Mathematisches Forschungsinstitut Oberwolfach

Report No. 7/2017

DOI: 10.4171/OWR/2017/7

Applications of Optimal Transportation in the Natural Sciences

Organised by
Jean-David Benamou, Le Chesnay
Virginie Ehrlacher, Marne-la-Vallée
Daniel Matthes, Garching

29 January – 4 February 2017

Abstract. The aim of this workshop was to gather a mixed group of experts and young researchers from different areas of applied mathematics in which optimal transport plays a central role. Applications in one of the classical areas of natural sciences, like physics, chemistry and (mathematical) biology were the main focus of the workshop.

Mathematics Subject Classification (2010): 28-XX, 35-XX, 81-XX, 92-XX.

Introduction by the Organisers

The conference was organized by Jean-David Benamou (INRIA), Daniel Matthes (TUM) and Virginie Ehrlacher (Ecole des Ponts Paristech). The program consisted of 30 lectures, given by senior and junior researchers.

The progress in the mathematical theory of optimal transportation is driven to a large extent by applications in natural sciences. Advanced understanding of the properties of transport maps and of Wasserstein gradient flows is going hand in hand with a better – typically more geometric – understanding of fluid-type models in physics, or more recently, in quantum chemistry, mathematical biology, and many other areas. On the one hand, it is now almost two decades since one of the most important results in optimal transport theory has been established, namely the "geometrization" of diffusion processes by using transport techniques for the analysis of the long-time asymptotics of particle systems. On the other hand, the current activity on the application of optimal of transport methods is intense and diverse, ranging from geometric optics to models for crowd motion. This was a very favorable moment for a meeting to discuss some recent achievements of the

theory in its applications to the natural sciences and to determine future directions of research.

The aim of the workshop was to bring together a mixed group of experts and young researchers from different areas of applied mathematics in which optimal transportation plays a role. Thematically, the talks illustrated the variety of applications of optimal transport theory in the natural sciences: physics, chemistry and biology. The diversity of the topics and participants stimulated a lot of fruitful discussion between the persons working in the different fields and gave rise to new collaborations, in particular for the younger generation of researchers.

In total, 52 scientists participated in this meeting; almost 40 came from countries other than Germany. The organizers and participants thank the Mathematisches Forschungsinstitut Oberwolfach for providing an inspiring setting for this conference. The abstracts are presented here in the chronological order of the lectures during the week.

Acknowledgement: The MFO and the workshop organizers would like to thank the National Science Foundation for supporting the participation of junior researchers in the workshop by the grant DMS-1049268, "US Junior Oberwolfach Fellows".

Workshop: Applications of Optimal Transportation in the Natural Sciences

Table of Contents

Felix Otto (joint with Michael Goldman)
Regularity for Monge-Ampère by Benamou-Brenier 345

Yann Brenier
Gradient flows for differential forms: example of the curve-shortening flow 347

Giuseppe Savaré (joint with Matthias Liero, Alexander Mielke)
Distances between nonnegative measures arising from optimal entropy-transport problems 350

Filippo Santambrogio
Estimates on the JKO scheme for the Fokker-Planck and Keller-Segel equations 353

Gero Friesecke
Density Functional Theory and Optimal Transport with Coulomb cost . . 355

Paola Gori-Giorgi
Optimal Transport and Density Functional Theory for the quantum many-body problem: some open questions 358

Mathieu Lewin
Multimarginal Optimal Transport with Coulomb Cost and the Uniform Gas 360

Eric A. Carlen (joint with Jan Maas)
Gradient flow and functional inequalities for quantum Markov semigroups 363

Ansgar Jüngel
New results on discrete Bakry-Emery methods and displacement convexity 364

Dejan Slepčev (joint with Jian-Guo Liu an Robert L. Pego)
Least action principles for incompressible flows and optimal transport between shapes 364

Matthias Erbar
A gradient flow to the Boltzmann equation 366

Eva Kopfer (joint with Karl-Theodor Sturm)
Heat flows on time-dependent metric measure spaces 368

Markus Mittnenzweig (joint with Alexander Mielke)
An entropic gradient structure for quantum Markov semigroups 371

Bao Q. Tang (joint with Klemens Fellner)
Convergence to equilibrium for renormalised solutions to reaction-diffusion systems 373

Francesco Saverio Patacchini (joint with José Alfredo Cañizo)
Discrete minimisers are close to continuum minimisers for the interaction energy 376

Gabriel Peyré (joint with J-D. Benamou, L. Nenna, G. Carlier, L. Chizat, M. Cuturi, J. Solomon, F-X. Vialard)
Computing certain invariants of topological spaces of dimension three .. 378

Carola-Bibiane Schönlieb (joint with Yoeri Boink, Christoph Brune, Martin Burger, Hendrik Dirks, Jocelyn Etienne, Lukas Lang, Sebastian Neumayer, Ozan Öktem)
Dynamic optimal transport and applications in imaging 380

Martin Rumpf (joint with Benjamin Berkels, Alexander Effland, Florian Schfer)
Discrete Geodesic Paths in the Space of Images 380

Brittany D. Froese
Generalised finite difference methods for Monge-Ampère equations 383

Clément Cancès (joint with Thomas O. Gallouët, Léonard Monsaingeon)
A variational approach for multiphase porous media flows 386

Gershon Wolansky
Multiphase Optimal Transport 388

David Kinderlehrer (joint with Patrick Bardsley, Katayun Barmak, Eva Eggeling, Maria Emelianenko, Yekaterina Epshteyn, Xin Yang Lu, Shlomo Ta'asan)
The gradient flow of microstructure 390

Martin Burger
Transport Metrics for Vlasov Hierarchies 392

Aldo Pratelli (joint with J. Bertrand, J. Louet, M. Puel, F. Zeisler)
On the transport problem with relativistic costs 395

Klemens Fellner (joint with B. Hughes, E. J. Hackett-Jones, K. A. Landman, G. Raoul)
Continuum and Discrete Models of Collective Aggregation Pattern 396

Christian Seis (joint with André Schlichting)
Optimal stability estimates for continuity equations 399

Mike Cullen
Applications of optimal transport to meteorology 402

Yao Yao (joint with José Carrillo, Katy Craig, Sabine Hittmeir, Inwon Kim, and Bruno Volzone)
Long time behavior of solutions to the 2D Keller-Segel equation with degenerate diffusion 404

Tryphon T. Georgiou (joint with Yongxin Chen and Allen Tannenbaum)
Matrix-valued Mass Transport: a Quantum Mechanical Approach 406

Brendan Pass (joint with Pierre-André Chiappori, Robert McCann)
Multi- to one-dimensional optimal transport 409

Abstracts

Regularity for Monge-Ampère by Benamou-Brenier

Felix Otto
(joint work with Michael Goldman)

The Monge-Ampère equation may be seen as the Euler-Lagrange equation for the potential of the Brenier map T in optimal transportation. However, the core of the existing regularity theory for this fully non-linear equation does not make much use of this variational interpretation. In this note, we report on an ϵ-regularity result for the Monge-Ampère equation that capitalizes on this variational nature. This type of regularity result was known by work of Figalli & Kim [2], relying on the series of works by Caffarelli. Our proof does *not* use the work by Caffarelli, and directly establishes $C^{1,\alpha}$-regularity of T. In establishing such an ϵ-regularity result, we follow the general strategy applied to many nonlinear variational problems like the one of minimal surfaces or of harmonic maps. Such results all make a "local flatness" assumption, which here assumes the form of a concentration of π near the diagonal in an averaged sense as measured by

$$F(r) := \frac{1}{r}\Big(\frac{1}{|B_r|}\int\int_{(B_r\times\mathbb{R}^d)\cup(\mathbb{R}^d\times B_r))}|x-y|^2\pi(dxdy)\Big)^{\frac{1}{2}}.$$

By affine invariance (see below), this may be replaced by a concentration near an affine subspace. For simplicity, in the following formulation of our main result we restrict to the case of densities f and g that are locally constant (and normalized to unity).

Proposition 1. Fix a positive exponent $\alpha < 1$. Let π be an optimal transference plan for two densities $f, g \le 1$. Let the radius R be such that

$$f = g = 1 \text{ on } B_R, \quad F(R) \ll 1.$$

Then there exists a continuously differentiable map T on $B_{\frac{R}{2}}$ such that

$$\text{supp}\pi \cap (B_{\frac{R}{2}} \times \mathbb{R}^d) = \{(x, T(x)) | x \in B_{\frac{R}{2}}\}$$

and

$$\frac{1}{R}\|T - \text{id}\|_{B_{\frac{R}{2}}} + \|DT - \text{id}\|_{B_{\frac{R}{2}}} + R^\alpha [DT]_{\alpha, B_{\frac{R}{2}}} \lesssim F(R),$$

where $\|\cdot\|_B$ denotes the supremum norm and $[\cdot]_{\alpha,B}$ the Hölder semi-norm over the set B. Here $\lesssim$ and $\ll$ refer to $d < \infty$ and $\alpha < 1$.

Like in [2], Proposition 1 in conjunction with Alexandrov's theorem on the almost everywhere twice differentiability of convex functions yields a partial regularity result for the Brenier map T.

Proposition 1 is the outcome of a Campanato iteration based on a one-step-improvement formulated in Proposition 2. Passing from Proposition 2 to Proposition 1 proceeds by comparing, on dyadically decreasing scales, π to affine transformations as parameterized by a vector b and a matrix Q with $\det Q = 1$:

$$E(r,b,Q) := \frac{1}{r}\Big(\frac{1}{|B_r|}\int\int_{(B_r\times\mathbb{R}^d)\cup(\mathbb{R}^d\times B_r)} |Qx+b-Q^{-*}y|^2\pi(dxdy)\Big)^{\frac{1}{2}},$$

where Q^{-*} denotes the inverse of the transpose Q^* of Q. Proposition 2 is in the spirit of DeGiorgi's approach to the regularity of minimal surfaces, and is the analogue of an "excess improvement by tilting" estimate. Because of the multiplicative nature (composition instead of addition) of the iteration, passing from Proposition 2 to Proposition 1 requires some book-keeping but otherwise is standard. The main structural property in this passage is the *affine* invariance of optimal transport.

Proposition 2. Let the transference plan π be optimal for the densities $f, g \le 1$; let the radius r be st

$$f = g = 1 \quad \text{on } B_r.$$

Then there exist a vector $b \in \mathbb{R}^d$ and a matrix Q with $\det Q = 1$ st for all $\theta \ll 1$

$$E(\theta r, b, Q) \lesssim \theta F(r) + \frac{1}{\theta^{1+\frac{d}{2}}} F^{\frac{d+2}{d+1}}(r) + F^2(r), \tag{1}$$

$$\frac{1}{r}|b| + |Q - \mathrm{id}| \lesssim F(r).$$

Here $\lesssim$ means $\le C$ up to a (generic) constant only depending on the dimension $d < \infty$.

Using the equivalence of the Lagrangian formulation of optimal transportation through a transference plan $\pi(dxdy)$ and the Eulerian formulation through a pair of density $\rho = \rho(t,x) \ge 0$ and flux $j = j(t,x) \in \mathbb{R}^d$ constrained by the (distributional) continuity equation $\partial_t \rho + \nabla \cdot j = 0$, Proposition 2 is a consequence of the following Proposition 3.

Proposition 3. Let (ρ, j) be an optimal Eulerian transference plan that connects densities $\rho_0, \rho_1 \le 1$ and

$$\rho_0 = \rho_1 = 1 \quad \text{on } B_1. \tag{2}$$

Then there exists a harmonic function ϕ in $B_{\frac{1}{2}}$ with

$$\int_0^1 \int_{B_{\frac{1}{2}}} \frac{1}{\rho}|j - \rho\nabla\phi|^2 \lesssim \Big(\int_0^1 \int_{B_1} \frac{1}{\rho}|j|^2\Big)^{\frac{d+2}{d+1}}, \tag{3}$$

$$\int_{B_{\frac{1}{2}}} |\nabla\phi|^2 \lesssim \int_0^1 \int_{B_1} \frac{1}{\rho}|j|^2.$$

The two main ingredients to pass from Proposition 3 to Proposition 2 are 1) the Benamou-Brenier-formula [1] relating the Eulerian and Lagrangian formulations, which we need in a local version, 2) the inner regularity estimates for harmonic functions, and 3) the linearization of the matrix exponential, which gives rise to the quadratic error term in (1). Because the exponent $\frac{d+2}{d+1}$ appearing on the rhs of (3) is strictly larger than one, estimate (3) may be read as follows: Provided the transportation cost $\int_0^1 \int_{B_1} \frac{1}{\rho}|j|^2 \ll 1$, the transportation velocity $\frac{1}{\rho}j$ is close (in the relative sense) to the gradient $\nabla\phi$ of a harmonic function. This reflects the fact that the Wasserstein distance behaves like the (homogeneous) H^{-1}-distance for densities close to one, cf (2). In this sense, Proposition 3 establishes in a quantitative way that the Laplace equation is the linearization of the Monge-Ampère equation.

Proposition 3 follows from constructing, on the space-time cylinder $(0,1)\times B_R$ for some suitable radius $R \in (\frac{1}{2}, 1)$, a competitor $(\tilde\rho, \tilde j)$ with identical flux boundary conditions $f := \nu \cdot j$. This competitor is a perturbation of $(1, \nabla\phi)$ where ϕ is the harmonic function with the flux boundary conditions $\bar f := \int_0^1 f dt$. It turns out that the perturbation $(s,q) := (\tilde\rho, \tilde j) - (1, \nabla\phi)$ may be constrained to be supported in a boundary layer $(0,1)\times(B_R - B_{R-r})$ of thickness $r \ll 1$ and to satisfy $|s| \le \frac{1}{2}$. In fact, (s,q) may be chosen to minimize $\int_0^1 \int_{B_R} |j|^2$ under these constraints, which allows to use the minimax formula to estimate $\int_0^1 \int_{B_R} |j|^2$. Ultimately, this reduces to a trace estimate for the cylindrical shell $(0,1)\times(B_R - B_{R-r})$; optimization in its thickness r yields the crucial exponent $\frac{d+2}{d+1}$. Proposition 3 uses McCann's displacement convexity to derive $\rho \le 1$ from the assumption $\rho_0, \rho_1 \le 1$.

References

[1] I. D. Benamou and Y. Brenier, *A numerical method for the optimal time-continuous mass transport problem and related problems*, in *Monge Ampére equation: applications to geometry and optimization (Deerfield Beach, FL, 1997)*, Contemp. Math. Amer. Math. Soc. **226**, Providence, RI, (1999), 1–11.

[2] A. Figalli, Y. H. Kim, *Partial regularity of Brenier solutions of the Monge-Ampère equation*, Discrete Contin. Dyn. Syst. **28** (2010), 559–565.

Gradient flows for differential forms: example of the curve-shortening flow

Yann Brenier

Optimal transport theory is largely about giving a Riemannian structure to the space of volume forms $\rho dx^1 \wedge \cdots dx^d$ in $\mathbb{R}^d$ and many gradient flows can be derived accordingly, following the seminal work of Otto and collaborators. Here we discuss the case of closed $d-1$ differential forms in $\mathbb{R}^d$, or, in other words, divergence-free vector fields. An elementary example of such forms is given by

$$B(x) = \int_{\mathbb{R}/\mathbb{Z}} \delta(x - X(s)) X'(s) ds, \tag{1}$$

where X is a smooth, non self-intersecting, loop X in $\mathbb{R}^d$. The transport equation for such forms is the so-called induction equation (commonly used in ideal Magnetohydrodynamics),

$$\partial_t B + \nabla \cdot (B \otimes v - v \otimes B) = 0,$$

where v is a time-dependent velocity field transporting the form. Then, we may monitor the steepest descent of a given functional $F[B]$ according to a given Hilbert norm $v \to ||v||_B$ (possibly depending on B) on the space of velocity fields v transporting B. Let us give a simple example, when

$$F[B] = \int_{\mathbb{R}^d} |B(x)|dx, \quad ||v||_B{}^2 = \int_{\mathbb{R}^d} |v(x)|^2 |B(x)|dx.$$

Thanks to the induction equation, we get, denoting $G = \nabla \cdot \frac{B \otimes B}{|B|}$,

$$\frac{d}{dt}F[B] = -\int v \cdot \ Gdx = \int \frac{(|B|v - G)^2}{2|B|}dx - \int \frac{|B|v^2}{2}dx - \int \frac{G^2}{2|B|}dx$$

and obtain the steepest descent as $P = G$, where P stands for $|B|v$, which, completed by the induction equation, provides a self-consistent evolution equation for B of (very) degenerate parabolic type:

$$\partial_t B + \nabla \cdot (\frac{B \otimes P - P \otimes B}{|B|}) = 0, \quad P = \nabla \cdot \frac{B \otimes B}{|B|}, \tag{2}$$

It turns out that, in the special case when B is of form (1), this exactly corresponds to the so-called short-curvening flow [4].

$$\partial_t X = \frac{1}{|\partial_s X|}\partial_s(\frac{\partial_s X}{|\partial_s X|}), \tag{3}$$

for a time-dependent loop, with $\partial_t X(t,s) = v(t, X(t,s))$. In a joint work with Xianglong Duan, we have recently analyzed equation (2), thanks to a relative entropy method borrowed from the theory of systems of first order conservation laws with a convex entropy [3]. We have obtained a concept of "dissipative solutions" related to the work of P.-L. Lions for the Euler equation of incompressible fluids [8] or to the work of L. Ambrosio, N. Gigli, G. Savaré [1] for the heat equation and, overall, quite similar to the one recently introduced in [2]. We also refer to the works of A. Tzavaras and collaborators [5], E. Feireisl and collaborators [6] for various concepts of "dissipative solutions".

Definition 1. *Let us fix $T > 0$ and denote $\mathbb{T}^d = (\mathbb{R}/\mathbb{Z})^d$. We say that (B, P) with*

$$B \in C([0,T], C(\mathbb{T}^d, \mathbb{R}^d)'_{w^*}), \quad P \in C([0,T] \times \mathbb{T}^d, \mathbb{R}^d)'$$

is a dissipative solution of (2) with initial data $B_0 \in C(\mathbb{T}^d, \mathbb{R}^d)'$ if and only if:

i) $B(0) = B_0, \quad \nabla \cdot B = 0$ *in sense of distributions;*

ii) B and P are bounded, respectively in the spaces $C^{1/2}([0,T], (C^1(\mathbb{T}^d))'_{w^*})$ and $C([0,T] \times \mathbb{T}^d, \mathbb{R}^d)'$, by constants depending only on T and $\int_{\mathbb{T}^d} |B_0|$;

iii) for all $\lambda > 0$, $\theta \in [0,T]$, for all trial functions (b^*, v^*, A) valued in $\mathbb{R}^d$, with $|A| \leq \lambda$ and ${b^*}^2 = 1$, for all $r \geq c^* + \frac{\lambda^2}{2}$, where c^* is a constant depending explicitly on (b^*, v^*), we have:

$$
\begin{aligned}
(4) \qquad & e^{-r\theta} \int_{\mathbb{T}^d} \eta(\theta) + \int_0^\theta e^{-r\sigma} \int_{\mathbb{T}^d} [(A - v^* - L_3) \cdot P \\
& + (r - c^* + \frac{{v^*}^2 - A^2}{2})\eta - (L_2 + b^* \frac{A^2 - {v^*}^2}{2}) \cdot B](\sigma) d\sigma \leq \int_{\mathbb{T}^d} \eta(0),
\end{aligned}
$$

where

$$
\eta = |B| - B \cdot b^*, \tag{5}
$$

$$
\mathrm{L}_2 = -\partial_\theta b^* - (v^* \cdot \nabla) b^* + (b^* \cdot \nabla) v^* + b^* {v^*}^2 + (\mathbb{I} - b^* \otimes b^*) \nabla (b^* \cdot v^*), \tag{6}
$$

$$
\mathrm{L}_3 = -v^* + (b^* \cdot \nabla) b^*. \tag{7}
$$

The "weak compactness" of such solutions (i.e. any sequence of dissipative solutions has accumulations points, in a suitable weak sense, and each of them is still a dissipative solution) follows almost immediately from the convexity of our formulation. We have also established a suitable "weak-strong" uniqueness principe for such solutions.

Acknowledgement. We are very grateful to Dmitry Vorotnikov for informing us about the possibility of deriving mean-curvature motion in co-dimension one as a gradient flow in optimal transportation style, in the spirit of [7]. This information was very influential for us.

References

[1] L. Ambrosio, N. Gigli, G. Savaré, *Gradient flows in metric spaces and in the space of probability measures,* Lectures in Mathematics ETH Zürich, Birkhäuser , 2008.

[2] Y. Brenier, *Topology preserving diffusion of divergence-free vector fields,* Comm. Math. Phys. 330 (2014) 757-770.

[3] C. M. Dafermos, *Hyperbolic conservation laws in continuum physics,* Springer, Berlin, 2000.

[4] K. Deckelnick, *Weak solutions of the curve shortening flow,* Calc. Var. 5 (1997) 489-510.

[5] S.Demoulini, D. Stuart, A. Tzavaras, *Weak-strong uniqueness of dissipative measure-valued solutions for polyconvex elastodynamics,* Arch. Ration. Mech. Anal. 205 (2012) 927-961.

[6] E. Feireisl, A. Novotný, *Singular Limits in Thermodynamics of Viscous Fluids,* Springer Science & Business Media, 2009.

[7] R. Jordan, D. Kinderlehrer, F. Otto, *The variational formulation of the Fokker-Planck equation,* SIAM J. Math. Anal. 29 (1998) 1-17.

[8] P.-L. Lions, *Mathematical topics in fluid mechanics. Vol. 1. Incompressible models,* Oxford Lecture Series in Mathematics and its Applications, 3. 1996.

Distances between nonnegative measures arising from optimal entropy-transport problems

Giuseppe Savaré
(joint work with Matthias Liero, Alexander Mielke)

Consider a couple of finite and nonnegative Borel measures $\mu_i \in \mathcal{M}(X_i)$, $i = 0, 1$, in the Polish topological spaces X_0, X_1 (the topology of each X_i is thus induced by a complete and separable distance). Optimal Entropy-Transport problems [3] deal with the minimization of convex functionals of the form

$$\mathscr{E}(\boldsymbol{\gamma}|\mu_0,\mu_1) := \int_{X_0\times X_1} \mathsf{c}(x_0,x_1)\,\mathrm{d}\boldsymbol{\gamma}(x_0,x_1) + \mathscr{F}_0(\gamma_0|\mu_0) + \mathscr{F}_1(\gamma_1|\mu_1), \tag{1}$$

depending on $\boldsymbol{\gamma} \in \mathcal{M}(X_0 \times X_1)$; here $\mathsf{c} : X_0 \times X_1 \to [0, +\infty]$ is a lower semicontinuous cost function, $\gamma_i = \pi^{X_i}_\sharp \boldsymbol{\gamma}$ are the marginals of $\boldsymbol{\gamma}$ on X_i and $\mathscr{F}_i : \mathcal{M}(X_i) \to [0, +\infty]$ are general entropy functionals

$$\mathscr{F}_i(\gamma_i|\mu_i) := \int_{X_i} F_i(\sigma_i)\,\mathrm{d}\mu_i + F_i'(\infty)\gamma_i^\perp(X_i), \quad \gamma_i = \sigma_i\mu_i + \gamma_i^\perp,$$

associated to convex and lower semicontinuous functions $F_i : [0, \infty) \to [0, \infty]$ satisfying $F(1) = 0$. We denoted by σ_i the density of γ_i w.r.t. μ_i and by $\gamma_i^\perp$ the singular part of the Lebesgue decomposition of γ_i; $F_i'(\infty) := \sup_{s>0} F_i(s)/s$.

The variational problem associated to (1)

$$\mathsf{ET}(\mu_0,\mu_1) := \min_{\boldsymbol{\gamma}\in\mathcal{M}(X_0\times X_1)} \mathscr{E}(\boldsymbol{\gamma}|\mu_0,\mu_1) \tag{2}$$

can be considered as a natural relaxation of the Optimal Transport problem [5]

$$\mathsf{OT}(\mu_0,\mu_1) := \min_{\boldsymbol{\gamma}\in\mathcal{M}(X_0\times X_1)} \Big\{ \int_{X_0\times X_1} \mathsf{c}(x_0,x_1)\,\mathrm{d}\boldsymbol{\gamma}(x_0,x_1) : \pi^{X_i}_\sharp\boldsymbol{\gamma} = \mu_i \Big\}, \tag{3}$$

where the marginals γ_i of $\boldsymbol{\gamma}$ are constrained to coincide with the given measures μ_i (and therefore one has to impose the balance mass condition $\mu_0(X_0) = \mu_1(X_1)$). In fact, (3) is a particular example of (2) when F_i are the indicator function of $\{s = 1\}$ (i.e. $F_i(1) = 0$ and $F_i \equiv +\infty$ in $[0, \infty) \setminus \{1\}$). More regular entropies satisfying $F_i(0) < \infty$, as the logarithmic or the Rény's ones associated to

$$F_i(s) := s\log s - (s-1), \quad F_i(s) := \frac{1}{p(p-1)}(s^p - p(s-1) + 1), \quad p \neq 0, 1,$$

induce a wider family of optimization problems, that always admit a minimizer even if $\mu_i(X_i)$ are different and the cost c takes the value $+\infty$.

As for Optimal Transport problems, (2) enjoys an equivalent dual formulation

$$\mathsf{ET}(\mu_0,\mu_1) = \sup\Big\{ \mathscr{D}(\phi_0,\phi_1|\mu_0,\mu_1) : \phi_i \in \mathrm{C}_b(X_i),\ \phi_0 \oplus \phi_1 \le \mathsf{c} \Big\}, \tag{4}$$

where $(\phi_0 \oplus \phi_1)(x_0, x_1) := \phi_0(x_0) + \phi_1(x_1)$ and

$$\mathscr{D}(\phi_0,\phi_1|\mu_0,\mu_1) := \sum_i \int_{X_i} F_i^\circ(\phi_i)\,\mathrm{d}\mu_i, \quad F_i^\circ(w) := -F_i^*(-w) = \inf_{s\ge 0} F_i(s) + ws.$$

Notice that $F_i^\circ : \mathbb{R} \to [-\infty, +\infty)$ are concave functions obtained by (the opposite) Legendre transform from F_i. The duality result (4) yields general optimality conditions and a third intrinsic formulation based on the 1-*homogeneous marginal perspective function* H

$$H(x_0, r_0; x_1, r_1) := \mathsf{ET}(r_0\delta_{x_0}, r_1\delta_{x_1}) = \inf_{\theta>0} \Big(\theta\, \mathsf{c}(x_0, x_1) + \sum_i r_i F_i(\theta/r_i)\Big),$$

representing the cost to connect two Dirac measures $\mu_i := r_i\delta_{x_i}$. We have

$$\mathsf{ET}(\mu_0, \mu_1) = \min_{\boldsymbol{\gamma}\in\mathcal{M}(X_0\times X_1)} \int H(x_0, \varrho_0(x_0); x_1, \varrho_1(x_1))\, \mathrm{d}\boldsymbol{\gamma} + \sum_i F_i(0)\mu_i^\perp(X_i),$$

where $\mu_i = \varrho_i\gamma_i + \mu_i^\perp$ is the Lebesgue decomposition of μ_i w.r.t. the marginals γ_i.

A particular interesting case arises when $X_0 = X_1 =: X$ is a (complete, separable) metric space endowed with the distance d, $F_i(s) = s\log s - (s-1)$, and

$$\mathsf{c}(x_0, x_1) = \ell(\mathsf{d}(x_0, x_1)), \quad \ell(d) := \begin{cases} \log\Big(1 + \tan^2(d)\Big) & \text{if } 0 \le d < \pi/2, \\ +\infty & \text{otherwise.} \end{cases}$$

The corresponding marginal perspective function H is given by the so-called squared cone distance

$$H(x_0, r_0^2; x_1, r_1^2) := r_0^2 + r_1^2 - 2r_0r_1\cos(\mathsf{d}(x_0, x_1) \wedge \pi/2),$$

and the corresponding *Logarithmic Entropy-Transport problem* $\mathsf{LET}(\mu_0, \mu_1)$ is the square of a complete and separable metric HK on $\mathcal{M}(X)$ which induces the usual weak topology in duality with continuous and bounded functions. Since HK results as a sort of infinitesimal inf-convolution of the Hellinger distance

$$\mathsf{H}^2(\mu_0, \mu_1) := \int_X \big(\sqrt{\varrho_0(x)} - \sqrt{\varrho_1(x)}\big)^2\, \mathrm{d}\gamma(x), \quad \mu_i = \varrho\gamma,$$

and the Kantorovich-Rubinstein-Wasserstein distance

$$\mathsf{W}_{\mathsf{d}}^2(\mu_0, \mu_1) := \min\Big\{ \int \mathsf{d}^2(x_0, x_1)\, \mathrm{d}\boldsymbol{\gamma} : \gamma \in \mathcal{M}(X\times X),\ \pi^i_\sharp\boldsymbol{\gamma} = \mu_i\Big\},$$

we called HK the *Hellinger-Kantorovich* distance; it has been independently introduced by S. Kondratyev, L. Monsaingeon, and D. Vorotnikov [2], and by L. Chizat, G. Peyré, B. Schmitzer, F.-X. Vialard [1], starting from its dynamical formulation. HK enjoys various equivalent characterizations [3, 4], which show its geometric relevance:

Static LET formulation:

$$\mathsf{HK}^2(\mu_0, \mu_1) = \min_{\boldsymbol{\gamma}\in\mathcal{M}(X\times X)} \int \ell(\mathsf{d}(x_0, x_1))\, \mathrm{d}\boldsymbol{\gamma} + \sum_i \big(\sigma_i\log\sigma_i - \sigma_i + 1\big)\, \mathrm{d}\gamma_i,\ \gamma_i = \sigma_i\mu_i.$$

Static dual formulation:

$$\mathsf{HK}^2(\mu_0, \mu_1) = \sup_{\phi_i\in \mathrm{C}_b(X)} \Big\{ \sum_i (1 - \mathrm{e}^{-\phi_i})\, \mathrm{d}\mu_i : \phi_0(x_0) + \phi_1(x_1) \le \ell(\mathsf{d}(x_0, x_1))\Big\}.$$

Dual conical Hopf-Lax formulation:

$$\frac{1}{2}\mathsf{HK}^2(\mu_0,\mu_1) = \sup\Big\{ \int \mathscr{P}_1(\xi)\,\mathrm{d}\mu_1 - \int \xi\,\mathrm{d}\mu_0 : \xi \in \mathrm{Lip}_b(X),\ \inf\xi > -1/2\Big\},$$

where

$$\mathscr{P}_t(\xi)(x) := \inf_{x'\in X} \frac{1}{t}\Big(1 - \frac{\cos^2(\mathsf{d}(x,x')\wedge\pi/2)}{1+2t\xi(x')}\Big). \tag{5}$$

Cone formulation:

$$\mathsf{HK}^2(\mu_0,\mu_1) = \min\Big\{ \int \Big(\varrho_0^2+\varrho_1^2 - 2\varrho_0\varrho_1\cos(\mathsf{d}(x_0,x_1)\wedge\pi/2)\,\mathrm{d}\boldsymbol{\gamma} : \mu_i = \varrho_i\gamma_i\Big\}.$$

Dynamic formulation "á la Benamou-Brenier" in $X=\mathbb{R}^d$:

$$\mathsf{HK}^2(\mu_0,\mu_1) = \min\Big\{ \int_0^1 \int \Big(|\boldsymbol{v}_t|^2 + \frac{1}{4}|w_t|^2\Big)\,\mathrm{d}\mu_t\,\mathrm{d}t : \mu\in \mathrm{C}([0,1];\mathcal{M}(\mathbb{R}^d)),$$
$$\mu_{t=i} = \mu_i,\quad \partial_t\mu_t + \nabla\cdot\big(\boldsymbol{v}_t\mu_t\big) = w_t\mu_t \text{ in } \mathcal{D}'(\mathbb{R}^d\times(0,1))\Big\}$$

Dynamic dual formulation in $X=\mathbb{R}^d$:

$$\mathsf{HK}^2(\mu_0,\mu_1) = \sup\Big\{ \int \xi_1\,\mathrm{d}\mu_1 - \int \xi_0\,\mathrm{d}\mu_0 : \xi\in \mathrm{C}^1(\mathbb{R}^d\times[0,1]),$$
$$\partial_t\xi_t + \frac{1}{2}|\mathrm{D}_x\xi_t(x)|^2 + 2\xi_t^2(x) \le 0 \text{ in } \mathbb{R}^d\times[0,1]\Big\}.$$

Moreover, a curve $(\mu_t)_{t\in[0,1]}\subset\mathcal{M}(\mathbb{R}^d)$ is a geodesic for HK if and only if there exists a solution $\xi_t = \mathscr{P}_t\xi_0$ of the Hamilton-Jacobi equation

$$\partial_t\xi_t + \frac{1}{2}|\mathrm{D}_x\xi_t(x)|^2 + 2\xi_t^2(x) = 0$$

given by (5) such that

$$\partial_t\mu_t + \nabla\cdot(\boldsymbol{v}_t\mu_t) = w_t\mu_t \quad\text{with}\quad \boldsymbol{v}_t = \mathrm{D}_x\xi_t,\ w_t = 4\xi_t.$$

References

[1] L. Chizat, G. Peyré, B. Schmitzer, and F.-X. Vialard, *An interpolating distance between optimal transport and Fisher–Rao.* arXiv:1506.06430v2, 2015.

[2] S. Kondratyev, L. Monsaingeon, and D. Vorotnikov, *A new optimal transport distance on the space of finite radon measures*, arXiv 1505.07746v2, 2015.

[3] M. Liero, A. Mielke, and G. Savaré, *Optimal entropy-transport problems and the Hellinger–Kantorovich distance*, arXiv 1508.07941, 2015.

[4] M. Liero, A. Mielke, and G. Savaré, *Optimal Transport in Competition with Reaction: The Hellinger–Kantorovich Distance and Geodesic Curves*, SIAM J. Math. Anal. 48 (2016) 2869–2911.

[5] C. Villani, *Optimal transport. Old and new*, volume 338 of *Grundlehren der Mathematischen Wissenschaften*, Springer-Verlag, Berlin, 2009.

Estimates on the JKO scheme for the Fokker-Planck and Keller-Segel equations

Filippo Santambrogio

Many evolution PDEs with mass conservation, often of parabolic type, can be expressed in the form of a gradient flow of a functional F defined on $\mathcal{P}(\Omega)$ with respect to the Wasserstein distance W_2 (see [1]). This in particular involves the iterated minimization scheme

$$\rho^\tau_{k+1} := \mathrm{argmin}_\rho F(\rho) + \frac{W_2^2(\rho, \rho^\tau_k)}{2\tau}, \tag{1}$$

where $\tau > 0$ is a small time-step. This iterated minimization scheme is called *JKO scheme* (after Jordan, Kinderleher and Otto, [4]). By suitably interpolating the sequence $(\rho^\tau_k)_k$ and letting $\tau \to 0$, we can prove (in fairly general situations), that we converge to a solution of

$$\partial_t \rho - \nabla \cdot \left(\rho \, \nabla \left(\frac{\delta F}{\delta \rho}(\rho) \right) \right) = 0. \tag{2}$$

Note that, given a functional $G : \mathcal{P}(\Omega) \to \mathbb{R}$ we call $\frac{\delta G}{\delta \rho}(\rho)$, if it exists, the unique (up to additive constants) function such that $\frac{d}{d\varepsilon} G(\rho + \varepsilon\chi)_{|\varepsilon=0} = \int \frac{\delta G}{\delta \rho}(\rho) d\chi$ for every perturbation χ such that, at least for $\varepsilon \in [0, \varepsilon_0]$, the measure $\rho + \varepsilon\chi$ belongs to $\mathcal{P}(\Omega)$. The function $\frac{\delta G}{\delta \rho}(\rho)$ is called *first variation* of the functional G at ρ.

The connection between the PDE and the minimization scheme comes from the optimality conditions of the latter, since we obtain (up to some technicalities)

$$\frac{\delta F}{\delta \rho}(\rho^\tau_{k+1}) + \frac{\varphi}{\tau} = const, \tag{3}$$

where φ is the Kantorovich potential associated with the transport from ρ^τ_{k+1} to ρ^τ_k for the cost $\frac{1}{2}|x-y|^2$. Indeed, we can prove that whenever we set $G(\rho) := \frac{1}{2} W_2^2(\rho, \nu)$, then we have $\delta G / \delta \rho = \varphi$ where φ is the corresponding Kantorovich potential.

In the talk, we considered functionals F of the following form

$$F(\rho) := \int f(\rho(x)) dx + \int V(x) \, d\rho(x),$$

where f is a convex, superlinear, and increasing function, and V is aeither a given and smooth potential, or depend on ρ in the following way: $V = -u_\rho$, where u_ρ is the solution of $-\Delta u = \rho$ with Dirichlet boundary conditions on $\partial\Omega$ (other equations and b.c. are possible). In the case $f(\rho) = \rho \ln(\rho)$ these two choices give the Fokker-Planck equation

$$\partial_t \rho - \Delta \rho - \nabla \cdot (\rho \nabla V) = 0$$

or the Keller-Segel system

$$\begin{cases} \partial_t \rho - \Delta \rho + \nabla \cdot (\rho \nabla u) = 0 \\ -\Delta u = \rho \end{cases}.$$

When f is not of entropy type, then instead of the linear diffusion term represented by the Laplacian we have a non-linear diffusion as in the porous medium equation $\partial_t \rho = \Delta(\rho^m)$ (see [7]).

The JKO minimization scheme allows for some iterative estimates on the minimizer ρ^τ_{k+1} in terms of the datum ρ^τ_k. Those which can be easily iterated when $\tau \to 0$ can be useful to prove existence or regularity properties of the solution of the corresponding equations. In particular, we proved the following results

- L^∞ estimate for the Fokker-Plank or porous medium equations with potential: for any f and $V \in C^{1,1}$, we have
$$||\rho^\tau_{k+1}||_{L^\infty} \leq (1 + C\tau)||\rho^\tau_k||_{L^\infty}.$$
- L^∞ estimate for a Keller-Segel equation with arbitrary diffusion (any f, any dimension), of the form
$$||\rho^\tau_{k+1}||_{L^\infty} \leq \frac{||\rho^\tau_k||_{L^\infty}}{1 - C\tau||\rho^\tau_k||_{L^\infty}}.$$
This estimate can be iterated but exploses in finite time, which is not surprising since in general solutions to Keller-Segel equations explose in finite time.
- L^p estimates for Fokker-Plank or porous medium equations with potential: for any f and $V \in C^{1,1}$, we have
$$||\rho^\tau_{k+1}||^p_{L^p} \leq (1 + C\tau)||\rho^\tau_k||^p_{L^p}$$
(this proof uses, for instance, the so-called flow interchange method, see [5])
- BV estimates for arbitrary f, when $V = 0$:
$$||\rho^\tau_{k+1}||_{BV} \leq ||\rho^\tau_k||_{BV}$$
(using a new inequality in optimal transport first introduced, exactly for BV purposes, in [3]: for any ρ_0, ρ_1 smooth enough, if φ_0 and φ_1 are the respective Kantorovich potentials, and H is arbitrary radial convex function, we have
$$\int \nabla H(\nabla\varphi_0) \cdot \nabla\rho_0 + \int \nabla H(\nabla\varphi_1) \cdot \nabla\rho_1 \geq 0;$$
the BV estimate is obtained by using $H(z) = |z|$).
- sort of $W^{1,p}$ estimates: defining
$$J_{p,V}(\rho) := \int \left| \frac{\nabla\rho}{\rho} + \nabla V \right|^p d\rho,$$
we have, for $f(\rho) = \rho \ln \rho$ and supposing $D^2V \geq \lambda I$,
$$(1 + \lambda\tau)J_p(\rho^\tau_{k+1}) \leq J_p(\rho^\tau_k)$$
(obtained by using $H(z) = |z|^p$ in the above inequality).

- A similar estimate for the Keller-Segel case, with liner diffusion ($f(\rho) = \rho \ln \rho$):

$$J_p(\rho^\tau_{k+1}) \leq J_p(\rho^\tau_k) + C\tau + C(F(\rho^\tau_k) - F(\rho^\tau_{k+1})).$$

The presented results are taken from various works (or on-going works) in collaboration with G. Carlier [2], G. De Philippis, A. Mészáros, B. Velichkov, [3], J. Carrillo and S. Di Marino. The easiest proofs are not new at all, and were already present, for instance, in Otto's work [6].

References

[1] L. Ambrosio, N. Gigli and G. Savaré, *Gradient flows in metric spaces and in the spaces of probability measures.* Lectures in Mathematics, ETH Zurich, Birkhäuser, 2005.

[2] G. Carlier, F. Santambrogio *A Variational Model for Urban Planning with Traffic Congestion*, ESAIM: COCV, **11** (2005), 595–613

[3] G. De Philippis, A. R. Mészáros, F. Santambrogio and B. Velichkov, *BV estimates in optimal transportation and applications,* Arch. Rati. Mech. An. **219** (2016), 829–860.

[4] R. Jordan, D. Kinderlehrer and F. Otto, *The variational formulation of the Fokker-Planck equation*, SIAM J. Math. Anal. **29** (1998), 1–17.

[5] D. Matthes, R. J. McCann, and G. Savaré. *A family of nonlinear fourth order equations of gradient flow type.* Comm. Partial Differential Equations, **34** (2009), 1352– 1397.

[6] F. Otto *Dynamics of Labyrinthine Pattern Formation in Magnetic Fluids: A Mean?Field Theory*, Archive for Rational Mechanics and Analysis **141** (1998), 63–103.

[7] F. Otto, *The geometry of dissipative evolution equations: The porous medium equation,* Comm. Partial Differential Equations, **26** (2001), 101–174.

Density Functional Theory and Optimal Transport with Coulomb cost

Gero Friesecke

I discussed recent developments in an exciting new application area of optimal transport, electronic structure. One is dealing with a multi-marginal problem, and with a Coulomb cost – not the usual positive power of the distance but a *negative* power. Some phenomena can be treated with existing theory, some have been understood by developing additional tools, and others give rise to open questions. The optimal transport problem is the following: find

$$\gamma^{opt} \in \operatorname{argmin}_{\substack{\gamma \in \mathcal{P}_{sym}(\mathbb{R}^{N\cdot d}) \\ \gamma \mapsto \mu}} \int_{\mathbb{R}^{Nd}} \sum_{1 \leq i < j \leq N} \frac{1}{|x_i - x_j|^\alpha} \, d\gamma(x_1, .., x_N) \quad (0 < \alpha < d).$$

The prototypical case arising in electronic structure is $d = 3$, $\alpha = 1$ (Coulomb cost). The minimization is over symmetric probability measures on $\mathbb{R}^{Nd}$ with equal marginals, given by an absolutely continuous probability measure $\mu = \rho/N$. Symmetric means $\gamma(A_1 \times \cdots \times A_N) = \gamma(A_{\sigma(1)} \times \cdots \times A_{\sigma(N)})$ for all permutations, and the marginal condition is

$$\gamma(\mathbb{R}^d \times \cdots \times A_i \times \cdots \times \mathbb{R}^d) = \int_{A_i} \mu \text{ for all } i.$$

(In physics notation, the marginal μ is denoted ρ/N, where ρ is the one-body density of the system, which integrates to the number N of particles.)

The following table contains a summary of some main results:

	$d=1$	$d=3$
$N=2$	unique minimizer, of Monge form [1, 2]	
$2<N<\infty$	unique min., Monge form [3, 4]	example of non-Monge min. [5]
$N=\infty$	unique minimizer, non-Monge [6]	

Here Monge form means $\gamma(x_1,..,x_N)$ = symmetrization of $\mu(x_1)\delta_{T_2(x_1)}(x_2)\cdots\delta_{T_N(x_1)}(x_N)$.

Let me explain how the marginal constraint in the optimal transport problem arises in a natural way from the idea of constrained search in the electronic structure community. A central goal there is to find the lowest eigenvalue E_0 and eigenstate Ψ_0 of the electronic Hamiltonian

$$H = \underbrace{-\frac{1}{2}\Delta_{\mathbb{R}^{3N}}}_{=:T} + \underbrace{\sum_{1\le i<j\le N}\frac{1}{|x_i-x_j|}}_{=:V_{ee}} + \underbrace{\sum_{i=1}^{N} v(x_i)}_{=:V_{ne}},$$

acting on N-electron wavefunctions $\psi \in L^2_{anti}((\mathbb{R}^3\times\mathbb{Z}_2)^N)$, where $|\Psi(x_1,s_1,..,x_N,s_N)|^2$ is the probability density that the electrons are at positions $x_1,..,x_N$ with spins $s_1,..,s_N$. Here $v:\mathbb{R}^3\to\mathbb{R}$ is the external potential which depends on the positions and charges of the atomic nuclei; it constitutes the only chemically specific (molecule-dependent) part of the Hamiltonian. An elementary but important observation going back to Hohenberg and Kohn is that the quadratic form of the electron-nuclei interaction depends only on the one-body density ρ, $\langle\Psi, V_{ne}\Psi\rangle = \int_{\mathbb{R}^3} v\,\rho$, where $\rho(x_1)/N = \sum_{s_1,..,s_N\in\mathbb{Z}_2}\int_{\mathbb{R}^{3(N-1)}}|\Psi|^2 dx_2..dx_N$. The Rayleigh-Ritz variational together with the idea of constrained search, i.e. splitting the minimization over wavefunctions into minimizing first over Ψ subject to given ρ, then over ρ yields

$$\begin{aligned} E_0 &= \min_\rho \min_{\Psi\mapsto\rho}\Big\langle\Psi,(T+V_{ee}+V_{ne})\Psi\Big\rangle \\ &= \min_\rho\Big(\underbrace{\inf_{\Psi\mapsto\rho}\langle\Psi,(T+V_{ee})\Psi\rangle}_{\text{universal part } F[\rho]\text{ Hohenberg-Kohn-Lieb-Levy fctnal}} + \underbrace{\int_{\mathbb{R}^3} v(x)\rho(x)\,dx}_{\text{chemically specific part}}\Big)\end{aligned}$$

Hence one can recover E_0 and the ground state density from a low-dimensional variational problem, $\min_\rho(F[\rho]+\int v\,\rho$, instead of a high-dimensional one. This fact is known as the "Hohenberg-Kohn theorem". Moreover the unknown piece F which needs to be approximated in practice is "universal" (molecule-independent).

The rigorous connection to optimal transport is the following. Start from a density ρ. Scale it via $\rho_\alpha(x) := \alpha^d\rho(\alpha x)$. Let $\Psi[\rho_\alpha]$ denote the minimizer of the variational problem in the definition of the Hohenberg-Kohn functional for the scaled density, i.e. the minimizer of $\langle\Psi,(T+V_{ee})\Psi\rangle$ subject to $\Psi\mapsto\rho_\alpha$. Finally,

scale back, i.e. $\Psi_0^{(\alpha)} := \alpha^{-Nd/2}\Psi[\rho_\alpha](\alpha^{-1}x_1, s_1, .., \alpha^{-1}x_N, s_N)$. One then has the following theorem [1]: for any ρ, in the dilute limit $\alpha \to 0$ we have

$$\sum_{s_1,..,s_N \in \mathbb{Z}_2} |\Psi_0^{(\alpha)}|^2 \rightharpoonup^* \gamma \text{ solution of OT with Coulomb cost, } \frac{1}{\alpha}F[\rho_\alpha] \to \text{ optimal cost}$$

(where the optimal plan γ is approached in the sense of weak* convergence of probability measures; for careful numerical comparisons of wave functions squared and optimal plans for small α see [8]). A more heuristic derivation of the limit problem which bypasses scaling arguments and directly studies the situation – implied by density scaling – that the interaction energy dominates the kinetic energy goes back to a fundamental paper by Seidl [3], who introduced the limit problem without being aware at the time that he was talking optimal transport.

Electron densities in practical simulations are insufficiently dilute to directly replace the Hohenberg-Kohn functional (or the electron-interaction part of it) by the optimal Coulomb cost. For instance, the binding curve of the hydrogen molecule turns out to be correct at long range, but far off at equilibrium [7], in contraposition to conventional density functionals, which are remarkably accurate near equilibrium but poor at long range. Nevertheless the fact that the functional and the associated many-body electron density are governed by OT theory in some limit is a very useful source of information in ongoing research to design more accurate functionals. See the talk by Paola Gori-Giorgi.

Another interesting question associated to our OT problem is what happens in the limit of large particle number N. In [6], together with Codina Cotar and Brendan Pass we introduced the infinite-marginal OT problem of minimizing the cost per particle pair,

$$C_\infty[\gamma] := \lim_{N\to\infty} \binom{N}{2}^{-1} \int \sum_{1\le i<j\le N} \ell(x_i - x_j) d\gamma(x_1, x_2, ...),$$

over symmetric probability measures in infinitely many variables, $\gamma \in \mathcal{P}_{sym}((\mathbb{R}^d)^\infty)$, subject to $\gamma \mapsto \mu \in \mathcal{P}(\mathbb{R}^d)$. It can be shown [6] that if the interaction potential ℓ has positive Fourier transform (as does the Coulomb cost), the optimizer is unique and given by the independent measure $\gamma = \mu \otimes \mu \otimes ...$ This is a consequence of the negative power cost (for positive power costs minimizers are trivially of Monge form, and hence far from independent) together with the fact – long kown in probability and statistics – that symmetry is a much stronger restriction on probability measures in infinitely many variables than might be naively expected. Technically, one uses the DeFinetti-Hewett-Savage theorem and a novel probabilistic interpretation of the above cost: if ν is a representing DeFinetti measure of a given probability measure γ, i.e. $\gamma = \int_{\mathcal{P}(\mathbb{R}^d)} Q \otimes Q \otimes \cdots d\nu(Q)$, and $\hat{\ell}$ is the Fourier transform of ℓ, then

$$C_\infty[\gamma] - C_\infty[\mu \otimes \mu \otimes ...] = (2\pi)^{-d} \int_{\mathbb{R}^d} \widehat{\ell}(z)\, var_{\nu(dQ)} \widehat{Q}(z)\, dz.$$

Since $\widehat{\ell} > 0$, the right hand side is minimized if and only if the variance of the random function $\hat{Q}$ vanishes, i.e. when $\nu = \delta_\mu$, establishing independence. An interesting open problem is to derive corrections to independence, expected at the next order: cost per particle instead of cost per particle pair.

References

[1] C. Cotar, G. Friesecke, and C. Klüppelberg, *Density functional theory and optimal transportation with Coulomb cost*, Comm. Pure Appl. Math., **66** (2013), 548–599, and arXiv:1104.0603 (2011)

[2] G. Buttazzo, L. De Pascale, and P. Gori-Giorgi, *Optimal-transport formulation of electronic density-functional theory*, Phys. Rev. A, **85** (2012), 062502

[3] M. Seidl, *Strong-interaction limit of density-functional theory*, Phys. Rev. A, **60** (1999), 4387–4395

[4] M. Colombo, L. De Pascale, and S. Di Marino, *Multimarginal optimal transport maps for 1-dimensional repulsive costs*, Canad. J. Math. **67** (2015), 350–368

[5] B. Pass, *Remarks on the semi-classical Hohenberg–Kohn functional*, Nonlinearity 26 (2013), 2731

[6] C.Cotar, G.Friesecke, B.Pass, *Infinite-body optimal transport with Coulomb cost*, Calc. Var. PDE **54**, no. 1 (2015), 717–742

[7] H.Chen, G. Friesecke, Ch.Mendl, *Numerical Methods for a Kohn–Sham Density Functional Model Based on Optimal Transport*, J. Chem. Theory Comput. 10 (2014), 4360–4368

[8] H.Chen, G. Friesecke, *Pair densities in density functional theory*, Multiscale Model. Simul. **13** (4) (2015), 1259–1289

Optimal Transport and Density Functional Theory for the quantum many-body problem: some open questions

Paola Gori-Giorgi

Electronic structure calculations are at the very heart of predictive computational materials science, chemistry and biochemistry. Their goal is to solve, in a reliable and computationally affordable way, the many-electron problem, a complex combination of quantum-mechanical and many-body effects. The most widely used approach, which achieves a reasonable compromise between accuracy and computational cost, is Kohn-Sham (KS) density-functional theory (DFT) [10]. Although exact in principle, practical implementations of KS-DFT must heavily rely on approximations for the so-called exchange-correlation functional. As illustrated in the talk of Gero Friesecke, it has been realized a few years ago [2, 6] that the semiclassical limit of the Hohenberg-Kohn (HK) functional [9, 11, 12] leads to a multimarginal optimal transport (OT) problem with cost function given by the Coulomb repulsive potential. This allows us to investigate the mathematical structure of the exchange-correlation functional in a well defined limiting case.

This talk focuses on some promising applications of the OT limit of DFT, in particular the fact that, when used as an approximation for the exchange-correlation energy functional, it is able to address prototypical challenging problems for which the currently available approximations fail [13, 14, 3, 19]. Unfortunately, the OT limit is still too expensive to be computed exactly for many particles (marginals) and used routinely in computational chemistry and materials science problems.

Nonetheless, its mathematical structure has already inspired new approximations [20, 21, 1, 18] totally different from traditional ones, based on *integrals* of the density rather than the usual local density, local density gradients, etc. In this context, it is clearly important to fully understand the structure of the OT limit of DFT and to study how it is approached from the quantum regime.

Some important open questions in this respect are

(1) How is the OT limit approached by the HK functional? In Ref. [8] a conjecture for the next leading order, the so called "zero-point energy" (ZPE) has been presented and computed. A recent comparison with very accurate numerical results for the exact HK functional in the case $N = 2$ and dimension $d = 1$ seems to confirm the validity of the expansion. Is it possible to prove rigorously the ZPE conjecture at least for dimension $d = 1$, and possibly for the physically relevant dimension $d = 3$?

(2) How does the quantum statistics effects enter in the OT limit? The OT limit is semiclassical and does not distinguish between bosons (totally symmetric wavefunctions with respect to permutation of the particles) and fermions with different spin states (specified antisymmetry with respect to permutation of the particles). In Ref. [7] it has been argued that the effects of quantum statistics enter at orders that go exponentially to zero as $\hbar$ goes to zero. Again, a very recent comparison with numerically accurate calculations seems to confirm the conjecture. Can this be proven?

(3) Is there always a deterministic (Monge) solution? The Monge solution is physically appealing because it reflects the intuition that the particles in the semiclassical limit become "strictly correlated" [15, 17], in the sense that the position of one of them determines all the others via maps with a cyclic structure. The question is still open [4] for the general $d = 3$ case with number of particles (marginals) $N > 2$.

(4) How to build *approximate* (even if not optimal) maps whose functional derivative has a gradient that satisfies the usual equation in terms of the gradient of the cost evaluated on the maps [16]? This last question is very important, because Kohn-Sham DFT needs the functional derivative (or Kantorovich potential) as an effective potential in the KS single-particle equations. For example, for the spherically-symmetric case it has been shown [16] that simple *approximate* maps [17] even if not always optimal [5] get numerically very close to the true minimal cost and have the expected functional derivative. Is there a route to build similar approximations in the general $d = 3$ case?

References

[1] H. Bahmann, Y. Zhou, and M. Ernzerhof. *J. Chem. Phys.*, 145:124104, 2016.
[2] G. Buttazzo, L. De Pascale, and P. Gori-Giorgi. *Phys. Rev. A*, 85:062502, 2012.
[3] H. Chen, G. Friesecke, and C. B. Mendl. *J. Chem. Theory Comput*, 10:4360–4368, 2014.
[4] M. Colombo and S. Di Marino. In *Annali di Matematica Pura ad Applicata*, pages 1–14. Springer, Berlin Heidelberg, 2013.
[5] M. Colombo and F. Stra. *Math. Models Methods Appl. Sci.*, 26:1025, 2016.

[6] C. Cotar, G. Friesecke, and C. Klüppelberg. *Comm. Pure Appl. Math.*, 66:548, 2013.
[7] P. Gori-Giorgi, M. Seidl, and G. Vignale. *Phys. Rev. Lett.*, 103:166402, 2009.
[8] P. Gori-Giorgi, G. Vignale, and M. Seidl. *J. Chem. Theory Comput.*, 5:743, 2009.
[9] P. Hohenberg and W. Kohn. *Phys. Rev.*, 136:B 864, 1964.
[10] W. Kohn and L. J. Sham. *Phys. Rev. A*, 140:1133, 1965.
[11] M. Levy. *Proc. Natl. Acad. Sci. U.S.A.*, 76:6062, 1979.
[12] E. H. Lieb. *Int. J. Quantum. Chem.*, 24:24, 1983.
[13] F. Malet and P. Gori-Giorgi. *Phys. Rev. Lett.*, 109:246402, 2012.
[14] C. B. Mendl, F. Malet, and P. Gori-Giorgi. *Phys. Rev. B*, 89:125106, 2014.
[15] M. Seidl. *Phys. Rev. A*, 60:4387, 1999.
[16] M. Seidl, S. Di Marino, A. Gerolin, L. Nenna, and P. Gori-Giorgi. *in preparation*, 2017.
[17] M. Seidl, P. Gori-Giorgi, and A. Savin. *Phys. Rev. A*, 75:042511, 2007.
[18] S. Vuckovic, T. J. P. Irons, A. Savin, A. M. Teale, and P. Gori-Giorgi. *J. Chem. Theory Comput.*, 12:2598, 2016.
[19] S. Vuckovic, L. O. Wagner, A. Mirtschink, and P. Gori-Giorgi. *J. Chem. Theory Comput.*, 11:3153, 2015.
[20] L. O. Wagner and P. Gori-Giorgi. *Phys. Rev. A*, 90:052512, 2014.
[21] Y. Zhou, H. Bahmann, and M. Ernzerhof. *J. Chem. Phys.*, 143:124103, 2015.

Multimarginal Optimal Transport with Coulomb Cost and the Uniform Gas

Mathieu Lewin

In the talk I have discussed a few open problems in statistical mechanics, for what can be called the *Riesz gas*. Consider a bounded domain $\Omega_N = N^{1/d}\Omega \subset \mathbb{R}^d$ with $|\Omega| = 1$. Place then N particles in Ω_N and optimize their positions, such as to minimize the Riesz energy

$$E(s,\Omega_N) := \min_{x_1,\dots,x_N\in\Omega_N} \sum_{1\le j<k\le N} V_s(x_j - x_k), \tag{1}$$

where

$$V_s(x) = \begin{cases} |x|^{-s} & \text{for } s > 0, \\ -\log|x| & \text{for } s = 0, \\ -|x|^{-s} & \text{for } s < 0. \end{cases}$$

We are interested in the limit $N \to \infty$ with $\rho := N/|\Omega_N|$ fixed and since V_s is homogeneous, we have assumed $\rho = 1$. The general case follows by scaling.

When $s > d$, the system is *stable*, which means that the energy behaves like N. Indeed, using the integrability of V_s at infinity, one can prove [10] that the limit

$$\lim_{N\to\infty} \frac{E(s,N)}{N} := e_{\mathrm{R}}(s) \tag{2}$$

exists and does not depend on Ω. One famous conjecture [1] is that the system crystallizes for every $s > d$ (the particles asymptotically arrange on a periodic lattice). This was proved in dimension $d = 1$ in [12], but a proof is still lacking in

higher dimensions. Assuming that the particles sit on a lattice $\mathcal{L}$, one finds

$$e_{\mathrm{R}}(s) = \frac{1}{2} \sum_{\ell \in \mathcal{L} \setminus \{0\}} \frac{1}{|\ell|^s} := \zeta_{\mathcal{L}}(s). \tag{3}$$

The right side is called the *Epstein Zeta function* and it is a natural generalization of the famous Riemann Zeta function to higher dimensions. It is known that it is minimized for the triangular lattice in dimension $d = 2$, and conjectured that the Face Centered Cubic (FCC) is optimal in dimension $d = 3$. For $s \to \infty$, one finds the sphere packing problem. See [1] for a detailed description of the existing results and open problems on $\zeta_{\mathcal{L}}$.

For $s < d$, the model is unstable. Due to the long range of V_s, the leading order grows much faster than N and depends on Ω. More precisely, this is a *mean-field limit* with

$$\lim_{N\to\infty} \frac{E(s,\Omega_N)}{N^{1+\frac{s}{d}}} = \frac{1}{2} \inf_{\rho \in \mathcal{P}(\Omega)} \int_\Omega \int_\Omega \frac{d\rho(x)\, d\rho(y)}{|x-y|^s},$$

the Riesz capacity of the set Ω. The optimal probability ρ_Ω diverges at the boundary of Ω for $s > d-2$ and it is supported on $\partial\Omega$ in the Coulomb case $s = d-2$. In addition, we have $\frac{1}{N}\sum_{j=1}^N \delta_{x_{j,N}} \rightharpoonup \rho_\Omega$ in the sense of measures, where the $x_{j,N}$ minimize (1), so the particles concentrate very much close to the boundary, which is a sign of instability.

To stabilize the system, it is convenient to add a neutralizing uniform background, leading to Wigner's *Jellium model*:

$$E_{\mathrm{Jel}}(s,\Omega_N) := \min_{x_1,\dots,x_N \in \Omega_N} \Bigg\{ \sum_{1\le j<k\le N} V_s(x_j - x_k) - \sum_{j=1}^N \int_{\Omega_N} V_s(x_j - y)\, dy + \frac{1}{2} \int_{\Omega_N}\int_{\Omega_N} V_s(x-y)\, dx\, dy \Bigg\}. \tag{4}$$

For $d-2 \le s < d$, the background is enough to ensure the existence of the limit

$$\lim_{N\to\infty} \frac{E_{\mathrm{Jel}}(s,N)}{N} := e_{\mathrm{R}}(s)$$

and its independence with respect to Ω. We have used the same notation $e_{\mathrm{R}}(s)$ for a reason that we explain in a moment.

It is again conjectured that the system crystallizes (this has only been proved for $s = -1$ in dimension $d = 1$ but here we concentrate on $s \ge 0$). If the particles are arranged on a lattice $\mathcal{L}$, it was proved in [2, 3, 5] that

$$e_{\mathrm{Jel}}(s) = \zeta_{\mathcal{L}}(s)$$

where $\zeta_{\mathcal{L}}(s)$ is the analytic extension of the Epstein Zeta function appearing in (3) (the latter has an extension to $\mathbb{C} \setminus \{d\}$ like for the usual 1D Zeta function). The fact that the Jellium problem is the analytic extension of the usual Riesz energy for $s > d$ (assuming crystallization) shows that the Jellium model is the most natural extension for $d-2 \le s < d$. Assuming that the system crystallizes, it is known that the particle must be placed on a triangular lattice in dimension $d = 2$,

and conjectured that it must be the FCC lattice for $3/2 \leq s < 3$ and the Body Centered Cubic (BCC) lattice for $1 \leq s \leq 3/2$ in dimension $d = 3$.

Let us finally introduce the *Uniform Electron Gas* which is a cornerstone of *Density Functional Theory* in Quantum Chemistry and Physics. For a given density ρ with $\int_{\mathbb{R}^d} \rho = N$ (an integer), we introduce the lowest Coulomb energy that can be reached by a probabilistic N-particle system having this density ρ, to which we subtract the average one-particle energy:

$$\mathcal{E}_s(\rho) := \inf_{\substack{\mathbb{P}\in\otimes_s^N \mathcal{P}(\mathbb{R}^d) \\ (e_1)\#\mathbb{P}=\rho/N}} \left\{ \int_{(\mathbb{R}^d)^N} \sum_{1\leq j<k\leq N} V_s(x_j - x_k)\, d\mathbb{P}(x_1, ..., x_N) \right\} - \frac{1_{0\leq s<d}}{2} \int_{\mathbb{R}^d}\int_{\mathbb{R}^d} V_s(x-y)\, d\rho(x)\, d\rho(y). \tag{5}$$

This is a multi-marginal optimal transport problem. We only fix the first marginal $(e_1)\#\mathbb{P}$ of the N-particle probability $\mathbb{P}$ since the latter is assumed to be symmetric.

The Lieb-Oxford inequality [6, 7] states that

$$\mathcal{E}_s(\rho) \geq -e_{\mathrm{LO}}(s) \int_{\mathbb{R}^d} \rho(x)^{1+\frac{s}{d}}\, dx. \tag{6}$$

This can be used to prove that the following limit exists

$$\lim_{N\to\infty} \frac{\mathcal{E}_s(1_{\Omega_N})}{N} = e_{\mathrm{UEG}}(s)$$

The quantity $e_{\mathrm{UEG}}(s)$ is the one which should be used in all the DFT programs (for $s = 1$ and $d = 3$), but it is commonly assumed that

$$e_{\mathrm{UEG}}(s) \quad \text{coincides} \quad e_{\mathrm{R}}(s), \quad \text{for} \quad s \in [d-2, \infty) \setminus \{d\}. \tag{7}$$

Indeed, if the system crystallizes, one can construct a trial state with constant density, by clamping the particles on the lattice and then average over the translations of this lattice. This amounts to choosing a Monge-type state where the transport map is only moving the particles according to translations of the lattice. Surprisingly, a calculation in [5] showed that this trial state has the right energy for $s > d-2$, but not for $s = d-2$ where a positive shift appears. Recent numerical computations in [11] for $s = 1$ in dimension $d = 3$ have shown that the Monge state described above is not at all optimal, leaving the question (7) widely open.

Since obviously $e_{\mathrm{LO}}(s) \geq -e_{\mathrm{UEG}}(s)$, the Uniform Electron Gas is also a way to get lower bounds on the best Lieb-Oxford constant (6). In [8, 9] it was even conjectured that the two must also be equal.

In dimension $d = 1$, recent works [4] in optimal transportation imply that the Monge state is indeed exact, and therefore we indeed have the equality of the UEG and Riesz energies for $s > -1$ and non-equality for $s = -1$ (Coulomb case), with the shift computed in [5].

References

[1] X. Blanc and M. Lewin, *The Crystallization Conjecture: A Review*, EMS Surv. Math. Sci., 2 (2015), pp. 219–306.

[2] D. Borwein, J. M. Borwein, and R. Shail, *Analysis of certain lattice sums*, J. Math. Anal. Appl., 143 (1989), pp. 126–137.

[3] D. Borwein, J. M. Borwein, and A. Straub, *On lattice sums and Wigner limits*, J. Math. Anal. Appl., 414 (2014), pp. 489–513.

[4] M. Colombo, L. De Pascale, and S. Di Marino, *Multimarginal optimal transport maps for one-dimensional repulsive costs*, Canad. J. Math., 67 (2015), pp. 350–368.

[5] M. Lewin and E. H. Lieb, *Improved Lieb-Oxford exchange-correlation inequality with gradient correction*, Phys. Rev. A, 91 (2015), p. 022507.

[6] E. H. Lieb and S. Oxford, *Improved lower bound on the indirect Coulomb energy*, Int. J. Quantum Chem., 19 (1980), pp. 427–439.

[7] E. H. Lieb, J. P. Solovej, and J. Yngvason, *Ground states of large quantum dots in magnetic fields*, Phys. Rev. B, 51 (1995), pp. 10646–10665.

[8] M. M. Odashima and K. Capelle, *How tight is the Lieb-Oxford bound?*, J. Chem. Phys., 127 (2007), p. 054106.

[9] E. Räsänen, S. Pittalis, K. Capelle, and C. R. Proetto, *Lower bounds on the exchange-correlation energy in reduced dimensions*, Phys. Rev. Lett., 102 (2009), p. 206406.

[10] D. Ruelle, *Statistical mechanics. Rigorous results*, Singapore: World Scientific. London: Imperial College Press , 1999.

[11] M. Seidl, S. Vuckovic, and P. Gori-Giorgi, *Challenging the lieb-oxford bound in a systematic way*, Molecular Physics, 114 (2016), pp. 1076–1085.

[12] W. J. Ventevogel and B. R. A. Nijboer, *On the configuration of systems of interacting particles with minimum potential energy per particle*, Phys. A, 99 (1979), pp. 569–580.

Gradient flow and functional inequalities for quantum Markov semigroups

Eric A. Carlen

(joint work with Jan Maas)

We study a class of ergodic quantum Markov semigroups on finite-dimensional unital C^*-algebras. These semigroups have a unique stationary state σ, and we are concerned with those that satisfy a quantum detailed balance condition with respect to σ. These are close, but non-commutative, analogs of the semigroups associated to Kolmogorov equations that can be written as gradient flow with respect to the 2-Wasserstein metric for the relative entropy, and as Felix Otto showed, this perspective on these evolution equations is very fruitful, especially with regard to functional inequalities governing rates of relaxation to equilibrium. In [1] is was shown that the infinite temperature Fermi oscillator semigroup, introduced by Leonard Gross in the 1970's, is gradient flow with respect to a natural non-commutative mass transportation metric for the relative entropy with respect to its invariant state, which is the normalized trace, and this was used to deduce certain functional inequalities.

The present work, [2], greatly extends [1]: We provide the appropriate non-commutative transportation metric with respect to which ergodic quantum Markov semigroups satisfying detailed balance are gradient flow for the relative entropy

with respect to their unique invariant state σ. In finite dimensions, this transportation metric is a smooth Riemannian metric on the set of density matrices, and is a non-commutative analog of the 2-Wasserstein metric. In several interesting cases we are able to show, in analogy with work of Otto on gradient flows with respect to the classical 2-Wasserstein metric, that the relative entropy is strictly and uniformly convex with respect to the Riemannian metric introduced here. As a consequence, we obtain a number of new inequalities for the decay of relative entropy for ergodic quantum Markov semigroups with detailed balance.

References

[1] E. A. Carlen and J. Maas, *An analog of the 2-Wasserstein metric in non-commutative probability under which the fermionic Fokker-Planck equation is gradient flow for the entropy*, Comm. Math. Phys. **331** (2014), 887–926.

[2] E. A. Carlen and J. Maas, *Gradient flow and entropy inequalities for quantum Markov semigroups with detailed balance*, arXiv:1609.01254, 2016.

New results on discrete Bakry-Emery methods and displacement convexity

Ansgar Jüngel

Several numerical schemes which preserve the entropy structure of the underlying evolution equations are discussed. Three different techniques are presented: the discrete gradient flow scheme of the 1D Fokker-Planck equation; the Bochner-Bakry-Emery method for stochastic processes; and a systematic summation-by-parts approach which also allows to treat nonlinear equations and which may lead to a general discrete Bakry-Emery method.

Least action principles for incompressible flows and optimal transport between shapes

Dejan Slepčev

(joint work with Jian-Guo Liu an Robert L. Pego)

Arnold observed in the 1960s, [1], that the Euler equations of incompressible fluid flow correspond formally to geodesic equations in a group of volume-preserving diffeomorphisms. Working in an Eulerian framework, we discuss [5] incompressible flows of shapes as critical paths for action (kinetic energy) along transport paths constrained to be shape densities (characteristic functions). The geodesic paths correspond to critical paths for the action

$$\mathcal{A} = \int_0^1 \int_{\mathbb{R}^d} \rho |v|^2 \, dx \, dt \,, \tag{1}$$

where $\rho = (\rho_t)_{t\in[0,1]}$ is a path of *shape densities* transported by a velocity field $v \in L^2(\rho \, dx \, dt)$ according to the continuity equation

$$\partial_t \rho + \nabla \cdot (\rho v) = 0 \,. \tag{2}$$

Here, saying that ρ_t is a shape density means that ρ_t is constrained to be a characteristic function of a fluid domain Ω_t:

$$\rho_t = \chi_{\Omega_t}\,, \quad t \in [0,1]. \tag{3}$$

The velocity is divergence free in the interior of Ω_t.

We study the action (1) subject to given endpoint conditions of the form

$$\rho_0 = \chi_{\Omega_0}, \quad \rho_1 = \chi_{\Omega_1}\,. \tag{4}$$

These conditions differ from Arnold-style conditions that fix the flow-induced volume-preserving diffeomorphism between Ω_0 and Ω_1, and correspond instead to fixing only the *image* of this diffeomorphism. It turns out that the geodesic equations that result are precisely the Euler equations for *potential flow* of an incompressible, inviscid fluid occupying domain Ω_t, with *zero pressure and zero surface tension* on the free boundary $\partial\Omega_t$. In short, the geodesic equations are classic water wave equations with zero gravity and surface tension.

The problem of minimizing this action exhibits an instability associated with microdroplet formation, with the following outcomes: Any two shapes of equal volume can be approximately connected by an Euler spray—a countable superposition of ellipsoidal droplet solutions of incompressible Euler equations with zero pressure on the droplet boundaries. The infimum of the action is the Wasserstein distance squared, and is almost never attained except in dimension 1. Every Wasserstein geodesic between bounded densities of compact support provides a solution of the (compressible) pressureless Euler system that is a weak limit of (incompressible) Euler sprays. Each such Wasserstein geodesic is also the unique minimizer of a relaxed least-action principle (introduced by Brenier [3]; see also [4]) for a two-fluid mixture theory corresponding to incompressible fluid mixed with vacuum.

References

[1] V. Arnold, *Sur la géométrie différentielle des groupes de Lie de dimension infinie et ses applications à l'hydrodynamique des fluides parfaits*, Ann. Inst. Fourier (Grenoble), 16 (1966), pp. 319–361.

[2] J.-D. Benamou and Y. Brenier, *A computational fluid mechanics solution to the Monge-Kantorovich mass transfer problem*, SIAM Numer. Math. (2000) pp. 375–393.

[3] Y. Brenier, *The least action principle and the related concept of generalized flows for incompressible perfect fluids*, J. Amer. Math. Soc., 2 (1989), pp. 225–255.

[4] Y. Brenier, *Minimal geodesics on groups of volume-preserving maps and generalized solutions of the Euler equations*, Comm. Pure Appl. Math., 52 (1999), pp. 411–452.

[5] J-G. Liu, R.L. Pego, and D. Slepčev, *Least action principles for incompressible flows and optimal transport between shapes*, arXiv preprint arXiv:1604.03387, (2016).

A gradient flow to the Boltzmann equation

Matthias Erbar

Since the pioneering work of Otto [8] it is well know that many diffusion equations can be cast as gradient flows in the space of probability measures with the relevant geometry being induced by the L^2 Wasserstein distance. Otto's approach has been widely used in the study of the trend to equilibrium, stability questions and construction of solutions.

In this talk, I presented a characterization of the spatially homogeneous Boltzmann equation as a gradient flow of the entropy based on [3]. Crucial for this is the identification of a novel geometry on the space of probability measures that takes the collision process between particles into account.

The spatially homogeneous Boltzmann equation is given as

$$\partial_t f = \int_{\mathbb{R}^d}\int_{S^{d-1}} \big[f'f'_* - ff_*\big]B(v-v_*,\omega)\mathrm{d}v_*\mathrm{d}\omega\ ,$$

where f is a probability density on $\mathbb{R}^d$. The shorthand f_*, f', f'_* stands for $f(v_*), f(v'), f(v'_*)$ with v, v_* and v', v'_* denoting the pre- and post-collisional velocities respectively related according to $v' = v - \langle v-v_*,\omega\rangle\omega$, $v'_* = v_* + \langle v-v_*,\omega\rangle\omega$ with $\omega \in S^{d-1}$. The collision kernel B encodes the microscopic details of the particle interaction and we assume it to be continuous and to satisfy $c^{-1} \le B \le c$ for some constant $c > 0$.

Boltzmann's H-Theorem asserts that the entropy $\mathcal{H}(f) = \int f\log f$ is non-increasing along solutions, i.e. $\frac{\mathrm{d}}{\mathrm{d}t}\mathcal{H}(f_t) = -D(f_t) \le 0$, where

$$D(f_t) = \frac{1}{4}\int \log\frac{f'f'_*}{ff_*}(f'f'_* - ff_*)B(v-v_*,\omega)\mathrm{d}\omega\mathrm{d}v_*\mathrm{d}v\ .$$

Our goal is to characterize the homogeneous Boltzmann equation as the evolution that decreases the entropy *as fast as possible*. This gradient flow structure rests on a novel geometry on the space of probability measures. Given probabilities f_0 and f_1 we solve (a suitable relaxation of) the minimization problem

$$\mathcal{W}_B(f_0,f_1)^2 = \inf\left\{\frac{1}{4}\int_0^1\int |\bar\nabla\psi_t|^2\Lambda(f_t)B(v-v_*,\omega)\mathrm{d}\omega\mathrm{d}v_*\mathrm{d}v\mathrm{d}t\right\}\ ,$$

where the infimum runs over all curves of probability densities $(f_t)_t$ connecting f_0 and f_1 and all functions $\psi : [0,1]\times\mathbb{R}^d \to \mathbb{R}$ related via

$$\partial_t f_t(v) + \frac{1}{4}\int \bar\nabla\psi_t\Lambda(f_t)B(v-v_*,\omega)\mathrm{d}\omega\mathrm{d}v_* = 0\ .$$

Here, we have set $\bar\nabla\phi = \phi' + \phi'_* - \phi - \phi_*$ and $\Lambda(f)$ is shorthand for $\Lambda\big(ff_*, f'f'_*\big)$, where $\Lambda(s,t) = (s-t)/(\log s - \log t)$ denotes the logarithmic mean.

It turns out that $\mathcal{W}_B$ defines an extended, separable and complete distance on the set $\mathcal{P}_*(\mathbb{R}^d)$ of probabilities with zero mean and unit variance. Moreover, each pair of densities at finite distance can be joined by a geodesic, i.e. an optimal curve $(f_t)_t$ in the problem above.

We obtain the following variational characterization of the Boltzmann equation

Theorem 1: *For any curve $(f_t)_{t\geq 0}$ of probability densities in $\mathcal{P}_*(\mathbb{R}^d)$ with $\mathcal{H}(f_0) < \infty$ we have that*

$$J_T(f) := \mathcal{H}(f_T) - \mathcal{H}(f_0) + \frac{1}{2}\int_0^T D(f_t) + |\dot{f}_t|^2_{\mathcal{W}_B}\mathrm{dt} \geq 0 \quad \forall \mathrm{T} \geq 0\ .$$

The solution $(f_t)_t$ to the homogeneous Boltzmann equation is the unique curve with $J_T(f) = 0$ for all T.

Here $|\dot{f}_t|_{\mathcal{W}_B}$ denotes the metric speed w.r.t. $\mathcal{W}_B$. In this sense, the Boltzmann equation is a steepest descent of the entropy, decreasing it as fast as possible. To motivate this result, note that for a smooth function E on $\mathbb{R}^n$ and any smooth curve x we have

$$E(x_T) - E(x_0) = \int_0^T \nabla E(x_t)\dot{x}_t\mathrm{dt} \geq -\frac{1}{2}\int_0^{\mathrm{T}} |\nabla \mathrm{E}|^2(\mathrm{x_t}) + |\dot{\mathrm{x}}_\mathrm{t}|^2\mathrm{dt}\ .$$

with equality if and only if x is a gradient flow curve of E, i.e. $\dot{x}_t = -\nabla E(x_t)$. In metric spaces this characterization can be used to define gradient flows by replacing $|\dot{x}_t|$with the metric speed of the curve and $|\nabla E|$ by an upper gradient. We refer to [1] for a detailed account on gradient flows in metric spaces. Thus, Theorem 1 characterizes the Boltzmann equation as the gradient flow of the entropy in the space $(\mathcal{P}_*(\mathbb{R}^d), \mathcal{W}_B)$.

As a first application, we obtain a time-discrete variational approximation scheme for the Boltzmann equation related to the implicit Euler scheme for the gradient flow structure. Given a time step $\tau > 0$ and an initial datum $f_0 \in \mathcal{P}_*(\mathbb{R}^d)$ with $\mathcal{H}(f_0) < \infty$ define iteratively

$$f_0^\tau = f_0\ , \quad f_{n+1}^\tau \in \operatorname*{argmin}_g \Big[\mathcal{H}(g) + \frac{1}{2\tau}\mathcal{W}_B(g, f_n^\tau)^2\Big]\ .$$

and let $f_t^\tau = f_n^\tau$ for $t \in ((n-1)\tau, n\tau]$ be the piecewise constant interpolation.

Theorem 2: *As τ goes to zero, f_t^τ converges weakly to the solution f_t of the Boltzmann equation with initial datum f_0.*

As a second application the gradient flow structure can be used to give a new and simple proof of the convergence of Kac's random walk, an N-particle stochastic dynamics, to the solution of the spatially homogeneous Boltzmann equation recovering results of Sznitman [10] (see also Mischler–Mouhot [6] and Norris [7] for quantitative results). It has been shown recently that Kac's random walk (in fact any continuous time Markov chain) has a gradient flow structure induced by a suitable transportation distance, see [4, 5, 2], i.e. a characterization similar to Theorem 1 holds. The crucial idea is to note that these gradient flow structures are consistent in the limit $N \to \infty$. Using the approach of Sandier–Serfaty [9] of evolutionary Γ-convergence this boils down to proving simple lim inf-estimates between the constituent elements of the gradient flow structure, the entropy, dissipation and metric speed, yielding the convergence of the gradient flows.

References

[1] L. Ambrosio, N. Gigli, G. Savaré, *Gradient flows in metric spaces and in the space of probability measures*, Lectures in Mathematics ETH, Birkhäuser, Zürich (2008).
[2] M. Erbar, *Gradient flows of the entropy for jump processes*, Ann. Inst. Henri Poincaré Probab. Stat. **50**(3) (2014), 920–945.
[3] M. Erbar, *A gradient flow approach to the Boltzmann equation*, arXiv 1603:0540.
[4] J. Maas, *Gradient flows for the entropy for finite Markov chains*, J. Funct. Anal. **261**(8) (2011), 2250–2292.
[5] A. Mielke, *A gradient flow structure for reaction-diffusion systems and for energy-drift-diffusion systems*, Nonlinearity **24** (2011), 1329–1346.
[6] S. Mischler, C. Mouhot, *Kac's program in kinetic theory*, Invent. Math. **193** (2013), 1–147.
[7] J. Norris, *A consistency estimate for Kac's model of elastic collisions in a dilute gas*, Ann. Appl. Probab. **26** (2016), 1029–1081.
[8] F. Otto, *The geometry of dissipative evolution equations: the porous medium equation*, Comm. Partial Differential Equations **26** (2001), 101–174.
[9] E. Sandier, S. Serfaty, *Gamma-convergence of gradient flows with applications to Ginzburg-Landau*, Comm. Pure Appl. Math. **57**(12) (2004), 1627–1672.
[10] A. Sznitman, *Équations de type de Boltzmann, spatialement homogènes*, Z. Wahrsch. Verw. Gebiete **66**(4) (1984), 559–592.

Heat flows on time-dependent metric measure spaces

Eva Kopfer

(joint work with Karl-Theodor Sturm)

Motivation. Let us start by describing the heat flow on a closed manifold equipped with a family of Riemannian tenors $(g_t)_{t\in[0,T]}$. We need to define two objects; firstly, the heat flow which solves

$$\partial_t u_t = \Delta_t u_t \qquad \text{on } (0,T)\times M$$

and the adjoint heat flow which solves

$$\partial_t v_t = -\Delta_t v_t - \frac{1}{2} tr(\partial_t g_t) v_t, \qquad \text{on } (0,T)\times M$$

where Δ_t denotes the Laplace Beltrami operator wrt the metric g_t. These objects are adjoint in the following sense

$$\int_0^T \int_M (\partial_t - \Delta_t) u_t \cdot v_t \, d\mathrm{vol}_t \, dt = \int_0^T \int_M u_t \cdot (-\partial_t - \Delta_t - \frac{1}{2} tr(\partial_t g_t)) v_t \, d\mathrm{vol}_t \, dt + \left[\int_M u_t v_t \, d\mathrm{vol}_t \right]_0^T .$$

Note that the adjoint heat equation has to be interpreted backwards in time.

We investigate tensors which evolve as a super-Ricci flow, cf. [1]. A smooth closed manifold M equipped with a family of Riemannian tensors $(g_t)_{t\in[0,T]}$ is said to be a super-Ricci flow if

$$\mathrm{Ric}(g_t) \geq -\frac{1}{2}\partial_t g_t$$

provided that g is differentiable in t.

Possible examples may include Ricci flows, i.e. $\mathrm{Ric}(g_t) = -\frac{1}{2}\partial_t g_t$ and, of course, the static case $\mathrm{Ric} \geq 0$. Let us recall the well-known fact that $\mathrm{Ric} \geq 0$ is equivalent to geodesic convexity of the relative entropy functional on the space of Borel probability measures equipped with the L^2-Kantorovich distance W

$$\mathrm{Ent}(\mu_t|\mathrm{vol}) \leq (1-t)\mathrm{Ent}(\mu_0|\mathrm{vol}) + t\mathrm{Ent}(\mu_1|\mathrm{vol}) \quad \forall W\text{-geodesics } \mu_t,$$

transport estimates for the heat flows $P_t\mu, P_t\nu$ on measures

$$W(P_t\mu, P_t\nu) \leq W(\mu, \nu),$$

and Bakry-Émery type gradient estimates

$$|\nabla P_t(u)|^2 \leq P_t(|\nabla u|^2).$$

In the same manner as Sturm, Lott and Villani defined lower Ricci curvature bounds for metric measure spaces, Sturm [3] introduced super-Ricci flows on metric measure spaces.

Definition 1. *We say that a family of metric measure spaces $(X, d_t, m_t)_{t\in[0,T]}$ is a super-Ricci flow if for each W_t-geodesic $(\mu_a)_{a\in[0,1]}$*

$$\partial_a \mathrm{Ent}(\mu_a|m_t)|_{a=1} - \partial_a \mathrm{Ent}(\mu_a|m_t)_{a=0} \geq -\frac{1}{2}\partial_t W_t(\mu^0, \mu^1)^2.$$

We give two examples as an illustration.

Consider the spherical suspension $\Sigma(S^2(1/\sqrt{3}) \times S^2(1/\sqrt{3}))$ over the product of the two-spheres with radius $1/\sqrt{3}$. It can be shown that outside of the north and south pole this space is a 5-dimensional (incomplete) Riemannian manifold with constant Ricci curvature $\mathrm{Ric}(g_0) = 4$. Then $g_t = (1-8t)g_0$ prescribes a Ricci flow and (M, g_t) shrinks homothetically to a point in finite time. We obtain a super-Ricci flow if we include the north and the south pole again and consider the associated sequence of metric measure spaces. Intuitively, the reason for this is justified by the fact that the optimal path measure assigns zero mass to geodesics which cross the poles.

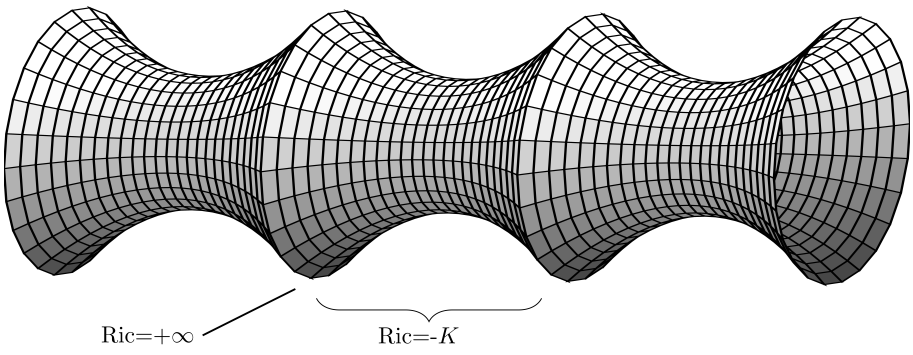


FIGURE 1. Surface of revolution of a piecewise hyperbolic space

An example of a surface is depicted in Figure 1. Under the evolution of a Ricci flow regions with negative Ricci curvature will inflate, while the edges will smooth out, whereas a super-Ricci flow may keep the edges but regions with negative curvature still have to inflate.

Main results. Let $(X, d_t, m_t)_{t\in[0,t]}$ be a family of metric measure spaces such that the map $t \mapsto \log d_t(x, y)$ is Lipschitz continuous and $m_t = e^{-f_t} m$ for some suitable function $f\colon [0, T] \times X \to \mathbb{R}$ presumed to be Lipschitz continuous in time and space. Moreover we assume that each (X, d_t, m_t) is a $\mathrm{RCD}(K, N)$ space, i.e. the space satisfies the curvature dimension condition $\mathrm{CD}(K, N)$ in the sense of Lott, Sturm and Villani and for each t the Cheeger energy defines a strongly local Dirichlet form $\mathcal{E}_t(u, v) = \int \Gamma_t(u, v)\, dm_t = -\int \Delta_t uv\, dm_t$.

Theorem 1 (cf. [2])**.** (1) *There exists a unique solution $(P_{t,s}u)_{t\geq s}$ to $\partial_t u_t = \Delta_t u_t$ on $(s, T) \times X$ with $u_s = u$ in a weak distributional sense.*
(2) *There exists a unique solution $(P^*_{t,s}v)_{t\geq s}$ to $\partial_s v_s = -\Delta_s v_s + \partial_s f_s v_s$ on $(0, t) \times X$ with $v_t = v$ in a weak distributional sense.*

Now we are ready to state our main theorem which can be thought of as a time-dependent version of lower Ricci curvature bounds $\mathrm{Ric} \geq 0$ and a generalization of the results obtained in [1].

Theorem 2 (cf. [2])**.** *The following are equivalent*

(1) $(X, d_t, m_t)_{t\in[0,t]}$ *is a super-Ricci flow.*
(2) *For each nonnegative u, h with the same mass* $W_s(P^*_{t,s}um_s, P^*_{t,s}hm_s) \leq W_t(um_t, hm_t)$
(3) *For each Lipschitz function u* $\Gamma_t(P_{t,s}u) \leq P_{t,s}(\Gamma_s u)$
(4) *"Bochner's inequality"* $\frac{1}{2}\Delta_t\Gamma_t(u) - \Gamma_t(u, \Delta_t u) \geq \frac{1}{2}\partial_t\Gamma_t(u)$.

References

[1] R. McCann and P. Topping, *Ricci flow, entropy and optimal transportation*, American Journal of Mathematics **132** (2010), 711–730.
[2] E. Kopfer and K.-T. Sturm *Super-Ricci flows for Metric Measure Spaces. II*, arXiv:1611.02570 (2016).
[3] K.-T. Sturm *Super-Ricci flows for Metric Measure Spaces.*, arXiv:1603.02193 (2016).

An entropic gradient structure for quantum Markov semigroups

Markus Mittnenzweig
(joint work with Alexander Mielke)

We describe a recently discovered entropic gradient structure for quantum Markov semigroups, see also the reports by Eric Carlen and Tryphon Georgiou on this topic. By a quantum Markov semigroup, also known as Lindblad equations, we mean an equation of the type

$$\dot\rho = \frac{\mathrm{i}}{\hbar}\big[\rho, H\big] \ + \mathcal{L}\rho \quad \text{with } \mathcal{L}A = \sum_{n=1}^{N^2-1} \gamma_n\big([Q_n, AQ_n^*] + [Q_nA, Q_n^*]\big). \tag{1}$$

for arbitrary operators $Q_n \in L(\mathfrak{h}, \gamma_n > 0$ and $\mathfrak{h}$ being some finite-dimensional complex Hilbert space. Such equations arise for instance in the fields of quantum optics and quantum information theory, where one needs to describe the influence of the environment on the quantum system, which is modeled by the dissipative term $\mathcal{L}\rho$. Here, we also ask the generator $\mathcal{L}$ to satisfy a detailed balance condition (DBC) with respect to the thermal equilibrium $\rho_\beta = Z_\beta^{-1}\exp(-\beta H)$, by which we mean that $\mathcal{L}\rho_\beta = 0$ and that $\mathcal{L}^*$ is symmetric with respect to the weighted operator inner product $(A, B) \mapsto \mathrm{tr}(A^*B\rho_\beta)$, also called GNS inner product. This class of Lindbladians arises naturally as the weak coupling limit of a quantum system coupled to a heat bath.

Moreover, the relative entropy (or free energy)

$$\mathcal{F} = \mathrm{tr}\big(\rho(\log\rho - \log\rho_\beta)\big)$$

is a Liapunov function for (1). Here we want to show more, namely that $\mathcal{L}$ can be written as

$$\mathcal{L} = -\mathbb{K}(\rho)\mathrm{D}\mathcal{F}(\rho)$$

where $\mathrm{D}\mathcal{F}(\rho) = \log\rho - \log\rho_\beta$ and $\mathbb{K}(\rho)$ is a positive symmetric operator, which we will simply call Onsager operators. Thus, our aim is the construction of an Onsager operator $\mathbb{K}$ which generalizes the Wasserstein operator $\mathbb{K}_{\mathrm{Wass}}(u) : \mu \mapsto -\mathrm{div}\big(\rho\nabla\mu\big)$ for the Fokker-Planck equation. The crucial point is that $\mathbb{K}(\rho)$ has to depend on ρ in a very specific way to obtain the relation

$$-\mathbb{K}(\rho)\big(\log\rho + \beta H\big) = \mathcal{L}\rho,$$

where the right-hand side is linear in ρ. In the Fokker-Plank equation this is achieved by the chain rule $u\nabla\big(\log u + V\big) = \nabla u + u\nabla V$. In the quantum case, a similar relation exists. it involves the use of the Kubo-Mori operator

$$\mathcal{C}\rho : L(\mathfrak{h}) \to L(\mathfrak{h}); \quad A \mapsto \mathcal{C}\rho A := \int_0^1 \rho^s\, A\, \rho^{1-s}\mathrm{d}s,$$

which satisfies for all $Q \in L(\mathfrak{h})$ the fundamental relation

$$\mathcal{C}\rho\,\big[Q, \log\rho\big] \ = \ \big[Q, \rho\big] \tag{2}$$

which we will call the miracle relation. The idea is, that in the noncommutative case, commutators replace derivatives, because they satisfy all the algebraic properties of derivatives, in particular the product rule $\nabla(ab) = (\nabla a)b + a(\nabla b)$.
However, we cannot directly apply the miracle relation (2) to (1) since we need to use $\log\rho - \log\rho_\beta$ as the thermodynamic driving force. But then, the expression $\mathcal{C}_\rho\big[Q, \log\rho_\beta\big]$ is nonlinear in ρ, so this approach will not give a gradient structure for linear Lindblad generators $\mathcal{L}\rho$. The solution to this problem is a suitable generalization of the miracle identity (2) as will be described in the following.

The starting point is a tensor-product representation of Lindblad operators. We set $\mathfrak{h}_1 = \mathfrak{h}$ and choose an arbitrary second Hilbert space $\mathfrak{h}_2$ assuming that $\mathfrak{h}_1$ and $\mathfrak{h}_2$ are both finite-dimensional. For an arbitrary Hermitian $\mathbb{Q} \in L(\mathfrak{h}_1 \otimes \mathfrak{h}_2)$ and a $\hat\sigma \in L(\mathfrak{h}_2)$ with $\hat\sigma = \hat\sigma^* > 0$ one sees that

$$\mathcal{L}\rho = -\mathrm{tr}_{\mathfrak{h}_2}\Big(\big[\mathbb{Q}, [\mathbb{Q}, \rho\otimes\hat\sigma]\big]\Big) \tag{3}$$

is indeed a Lindblad operator. Here $\mathrm{tr}_{\mathfrak{h}_2}$ is the partial trace over $\mathfrak{h}_2$, i.e. $\mathrm{tr}_{\mathfrak{h}_2}(A\otimes B) = A\,\mathrm{tr}(B)$. Moreover, it can be shown easily that this $\mathcal{L}$ satisfies the DBC with respect to ρ_β, if the commutation relation

$$\big[\,\mathbb{Q}\,,\,\rho_\beta\otimes\hat\sigma\,\big] \;=\; 0$$

holds in addition. Under this condition it is then straightforward to show the following generalization of the miracle identity:

$$\mathcal{C}_{\rho\otimes\hat\sigma}\big[\,\mathbb{Q}\,,\,(\log\rho+\beta H)\otimes\mathbf{1}_{\mathfrak{h}_2}\,\big] = \big[\,\mathbb{Q}\,,\,\rho\otimes\hat\sigma\,\big]. \tag{4}$$

Indeed, it suffices to use the fact that $\mathbb{Q}$ also commutes with

$$\log(\rho_\beta\otimes\hat\sigma) = -\beta H\otimes\mathbf{1}_{\mathfrak{h}_2} + \mathbf{1}_{\mathfrak{h}_1}\otimes\log\hat\sigma$$

and then apply the classical miracle identity (2). With this, we can define the Onsager operator

$$\mathbb{K}(\rho)\xi = \mathrm{tr}_{\mathfrak{h}_2}\Big(\big[\mathbb{Q}, \mathcal{C}_{\rho\otimes\hat\sigma}[\mathbb{Q}, \xi\otimes\mathbf{1}_{\mathfrak{h}_2}]\big]\Big),$$

which is a symmetric and positive semidefinite operator and satisfies the desired relation

$$-\mathbb{K}(\rho)\big(\log\rho + \beta H\big) = -\mathrm{tr}_{\mathfrak{h}_2}\Big(\big[\mathbb{Q}, [\mathbb{Q}, \rho\otimes\hat\sigma]\big]\Big) = \mathcal{L}\rho.$$

Finally let us mention, that the following generalization $\mathcal{D}_\rho^\alpha$ of the Kubo-Mori operator lies at the core of this gradient structure:

$$\mathcal{D}_\rho^\alpha A := \mathrm{e}^{-\alpha/2}\int_0^1 \mathrm{e}^{s\alpha}\rho^s A\,\rho^{1-s}\mathrm{d}s$$

This operator satisfies the following generalization of the identity (2):

$$\mathcal{D}_\rho^\alpha\big([Q, \log\rho] - \alpha Q\big) = \mathrm{e}^{-\alpha/2}Q\rho - \mathrm{e}^{\alpha/2}\rho Q.$$

If we assume additionally that $[Q, H] = \alpha Q$, then again we obtain a noncommutative variant of the chain rule:

$$\mathcal{D}_\rho^\alpha\big([Q, \log\rho + \beta H]\big) = \mathrm{e}^{-\beta\alpha/2}Q\rho - \mathrm{e}^{\beta\alpha/2}\rho Q.$$

References

[1] M. Mittnenzweig, A. Mielke, *Computing certain invariants of topological spaces of dimension three*, preprint, arXiv:1609.05765 (2016).
[2] E. A. Carlen, J. Maas *Gradient flow and entropy inequalities for quantum Markov semigroups with detailed balance*, preprint, arXiv:1609.01254 (2016).
[3] Y. Chen, T. T. Georgiou, A. Tannenbaum *Matrix Optimal Mass Transport: A Quantum Mechanical Approach*, preprint, arXiv:1610.03041 (2016).

Convergence to equilibrium for renormalised solutions to reaction-diffusion systems

Bao Q. Tang
(joint work with Klemens Fellner)

Consider N chemical substances $S_1, \dots, S_N$ reacting in R reactions of the form

$$y_{r,1}S_1 + \dots + y_{r,N}S_N \xrightarrow{k_r} y'_{r,1}S_1 + \dots + y'_{r,N}S_N \quad \text{or shortly} \quad y_r \xrightarrow{k_r} y'_r \tag{1}$$

for all $r = 1, \dots, R$, where $y_r, y'_r \in \mathbb{N}^N$ are stoichiometric coefficients, and $k_r > 0$ is the reaction rate constant . Denote by u_i the concentration of S_i and $u = (u_1, \dots, u_N)$ the vector of concentrations, the considered reaction-diffusion system is of the form

$$\begin{aligned} \partial_t u_i - d_i \Delta u_i &= f_i(u), && x \in \Omega, \\ \nabla u_i \cdot \nu &= 0, && x \in \partial\Omega, \\ u_i(x,0) &= u_{i,0}(x), && x \in \Omega, \end{aligned} \tag{2}$$

where Ω is a bounded domain in $\mathbb{R}^n$ with smooth boundary $\partial\Omega$, $d_i > 0$ is the diffusion coefficient, and the nonlinearities $f_i : \mathbb{R}^N \to \mathbb{R}$ are defined through the reactions (1) by using *law of mass action*,

$$f_i(u) = \sum_{r=1}^{R} k_r (y'_{r,i} - y_{r,i}) u^{y_r} \qquad \text{with} \qquad u^{y_r} = \prod_{i=1}^{N} u_i^{y_{r,i}}.$$

The aim of this work is to study the convergence to equilibrium for solutions to (2) with the *complex balance condition*. A spatial homogeneous state $u_\infty \in [0,\infty)^N$ is called a complex balanced equilibrium for (2) if the total-inflow and total-outflow are balanced at any complex, that is for any $y \in \{y_r, y'_r : r = 1, \dots, R\}$ the following holds

$$\sum_{\{r:\, y_r = y\}} k_r u_\infty^{y_r} = \sum_{\{r:\, y'_r = y\}} k_r u_\infty^{y_r}.$$

System (2) is said to satisfy the complex balance condition, and consequently is called a complex balanced system, if it has a complex balanced equilibrium. Moreover, it is well-known that for each initial data $u_0 = (u_{1,0}, \dots, u_{N,0})$ there exists *a unique strictly positive equilibrium* $u_\infty \in (0,+\infty)^N$ and possibly additional *boundary equilibria* $u^* \in \partial[0,+\infty)^N$.

It is remarked that the global existence of (classical, strong, weak) solutions to (2) is widely open in general since the nonlinearities can have arbitrary polynomial

growth (in terms of concentrations). Recently, it was proved in [5] that under certain physical assumptions, system (2) possesses a global renormalised solution. Our aim is to prove that under the complex balance condition and (2) has no boundary equilibrium, then any renormalised solution converges exponentially to the complex balanced equilibrium. Moreover, the proof is explicit up to a finite dimensional inequality.

The main tool used is the so-called *entropy method*, which was initiated in kinetic theory, and has recently proved very useful in investigating reaction-diffusion systems. If (2) is complex balanced, and thus possesses a unique strictly positive complex balanced equilibrium u_∞, then the relative entropy functional defined by

$$E(u|u_\infty) = \sum_{i=1}^{N} \int_\Omega \left(u_i \log \frac{u_i}{u_{i,\infty}} - u_i + u_{i,\infty} \right) dx$$

is decreasing along any trajectory of (2), with the entropy-dissipation functional

$$0 \le D(u) := -\frac{d}{dt} E(u|u_\infty) = \sum_{i=1}^{N} \int_\Omega d_i \frac{|\nabla u_i|^2}{u_i} dx + \sum_{r=1}^{R} k_r u_\infty^{y_r} \int_\Omega \Psi \left(\frac{u^{y_r}}{u_\infty^{y_r}}; \frac{u^{y'_r}}{u_\infty^{y'_r}} \right) dx$$

in which $\Psi(x;y) = x\log(x/y) - x + y \ge 0$. The main goal is to establish an *entropy entropy-dissipation inequality* of the form

$$D(u) \ge \lambda E(u|u_\infty) \tag{3}$$

for some $\lambda > 0$. Once this is proved, then thanks to a Gronwall inequality one obtains first the decay to zero of relative entropy, and consequently the convergence of solutions to the equilibrium u_∞, thanks to a Csiszár-Kullback-Pinsker inequality. Note that this functional inequality only holds when one takes into account all the conservation laws of (2). Many works were been carried out to prove (3) in special cases, see e.g. [1, 2, 3, 4] and references therein. The mentioned works proved (3) either explicitly in special cases or implicitly in the general case. This work fills in the gap, i.e. (3) is proved in general case explicitly, up to an explicit finite dimensional inequality.

The main ideas to prove (3) are stated in the following steps:

step 1: By using the additivity of the relative entropy, we have

$$E(u|u_\infty) = E(u|\overline{u}) + E(\overline{u}|u_\infty)$$

where $\overline{u} = (\overline{u}_1, \dots, \overline{u}_N)$ with $\overline{u}_i = \frac{1}{|\Omega|}\int_\Omega u_i dx$. The Logarithmic Sobolev inequality easily implies, for some $\lambda_1 > 0$,

$$\frac{1}{2} D(u) \ge \lambda_1 E(u|\overline{u})$$

step 2: By elementary inequality and the bound of solution in L^1-norm, one obtains first

$$E(\overline{u}|u_\infty) \le K_1 \sum_{i=1}^{N} \left[\sqrt{\frac{\overline{u}_i}{u_{i,\infty}}} - 1 \right]^2 .$$

On the other hand, by using a domain composition technique the entropy-dissipation is estimated below as

$$\frac{1}{2}D(u) \geq K_2 \sum_{r=1}^{R} \left[\sqrt{\frac{\overline{u}}{u_\infty}}^{y_r} - \sqrt{\frac{\overline{u}}{u_\infty}}^{y'_r} \right]^2 .$$

Note that all estimates in **step 1** and **step 2** are explicit.

step 3: It remains in this step to prove the finite dimensional inequality

$$\sum_{r=1}^{R} \left[\sqrt{\frac{\overline{u}}{u_\infty}}^{y_r} - \sqrt{\frac{\overline{u}}{u_\infty}}^{y'_r} \right]^2 \geq K_3 \sum_{i=1}^{N} \left[\sqrt{\frac{\overline{u}_i}{u_{i,\infty}}} - 1 \right]^2$$

under the conservation laws of (2) and the assumption that (2) *does not have boundary equilibria*. This inequality is then proved in general by a contradiction argument, and hence implicit. However, for concrete systems, for instance

$$\alpha_1 S_1 + \ldots + \alpha_N S_N \leftrightharpoons \beta_1 S_1 + \ldots + \beta_N S_N$$

or

$$\alpha_1 S_1 \to \alpha_2 S_2 \to \ldots \to \alpha_N S_N \to \alpha_1 S_1,$$

the finite dimensional inequality can be *explicitly* proved.

In the case when (2) has (possibly many) boundary equilibria, the entropy entropy-dissipation inequality (3) does not hold in general. In such a case, we look for a weaker version, that is, along any trajectory $u(t)$ of (2) it holds

$$D(u(t)) \geq \lambda(t) E(u(t)|u_\infty) \qquad \text{for all} \qquad t > 0,$$

where $\lambda(t)$ is *a function of time* such that $\int_0^{+\infty} \lambda(\tau) d\tau = +\infty$. This generalised entropy entropy-dissipation estimate is then verified in a special case

$$S_1 \to \alpha S_2 + S_3 \to (\alpha + 1) S_2 \to S_1$$

in which $\alpha \geq 1$ is a constant. The convergence to equilibrium for general systems having boundary equilibria remains as an important open question.

References

[1] Laurent Desvillettes and Klemens Fellner. *Exponential decay toward equilibrium via entropy methods for reaction–diffusion equations.* Journal of mathematical analysis and applications **319**.1 (2006): 157-176.

[2] Laurent Desvillettes and Klemens Fellner. *Entropy methods for reaction-diffusion equations: slowly growing a-priori bounds.* Revista Matemtica Iberoamericana **24**.2 (2008): 407-431.

[3] Laurent Desvillettes, Klemens Fellner and Bao Q. Tang, *Trend to equilibrium for reaction-diffusion systems arising from complex balanced chemical reaction networks*, to appear in SIAM Journal on Mathematical Analysis.

[4] Alexander Mielke, Jan Haskovec and Peter A. Markowich. *On uniform decay of the entropy for reaction–diffusion systems.* Journal of Dynamics and Differential Equations **27**.3-4 (2015): 897-928.

[5] Julian Fischer, *Global existence of renormalized solutions to entropy-dissipating reaction–diffusion systems*, Archive for Rational Mechanics and Analysis **218**.1 (2015), 553-587.

Discrete minimisers are close to continuum minimisers for the interaction energy

Francesco Saverio Patacchini
(joint work with José Alfredo Cañizo)

Consider $N \in \mathbb{N}$ particles $x_1, \dots, x_N$ in $\mathbb{R}^d$ interacting via a potential $W\colon \mathbb{R}^d \to (-\infty, +\infty]$. If the particles have equal masses $1/N$ then, if one writes $\boldsymbol{X} = (x_1, \dots, x_N)$, their total interaction energy is given by

$$E_N(\boldsymbol{X}) := \frac{1}{2N^2} \sum_{i=1}^{N} \sum_{\substack{j=1 \\ j \neq i}}^{N} W(x_i - x_j).$$

We call E_N the *discrete interaction energy* of the configuration $\boldsymbol{X}$. A natural question regards the existence and shape of *global* minimisers of this interaction energy among all possible particle configurations as $N \to \infty$; we refer to them as *discrete minimisers*. We show that, for large N, these minimisers are closely related to those of the *continuum interaction energy* defined by

$$E(\rho) = \frac{1}{2} \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} W(x - y)\, d\rho(x)\, d\rho(y)$$

for any $\rho \in \mathcal{P}(\mathbb{R}^d)$, where $\mathcal{P}(\mathbb{R}^d)$ is the set of Borel probability measures on $\mathbb{R}^d$. The existence of *continuum minimisers*, i.e., global minimisers of E, have been shown to be almost equivalent to the *instability* of the potential W [2, 6]: we say that W is *unstable* if there exists $\rho \in \mathcal{P}(\mathbb{R}^d)$ such that $E(\rho) < \frac{1}{2} \lim_{|x| \to \infty} W(x)$, whenever this limit exists. This concept of stability (or instability) is very close to the classical concept of H-stability as used in statistical mechanics [10]. In [2, 6] it was proved that, under some technical assumptions, continuum minimisers exist if and only if there exists a probability measure ρ with $E(\rho) \leq \frac{1}{2} \lim_{\infty} W$; that is, if and only if W is unstable or there is ρ with $E(\rho) = \frac{1}{2} \lim_{\infty} W$. Quite naturally this notion of instability plays a central role in the study of discrete minimisers and, leaving out for now the technical assumptions on W, our main result is the following [3]: for every N there exists a discrete minimiser $\boldsymbol{X}_N$ and

(1) if W is unstable, then the diameter of $\boldsymbol{X}_N$ is uniformly bounded for all N and $(\boldsymbol{X}_N)_N$, that is, its empirical measure, has a subsequence which converges in the weak sense, up to translations, to a minimiser of the continuum energy E as $N \to \infty$;
(2) if W is strictly stable, then the diameter of $\boldsymbol{X}_N$ tends to ∞ as $N \to \infty$.

The main assumptions on W in this result are that it is locally integrable, and that it is either continuous and bounded or has a singularity behaving like $|x|^{2-\beta}$ close to the origin for some $2 < \beta < d$ (i.e., W is β-repulsive for $2 < \beta < d$). The *power-law* potentials

$$W(x) = \frac{|x|^a}{a} - \frac{|x|^b}{b}, \quad \text{with} \quad \begin{cases} 0 < b < a & \text{when } d \in \{1, 2\}, \\ 2 - d < b < a, \quad b \neq 0, & \text{when } d \geq 3, \end{cases}$$

and the *Morse* potentials

$$W(x) = C_{\mathrm{r}} e^{-|x|/\ell_{\mathrm{r}}} - C_{\mathrm{a}} e^{-|x|/\ell_{\mathrm{a}}}, \quad \text{with } \ell_{\mathrm{r}} < \ell_{\mathrm{a}} \text{ and } C_{\mathrm{r}}/C_{\mathrm{a}} < (\ell_{\mathrm{a}}/\ell_r)^d,$$

are all included in our main result and are unstable.

Understanding the shape of discrete minimisers when the number of particles is very large is of great interest in statistical mechanics. For physically relevant potentials such as the Lennard-Jones potential $W(x) = |x|^{-12} - |x|^{-6}$ the conjectured behaviour is that *crystallisation* takes place as $N \to \infty$. Even if showing a crystallisation property is remarkably hard, one can make a weaker observation: for certain potentials, including Lennard-Jones, the diameter of ground states seems to increase without bound as $N \to \infty$, while for others the diameter seems to tend to a fixed value. This is part of the content of our result, whose main restriction in this setting is that it requires the potential W to be less singular than $|x|^{2-d}$ at $x = 0$. When the singularity is stronger, between $|x|^{-d}$ and $|x|^{2-d}$, we expect our main result still to be true, although we are unable to show it. Hence our statement does not say anything about the Lennard-Jones case, but does show that minimisers grow in diameter without bound for a range of stable potentials with a possible singularity at $x = 0$.

A more recent motivation for our result comes from the field of collective behaviour, where shapes of self-organised structures in some individual-based models exhibit very interesting phenomena and are closely related to those of discrete minimisers. In this context, models aim at capturing the behaviour of a large number of individuals, with applications to fish, cattle, birds, ants, and crowds of people. In very simplified models, interaction through a potential reflects a tendency in individuals to avoid close contact while staying close to the group. In this field there is an interest in the shape of minimisers for potentials which are very different from those found in physics, including potentials with a mild or no singularity at the origin or which tend to infinity at large distances. The paper [4] is the first example we know of where the link was made between the stability properties of the potential and the size of stationary states. It was observed that their size increases with N for stable potentials while it does not for unstable ones. This is precisely the behaviour which our result justifies rigorously.

The general strategy to prove our main result in the singular case (the continuous and bounded case being much easier) is to draw a parallel discrete version of several results which have been recently obtained for continuum minimisers. A first one is the *regularity* of continuum minimisers. If the potential W is β-repulsive with $2 < \beta < d$, then one obtains that any minimiser ρ is in the *Morrey space* of measures which satisfy

$$\rho(B_r) \leq C r^{\beta} \quad \text{for any ball } B_r \text{ of radius } r > 0$$

for some $C > 0$ independent of the ball B_r. This implies that any component of the support of ρ which has Hausdorff dimension n satisfies $n \geq \beta$, as already noted in [1]. An analogue of this regularity is needed for discrete minimisers: we prove

that there exists a constant $C > 0$ depending only on W such that any minimiser $\boldsymbol{X} = (x_1, \dots, x_N)$ satisfies

$$\mu_{\boldsymbol{X}}\big(B_r(x_i)\big) \leq Cr^\beta + \tfrac{1}{N} \quad \text{for any } r > 0 \text{ and } i \in \{1, \dots, N\},$$

where $\mu_{\boldsymbol{X}}$ is the empirical measure associated to $\boldsymbol{X}$. This motivates an interesting definition of *empirical* Morrey measures which serves as a discrete version of the Morrey spaces. Another important property of continuum minimisers is that they satisfy the following *Euler–Lagrange* condition: if ρ is a continuum minimiser then

$$W * \rho(x) = 2E(\rho) \quad \text{for } \rho\text{-almost every } x \in \mathbb{R}^d.$$

The quantity corresponding to $W * \rho(x)$ in the discrete case, for a particle distribution $\boldsymbol{X} = (x_1, \dots, x_N)$, is $P_i(\boldsymbol{X}) := \frac{1}{N} \sum_{j=1, j\neq i}^N W(x_i - x_j)$ for all $i \in \{1, \dots, N\}$. Interestingly, for a discrete minimiser this does *not* seem to be constant at all sites i, but we show a bound on its variation across sites which decays asymptotically as $N \to \infty$: there exist $A > 0$ and $0 < k \leq 1$ such that any minimiser $\boldsymbol{X}$ satisfies

$$|P_i(\boldsymbol{X}) - P_j(\boldsymbol{X})| \leq AN^{-k} \quad \text{for all } i, j \in \{1, \dots, N\}. \tag{1}$$

Again another property is that continuum minimisers are known to be *compactly supported* if W is increasing at long range [2]. Analogously, we can give a uniform bound on the diameter of discrete minimisers. Finally, our proof of convergence of minimisers contains the fact that the discrete energy Γ-*converges* to the continuum energy in the weak topology.

References

[1] D. Balagué, J. A. Carrillo, T. Laurent and G. Raoul, *Dimensionality of local minimizers of the interaction energy*, Arch. Ration. Mech. Anal. **209** (2013), 1055–1088.

[2] J. A. Cañizo, J. A. Carrillo and F. S. Patacchini, *Existence of compactly supported global minimisers for the interaction energy*, Arch. Ration. Mech. Anal. **217** (2015), 1197–1217.

[3] J. A. Cañizo and F. S. Patacchini, *Discrete minimisers are close to continuum minimisers for the interaction energy*, ArXiv preprint 1612.09233 (2017).

[4] M. R. D'Orsogna, Y.-L. Chuang, A. L. Bertozzi and L. S. Chayes, *Self-propelled particles with soft-core interactions: patterns, stability, and collapse*, Phys. Rev. Lett. **96** (2006), 104302.

[5] D. Ruelle, *Statistical Mechanics: Rigorous Results*, New York: W.A. Benjamin (1969).

[6] R. Simione, D. Slepčev and I. Topaloglu, *Existence of ground states of nonlocal-interaction energies*, J. Stat. Phys. **159** (2015), 972–986.

Computing certain invariants of topological spaces of dimension three

Gabriel Peyré

(joint work with J-D. Benamou, L. Nenna, G. Carlier, L. Chizat, M. Cuturi, J. Solomon, F-X. Vialard)

Optimal transport (OT) has become a fundamental mathematical tool at the interface between calculus of variations, partial differential equations and probability. It took however much more time for this notion to become mainstream in numerical applications. This situation is in large part due to the high computational

cost of the underlying optimization problems. There is however a recent wave of activity on the use of OT-related methods in fields as diverse as computer vision, computer graphics, statistical inference, machine learning and image processing. In this talk, I will review an emerging class of numerical approaches for the approximate resolution of OT-based optimization problems. These methods make use of an entropic regularization of the functionals to be minimized, in order to unleash the power of optimization algorithms based on Bregman-divergences geometry (see [2] for a theoretical analysis and a litterature review). This results in fast, simple and highly parallelizable algorithms, in sharp contrast with traditional solvers based on the geometry of linear programming. For instance, they allow for the first time to compute barycenters (according to OT distances) of probability distributions discretized on computational 2-D and 3-D grids with millions of points [1]. This offers a new perspective for the application of OT in machine learning (to perform clustering or classification of bag-of-features data representations) and imaging sciences (to perform color transfer or shape and texture morphing [6]). These algorithms also enable the computation of gradient flows for the OT metric, and can thus for instance be applied to simulate crowd motions with congestion constraints [4]. We will also discus various extensions of classical OT, such as handling unbalanced transportation between arbitrary positive measures [3] (the so-called Hellinger-Kantorovich/Wasserstein-Fisher-Rao problem), and the computation of OT between different metric spaces (the so-called Gromov-Wasserstein problem) [7, 5].

References

[1] J-D. Benamou, G. Carlier, M. Cuturi, L. Nenna, G. Peyré. *Iterative Bregman Projections for Regularized Transportation Problems*. SIAM Journal on Scientific Computing, **37**(2), pp. A1111-A1138, 2015.

[2] G. Carlier, V. Duval, G. Peyré, B. Schmitzer. *Convergence of Entropic Schemes for Optimal Transport and Gradient Flows*. to appear in SIAM Journal on Mathematical Analysis, 2017.

[3] L. Chizat, G. Peyré, B. Schmitzer, F-X. Vialard. *Scaling Algorithms for Unbalanced Transport Problems*. Preprint Arxiv:1607.05816, 2016.

[4] G. Peyré. *Entropic Approximation of Wasserstein Gradient Flows*. SIAM Journal on Imaging Sciences, **8**(4), pp. 2323-2351, 2015.

[5] G. Peyré, M. Cuturi, J. Solomon. *Gromov-Wasserstein Averaging of Kernel and Distance Matrices*. In Proc. ICML'16, pp. 2664-2672, 2016.

[6] J. Solomon, F. de Goes, G. Peyré, M. Cuturi, A. Butscher, A. Nguyen, T. Du, L. Guibas. *Convolutional Wasserstein Distances: Efficient Optimal Transportation on Geometric Domains*. ACM Transactions on Graphics (Proc. SIGGRAPH 2015), **34**(4), pp. 66:1-66:11, 2015.

[7] J. Solomon, G. Peyré, V. Kim, S. Sra. *Entropic Metric Alignment for Correspondence Problems*. ACM Transactions on Graphics (Proc. SIGGRAPH 2016), **35**(4), pp. 72:1-72:13, 2016.

Dynamic optimal transport and applications in imaging

Carola-Bibiane Schönlieb
(joint work with Yoeri Boink, Christoph Brune, Martin Burger, Hendrik Dirks, Jocelyn Etienne, Lukas Lang, Sebastian Neumayer, Ozan Öktem)

The problem of optimal transport and its dynamic formulation due to Benamou-Brenier is classical by now and since the seminar paper [1] has resulted in the spread of optimal transport into various applications. In this talk I focused on two applications of optimal transport in imaging: (1) generalized optimal transport for the interpolation of images [3] and for indirect image matching [this is work in progress with S. Neumayer and O. Öktem]; (2) optimal transport for motion estimation in videos [this includes joint work with M. Burger and H. Dirks [2]; work in progress with C. Brune and Y. Boink on joint optimal transport and segmentation; and work in progress with Jocelyn Etienne and L. Lang on convective optimal transport for simulating dynamics of cell boundaries] .

References

[1] J. D. Benamou and Y. Brenier, Y. A computational fluid mechanics solution to the Monge-Kantorovich mass transfer problem. Numerische Mathematik, 84(3), 375-393, 2000.
[2] M. Burger, H. Dirks and C.-B. Schönlieb, A Variational Model for Joint Motion Estimation and Image Reconstruction, submitted, arXiv:1607.03255, 2016.
[3] J. Maas, M. Rumpf, C.-B. Schönlieb, and S. Simon, A generalized model for optimal transport of images including dissipation and density modulation, ESAIM: Mathematical Modelling and Numerical Analysis, 49(6), 1745-1769, 2015.

Discrete Geodesic Paths in the Space of Images

Martin Rumpf
(joint work with Benjamin Berkels, Alexander Effland, Florian Schfer)

In this extended abstract the space of images is considered as a Riemannian manifold using the metamorphosis approach [2, 4, 5], where the underlying Riemannian metric simultaneously measures the cost of image transport and intensity variation. A robust and effective variational time discretization of geodesics paths is proposed and a variational scheme for a time discrete exponential map is investigated as well. The approach requires the definition of a discrete path energy consisting of a sum of consecutive image matching functionals over a set of image intensity maps and pairwise matching deformations. For square-integrable input images the existence of discrete and interpolating geodesic paths defined as minimizers of this variational problem is shown.

Metamorphosis. The metamorphosis model [2, 5] generalizes the flow of diffeomorphism approach allowing for intensity variations along transport paths and associates a corresponding cost functional with these variations. In [4], Trouvé and Younes rigorously analyzed the local geometry of the resulting Riemannian manifold and proved the existence of geodesic curves for square-integrable images

and the (local) existence as well as the uniqueness of solutions of the initial value problem for the geodesic equation.

Let us suppose that the image domain $D \subset \mathbb{R}^n$ for $n \in \{2,3\}$ has Lipschitz boundary. For a flow of diffeomorphisms $\phi(t) : \bar{D} \to \mathbb{R}^n$ for $t \in [0,1]$ driven by the *Eulerian velocity* $v(t) = \dot{\phi}(t) \circ \phi^{-1}(t)$ we take into account a quadratic form L subject to certain growth and consistency conditions, which can be considered as a Riemannian metric on the space of diffeomorphisms and thus on the space of diffeomorphic transformations $u(t) = u_0 \circ \phi^{-1}(t)$ of a given reference image u_0. Now, the metamorphosis approach allows for image intensity variations along motion paths and penalizes the integral over the squared material derivative. In fact, the *path energy* in the metamorphosis model for an image curve $u \in L^2((0,1), L^2(D))$ and $\delta > 0$ is defined as

$$\mathcal{E}[u] := \int_0^1 \inf_{(v,z)} \int_D L[v,v] + \frac{1}{\delta} z^2 \,\mathrm{d}x \,\mathrm{d}t\,, \tag{1}$$

where the infimum is taken over all pairs (v,z) which fulfill the transport equation $\frac{D}{\partial t} u = \dot{u} + v \cdot \nabla u = z$. Here, we consider $L[v,v] := Dv : Dv + \gamma \Delta^m v \cdot \Delta^m v$ with $\gamma > 0$ and $2m > 1 + \frac{n}{2}$. To formulate this rigorously, one has to take into account the weak material derivative $z \in L^2((0,1), L^2(D))$ defined via the equation

$$\int_0^1 \int_D \eta z \,\mathrm{d}x \,\mathrm{d}t = -\int_0^1 \int_D (\partial_t \eta + \operatorname{div}(v\eta)) u \,\mathrm{d}x \,\mathrm{d}t$$

for all $\eta \in C_c^\infty((0,1) \times D)$. Geodesic curves are defined as minimizers of the path energy (1).

Time discrete pathenergy and discrete geodesics. In what follows we consider the time discretization of the path energy (1) proposed in [1]. To this end, we define for arbitrary images $u, \tilde{u} \in L^2(D)$ the *discrete matching energy*

$$\mathcal{W}[u,\tilde{u}] := \min_{\phi \in \mathcal{A}} \left\{ \mathcal{W}^D[u,\tilde{u},\phi] := \int_D |D\phi - \mathrm{Id}|^2 + \gamma |\Delta^m \phi|^2 + \frac{1}{\delta}(\tilde{u} \circ \phi - u)^2 \,\mathrm{d}x \right\}, \tag{2}$$

which is composed of a rescaled thin plate regularization term (first two terms) and a quadratic $L^2(D)$-mismatch measure (cf [1, (6.2)]). The *set of admissible deformations* $\mathcal{A}$ is defined as $\mathcal{A} := \left\{ \phi \in H^{2m}(D,D) : \phi - \mathrm{Id} \in H_0^{2m}(D,D) \right\}$. Now, we consider discrete curves $\mathbf{u} = (u_0, \ldots, u_K) \in (L^2(D))^{K+1}$ in image space and define a *discrete path energy* as the sum of pairwise matching functionals $\mathbf{W}$ evaluated on consecutive images of these discrete curves as follows

$$\mathcal{E}_K[\mathbf{u}] := K \sum_{k=1}^{K} \mathbf{W}[u_{k-1}, u_k]\,. \tag{3}$$

We refer to [3] for the introduction of such a variational time discretization on shape manifolds.

Let $u_A, u_B \in L^2(D)$ and $K \geq 1$. A *discrete geodesic* connecting u_A and u_B is a discrete curve in image space that minimizes $\mathcal{E}_K$ over all discrete curves $\mathbf{u} = (u_0, \ldots, u_K) \in (L^2(D))^{K+1}$ with $u_0 = u_A$ and $u_K = u_B$. For this discretization

Mosco-convergence of the underlying discrete path energy to the continuous path energy can be proved. This includes a diffeomorphism property for the induced transport and the existence of a square-integrable weak material derivative in space and time. For details we refer [1].

The time discrete exponential map. Let us briefly recall the definition of the continuous exponential map on a Riemannian manifold. Let $u : [0,1] \to L^2(D)$ be the unique geodesic curve for a prescribed initial position $u(0) = u_A$ and an initial velocity $\dot{u}(0) = v$ on a Riemannian manifold. The exponential map is then defined as $\mathrm{Exp}_{u_A}(v) = u(1)$. Furthermore, one easily checks that $\mathrm{Exp}_{u_A}(\frac{k}{K}v) = u(\frac{k}{K})$ for $0 \leq k \leq K$. Now, we ask for a time discrete counterpart of the exponential map in the metamorphosis model. To this end, we consider an image u_0 as the initial data and a second image u_1 such that $\zeta_1 = u_1 - u_0$ represents a small variation of the image u_0. For varying values of $K \geq 2$ we now ask for a discrete geodesic $(u_0, u_1, u_2, \ldots, u_K)$ described as the minimizer of the discrete path energy (3). Next, we define $\mathrm{Exp}^k_*(\,\cdot\,)$ as the discrete counterpart of $\mathrm{Exp}_*(\frac{k}{K}\,\cdot\,)$, i.e. we set

$$\mathrm{Exp}^k_{u_0}(\zeta_1) := u_k$$

for $k = 1, \ldots, K$. Taking into account $k = 2$ we immediately observe that the sequence of discrete exponential maps $(\mathrm{Exp}^k_{u_0}(\zeta_1))_{k=1,\ldots}$ can iteratively be defined as follows

$$\mathrm{Exp}^k_{u_0}(\zeta_1) = u_k := \mathrm{Exp}^2_{u_{k-2}}(\zeta_{k-1}) \tag{4}$$

for $k \geq 2$, where $\zeta_{k-1} = u_{k-1} - u_{k-2}$, and for the sake of completeness we define $\mathrm{Exp}^0_{u_0}(\zeta_1) = u_0$ and $\mathrm{Exp}^1_{u_0}(\zeta_1) = u_1 = u_0 + \zeta_1$. Thus, it essentially remains to compute Exp^2 for a given input image u_{k-2} and an image variation $\zeta_{k-1} = u_{k-1} - u_{k-2}$. For a detailed discussion of the discrete exponential map in the simpler model of Hilbert manifolds we refer to [3]. The particular challenge here is that the matching energy $\mathcal{W}$ cannot be evaluated directly, but requires to solve the variational problem (2) for the matching deformation.

There are two major restrictions regarding the input images u_0 and u_1: Firstly, the existence and uniqueness result for the discrete exponential map will require weakly differentiable input images. Secondly, the initial variation $\zeta_1 = u_1 - u_0$ is assumed to be sufficiently small in $L^2(D)$ in order to ensure the existence of the initial deformation ϕ_1 and guarantee the convergence of a suitable fixed point algorithm. We will also see that for fixed K the variations $u_{k-1} - u_{k-2}$ for $k \leq K$ will remain small provided that ζ_1 is small. Thus, for fixed K the discrete exponential map $\mathrm{Exp}^k_{u_0}(\,\cdot\,)$ will be well-posed for sufficiently small initial variation ζ_1.

For a given weakly differentiable initial image and an initial image variation, the exponential map allows to compute a discrete geodesic extrapolation path in the space of images. One can show that a time step of this shooting method can be formulated in the associated deformations only. For sufficiently small time steps local existence and uniqueness can be established.

Acknowledgements. A. Effland and M. Rumpf acknowledge support of the Hausdorff Center for Mathematics and the Collaborative Research Center 1060 funded by the German Research Foundation.

References

[1] B. Berkels, A. Effland, and M. Rumpf. Time discrete geodesic paths in the space of images. *SIAM J. Imaging Sci.*, 8 (3):1457–1488, 2015.
[2] M. I. Miller and L. Younes. Group actions, homeomorphisms, and matching: a general framework. *Int. J. Comput. Vis.*, 41(1–2):61–84, 2001.
[3] M. Rumpf and B. Wirth. Variational time discretization of geodesic calculus. *IMA J. Numer. Anal.*, 35(3):1011–1046, 2015.
[4] A. Trouvé and L. Younes. Local geometry of deformable templates. *SIAM J. Math. Anal.*, 37(1):17–59, 2005.
[5] A. Trouvé and L. Younes. Metamorphoses through Lie group action. *Found. Comput. Math.*, 5(2):173–198, 2005.

Generalised finite difference methods for Monge-Ampère equations

Brittany D. Froese

We consider the development and analysis of numerical methods for equations of Monge-Ampère type,

$$\begin{cases} \det(A(x,u(x),\nabla u(x)) + D^2u(x)) = G(x,u(x),\nabla u(x)) \\ A(x,\nabla u(x)) + D^2u(x) \geq 0. \end{cases} \tag{1}$$

1. Generalised finite difference schemes

First, we rely on Hadamard's inequality to find an alternative expression for the determinant of a symmetric, positive-definite matrix.

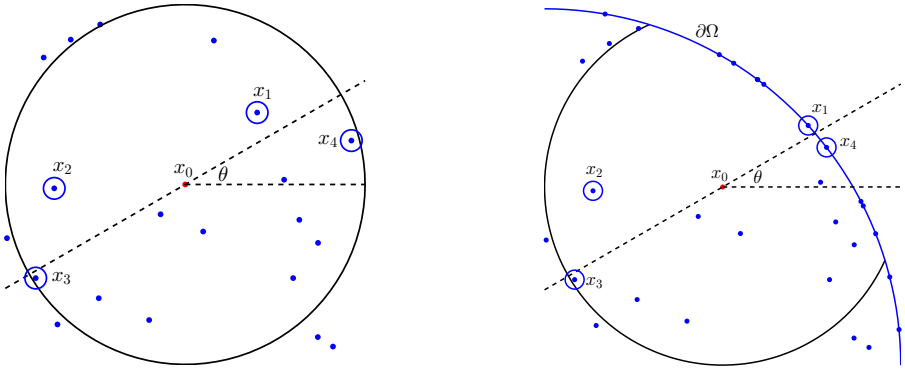
$$\det{}^+(M) = \min_{(\nu_1,\dots,\nu_d)} \left\{ \prod_{j=1}^{d} \max\{\nu_j^T M \nu_j, 0\} + \min\{\nu_j^T M \nu_j, 0\} \right\} \tag{2}$$

over orthonormal $\nu_1, \dots, \nu_d$. Substituting $M = D^2u$, we obtain a globally elliptic extension of the Monge-Ampère equation in terms of the second directional derivatives of u in the directions ν_j [FO11]. Consistent, montone schemes for the second directional derivatives $u_{\nu\nu}$ can then be combined to create an appropriate scheme for general Monge-Ampère type equations.

Next we consider a discrete set of discretisation points $\mathcal{G}^h$, where h is the spatial resolution of the point cloud. To approximate the PDE at a point $x_0 \in \mathcal{G}^h$, we consider all discretisation points living within a distance $\sqrt{h}$ of x_0. We select four points x_1, x_2, x_3, x_4 in this search ball that best align with the direction ν, each living in a different quadrant (Figure 1). We then approximate derivatives by

$$u_{\nu\nu}(x_0) = \sum_{j=1}^{4} a_j(u(x_j) - u(x_0)) + \mathcal{O}(\sqrt{h}), \quad a_j \geq 0, \tag{3}$$

wher



(a) (b)

FIGURE 1. A finite difference stencil chosen from a point cloud (a) in the interior and (b) near the boundary, where higher resolution is needed.

2. Convergence

Construction of consistent, monotone approximations of Monge-Ampère operators is not sufficient to guarantee convergence to the viscosity solution of the PDE. The difficulty here is that the Barles-Souganidis convergence result [BS91] only applies to PDEs that satisfy a strong form of the comparison principle, which may not be valid for the PDE in question.

As an example, we consider the Dirichlet problem for the equation of prescribed Gaussian curvature.

$$\begin{cases} \det{}^+(D^2u(x)) = \kappa(x)(1+|\nabla u(x)|^2)^{(d+2)/2}, & x \in \Omega \\ u(x) = g(x), & x \in \partial\Omega. \end{cases} \tag{4}$$

Even in the simple one-dimensional setting with constant curvature $\kappa(x) = 1$ and boundary data $u(0) = -1$, $u(1) = 1$, no continuous solution exists. Consequently, the boundary conditions need to be interpreted in a weak sense, the standard comparison principle fails (i.e. sub-solutions can lie above super-solutions), and the Barles-Souganidis convergence proof fails.

We prove an interior comparison principle for this problem.

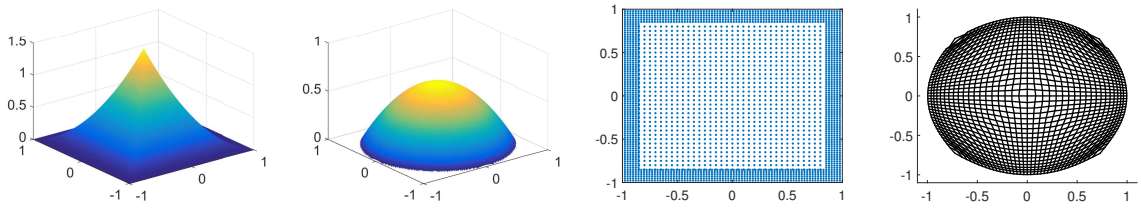
Theorem 1 (Interior Comparison Principle)**.** *Let u be a sub-solution and v a super-solution of* (4)*. Then $u(x) \leq v(x)$ at interior points $x \in \Omega$.*

This result allows us to modify the Barles-Souganidis proof in order to obtain convergence in the interior of the domain [Fro16].

Establishing a similar convergence result for more general Monge-Ampère equations and boundary conditions (e.g. the second boundary value problem) remains an important open problem.

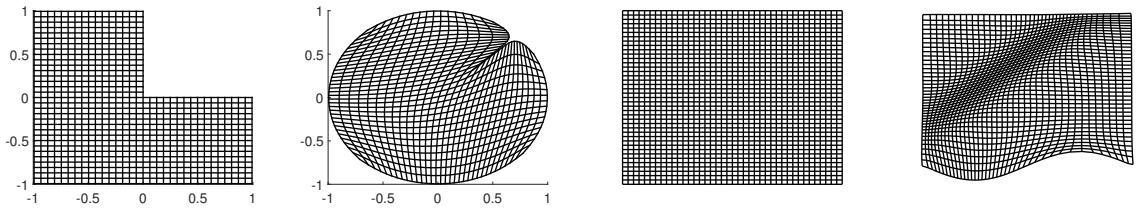
3. Computational Examples

We conclude by illustrating several optimal transportation maps computed with this method. See Figures 2-4.



(a) (b) (c) (d)

FIGURE 2. (a) Source and (b) target densities that vanish on the boundary. (c) Non-uniform mesh that resolves boundary layers and (d) computed quadratic optimal transport map.



(a) (b) (c) (d)

FIGURE 3. Computed quadratic optimal transport maps from (a) non-convex to (b) convex or (c) convex to (d) non-convex.

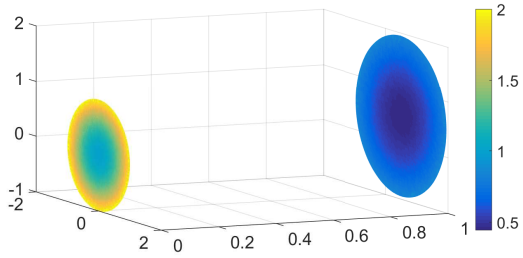


FIGURE 4. Computed L^1 optimal transport map between measures supported on parallel planes.

References

[BS91] G. Barles and P. E. Souganidis. Convergence of approximation schemes for fully nonlinear second order equations. *Asymptotic Anal.*, 4(3):271–283, 1991.

[FO11] B. D. Froese and A. M. Oberman. Convergent finite difference solvers for viscosity solutions of the elliptic Monge-Ampère equation in dimensions two and higher. *SIAM J. Numer. Anal.*, 49(4):1692–1714, 2011.

[Fro15] B. D. Froese. Meshfree finite difference approximations for functions of the eigenvalues of the Hessian. 2015. Submitted.

[Fro16] B. D. Froese. Convergent approximation of non-continuous surfaces of prescribed Gaussian curvature. 2016. Submitted.

A variational approach for multiphase porous media flows

Clément Cancès

(joint work with Thomas O. Gallouët, Léonard Monsaingeon)

Since Otto's seminal paper [4], it has been understood that the equation governing the motion of a gaz within a porous medium can be interpreted as the gradient flow of the osmotic energy in the metric space of probability measures endowed with some Wasserstein distance. The goal of our work [2] is to extend this beautiful interpretation to more complex models used to model flows in the subsurface. In order to ease the presentation, we make some simplifying assumptions, assuming for instance the homogeneity and the isotropy of the porous medium Ω. We refer to [1, 2] for a more general presentation.

In what follows, Ω denotes an open convex subset of $\mathbb{R}^d$. We consider the flow of $N+1$ incompressible and immiscible phases in Ω. For $i \in \{0, \dots, N\}$, we denote by $s_i(\boldsymbol{x}, t) \in [0, 1]$ the proportion of the phase i in the fluid at time $t \geq 0$ and position $\boldsymbol{x} \in \Omega$, and by $\boldsymbol{s} = (s_0, \dots, s_N)$. Since the fluid if fully composed by the $N+1$ phases, we get the constraint

$$\sum_{i=0}^{N} s_i = 1 \qquad \text{a.e. in } \Omega \times \mathbb{R}_+. \tag{1}$$

In what follows, we denote by $\boldsymbol{\mathcal{X}} = \left\{ \boldsymbol{s} \in L^1(\Omega; \mathbb{R}_+^{N+1}) \;\middle|\; (1) \text{ holds a.e.} \right\}$. On the other hand, each phase is convected by its own speed $\boldsymbol{v}_i$, i.e.,

$$\partial_t s_i + \nabla \cdot s_i \boldsymbol{v}_i = 0 \quad \text{in } \Omega \times (0, \infty). \tag{2}$$

The phase velocities $\boldsymbol{v}_i$ are prescribed by Darcy's law

$$\boldsymbol{v}_i = -\nabla(p_i + \Psi_i), \qquad i \in \{0, \dots, N\}, \tag{3}$$

p_i being the (unknown) pressure of phase i, and Ψ_i its gravitational potential. The phase pressures are linked to the fluid composition through N capillary pressure relations

$$p_i - p_0 = \pi_i(\boldsymbol{s}), \qquad i \in \{1, \dots, N\}, \tag{4}$$

where the capillary functions π_i are supposed to derive from a smooth uniformly convex potential

$$\Pi : \left\{ (s_1, \dots, s_N) \in \mathbb{R}_+^N \;\middle|\; \sum_{i\geq 1} s_i \leq 1 \right\} \to \mathbb{R}$$

in the sense that $\pi_i = \partial_{s_i}\Pi$. No-flux boundary conditions are imposed, i.e.,

$$s_i \boldsymbol{v}_i \cdot \boldsymbol{n} = 0 \quad \text{on } \Omega \times (0,\infty), \tag{5}$$

as well as the initial saturation configuration

$$\boldsymbol{s}_{|t=0} = \boldsymbol{s}^0 \in \boldsymbol{\mathcal{X}}. \tag{6}$$

We endow the set

$$\boldsymbol{\mathcal{A}} = \left\{ \boldsymbol{s} \in L^1(\Omega; \mathbb{R}_+^{N+1}) \;\middle|\; \int_\Omega s_i \mathrm{d}\boldsymbol{x} = \int_\Omega s_i^0 \mathrm{d}\boldsymbol{x} \text{ for } i \in \{0, \dots, N\} \right\}$$

(where the solution to (1)–(6) lives) with the squared Wasserstein distance

$$\boldsymbol{W}^2(\boldsymbol{s}, \widehat{\boldsymbol{s}}) = \sum_{i=0}^{N} W^2(s_i, \widehat{s}_i), \qquad \forall \boldsymbol{s}, \widehat{\boldsymbol{s}} \in \boldsymbol{\mathcal{A}}.$$

where W denotes the usual Wasserstein distance with quadratic cost. The energy corresponding to saturation configurations fulfilling the constraint (1) is defined by

$$\mathcal{E}(\boldsymbol{s}) = \int_\Omega \left(\Pi(\boldsymbol{s}) + \sum_{i=0}^{N} s_i \Psi_i \right) \mathrm{d}\boldsymbol{x}, \qquad \forall \boldsymbol{x} \in \boldsymbol{\mathcal{X}}.$$

Then we interpret the system (1)–(6) as the generalized gradient flow of the singular energy $\mathcal{E}$ with respect to the Wasserstein metric $\boldsymbol{W}$. This claim was formally established thanks to heuristic arguments in our note [1]. In order to make it rigorous, we prove in [2] that the minimizing movement scheme *à la* Jordan-Kinderlehrer-Otto [3] converges. More precisely, for any given time step $\tau > 0$, we consider the sequence $(\boldsymbol{s^n})_{n\geq 0}$ defined by

$$\boldsymbol{s}^n = \operatorname*{Argmin}_{\boldsymbol{s} \in \boldsymbol{\mathcal{X}} \cap \boldsymbol{\mathcal{A}}} \left(\frac{\boldsymbol{W}^2(\boldsymbol{s}, \boldsymbol{s^{n-1}})}{2\tau} + \mathcal{E}(\boldsymbol{s}) \right)$$

and the corresponding piecewise constant function $\boldsymbol{s}^\tau : \mathbb{R}_+ \to \boldsymbol{\mathcal{X}} \cap \boldsymbol{\mathcal{A}}$ by

$$\boldsymbol{s}^\tau(0) = \boldsymbol{s}^0, \qquad \boldsymbol{s}^\tau(t) = \boldsymbol{s}^n \text{ if } t \in ((n-1)\tau, n\tau].$$

We have now introduced all the necessary material to state our main result:

There exist $\boldsymbol{s} \in C(\mathbb{R}_+; \boldsymbol{\mathcal{A}})$ *and* $\boldsymbol{p} = (p_i)_{0\leq i\leq N} \in L^2_{loc}(\mathbb{R}_+; H^1(\Omega))$ *fulfilling* (1)–(6) *such that, up to the extraction of an unlabeled subsequence,* $\boldsymbol{s}^\tau$ *converges almost everywhere and in* $C(\mathbb{R}_+, \boldsymbol{\mathcal{A}})$ *towards* $\boldsymbol{s}$.

The interests of this result are twofold: it provides both a new existence result and a variational interpretation for a well-established model for subsurface flows.

References

[1] C. Cancès, T. O. Gallouët, and L. Monsaingeon, *The gradient flow structure for incompressible immiscible two-phase flows in porous media*, C. R. Acad. Sci. Paris, Srie I, 353 (2015), 985–989.

[2] C. Cancès, T. O. Gallouët, and L. Monsaingeon, *Incompressible immiscible multiphase flows in porous media: a variational approach*, HAL preprint hal-01345438 (2016).

[3] R. Jordan, D. Kinderlehrer, and F. Otto, *The variational formulation of the Fokker-Planck equation*, SIAM J. Math. Anal., 29 (1998), 1–17.

[4] F. Otto, *The geometry of dissipative evolution equations: the porous medium equation*, Comm. Partial Differential Equations, 26 (2001), 101–174.

Multiphase Optimal Transport

Gershon Wolansky

The subject of multiphase transport deals with a transport of a given vector-valued measure into another, such that the *same* transport plan applies to all components. Possible applications are computer vision (transportation distance between colored images), microeconomics (exchange of multiple goods between agents), incompressible flows (when on of the components in each space is the Lebesgue measure), etc..

Given vector ($\mathbb{R}^k_+$) valued, atomless measures $(X,\vec\sigma)$, $(Y,\vec\eta)$: $\vec\sigma := (\sigma_1,\dots\sigma_k)$ are k positive measures on a sigma algebra $(X,\mathcal{A}_X)$. Likewise, $\vec\eta := (\eta_1,\dots\eta_k)$ are k positive measures on a sigma algebra $(Y,\mathcal{A}_Y)$.

A multiphase transport is defined by a measurable family $x\in X\to P_x(dy)\in \mathcal{A}_X(Y)$ such that

$$\int_X P_x(dy)\sigma_i(dx)=\eta_i(dy)\quad 1\le i\le k\ .$$

If such a transport exists then $(\vec\sigma,X)$ is said to dominate $(\vec\eta,Y)$: $(\vec\sigma,X)\succeq(\vec\eta,Y)$. An equivalent condition [1]:

$$\int_X F\left(\frac{d\vec\sigma}{d|\sigma|}\right)d|\sigma|\ge\int_Y F\left(\frac{d\vec\eta}{d|\eta|}\right)d|\eta|$$

for any convex $F:\mathbb{R}^k\to\mathbb{R}$. The necessary condition is $\vec\sigma(X)=\vec\eta(Y)$ (corresponds to linear F). In general, this condition is not sufficient (unless $k=1$).

Open Question: *If both* $(\vec\sigma,X)\succeq(\vec\eta,Y)$ *and* $(\vec\eta,Y)\succeq(\vec\sigma,Y)$, *then there exist* deterministic *transports* $T:X\to Y$, $S:Y\to X$ *such that* $T_\#\sigma_i=\eta_i$ *and* $S_\#\eta_i=\sigma_i$ *for* $i=1,\dots N$*?*

Let $S^N_{\vec\sigma}\subset\mathbb{R}^{N\times k}$ be the subset of all $k-$valued measures on a finite set Y of cardinality N which are dominated by $\vec\sigma$.

Theorem $S^N_{\vec\sigma}$ *is a compact and convex set in* $\mathbb{R}^{N\times k}$.

We define $(X,\vec\sigma)\succeq_N(Y,\vec\eta)$ iff $S^N_{\vec\eta}\subset S^N_{\vec\sigma}$.

Theorem $(X,\vec{\sigma}) \succeq (Y,\vec{\eta})$ *iff* $(X,\vec{\sigma}) \succeq_N (Y,\vec{\eta})$ *for any* $N \in \mathbb{N}$.

Let $\theta \in C(X \times Y)$. A multi-$(\vec{\sigma},\vec{\eta})$ transport $\{P_x(dy); x \in X\}$ is said to be a (maximizing) optimal iff

$$\int_X \int_Y P_x(dy)\theta(x,y)\sigma(dx) \geq \int_X \int_Y \bar{P}_x(dy)\theta(x,y)\sigma(dx)$$

for any multi-$(\vec{\sigma},\vec{\eta})$ transport $\{\bar{P}_x\}$. If $(X,\vec{\sigma}) \succeq (Y,\vec{\eta})$ are Borel and X is a compact space then there exists an optimal multi transport. In case Y is of finite cardinality it reduces to an *optimal partition.*

Let $\theta \in C(X,Y)$ be given by $\theta(x,y) = \inf_{z\in Z} \theta_1(x,z) + \theta_2(y,z)$, where Z is a separable space and θ_1 (resp. θ_2) is continuous on (X,Z) (resp. (Y,Z)). Assume

$$\sigma\left(x \in X; \vec{p}\cdot d\vec{\sigma}/d|\vec{\sigma}|(x) = \theta_1(x,z) - \theta_2(y,z)\right) = 0$$

for any $\vec{p} \neq 0 \in \mathbb{R}^k$ and any $z \in Z$. Let $\hat{Z} := \{z_1, z_2 \ldots\} \subset Z$ be an enumeration of a dense set in Z and $Z_N := \{z_1, \ldots z_N\}$, we define $\theta_N(x,y) := \max_{i\in Z_N} \theta_1(x,z_i) + \theta_2(y,z_i)$, and consider the optimal multi-phase transport with respect to θ_N. In the dual formulation we define

$$\Xi^N_{\vec{\sigma}}(\vec{P}) := \int_X [\max_{i\in N} \vec{p_i}\cdot d\vec{\sigma} d|\vec{\sigma}|(x) + \theta_1(x,z_i)]d|\sigma| \ ,$$

$$\Xi^N_{\vec{\eta}}(\vec{P}) := \int_X [\max_{i\in N} \vec{p_i}\cdot d\vec{\eta}/d|\vec{\eta}|(x) + \theta_2(x,z_i)]d|\eta| \ .$$

where $\vec{P} \in \mathbb{R}^{N\times k} := (\vec{p_1}, \ldots \vec{p}_N)$, $\vec{p_i} \in \mathbb{R}^k$.

Theorem *There exists a unique minimizer to* $\vec{P} \to \Xi^N_{\vec{\sigma}}(\vec{P}) + \Xi^N_{\vec{\eta}}(-\vec{P})$ *in* $\mathbb{R}^{N\times k}$. *If, in addition,* $(X,\vec{\sigma}) \succeq_N (Y,\vec{\eta})$ *then this minimizer induces an optimal* θ_N *transport form* $(X,\vec{\sigma})$ *to* $(Y,\vec{\eta})$*:*

$$\min_{\vec{P}\in\mathbb{R}^{N\times k}} \Xi^N_{\vec{\sigma}}(\vec{P}) + \Xi^N_{\vec{\eta}}(-\vec{P}) = \max_{\{A_i,B_i\}} \sum_{i=1}^N \int_{A_i} \theta_1(x,z_i)d|\sigma|(x) + \int_{B_i} \theta_2(y,z_i)d|\eta|(y)$$

where A_i, B_i *are mutually disjoint measurable partitions of* X *(resp.* Y*) such that* $\vec{\sigma}(A_i) = \vec{\eta}(B_i)$ *for* $i = 1, \ldots N$. *In particular,*

$$P^N_x(dy) := \sum_{i=1}^N 1_{A_i}(x)\,(1_{B_i}(y)/|\eta|(B_i))\,d|\eta|(y)$$

is an optimal multiphase transport with respect to θ_N.

The main result is:

Theorem: *If* $(X,\vec{\sigma}) \succeq (Y,\vec{\eta})$ *then the weak-* limit* $P^N \to P$ *exists, and* P *is an optimal multiphase transport for* θ.

References

[1] Blackwell D., *Comparison of Experiments*, Proc. Second Berkeley Symp. on Math. Statist. and Prob. (Univ. of Calif. Press, 1951), 93–102

The gradient flow of microstructure

David Kinderlehrer

(joint work with Patrick Bardsley, Katayun Barmak, Eva Eggeling, Maria Emelianenko, Yekaterina Epshteyn, Xin Yang Lu, Shlomo Ta'asan)

Cellular networks are ubiquitous in nature. They exhibit behavior on many different length and time scales and are generally metastable. Most technologically useful materials are polycrystalline microstructures composed of a myriad of small monocrystalline grains separated by grain boundaries, and thus comprise cellular networks. The energetics and connectivity of the grain boundary network plays a crucial role in determining the properties of a material across a wide range of scales. A central problem is to develop technologies capable of producing an arrangement of grains that provides for a desired set of material properties. The traditional focus has been on distributions of geometric features, like cell size, and a preferred distribution of grain orientations, termed texture. Attaining these gives the configuration order in a statistical sense. More recent mesoscale experiment and simulation permit harvesting large amounts of information about both geometric features and crystallography of the boundary network itself, [3, 2, 16, 18, 19]. This has led us to the notion of the Grain Boundary Character Distribution (GBCD). The GBCD is an empirical distribution of the relative length (in 2D) or area (in 3D) of interface with a given lattice misorientation and grain boundary normal. It is a leading candidate to characterize texture of the boundary network.

In the special situation where given interfacial energy depends only on lattice misorientation, our simulations show that the steady state GBCD and the interfacial energy density are related via a Boltzmann distribution. This is among the simplest non-random distributions, corresponding to independent trials with respect to the density. Such a straightforward dependence between the character distribution and the interfacial energy offers evidence that the GBCD is a material property. For this GBCD statistic, we develop a theory that relies on mass transport and entropy, see [10, 7, 8, 11, 13, 12] for a more detailed discussion and [9] for a discussion directed to materials researchers. To further develop this, we seek to identify it as a gradient flow in the sense of De Giorgi as developed by Ambrosio, Gigli, and Savaré, [5, 4]. In this way, the empirical texture statistic is revealed as a solution of a Fokker-Planck type equation whose evolution is determined by weak topology kinetics and whose limit behavior is the observed Boltzmann distribution for the prescribed interfacial energy density, [6]. To achieve this we must determine first an appropriate dissipation relation, for which we introduce the discrete iteration principle first noted in [15]. Further viewing the GBCD as samples of a process motivates us to adjust the time scale in a nontrivial manner, precisely

matching the evolution of a Fokker Planck Equation. That the simulation time scale must be coordinated to the physical time scale is a challenging yet persistent problem. It arises even in the simulation of the simplest systems, like the Ehrenfest Urn. This identification as a gradient flow is tantamount to exhibiting the harvested statistic as the iterates in a JKO implicit scheme, [15].

Enroute to the GBCD results, we also study a simpler 1D model coarsening system that shares many qualitative features with the GBCD, omitted in this report. The development exposes the question of how to understand the circumstances under which a harvested empirical statistic is a property of the underlying process. Both the GBCD and the simpler 1D model coarsening system are 'gradient flows found in nature': they arise from complex systems, perhaps even outside of mathematics. We gave brief attention to two other such systems in our Oberwolfach discussion. These are the Ehrenfest Urn and the Wright-Fisher formulation of (neutral) genetic drift, in collaboration with Laurent Dietrich and Léonard Monsaingeon. For the latter, we have yet to achieve adequate success.

The GBCD viewpoint on texture and texture dependent properties is an active area of materials research [1, 14]. Viewing the evolution of complex systems as gradient flows is also discussed in the very interesting work [17].

References

[1] Daniel Abou-Ras, Norbert Schaefer, Thorsten Rissom, Madeleine N. Kelly, Jakob Haarstrich, Carsten Ronning, Gregory S. Rohrer, and Anthony D. Rollett. Grain-boundary character distribution and correlations with electrical and optoelectronic properties of CuInSe2 thin films. *ACTA MATERIALIA*, 118:244–252, OCT 1 2016.

[2] B.L. Adams, D. Kinderlehrer, I. Livshits, D. Mason, W.W. Mullins, G.S. Rohrer, A.D. Rollett, D. Saylor, S Ta'asan, and C. Wu. Extracting grain boundary energy from triple junction measurement. *Interface Science*, 7:321–338, 1999.

[3] BL Adams, D Kinderlehrer, WW Mullins, AD Rollett, and S Ta'asan. Extracting the relative grain boundary free energy and mobility functions from the geometry of microstructures. *Scripta Materiala*, 38(4):531–536, Jan 13 1998.

[4] Luigi Ambrosio and Nicola Gigli. *Modelling and Optimisation of Flows on Networks: Cetraro, Italy 2009, Editors: Benedetto Piccoli, Michel Rascle*, chapter A User's Guide to Optimal Transport, pages 1–155. Springer Berlin Heidelberg, Berlin, Heidelberg, 2013.

[5] Luigi Ambrosio, Nicola Gigli, and Giuseppe Savaré. *Gradient flows in metric spaces and in the space of probability measures*. Lectures in Mathematics ETH Zürich. Birkhäuser Verlag, Basel, second edition, 2008.

[6] Patrick Bardsley, Katayun Barmak, Eva Eggeling, Yekaterina Epshteyn, David Kinderlehrer, and Shlomo Ta'asan. Material microstructure: Texture evolution in two dimensions. *in preparation*, 2016.

[7] K. Barmak, E. Eggeling, M. Emelianenko, Y. Epshteyn, D. Kinderlehrer, R.Sharp, and S.Ta'asan. Predictive theory for the grain boundary character distribution. In Proc. Recrystallization and Grain Growth IV, 2012.

[8] K. Barmak, E. Eggeling, M. Emelianenko, Y. Epshteyn, D. Kinderlehrer, R. Sharp, and S. Ta'asan. Critical events, entropy, and the grain boundary character distribution. *Phys. Rev. B*, 83:134117, Apr 2011, Editors' Suggestion.

[9] K. Barmak, E. Eggeling, M. Emelianenko, Y. Epshteyn, D. Kinderlehrer, R. Sharp, and S. Ta'asan. Recent developments in material microstructure: a theory of coarsening. In *Symposium NN - Mathematical and Computational Aspects of Materials Science*, volume 1753 of *MRS Proceedings*, 2015.

[10] K. Barmak, E. Eggeling, M. Emelianenko, Y. Epshteyn, D. Kinderlehrer, and S. Ta'asan. Geometric growth and character development in large metastable networks. *Rend. Mat. Appl. (7)*, 29(1):65–81, 2009.
[11] Katayun Barmak, Eva Eggeling, Maria Emelianenko, Yekaterina Epshteyn, David Kinderlehrer, Richard Sharp, and Shlomo Ta'asan. An entropy based theory of the grain boundary character distribution. *Discrete Contin. Dyn. Syst.*, 30(2):427–454, 2011.
[12] Katayun Barmak, Eva Eggeling, Maria Emelianenko, Yekaterina Epshteyn, David Kinderlehrer, Richard Sharp, and Shlomo Ta'asan. A theory and challenges for coarsening in microstructure. In *Analysis and numerics of partial differential equations*, volume 4 of *Springer INdAM Ser.*, pages 193–220. Springer, Milan, 2013.
[13] Katayun Barmak, Eva Eggeling, Maria Emelianenko, Yekaterina Epshteyn, David Kinderlehrer, Richard Sharp, and Shlomo Ta'asan. Materials microstructure: entropy and curvature driven coarsening. volume 1881, pages 71–91. RIMS, Research Institute for Mathematical Sciences, U. Kyoto, 2014.
[14] K. J. Ganesh, A. D. Darbal, S. Rajasekhara, G. S. Rohrer, K. Barmak, and P. J. Ferreira. Effect of downscaling nano-copper interconnects on the microstructure revealed by high resolution TEM-orientation-mapping. *NANOTECHNOLOGY*, 23(13), APR 6 2012.
[15] R Jordan, D Kinderlehrer, and F Otto. The variational formulation of the fokker-planck equation. *SIAM J. Math. Analysis*, 29(1):1–17, Jan 1998.
[16] D Kinderlehrer, I Livshits, GS Rohrer, S Ta'asan, and P Yu. Mesoscale simulation of the evolution of the grain boundary character distribution. *Recrystallization and grain growth, pts 1 and 2*, 467-470(Part 1-2):1063–1068, 2004.
[17] Tim Laux and Felix Otto. Convergence of the thresholding scheme for multi-phase mean-curvature flow. *Calc. Var. Partial Differential Equations*, 55(5):Paper No. 129, 74, 2016.
[18] GS Rohrer. Influence of interface anisotropy on grain growth and coarsening. *Annual Review of Materials Research*, 35:99–126, 2005.
[19] Anthony D. Rollett, S.-B. Lee, R. Campman, and G. S. Rohrer. Three-dimensional characterization of microstructure by electron back-scatter diffraction. *Annual Review of Materials Research*, 37:627–658, 2007.

Transport Metrics for Vlasov Hierarchies

Martin Burger

Strong progress has been made in the geometric understanding of macroscopic evolution equations for probability densities in the last decades, which often lead to gradient structures in transport metrics (cf. e.g. [2, 6]). Since frequently microscopic evolution laws can be formulated in a similar way, it is a natural question how to proceed from the microscopic to the macroscopic structures. Strong progress has been made recently e.g. via large deviation functionals (cf. [1, 3]), but we want instead take a standard route in statistical mechanics and consider equation hierarchies for marginals and define a (at this point formal) gradient flow structure on them.

1. Microscopic Models and Gradient Structure

For simplicity consider a system of N indistinguishable random particles on the d-dimensional torus $\mathbb{T}$, with positions $X_i(t) \in \mathbb{T}$ evolving via

$$\frac{dX_i}{dt} = -\nabla V(X_i) - \frac{1}{N}\sum_{j\neq i} \nabla W(X_i - X_j).$$

The joint probability measure μ_N on $\mathbb{T}^N$ evolves via Liouville equation (in the sense of a measure valued weak solution)

$$\partial_t \mu_N = \sum_{i=1}^{N} \nabla_{x_i} \cdot \left(\mu_N \nabla V(x_i) + \frac{1}{N} \sum_{j \neq i} \mu_N \nabla W(x_i - x_j) \right).$$

With the above scaling, the Liouville equation is a gradient flow in the 2-Wasserstein metric on $\mathcal{P}_2(\mathbb{T}^N)$, the space of probability measures with finite second moment on $\mathbb{T}^N$. The energy functional is given by

$$\mathcal{E}_N[\mu^N] = \int \left(\sum_{i=1}^{N} V(x_i) + \frac{1}{2N} \sum_{i=1}^{N} \sum_{j \neq i} W(x_i - x_j) \right) \, d\mu_N.$$

Following [2, 6] the metric can be written as

$$d(\mu_N^0, \mu_N^1)^2 = \inf \int_0^1 \int_{\mathbb{T}^N} |\nabla \lambda_N|^2 \, d\mu_N$$

with the infimum carried out over all vector fields λ and measures $\mu_N \in C(0, T; \mathcal{P}_2(\mathbb{T}^N))$ such that $\mu_N(t = 0) = \mu_N^0$, $\mu_N(t = 1) = \mu_N^1$, and

$$\partial_t \mu_N = \sum_{i=1}^{N} \nabla_{x_i} \cdot (\mu_N \nabla_{x_i} \lambda_N)$$

in a weak sense on $\mathbb{T}^N \times (0, 1)$. In order to keep notation simple we will assume in the following that all measures are absolutely continuous with respect to the Lebesgue measure and denote the probability density with f^N, by abuse of notation we will use the Wasserstein metric as a metric on probability densities.

2. The Vlasov Hierarchy

In order to perform mean-field limits in systems such as the Liouville equation above, it is a common approach to consider the marginals of f_N (cf. [4] and references therein), defined by

$$f_{N:k} = \int \dots \int f_N \, dx_{k+1} \dots \, dx_N,$$

which allow for a more suitable limit, since each $f_{N:k}$ is a density on the domain $\mathbb{T}^k$ of fixed dimension. Integrating the Liouville equation one finds the Vlasov (or BBGKY) hierarchy of equations satisfied by the marginals for $1 \leq k \leq N$:

$$\begin{aligned} \partial_t f_{N:k} = & \sum_{i=1}^{k} \nabla_{x_i} \cdot \left(f_{N:k} \nabla V(x_i) + \frac{1}{N} \sum_{j \neq i, j \leq k} f_{N:k} \nabla W(x_i - x_j) \right) + \\ & \sum_{i=1}^{k} \nabla_{x_i} \cdot \left(\frac{N-k}{N} \int f_{N:k+1} W(x_i - x_{k+1}) \, dx_{k+1} \right) \end{aligned}$$

As $N \to \infty$ we may expect convergence to the infinity BBGKY-type hierarchy with marginals $f_{\infty,k}$, $k \geq 1$, characterized by

$$\partial_t f_{\infty:k} = \sum_{i=1}^{k} \nabla_{x_i} \cdot \left(f_{\infty:k} \nabla V(x_i) + \int f_{\infty:k+1} \nabla W(x_i - x_{k+1}) \, dx_{k+1} \right).$$

We refer e.g. to [5] for quantitative convergence results.

3. Gradient Structure for the Hierarchy

In order to understand the gradient structure present in the Vlasov hierarchy, we first rewrite the energy in terms of the hierarchy, we find

$$\mathcal{E}_N[\mu^N] = N \int V(x_1) f_{N:1}(x_1) \, dx_1 + \frac{N-1}{2} \int \int W(x_1 - x_2) f_{N:2}(x_1, x_2) \, dx_1 dx_2.$$

Thus, rescaling by $\frac{1}{N}$ in the limit $N \to \infty$, we obtain a limiting energy functional on the hierarchy

$$\mathcal{F} = \int V(x) f_{\infty:1}(x) \, dx + \frac{1}{2} \int \int W(x-y) f_{\infty:2}(x,y) \, dxdy.$$

This suggests to rewrite the hierarchy as

$$\partial_t f_{\infty:k} = \sum_{i=1}^{k} \nabla_{x_i} \cdot \left(f_{\infty:k} \nabla_{x_i} \frac{\partial \mathcal{F}}{\partial f_{\infty:1}}(x_i) + 2 \int f_{\infty:k+1} \nabla_{x_i} \frac{\partial \mathcal{F}}{\partial f_{\infty:2}}(x_i, x_{k+1}) \, dx_{k+1} \right),$$

which indicates some kind of gradient structure in the hierarchy.The fact that $\mathcal{F}$ only depends on the first two marginals shadows the general structure given by

$$\partial_t f_{\infty:k} = \sum_{i=1}^{k} \nabla_{x_i} \cdot \Big(\sum_{m=1}^{\infty} m \int \dots \int f_{\infty:k+m-1} \nabla_{x_i} \frac{\partial \mathcal{F}}{\partial f_{\infty:m}}(x_i, x_{k+1}, \dots, x_{k+m-1}) \, dx_{k+1} \dots \, dx_{k+m-1} \Big).$$

The further terms in the asymptotic can be checked using interaction terms on triplet, quadruples and so on, but we omit the tedious computations at this point. This observation induces a formal gradient flow structure on the hierarchy via

$$d(f_\infty^0, f_\infty^1)^2 = \inf \left(\frac{1}{2} \int_0^1 \langle \lambda, \mathcal{K}_f \lambda \rangle \, dt \right),$$

where the infimum is taken over all $f_\infty = (f_{\infty:k})_{k=1}^\infty$ with initial value f_∞^0 and final value f_∞^1 and potentials $\lambda = (\lambda_k)_{k=1}^\infty$ such that $\partial_t f_\infty = \mathcal{K}_f \lambda$. The formally self-adjoint operator $\mathcal{K}_f = (\mathcal{K}_f^k)_{k=1}^\infty$ is given by

$$\mathcal{K}_f^k \lambda = -\sum_{i=1}^{k} \nabla_{x_i} \cdot \Big(\sum_{m=1}^{\infty} m \int \dots \int f_{\infty:k+1} \nabla_{x_i} \lambda_m(x_i, x_{k+1}, \dots, x_{k+m-1}) \, dx_{k+1} \dots \, dx_{k+m-1} \Big)$$

with the scalar product

$$\langle \varphi, \mathcal{K}_f \lambda \rangle = \sum_{k=1}^{\infty} \int \dots \int \varphi_k(x_1, \dots, x_k)(\mathcal{K}_f^k \lambda)(x_1, \dots, x_k)\, dx_1 \dots\, dx_k.$$

The transport metric for the hierarchy remains an interesting subject for a rigorous analyis and is expected to provide novel insights into continuum limits. The mean-field limit can be interpreted as a confinement of the gradient flow to the closed submanifold of product measures. Even more interestingly, preservation of the gradient structure might allow to select meaningful higher-order closures.

References

[1] S. Adams, N. Dirr, M. Peletier, J. Zimmer, *From a large-deviations principle to the Wasserstein gradient flow: a new micro-macro passage*, Communications in Mathematical Physics, **307** (2011), 791–815.

[2] L. Ambrosio, N. Gigli, G. Savare, *Gradient Flows in Metric Spaces and in the Space of Probability Measures*, Springer Science and Business Media (2008).

[3] M. Duong, A. Lamacz, M. Peletier, U. Sharma, *Variational approach to coarse-graining of generalized gradient flows*, Preprint, arXiv:1507.03207 (2015).

[4] F. Golse, *On the dynamics of large particle systems in the mean field limit*, in:Macroscopic and Large Scale Phenomena: Coarse Graining, Mean Field Limits and Ergodicity, Springer International Publishing, (2016). 1–144.

[5] F. Golse, C.Mouhot, V.Ricci, *Empirical measures and Vlasov hierarchies*, Kinetic and Related Models **6** (2013), 919–943.

[6] M. Liero, A. Mielke *Gradient structures and geodesic convexity for reaction-diffusion systems*, Phil. Trans. R. Soc. A **371** (2005), 20120346.

On the transport problem with relativistic costs

Aldo Pratelli

(joint work with J. Bertrand, J. Louet, M. Puel, F. Zeisler)

The so-called "transport problem with relativistic costs" is a particular case of the general mass transportation problems. Basically, we are given a closed, bounded, convex set with non-empty interior $\mathcal{K} \subseteq \mathbb{R}^N$, and a function $h : \mathbb{R}^N \to [0, +\infty]$ with $h(0) = 0$, which is constantly $+\infty$ on $\mathbb{R}^N \setminus \mathcal{K}$, and it is bounded and strictly convex in $\mathcal{K}$. Then, for every $t > 0$ one considers that mass transport problem with the cost

$$c_t(x, y) = h\left(\frac{y - x}{t}\right).$$

Such a cost is called *relativistic*, while it is called *highly relativistic* if, in addition, the slope of h explodes on $\partial\mathcal{K}$. The model case, first considered by Brenier in [3], deals with $\mathcal{K}$ being the unit ball of $\mathbb{R}^N$, and $h(z) = 1 - \sqrt{1 - |z|^2}$ in $\mathcal{K}$, which is immediately seen to be a highly relativistic cost.

It is easy to observe that, given two probability measures with compact support μ and $\nu \neq \mu$, there exists some $T > 0$, called *critical time*, such that the minimal cost $\mathcal{C}(t)$ to transport μ onto ν with respect to the cost c_t is infinite for $t < T$, and bounded for $t \geq T$; in addition, for each $t \geq T$ there exists a unique optimal

transport plan (actually, a transport map), that we call γ_t. The function $t \mapsto \mathcal{C}(t)$ is obviously decreasing.

Following earlier partial results found in [5, 2], in the recent papers [1, 4] the following properties were proved; most of the assumptions are sharp, thanks to explicit counterexamples.

Theorem. Let $\mu \neq \nu$ be two probability measures with compact support in $\mathbb{R}^N$, and assume that c_t is a relativistic cost function, and that $\mu \ll \mathcal{L}$. Then,

(i) The function $t \mapsto \mathcal{C}(t)$ is continuous on $[T, +\infty)$.

(ii) For every supercritical time $t > T$, the optimal plan γ_t satisfies $\gamma_t\big(\big\{(x,y) : \frac{y-x}{t} \in \Theta\big\}\big) = 0$, where

$$\Theta := \Big\{ v \in \partial\mathcal{K} :\ D_{-v}h(v) = -\infty \Big\},$$

and $D_{-v}h(v) \in [-\infty, +\infty)$ is the slope of h at the point v in the direction $-v$. In particular, if c_t is highly relativistic then $\gamma_t(\partial\mathcal{K}) = 0$.

(iii) If c_t is highly relativistic and μ has connected support, then there exists a Kantorovich potential φ_t for every supercritical time $t > T$.

References

[1] J. Bertrand, A. Pratelli & M. Puel, Kantorovich potentials and continuity of total cost for relativistic cost functions, to appear on JMPA (2017).

[2] J. Bertrand, M. Puel, The optimal mass transport problem for relativistic costs, Calc. Var. PDEs **46** (2013), no. 1–2, 353–374.

[3] Y. Brenier, Extended Monge-Kantorovich theory, in "Optimal Transportation and Applications", Lecture Notes in Mathematics, LNM **1813**, Springer (2003), 91–121.

[4] J. Louet, A. Pratelli & F. Zeisler, On the continuity of the total cost in the mass transport problem with relativistic cost functions, preprint (2016).

[5] R. McCann, M. Puel, Constructing a relativistic heat flow by transport time steps, Ann. Inst. H. Poincaré Anal. Non Linéaire **26** (2009), no. 6, 2539–2580.

Continuum and Discrete Models of Collective Aggregation Pattern

Klemens Fellner

(joint work with B. Hughes, E. J. Hackett-Jones, K. A. Landman, G. Raoul)

The talk discusses the behaviour of solutions of the following non-local Fokker-Planck type equation

$$\partial_t \rho = \partial_x(\rho\, \partial_x[a(\rho) + W * \rho + V]) \tag{1}$$

where we focus in particular on the one-dimensional case $x \in \mathbb{R}$.

The above equation (1), which is also called non-local interaction equation, or aggregation equation provides a continuum picture of the collective motion of an ensemble of particles or individuals subject to diffusion (linear diffusion for $a(\rho) = \ln(\rho)$), an external potential $V(x)$ and, most importantly, an interaction potential W, which is assumed to be an even function.

As a first structural property, eq. (1) conserves the total mass, which shall thus be considered normalised, i.e. mass $\int_{\mathbb{R}} \rho = 1$ and $\rho(x)$ can be interpreted as a

probability density. Moreover, it is well known that depending on the interaction potential W, the evolution (1) may lead to the development of measures in finite time even for smooth initial data. A particular example is the famous Keller-Segel model, where $K \sim \log$ in 2D.

Thus, a suitable solutions concept of (1) is based on its gradients flow structure in probability spaces with bounded second moment and the associated JKO scheme with respect to the Wasserstein metric, see e.g. [1, 4]. In the one-dimensional case, introducing a change of variables in terms of the pseudo-inverse of the accumulated probability density, i.e. $u(z) = \inf\{x \in \mathbb{R};\ \int_{-\infty}^{x} \rho\, dx > z\}$, for $z \in [0,1]$ (see e.g. [10, 3]) allows to transform atomic parts of ρ into constants parts of u and the p-Wasserstein norm into a standard L^p-norm. Moreover, when only considering the interaction potential W, eq. (1) transforms into

$$\partial_t u(z) = \int_0^1 W'(u(\zeta) - u(z))\, d\zeta\,, \qquad z \in [0,1]. \tag{2}$$

In [6, 7], the authors studied (2) subject to locally repulsive, globally confining/attractive interaction potentials W. Such interaction potentials appear in many application backgrounds: smooth power-like interaction potentials in cell-biology, various Morse potentials in models of flocking and swarming, and also the Lennard Jones potential in physics, see e.g. [11, 12, 14, 15].

Besides linear and local nonlinear stability of stationary states, it was shown in [6, 7] that an increasing "singularity" of the repulsive part leads to the formation of arbitrary many Dirac measures as stationary states, which are highly non-unique and depend in a complicated way on the interaction potential and the initial data.

In the limit towards a singular repulsive Newtonian potential, these multitude of atomic stationary states converges weakly (in the sense of signed measures) towards a unique bounded and compactly supported stationary state, which are moreover Hölder continuous for even more singular repulsive parts [5].

To the author, it is this complicated relationship between interaction potential, initial data and aggregation pattern, which is particularly intriguing. For doubly singular interaction potentials, where also attractive parts are Newtonian, a comparison of (2) with discrete stochastic lattice models showed a complete qualitative correspondence between continuum and stochastic discrete models, see [8]: Both models predicted aggregation pattern, which combine sharp aggregates with smoothly distributed parts, both models showed similar slow-fast dynamics in the formation of metastable aggregation pattern, which take then much larger times to finally converge to stationary states.

In [9], motility effects like diffusion were discussed and for some particular interaction potentials, it was possible to derive sharp thresholds between uni- and bimodal stationary states, according to the ration of diffusion versus aggregation parameters.

In higher space dimensions, the evolution of aggregation equation (1) leads to a stunning variety of aggregation pattern, see e.g. [2]. In the current interdisciplinary research project "Mathematics and Arts: Towards a balance between

artistic intuition and mathematical complexity", which is located at the University of Graz, a team of two mathematicians and two artists aims to research mathematical structures and dynamical systems, which offer "contact points" to non-mathematicians like composers, designers, artists, etc. to insert their creativity and intuition at the "centre" of a mathematical system with inspiring complexity.

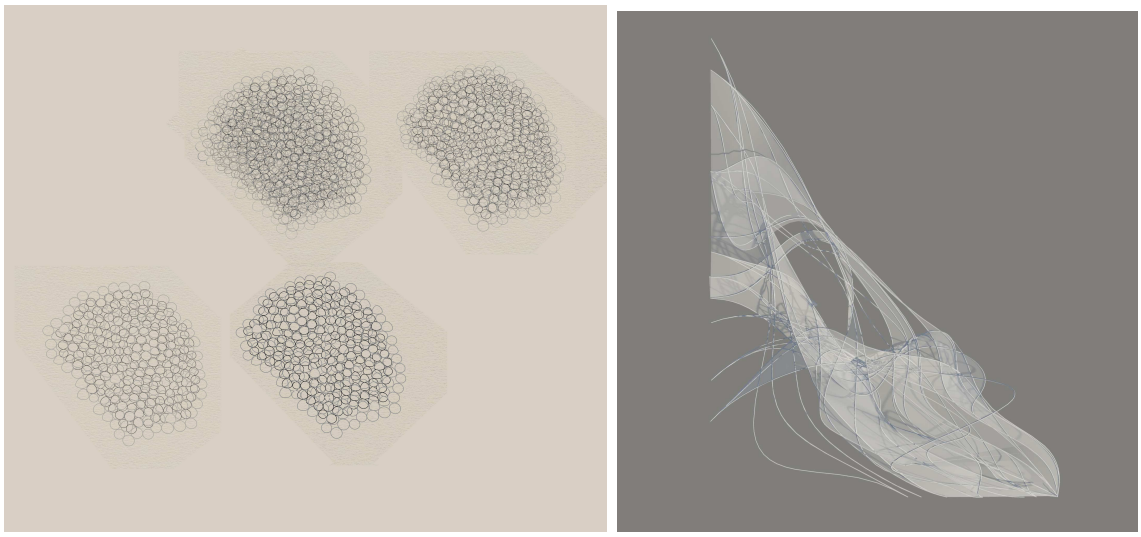


FIGURE 1. Two studies for [13]: layers of arranged aggregation pattern and juxtaposition of lines and surfaces derived from bi-modality plots.

An example can be visited at the Research Catalogue site "The intricacy of self-similarity", [13]. The Research Catalogue is a timely platform for artistic research and offers a flexible online exhibition of texts, graphical contents, sound- and video materials. The site "The intricacy of self-similarity" collects artistic material in terms of graphical exercises such as layers of pattern of arranged aggregation pattern redrawn from [2], see also Figure 1. A second exercise explores plots of [9] concerning the above mentioned bi-modality threshold. Both exercises lead finally towards a soundscape pattern and the viewer/listener is invited to play with the various layers in relation to the graphical material.

REFERENCES

[1] L. Ambrosio, N. Gigli and G. Savaré, *Gradient flows in metric spaces and in the space of probability measures*. Second edition. Lectures in Mathematics ETH Zürich. Birkhäuser Verlag, Basel, (2008).

[2] J. von Brecht, D. Uminsky, T. Kolokolnikov, A. L. Bertozzi, *Predicting pattern formation in particle interactions* M3AS, **22**, (2012) 1140002.

[3] M. Burger and M. Di Francesco, *Large time behavior of nonlocal aggregation models with non-linear diffusion*, Netw. Heterog. Media **3** (2008) 749–785.

[4] J.A. Carrillo, M. Di Francesco, A. Figalli, T. Laurent and D. Slepčev, *Global-in-time weak measure solutions, finite-time aggregation and confinement for nonlocal interaction equations*, Duke Mathematical Journal, **156**, 229 (2011) 229–271.

[5] J.A. Carrillo, M.G. Delgadino, A. Mellet, *Regularity of local minimizers of the interaction energy via obstacle problems* Comm. Math. Phys. **343** no. 3 (2016) 747–781.
[6] K. Fellner, G. Raoul, *Stable stationary states of non-local interaction equations*, Mathematical Models and Methods in Applied Sciences **20** (2010) 2267–2291.
[7] K. Fellner, G. Raoul, *Stability of stationary states of non-local interaction equations*, Mathematical and Computer Modelling **53** (2011) 1436–1450.
[8] E.J. Hackett-Jones, K.A. Landman and K. Fellner, *Aggregation patterns from non-local interactions: discrete stochastic and continuum modelling*, Physical Review E **85** (2012) 041912.
[9] B. Hughes, K. Fellner, *Continuum models of cohesive stochastic swarms: the effect of motility on aggregation patterns*, Physica D, **260** (2013) 26–48.
[10] H. Li and G. Toscani, *Long-time asymptotics of kinetic models of granular flows*, Arch. Ration. Mech. Anal. **172**(3) (2004) 407–428.
[11] K. Kang, B. Perthame, A. Stevens and J.J.L. Velázquez, An integro-differential equation model for alignment and orientational aggregation, *J. Differential Equations* **246**(4) (2009) 1387–1421.
[12] I. Primi, A. Stevens, J. J.L. Velazquez, *Mass-Selection in Alignment Models with Non-Deterministic Effects*, Comm. Partial Differential Equations **34**(5) (2009).
[13] The intricacy of self-similarity, `https://www.researchcatalogue.net/view/315179/315180/0/0`.
[14] C.M. Topaz and A.L. Bertozzi, Swarming patterns in a two-dimensional kinematic model for biological groups, *SIAM J. Appl. Math. 65* (2004) 152–174.
[15] F. Theil, A proof of crystallization in two dimensions, *Comm. Math. Phys.* **262**(1) (2006) 209–236.

Optimal stability estimates for continuity equations

Christian Seis
(joint work with André Schlichting)

The continuity equation is one of the most elementary and at the same time most important equations of motion with applications in fluid dynamics, thermodynamics, engineering, biology, social sciences. It describes the conservative transport of a quantity ρ by a vector field u. In a bounded (say, convex Lipschitz) domain Ω in $\mathbf{R}^d$, this equation takes the form

$$\begin{cases} \partial_t \rho + \nabla \cdot (u\rho) = 0 & \text{in } (0,T)\times\Omega, \\ \rho(0,\cdot) = \rho^0 & \text{in } \Omega. \end{cases} \tag{1}$$

In many problems, both the advecting velocity field and the transported quantity are non-smooth functions. In the following we will be concerned with the low regularity setting investigated by DiPerna and Lions [5], that is, we suppose that

$$\rho \in L^\infty((0,T);L^q(\Omega)), \quad \rho^0 \in L^q(\Omega), \tag{2}$$

and

$$u \in L^1((0,T);W^{1,p}(\Omega)) \quad \text{with} \quad (\nabla\cdot u)^- \in L^1((0,T);L^\infty(\Omega)), \tag{3}$$

where $p,q \in (1,\infty)$ are such that $1/p + 1/q = 1$. To ensure that (1) is mass preserving, we moreover assume that $u\cdot\nu = 0$ on $\partial\Omega$. In their ground breaking paper, DiPerna and Lions show existence, uniqueness and qualitative stability

of distributional solutions in the class (2), (3), using their so-called theory of renormalized solutions. In [8], *quantitative* stability estimates were established for the first time. The prototype estimate is the following:

Theorem 1 ([8])**.** *If $u_1, u_2, \dots$ are an approximating vector fields in the sense that*

$$\delta_k(t) := \|u - u_k\|_{L^1((0,t);L^p(\Omega))} \overset{k\to\infty}{\longrightarrow} 0$$

uniformly in t and $\rho_1, \rho_2, \dots$ are the associated solutions with initial datum ρ^0, then

$$\sup_{[0,T]} D_{\delta_k}(\rho, \rho_k) \lesssim \|\nabla u\|_{L^1(L^p)} + 1 \tag{4}$$

uniformly in $k \in \mathbf{N}$.

Here D_δ is a Kantorovich–Rubinstein distance of the form

$$D_\delta(\rho_1, \rho_2) = \inf_{\pi \in \Pi((\rho_1-\rho_2)^+, (\rho_1-\rho_2)^-)} \iint \log\left(\frac{|x-y|}{\delta} + 1\right) d\pi(x,y).$$

Constance depending on ρ^0 or $\nabla \cdot u$ are suppressed in (4).

Before discussing the statement in the theorem, we will make a connection with previously known estimates: The estimate in (4) is a natural extension of stability estimates for Lagrangian flows describing the motion of single particles. Indeed, in the case of Lipschitz vector fields $u, u_k \in L^1(W^{1,\infty})$ a straightforward computation for the distance of the flows ϕ and ϕ_k defined by

$$\partial_t \phi = u(\cdot, \phi), \quad \partial_t \phi_k = u_k(\cdot, \phi_k), \quad \phi(0,x) = \phi_k(0,x) = x,$$

shows that

$$\sup_{[0,T]} \sup_{\Omega} \log\left(\frac{|\phi - \phi_k|}{\delta_k} + 1\right) \le \|\nabla u\|_{L^1(L^\infty)} + 1,$$

where $\delta_k = \|u - u_k\|_{L^1(L^\infty)}$. That is, *the logarithmic relative distance of two particles transported by different flows is controlled by the velocity gradient.* Here "relative" refers to the distance of the particles relative to the distance δ_k of the velocity fields. This estimate was generalized to the DiPerna–Lions setting by Crippa and De Lellis [1]. It holds

$$\sup_{[0,T]} \int_\Omega \log\left(\frac{|\phi - \phi_k|}{\delta_k} + 1\right) dx \lesssim \|\nabla u\|_{L^1(L^p)} + 1. \tag{5}$$

The prototype estimate (4) in Theorem 1 is of the same form: The velocity gradient controls the logarithmic relative distance of the densities ρ and ρ_k. In the case of the transport of a single particle, estimate (4) reduces to (5).

Because Kantorovich–Rubinstein distances metrize weak convergence, the statement in Theorem 1 shows that

$$\rho_k \longrightarrow \rho \quad \text{weakly with rate at most } \mathcal{O}(\delta_k).$$

This is optimal as can be seen in the following example partially taken from [4]:

Example 1. *Consider the vector fields* $u_k(x) = \sin(2\pi kx)/2\pi k$ *on the interval* $\Omega = [0,1]$*, and the initial datum* $\rho^0 = 1$*. Then* ρ_k *is an oscillating solution converging weakly to* $\rho = 1$*. In particular*

$$\|\rho - \rho_k\|_{L^1(L^1)} \not\to 0.$$

Moreover, the $L^1(L^p)$ *distance of the vector fields* u_k *and* $u = 0$ *(to which* u_k *is uniformly converging) is of order* t/k*. A scaling argument then yields*

$$\forall t, k: \quad D_{\delta_k(t)}(\rho(t), \rho_k(t)) \lesssim 1.$$

This example thus shows that, firstly, one cannot expect to prove stability estimates for continuity equations in strong (Lebesgue) norms. Secondly, the order δ_k of weak convergence is optimal.

The approach from [8, 9] turns out to be a helpful tool for estimating the numerical approximation error of the upwind finite volume scheme. In [6] we investigate the explicit scheme on Cartesian meshes. Using a probabilistic interpretation of the upwind scheme analyzed earlier in [2, 3], we show that the rate of (weak) convergence is at most $1/2$ in the mesh size h:

Theorem 2. *Let* ρ_h *denote the approximate solution given by the upwind scheme on a mesh of size* h*. Then*

$$\sup_{[0,T]} D_{\sqrt{h}}(\rho, \rho_h) \lesssim 1$$

uniformly in h*.*

Also this result is optimal, as can be seen by the following example.

Example 2. *Consider a constant vector field* $u = U > 0$ *and the initial datum*

$$\rho^0(x) = \begin{cases} x^{-s} & \text{for } x > 0, \\ 0 & \text{for } x \le 0. \end{cases}$$

Then the L^1 *distance of exact and approximate solution is bounded below as follows:*

$$\|\rho - \rho_h\|_{L^1(L^1))} \gtrsim h^{1-s}.$$

Moreover, it holds

$$\forall t: \quad W_1(\rho, \rho_h) \gtrsim \sqrt{h}^{2-s},$$

if W_1 *is the (standard)* 1*-Wasserstein distance.*

In the limit $s \to 1$, this example shows that, firstly, for general merely integral initial data, no convergence rates can be obtained in strong (Lebesgue) norms. Secondly, the rate of weak convergence is at least $1/2 - \varepsilon$, which almost matches the upper bound from Theorem 2, though with a different measure of weak convergence.

In [7], we study the implicit upwind scheme and extend the result from [6] to the case of arbitrary mesh geometries. The new result moreover relies on a priori estimates on the scheme, so-called "weak BV estimates" rather than probabilistic methods.

References

[1] G. Crippa and C. De Lellis. Estimates and regularity results for the DiPerna-Lions flow. *J. Reine Angew. Math.*, 616:15–46, 2008.

[2] F. Delarue and F. Lagoutière. Probabilistic analysis of the upwind scheme for transport equations. *Arch. Ration. Mech. Anal.*, 199(1):229–268, 2011.

[3] F. Delarue, F. Lagoutière, and N. Vauchelet. Convergence order of upwind type schemes for transport equations with discontinuous coefficients. Preprint arxiv:1602.05746, 2016.

[4] C. De Lellis, P. Gwiazda, and A. Świerczewska-Gwiazda. Transport equation with integral terms. Preprint arXiv:1602.03193, 2016.

[5] R. J. DiPerna and P.-L. Lions. Ordinary differential equations, transport theory and Sobolev spaces. *Invent. Math.*, 98(3):511–547, 1989.

[6] A. Schlichting and C. Seis. Convergence rates for upwind schemes with rough coefficients. To appear in *SIAM J. Numer. Anal.*, 2017.

[7] A. Schlichting and C. Seis. Analysis of the implicit upwind finite volume scheme with rough coefficients. *Submitted, arXiv:1702.02534*, 2017.

[8] C. Seis. A quantitative theory for the continuity equation. To appear in *Ann. Inst. H. Poincaré Anal. Non Linéaire*, 2017.

[9] C. Seis. Optimal stability estimates for continuity equations. To appear in *Proc. Roy. Soc. Edinburgh Sect. A*, 2017.

Applications of optimal transport to meteorology

Mike Cullen

I reviewed progress of three applications of optimal transport methods to meteorology.

The first application is to the semi-geostrophic equations. These model the large-scale behaviour of the atmosphere and oceans. Much previous work is described in [4]. The equations start from the shallow atmosphere and hydrostatic approximations to the governing equations, which are widely used in operational prediction. They then introduce the geostrophic momentum approximation, which restricts the solutions to those dominated by the Earth's rotation. No other approximations are made.

All previous rigorous work has, in addition, assumed a spatially uniform rotation rate. This allows a change of variables to be made and a new conserved quantity to be derived which is the mass as a function of the new variables. This extra conservation property results from the symmetry of the problem about the axis of rotation. The solutions can then by found by optimal transport between the new variables and the original coordinates. The cost function is the energy, and the solutions can be written in terms of a convex potential function. The convexity property is fundamental in showing that a sequence of approximate solutions converges, and thus in proving the existence of solutions.

However, in real applications, the strong radial symmetry of flows in the Earth's atmosphere overrides the symmetry about the Earth's axis of rotation, and so only the vertical component of the rotation vector is retained in the equations. This is a function of position, and so the conservation property and associated change of variables is lost. The variable rotation is fundamental in predicting the evolution of large-scale atmospheric disturbances (Rossby waves), so it is essential to show that

the semi-geostrophic equations can still be solved with variable rotation. Previous work on this problem has shown that there still appears to be enough structure in the equations for this to be possible. In particular, the solutions can be regarded as local energy minimisers, and satisfy a stability condition on the pressure which is equivalent to convexity of the potential in the uniform rotation case.

Very recent work described in [3] has exploited this structure to prove short time existence of smooth solutions. However, the procedure for advancing the solution in time loses spatial regularity, so a fixed point argument cannot be used to prove existence of a solution. Instead, for a short time while the Lagrangian particle positions can be smoothly mapped back to their initial positions, the implicit function theorem can be used to prove that a solution exists. The challenge is then to extend this result to long time existence of weak solutions. This is likely to require optimal transport techniques. Some ideas were mentioned in the lecture.

The second problem that I discussed is the partitioning of an atmospheric flow into a zonal mean and perturbations. This is important in studying the constraints on the evolution of large-scale anomalies which are responsible for abnormal weather types. A method for doing this has been developed by [7]. Recently, it has been shown that their technique can be written as an optimal transport problem, where the mass is calculated as a function of angular momentum and potential temperature, and then mapped back to physical space by minimising the energy. This procedure ensures hydrostatic balance and the gradient wind balance between zonal wind and the north-south pressure gradient. Optimal transport has been used successfully in constructing an axisymmetric vortex with a free boundary in an ambient rotating fluid at rest by [5] and [6]. I decsribed how their method could be extended to the problem solved by [7].

The third problem extends previous work by including the effects of moisture. Moisture interacts strongly with the dynamics through latent heat release when fluid parcels become saturated. This process also generates clouds and precipitation and can thus describe actual weather conditions. In semi-geostrophic dynamics, the large-scale flow represents a global energy minimising state. However, once moisture is included, this is no longer true, since moisture only interacts with the solution when saturation occurs. Thus a given state will only represent a local energy minimiser, and may lose stability as the flow evolves, If this happens, there may be a rapid evolution to a new stable state, in which large amounts of precipitation may be generated. Such events are responsible for much severe weather.

In [1] and [2] a one dimensional model was used to study this process. A stable state is represented by a column of potential temperature and moisture such that potential temperature is nondecreasing in height and the moisture is no greater than its saturation value. This column can be destabilised by moving it upwards in the atmosphere. In the simple model this was represented by imposing a time dependence on the saturation value of the moisture content. As a result, parcels may become supersaturated, and so increase their potential temperature through latent heat release. It was then shown that a well-defined evolution of a

discrete version of the problem could be found by moving supersaturated parcels to new stable positions higher up in the column and moving the others down to compensate. However, no parcel is allowed to rise past a parcel with higher potential temperature. In [1] it is rigorously proved that the solutions converge as the discretisation is refined, though the resulting solution can only be expressed as a probability measure. This is because, though the potential temperature is required to be nondecreasing, there is no control on the moisture distribution in a stable column. A key aspect of the proof is showing that the rate of upward mass transfer is bounded, even though parcels can move upwards a finite distance in infinitesimal time.

References

[1] Bin Cheng, Jingrui Cheng, Michael Cullen, John Norbury, and Matthew Turner *A rigorous treatment of moist convection in a single column.* SIAM J. Math. Anal., (2016) submitted. arXiv:1608.05822

[2] Bin Cheng, Michael Cullen, Gavin Esler, John Norbury, Matthew Turner, Jacques Vanneste and Jingrui Cheng, *A Model for Moist Convection in an Ascending Atmospheric Column.* Quart. J. Roy. Meteorol. Soc., (2016) submitted.

[3] Jingrui Cheng, Michael Cullen, and Mikhail Feldman *Semi-geostrophic system with variable Coriolis parameter.* Arch. Rat. Mech. Anal., (2016) submitted. arXiv:1608.05818

[4] M. Cullen. *A Mathematical Theory of Large-Scale Atmospheric Flow.* Imperial College Press. (2006) 259pp.

[5] M. Cullen and M. Sedjro. *On a model of forced axisymmetric flows.* SIAM Journal on Mathematical Analysis. **46-6** (2014), pp. 3983–4013, doi 10.1137/130944515.

[6] M. Cullen, W. Gangbo,and M. Sedjro. *A mathematically rigorous analysis of forced axisymmetric flows in the atmosphere.* Quart. J. Roy. Meteorol. Soc, **141**, (2015) 1836–1844.

[7] J. Methven and P. Berrisford. *The slowly evolving background state of the atmosphere.* Quart. J. Roy. Meteorol. Soc. **141**, (2015), 2237–2258.

Long time behavior of solutions to the 2D Keller-Segel equation with degenerate diffusion

Yao Yao

(joint work with José Carrillo, Katy Craig, Sabine Hittmeir, Inwon Kim, and Bruno Volzone)

In this talk I discussed the Keller-Segel equation with degenerate diffusion, which models the collective motion of cells attracted by a self-emitted chemical substance. It is given by

$$\rho_t = \Delta \rho^m - \nabla \cdot (\rho \nabla (\mathcal{N} * \rho)) \quad \text{in } \mathbb{R}^d \times [0, T), \tag{1}$$

where $m > 1$, $d \geq 2$, and $\mathcal{N}$ is the Newtonian potential in $\mathbb{R}^d$. Here the degenerate diffusion term $\Delta \rho^m$ models the anti-overcrowding effect of cells. It is well known that the behavior of solutions depend on the balance of the two terms on the right hand side. Namely, for $m > 2 - 2/d$ (the subcritical regime), the diffusion term dominates at high density, hence all solutions exist globally in time; whereas for

$1 \leq m < 2 - 2/d$ (the supercritical regime) the solution may blow-up in finite time [7, 1].

Although the blow-up/global existence criterion for (1) is well-studied, much less is known about the asymptotic behavior of solutions. In particular, in the subcritical regime $m > 2 - 2/d$, it was unknown whether solutions converge to a stationary solution, or can a part of mass dissipate to infinity as the time goes to infinity.

To understand the asymptotic behavior, it would be helpful to start with studying the existence and uniqueness of stationary solutions to (1) for any given mass. Formally, (1) can be understood as the gradient flow (in the space of measures endowed by 2-Wasserstein distance) of the free energy functional

$$E_m[\rho] = \frac{1}{m-1}\int \rho^m dx + \frac{1}{2}\rho(\rho * \mathcal{N})dx. \tag{2}$$

Using concentration-compactness arguments, Lions [6] showed that in the subcritical regime, (2) has a global minimizer for any given mass, and such global minimizer is automatically a stationary solution of (1). By Riesz rearrangement inequality, such global minimizer must be radially decreasing after a translation. Regarding uniqueness, for any given mass, Lieb and Yau [5] proved uniqueness of stationary solution among radial functions. However, it was unknown whether there can be non-radial stationary solutions.

In the joint work [2] with Carrillo, Hittmeir and Volzone, we prove that every $L^1 \cap L^\infty$ stationary solution to (1) must be radially symmetric. (Indeed, we show that this is also true if $\mathcal{N}$ is replaced by any attracting interaction kernel.) This is done by combining Steiner symmetrization techniques with some a priori regularity estimates on stationary solutions. Combining the radial symmetry result with [5], we obtain the uniqueness of stationary solutions of (1) (for any given mass) up to a translation.

Regarding the long time asymptotic behavior of solutions, we show that for the 2D Keller-Segel equation with degenerate diffusion $m > 1$, all solutions (whose initial data has finite second moment) must converge to a translation of the global minimizer as time goes to infinity. We show this by combining the uniqueness result with some a priori estimate on the growth of second moment (which is only valid in 2D). Since the convergence is done by a compactness argument, our argument does not give us any convergence rate as $t \to \infty$.

In the talk I also discussed a joint work with Craig and Kim [4], where we study the "$m = \infty$" limit of (1), and the equation formally becomes an aggregation equation with a density constraint. Here our solution is defined via a gradient flow, rather than solutions of a PDE. Namely, we consdier the gradient flow $\rho(\cdot, t)$ of the energy functional $E_\infty[\rho]$ in the space of probability measures endowed with 2-Wasserstein distance, where

$$E_\infty[\rho] := \begin{cases} \int \rho(\rho * \mathcal{N}) & \text{if } \|\rho\|_\infty \leq 1, \\ +\infty & \text{if } \|\rho\|_\infty > 1. \end{cases} \tag{3}$$

Note that (3) is the limit of (2) as $m \to \infty$. The existence, uniqueness and stability of such gradient flow solutions of (3) have been established by Craig [3].

In [4], we show that if the initial data is of "patch type" (i.e. $\rho_0 = 1_{\Omega_0}$), then the gradient flow solution $\rho(\cdot, t)$ remains as a patch for all times, with its boundary velocity determined by Hele-Shaw type free boundary problem. In addition, in two dimensions, we show that $\rho(\cdot, t)$ converges to the characteristic function of a disk 1_B (which is the global minimizer of (3)) as the time goes to infinity, where B is a disk with same area and center of mass as Ω_0. In addition, we show that the free energy has the convergence rate

$$0 \leq E_\infty[\rho(t)] - E_\infty[1_B] \leq C(\rho_0)t^{-1/6} \text{ for all } t \geq 0.$$

Both of the two asymptotic convergence results [2, 4] are in two dimensions only, since the exact formula of Newtonian potential in two dimensions plays an important role in controlling the second moment. For dimension $d > 2$, the asymptotic behavior of solutions to (1) is still an open problem.

References

[1] A. Blanchet, J. Carrillo and P. Laurençot, Critical mass for a Patlak-Keller-Segel model with degenerate diffusion in higher dimensions, *Calc. Var. Part. Diff. Eq.*, 35 (2009): 133–168.

[2] J. Carrillo, S. Hittmeir, B. Volzone and Y. Yao, Nonlinear aggregation-diffusion equations: radial symmetry and long time asymptotics, submitted, arXiv:1603.07767.

[3] K. Craig, Nonconvex gradient flow in the Wasserstein metric and applications to constrained nonlocal interactions, to appear in *Proc. London Math. Soc.*, arXiv:1512.07255.

[4] K. Craig, I. Kim and Y. Yao, Congested aggregation via Newtonian interaction, submitted, arXiv:1603.03790.

[5] E. H. Lieb and H. T. Yau, The Chandrasekhar theory of stellar collapse as the limit of quantum mechanics, *Comm. Math. Phys.*, 112 (1987): 147–174.

[6] P.-L. Lions, The concentration-compactness principle in the calculus of variations. The locally compact case. I, *Ann. Inst. H. Poincaré Anal. Non Linéaire*, 1 (1984): 109–145.

[7] Y. Sugiyama, Global existence in sub-critical cases and finite time blow-up in super-critical cases to degenerate Keller-Segel systems, *Diff. Int. Eqns.* 19 (2006): 841–876.

Matrix-valued Mass Transport: a Quantum Mechanical Approach

Tryphon T. Georgiou

(joint work with Yongxin Chen and Allen Tannenbaum)

We present a possible generalization of the Wasserstein 2-metric, originally defined on the space of scalar probability densities, to the space of Hermitian matrices and more generally to the space of matrix-valued distributions.

Our approach in [1, 2] was influenced by [3] and relies on a suitable continuity equation in the non-commutative spaces of matrices. It is substantially different from our earlier attempt in [4]. We invoke certain notions for the gradient and divergence operators on spaces of matrices that are explained below. Accordintly, the continuity equation in the space of Hermitian matrices $\mathcal{H}$ (of a given dimension) takes the familiar form

$$\dot{\rho} = \nabla_L^*(\rho \circ v) \tag{1}$$

in complete analogy with the continuity equation on scalar densities.

Here, $\rho(t) \in \mathcal{H}$ is positive-semidefinite while the velocity v is a vector of skew-Hermitian matrices (denoted by $\mathcal{S}$, of the same dimension as $\mathcal{H}$, i.e., $v \in \mathcal{S}^N$) while $\rho \circ v$ denotes one of several possible choices of "non-commutative" multiplication. We have considered specifically the following two choices:

$$\text{(i)}\ \ \rho \circ v = \frac{1}{2}(\rho v + v\rho) \text{ and (ii)}\ \ \rho \circ v = \int_0^1 \rho^s v \rho^{1-s} ds.$$

In the above, for $\rho \in \mathcal{H}$ and $v \in \mathcal{S}^N$,

$$v\rho := \begin{bmatrix} v_1\rho \\ \vdots \\ v_N\rho \end{bmatrix}, \text{ and } \rho v := \begin{bmatrix} \rho v_1 \\ \vdots \\ \rho v_N \end{bmatrix}.$$

The gradient operator with respect to $L \in \mathcal{H}^N$ is defined as

$$\nabla_L : \mathcal{H} \to \mathcal{S}^N,\ X \mapsto \begin{bmatrix} L_1X - XL_1 \\ \vdots \\ L_NX - XL_N \end{bmatrix}$$

and, accordingly, with respect to the standard inner product $\langle X, Y\rangle = \text{trace}(X^*Y)$, the divergence operator is

$$\nabla_L^* : \mathcal{S}^N \to \mathcal{H},\ Y = \begin{bmatrix} Y_1 \\ \vdots \\ Y_N \end{bmatrix} \mapsto \sum_k^N L_kY_k - Y_kL_k.$$

A Wasserstein distance can now be defined as the solution to a Benamou-Brenier *minimum-action-integral* problem [5]:

$$W_{2,a}(\rho_0, \rho_1)^2 := \min_{\rho, v} \int_0^1 \langle v, \rho \circ v\rangle dt, \tag{2}$$
$$\dot\rho = \frac{1}{2}\nabla_L^*(\rho \circ v),$$
$$\rho(0) = \rho_0,\ \rho(1) = \rho_1.$$

In this, ρ_0 and ρ_1 are assumed positive definite having trace 1. The optimization is over $\rho(\cdot)$ with similar properties and over $v \in \mathcal{S}^N$.

The matrix-continuity equation (1) preserves positive definiteness and trace. Suitable generalizations for flows between boundary values of unequal trace as well as for flows between matrix-valued distributions were considered in [1, 6]. Transport along the spatial coordinate $x \in \mathbb{R}^m$ is effected by a term $\nabla_x \cdot (\rho \circ w)$ in the continuity equation, with $w \in \mathcal{H}^m$, i.e.,

$$\dot\rho = \nabla_L^*(\rho \circ v) - \nabla_x \cdot (\rho \circ w)$$

Likewise, variations in the trace are effected through an added term in the continuity equation which is duly penalized in the corresponding optimization problem, see [6].

For the choice $\rho \circ v = \frac{1}{2}(\rho v + v\rho)$, the computation of the Wasserstein distance is feasible from a numerical standpoint, hence the interest in this choice in our work [1]. This is accomplished by turning (2) into a convex optimization problem in a new set of variables (ρ, u) where $u = \rho v$, i.e., "mass + momentum" instead of "mass + velocity," following [5].

However, on the other hand, the choice $\rho \circ v = \int_0^1 \rho^s v \rho^{1-s} ds$ in the continuity equation leads to the rather remarkable result. *The gradient flow of the von Neumann entropy*

$$S(\rho) := -\operatorname{trace}(\rho \log \rho)$$

is precisely the Lindblad equation of open quantum systems. In turn, in light of the above expressions for divergence and gradient operators, the *Lindblad equation* can be conveniently written in the form

$$\dot{\rho} = -\nabla_L^* \nabla_L \rho =: \Delta_L \rho,$$

reminiscent of the *heat equation.* This result generalizes to the matrix case the well known discovery of [7] that for *scalar* densities the gradient flow of the entropy is precisely the heat equation. The matrix-valued generalization was derived in [1] by making use of a key identity in [3] that relates the gradient of ρ to the gradient of the logarithm of ρ. At about the same time as our work was being reported in [1], two other closely related approaches had been formulated independently and simultaneously by Carlen and Maas [8] and by Mittnenzweig and Mielke [9], who derived this same result as well.

Our interest in this non-commutative counterpart of optimal transport, where density matrices ρ (i.e., Hermitian matrices that are positive-definite and have unit trace) replace probability distributions, stems from control and signal processing applications. Such applications include diffusion tensor imaging (DTI), in which one wants to interpolate spatially varying tensor fields, and spectral analysis of multivariable time series, in which one may need to track changes in matrix-valued power spectral distributions [4]. Other possible generalizations of the theory of transport are motivated by corresponding flows on discrete spaces that include graphs and networks, see [10, 11].

References

[1] Y. Chen, T.T. Georgiou, and A. Tannenbaum, "Matrix optimal mass transport: a quantum mechanical approach," https://arxiv.org/abs/1610.03041 (2016).

[2] Y. Chen, W. Gangbo, T.T. Georgiou, and A. Tannenbaum, "On the Matrix Monge-Kantorovich Problem," https://arxiv.org/abs/1701.02826

[3] E. Carlen and J. Maas, "An analog of the 2-Wasserstein metric in non-commutative probability under which the fermionic Fokker-Planck equation is gradient flow for the entropy," *Commun. Math. Phys.* **331:**, 887-926 (2014), https://arxiv.org/abs/1203.5377

[4] L. Ning, T. Georgiou, and A. Tannenbaum, "On matrix–valued Monge-Kantorovich optimal mass transport," *IEEE Transactions on Automatic Control* **60(2):** 373-382 (2015), https://arxiv.org/abs/1304.3931

[5] J.-D. Benamou and Y. Brenier, "A computational fluid mechanics solution to the Monge–Kantorovich mass transfer problem," *Numerische Mathematik* **84**: 375-393 (2000).

[6] Y. Chen, T.T. Georgiou, and A. Tannenbaum, "Interpolation of Density Matrices and Matrix-Valued Measures: The Unbalanced Case," https://arxiv.org/abs/1612.05914.
[7] R. Jordan, D. Kinderlehrer, and F. Otto, "The variational formulation of the Fokker-Planck equation," *SIAM J. Math. Anal.* **29** (1998), pp. 1-17.
[8] E. Carlen and J. Maas, "Gradient flow and entropy inequalities for quantum Markov semigroups with detailed balance," https://arxiv.org/abs/1609.01254, 2016.
[9] M. Mittnenzweig and A. Mielke, "An entropic gradient structure for Lindblad equations and GENERIC for quantum systems coupled to macroscopic models," https://arxiv.org/abs/1609.05765, 2016.
[10] Y. Chen, T.T. Georgiou, and A. Tannenbaum, "Vector-Valued Optimal Mass Transport," https://arxiv.org/pdf/1611.09946.pdf
[11] Y. Chen, T.T. Georgiou, M. Pavon, and A. Tannenbaum, "Thermodynamics of efficient-robust transport over networks," https://arxiv.org/pdf/1701.07625.pdf

Multi- to one-dimensional optimal transport

Brendan Pass
(joint work with Pierre-André Chiappori, Robert McCann)

Given probability measures μ on $X \subseteq \mathbb{R}^m$ and ν on $Y \subseteq \mathbb{R}^n$, and a surplus function $s : X \times Y \to \mathbb{R}$, consider the Monge-Kantorovich problem of maximizing

$$\int_{X\times Y} s(x,y)d\gamma(x,y)$$

among probability measures γ on $X \times Y$ whose marginals are μ and ν. Most of the research on this problem to date has focused on the case where the dimensions m of the source and n of the target are equal. This progress is reviewed in detail in, for example, [2], but let us recall here that, under mild conditions on s, μ and ν, there is a unique solution γ, which is concentrated on the graph of a function $T : X \to Y$. In addition, under much stronger conditions, recent years have seen the development of a deep regularity theory for the optimal map T. We also note that in the simplest case $m = n = 1$, it is possible to solve the problem explicitly; for a supermodular s, T is the unique monotone increasing function pushing μ forward to ν.

On the other hand, problems where the dimensions of the source and target differ have not received much attention in the literature; the existence, uniqueness and graphical structure of the optimizer can be established exactly as in the equal dimensional setting, but little else is known. However, these problems are very relevant to applications in economic theory; in particular, the problem of matching distributions of, say, women and men on the marriage market under transferable utility can be recast as an optimal transport problem and there is no reason that the dimensions (which represent the number of characteristics used to distinguish between agents on different sides of the market) should coincide in these models.

In this abstract, we report on recent work on the case where $m > n = 1$ [1]. We introduce a *nestedness* condition under which it is possible to solve this problem uniquely, and essentially explicitly. This condition is somewhat unusual in optimal transport theory in that it depends on the *interaction* of the surplus and the

marginals in a subtle way. The basic idea is, for each $y \in \mathbb{R}$, to choose the $k = k(y)$ which *splits the population proportionately*; that is, $\mu(X_{\leq}(y,k)) = \nu(-\infty, y)$, where $X_{\leq}(y,k) := \{x \in X : \frac{\partial s}{\partial y}(x,y) \leq k\}$. We say the model (s, μ, ν) is *nested* if these sublevel sets evolve monotonically with y: $X_{\leq}(y_0, k(y_0)) \subseteq X_{<}(y_1, k(y_1))$ whenever $y_0 < y_1$ ($X_{<}(y,k)$ is defined analogously to $X_{\leq}(y,k)$). Under this condition, we show that the map defined by sending each point x in $X_{=}(y, k(y)) := \{x \in X : \frac{\partial s}{\partial y}(x,y) = k(y)\}$ to y is well defined and optimal.

We go on to establish a number of equivalent conditions for nestedness; it is equivalent to the uniqueness of the population splitting $X_{=}(y, k(y))$ passing through each x (existence of such a y for each x is straightforward to establish in much greater generality). This asserts that nestedness is essentially a sharp condition for the construction of the optimal map outlined above to be well defined; if it failed, there would exists some x sitting in two distinct population splitting hypersurfaces $X_{=}(y_0, k(y_0))$ and $X_{=}(y_1, k(y_1))$, in which case our procedure would ambiguously assign both $T(x) = y_0$ and $T(x) = y_1$.

On the other hand, the outward unit normal speed of $X_{\leq}(y, k(y))$ as y varies is $(k'(y) - s_{yy}(y,x))/|D_x s_y|$, and one can prove that nestedness implies that this must always be nonnegative, and strictly positive for at least one x in each $X_{\leq}(y, k(y))$. Conversely, strict positivity everywhere is sufficient (but not necessary) to imply nestedness. We also establish the integral-differential equation for $k(y)$,

$$\int_{X(y,k(y))} \frac{k'(y) - s_{yy}(x,y)}{|D_x s_y|} f(x), d\mathcal{H}^{m-1}(x)$$

expressing the mass balance condition.

Finally, we address regularity. In the nested case, we first show that the optimal map is continuous. Higher order regularity of the Kantorovich potential $v(y) = k'(y)$ on the uni-dimensional side of the market can be established by exploiting level set dynamics, under natural conditions on the data. Under the strengthened variant of nestedness, $(k'(y) - s_{yy}(y,x))/|D_x s_y| > 0$, we can use this result to establish higher regularity of the optimal map and the other Kantorovich potential $u(x) : \max_{y \in Y} \big(s(x,y) - v(y)\big)$. It is worth noting that one cannot expect a general regularity theory in this setting; an earlier result of the present author asserts that one cannot have regularity for all (nice) marginals μ and ν unless s is of pseudo-index form, $s(x,y) = b(I(x), y) + \alpha(x)$, where $I : X \to \mathbb{R}$ and $b : I(X) \times Y \to \mathbb{R}$, in which case the problem reduces to a strictly uni-dimensional optimal transport problem with surplus b. In a follow up with R. McCann to the present work, we show that, under certain topological conditions on s, X and Y, if both Kantorovich potentials are C^2, the model *must* be nested, indicating that nestedness is a natural (and necessary) hypothesis for higher regularity.

References

[1] P.-A. Chiappori, R.J. McCann and B. Pass, *Multi- to one-dimensional optimal transport*, To appear in Comm. Pure Appl. Math.

[2] C. Villani, *Optimal transport: old and new*, volume 338 of Grundlehren der mathematischen Wissenschaften. Springer, New York, 2009.

Reporter: Virginie Ehrlacher

Participants

Dr. Jean-David Benamou
INRIA Rocquencourt
Domaine de Voluceau
B. P. 105
78153 Le Chesnay Cedex
FRANCE

Prof. Dr. Adrien Blanchet
Toulouse School of Economics
21, Allee de Brienne
31000 Toulouse Cedex
FRANCE

Prof. Dr. Yann Brenier
C M L S
École Polytechnique
Plateau de Palaiseau
91128 Palaiseau Cedex
FRANCE

Prof. Dr. Martin Burger
Institut für Numerische und
Angewandte Mathematik
Universität Münster
Einsteinstrasse 62
48149 Münster
GERMANY

Prof. Dr. Giuseppe Buttazzo
Dipartimento di Matematica "L.Tonelli"
Universita di Pisa
Largo Bruno Pontecorvo, 5
56127 Pisa
ITALY

Dr. Clément Cancès
INRIA Lille-Nord Europe
Parc Scientifique de la Haute Borne
Bat A.
40, avenue Halley
59650 Villeneuve-d'Ascq
FRANCE

Prof. Dr. Eric A. Carlen
Department of Mathematics
Rutgers University
Hill Center, Busch Campus
110 Frelinghuysen Road
Piscataway, NJ 08854-8019
UNITED STATES

Prof. Dr. Guillaume Carlier
CEREMADE
Université Paris Dauphine
Place du Marechal de Lattre de Tassigny
75775 Paris Cedex 16
FRANCE

Prof. Dr. Jose Antonio Carrillo de la Plata
Department of Mathematics
Imperial College London
Huxley Building
London SW7 2AZ
UNITED KINGDOM

Prof. Dr. Mike Cullen
Met Office
Fitzroy Road
Exeter EX1 3PB
UNITED KINGDOM

Prof. Dr. Laurent Desvillettes
CMLA
ENS Cachan
61, Avenue du President Wilson
94235 Cachan Cedex
FRANCE

Prof. Dr. Jean Dolbeault
CEREMADE
Université Paris Dauphine
Place du Marechal de Lattre de Tassigny
75775 Paris Cedex 16
FRANCE

Dr. Virginie Ehrlacher
CERMICS - ENPC
Cite Descartes, Champs-sur-Marne
6 et 8 avenue Blaise Pascal
77455 Marne-la-Vallée Cedex 2
FRANCE

Dr. Matthias Erbar
Mathematisches Institut
Universität Bonn
Endenicher Allee 60
53115 Bonn
GERMANY

Max Fathi
Department of Mathematics
University of California, Berkeley
970 Evans Hall
Berkeley CA 94720-3840
UNITED STATES

Prof. Dr. Klemens Fellner
Institut für Mathematik
Karl-Franzens-Universität Graz
Heinrichstrasse 36
8010 Graz
AUSTRIA

Prof. Dr. Gero Friesecke
Fakultät für Mathematik
Technische Universität München
Boltzmannstrasse 3
85748 Garching bei München
GERMANY

Prof. Dr. Brittany D. Froese
Department of Mathematics
New Jersey Institute of Technology
606 Cullimore Hall
Newark, NJ 07102-1982
UNITED STATES

Prof. Dr. Tryphon T. Georgiou
Department of Mathematics
University of California, Davis
Davis, CA 95616-8633
UNITED STATES

Prof. Dr. Paola Gori-Giorgi
Department of Chemistry
Vrije University
De Boelelaan 1081
1081 HV Amsterdam
NETHERLANDS

Prof. Dr. Ansgar Jüngel
Institut für Analysis & Scientific Computing
Technische Universität Wien
Wiedner Hauptstrasse 8 - 10
1040 Wien
AUSTRIA

Prof. Dr. Inwon C. Kim
Department of Mathematics
University of California at Los Angeles
405 Hilgard Avenue
Los Angeles, CA 90095-1555
UNITED STATES

Prof. Dr. David Kinderlehrer
Department of Mathematical Sciences
Carnegie Mellon University
Pittsburgh, PA 15213-3890
UNITED STATES

Eva Kopfer
Institut für Angewandte Mathematik
Universität Bonn
Endenicher Allee 60
53115 Bonn
GERMANY

Prof. Dr. Mathieu Lewin
CEREMADE
Université Paris Dauphine
Place du Marechal de Lattre de Tassigny
75775 Paris Cedex 16
FRANCE

Prof. Dr. Daniel Matthes
Zentrum Mathematik
Technische Universität München
Boltzmannstrasse 3
85747 Garching bei München
GERMANY

Prof. Dr. Robert J. McCann
Department of Mathematics
University of Toronto
40 St George Street
Toronto ON M5S 2E4
CANADA

Dr. Quentin Mérigot
CEREMADE
Université Paris Dauphine and CNRS
Place du Marechal de Lattre de Tassigny
75775 Paris Cedex 16
FRANCE

Markus Mittnenzweig
Weierstraß-Institut für
Angewandte Analysis und Stochastik
Mohrenstrasse 39
10117 Berlin
GERMANY

Prof. Dr. Adam Oberman
Department of Mathematics and
Statistics
McGill University
805, Sherbrooke Street West
Montreal QC H3A 0G4
CANADA

Dr. Carlo Orrieri
Dipartimento di Matematica
Universita di Roma "La Sapienza"
Piazzale Aldo Moro, 5
00195 Roma
ITALY

Prof. Dr. Felix Otto
Max-Planck-Institut für Mathematik
in den Naturwissenschaften
Inselstrasse 22 - 26
04103 Leipzig
GERMANY

Prof. Dr. Brendan Pass
Department of Mathematical Sciences
University of Alberta
Edmonton AB T6G 2R3
CANADA

Dr. Francesco S. Patacchini
Department of Mathematics
Imperial College London
Huxley Building
180 Queen's Gate
London SW7 2AZ
UNITED KINGDOM

Prof. Dr. Mark A. Peletier
Department of Mathematics
Eindhoven University of Technology
P.O.Box 513
5600 MB Eindhoven
NETHERLANDS

Prof. Dr. Gabriel Peyré
CEREMADE
Université Paris Dauphine
Place du Marechal de Lattre de Tassigny
75016 Paris Cedex
FRANCE

Simon Plazotta
Zentrum Mathematik
Technische Universität München
Boltzmannstrasse 3
85747 Garching bei München
GERMANY

Prof. Dr. Aldo Pratelli
Department Mathematik
FAU Erlangen-Nürnberg
Cauerstraße 11
91058 Erlangen
GERMANY

Dr. Gael Raoul
Departement de Mathématiques et Applications
École Normale Superieure
45, rue d'Ulm
75230 Paris Cedex 05
FRANCE

Prof. Dr. Martin Rumpf
Institut für Numerische Simulation
Universität Bonn
Endenicher Allee 60
53115 Bonn
GERMANY

Prof. Dr. Filippo Santambrogio
Laboratoire de Mathématiques d'Orsay
Université Paris-Sud
Centre d'Orsay
Batiment 425
91405 Orsay Cedex
FRANCE

Prof. Giuseppe Savare
Dipartimento di Matematica
"F. Casorati"
Universita di Pavia
Via Ferrata, 1
27100 Pavia
ITALY

Prof. Dr. Carola-Bibiane Schoenlieb
Department of Applied Mathematics and Theoretical Physics (DAMTP)
Centre for Mathematical Sciences
Wilberforce Road
Cambridge CB3 OWA
UNITED KINGDOM

Dr. Christian Seis
Institut für Angewandte Mathematik
Universität Bonn
Endenicher Allee 60
53115 Bonn
GERMANY

Prof. Dr. Dejan Slepcev
Department of Mathematical Sciences
Carnegie Mellon University
5000 Forbes Avenue
Pittsburgh, PA 15213-3890
UNITED STATES

Prof. Dr. Karl-Theodor Sturm
Institut für Angewandte Mathematik
Universität Bonn
Endenicher Allee 60
53115 Bonn
GERMANY

Dr. Quoc Bao Tang
Institut für Mathematik
Karl-Franzens-Universität Graz
Heinrichstrasse 36
8010 Graz
AUSTRIA

Prof. Dr. Max von Renesse
Mathematisches Institut
Universität Leipzig
Augustusplatz 10
04109 Leipzig
GERMANY

Dr. Irène Waldspurger
Departement de Mathématiques et Applications
École Normale Superieure
45, rue d'Ulm
75005 Paris Cedex
FRANCE

Dr. Gershon Wolansky
Department of Mathematics
Technion
Israel Institute of Technology
Haifa 32000
ISRAEL

Prof. Dr. Yao Yao
School of Mathematics
Georgia Institute of Technology
686 Cherry Street
Atlanta, GA 30332-0160
UNITED STATES

Prof. Dr. Johannes Zimmer
Department of Mathematical Sciences
University of Bath
Bath BA2 7AY
UNITED KINGDOM