R - B_{R-r}} |\partial_t \zeta|.$$

Reminiscent of Alberti&Choksi&O. '09.

Analogies to minimal surfaces (Schoen&Simon '82)

Approximate minimal surface by harmonic graph /
approximate displacement by harmonic gradient.

Use: Object is minimizing under compact perturbations.
Don't use: Euler-Lagrange equation (= first variation).

Mismatch of type of boundary condition
for construction of harmonic competitor:
graph vs. non-graph / time-averaged vs. time-resolved;
Lower-dimensional isoperimetric estimate:
error is of higher-order (choice of good radius).

Use of strict convexity to convert energy gap
into distance ("approximate orthogonality");
need to smooth out boundary data.

From Lemma 1 to Theorem 1

Inner regularity for ϕ. Eulerian $(\rho, j) \rightsquigarrow$ Lagrangian T.

For $\quad Q \in \mathbb{R}^{d\times d}$, $b \in \mathbb{R}^d \quad$ consider

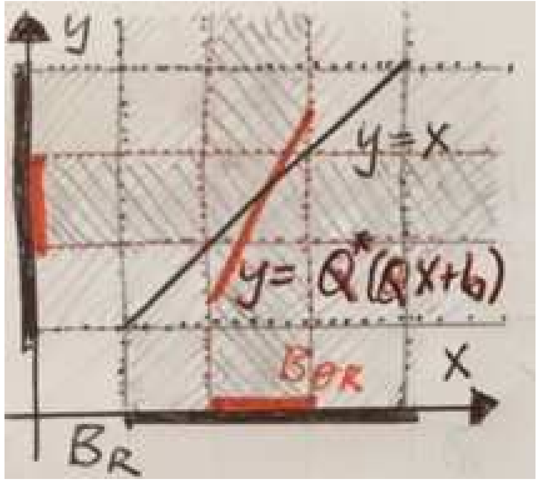
$$E(R,Q,b) := \frac{1}{R^{d+2}}\int_{B_R}|Q^{-*}T-(Qx+b)|^2dx$$

$E(R) := E(R,\mathrm{id},0) \ll 1 \Longrightarrow$

$\exists\, Q, b \quad \forall\, 0<\theta\ll 1$

$$E(\theta R, Q, b) \lesssim \theta^2 E(R) + \theta^{-(1+\frac{d}{2})}E^{\frac{d+2}{d+1}}(R)$$

$$|Q-\mathrm{id}|^2 + \frac{1}{R^2}|b|^2 \lesssim E(R)$$



Affine change of variables $\hat{x} = Qx+b$, $\hat{y} = Q^{-*}y$

preserves optimality of transport. Campanato iteration.

Changes for rough μ, λ, flux boundary data

Recall: $D := W^2_{B_6}(\mu,\kappa)+(\kappa\text{-}1)^2, \quad E := \int_{(B_6\times\mathbb{R}^d)\cup(\mathbb{R}^d\times B_6)} |y-x|^2\pi(dxdy).$

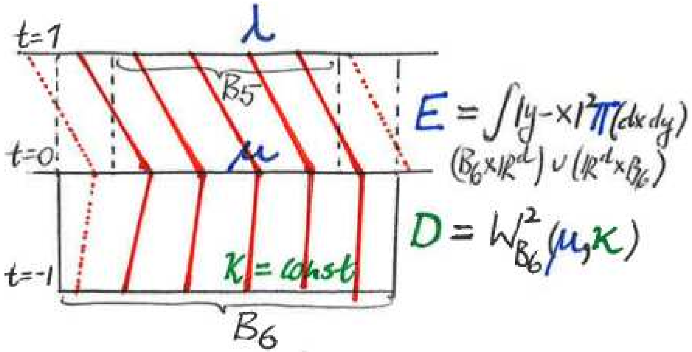
Issue: $\bar{f} = \int_0^1 f\,dt, \quad f = \nu\cdot j$ measure $\notin L^2$.

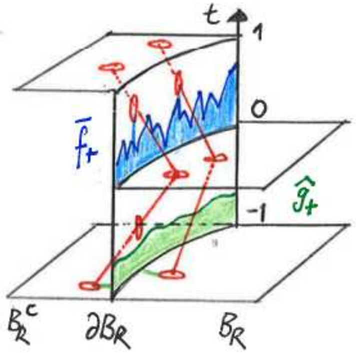
Replace $\bar{f}$ by $\hat{g}\in L^2$ in

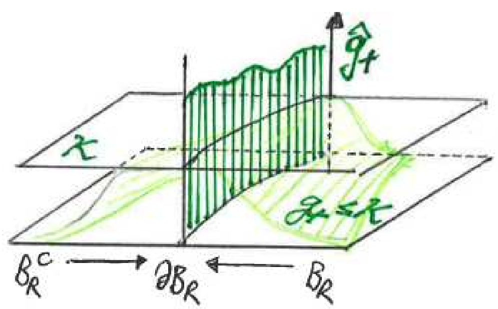
$$\triangle\phi = c \text{ in } B_R, \qquad \nu\cdot\nabla\phi = \hat{g} = \hat{g}_+ - \hat{g}_- \text{ on } \partial B_R$$

Lemma (nonlin. L^2-approx.) $\exists\, R\in(3,4)$, $\hat{g}_\pm$ on ∂B_R s. t.

$$\int_{\partial B_R} \hat{g}_\pm^{\,2} \lesssim E+D, \qquad W^2_{\partial B_R}(\hat{g}_\pm,\bar{f}_\pm) \lesssim (E+D)^{1+\frac{1}{d+2}}$$







Changes for rough μ, λ, L^∞-bound

Recall: $D := W^2_{B_6}(\mu,\kappa)+(\kappa\text{-}1)^2, \quad E := \int_{(B_6\times\mathbb{R}^d)\cup(\mathbb{R}^d\times B_6)} |y-x|^2\pi(dxdy).$

Lemma (nonlinear L^2-approximation)

$\exists R \in (3,4)$, $\hat{g}_\pm$ on ∂B_R s. t.

$$\int_{\partial B_R} \hat{g}_\pm^2 \lesssim E+D, \quad W^2_{\partial B_R}(\hat{g}_\pm, \bar{f}_\pm) \lesssim (E+D)^{1+\frac{1}{d+2}}$$

Superlinear exponent comes from

Lemma (L^∞-estimate)

For $(x_0, x_1) \in \mathrm{supp}\pi \cap (B_5 \times \mathbb{R}^d) \cup (\mathbb{R}^d \times B_5)$

$$|x_0 - x_1| \lesssim (E+D)^{\frac{1}{d+2}}$$

Follows from monotonicity of suppπ

Summary and outlook

Harmonic approximation $\Longrightarrow$ One-Step improvement
$\overset{\text{Campanato iteration}}{\Longrightarrow}$ small or large-scale-regularity

Applications: ϵ-regularity at boundary for $\partial\Omega \in C^{1,\alpha}$

[Savin-Yu, Chen-Figalli, Chen-Liu-Wang],

Matching Poisson to Lebesgue

[Sturm-Huesmann, Parisi et.al, Ambrosio et.al, Ledoux, Talagrand]

general cost functions $c \in C^{2,\alpha}$ with Prod'homme & Ried,

Next step:

thermodynamic limit for Matching in $d = 2$.