Mathematisches Forschungsinstitut Oberwolfach

Report No. 54/2017

DOI: 10.4171/OWR/2017/54

Variational Methods for Evolution

Organised by
Alexander Mielke, Berlin
Mark Peletier, Eindhoven
Dejan Slepcev, Pittsburgh

12 November – 18 November 2017

Abstract. Many evolutionary systems, as for example gradient flows or Hamiltonian systems, can be formulated in terms of variational principles or can be approximated using time-incremental minimization. Hence they can be studied using the mathematical techniques of the field of calculus of variations. This viewpoint has led to many discoveries and rapid expansion of the field over the last two decades. Relevant applications arise in mechanics of fluids and solids, in reaction-diffusion systems, in biology, in many-particle models, as well as in emerging uses in data science.

This workshop brought together a broad spectrum of researchers from calculus of variations, partial differential equations, metric geometry, and stochastics, as well as applied and computational scientists to discuss and exchange ideas. It focused on variational tools such as minimizing movement schemes, Gamma convergence, optimal transport, gradient flows, and large-deviation principles for time-continuous Markov processes.

Mathematics Subject Classification (2010): 49-06 (Calc. of Var.), 35-05 (PDEs),70-06 (mechanics of particles and systems), 58E30 (Variational principles), 60F10 (Large deviations), 82C05 (Classical dynamic and nonequilibrium statistical mechanics (general)).

Introduction by the Organisers

Variational approaches to evolution systems provide a rich and very active field of mathematical research combining several previously only loosely connected branches of mathematics. This includes studies of geometric aspects like metric structures in infinite dimensional function spaces and of the geometry of relevant energy landscapes, like their geodesic lambda-convexity. The connection with theory of optimal transport provides new insights into metric spaces on the one hand

and provides new tools for the theory of partial differential equations on the other hand. Another surprising connection links analysis and stochastics to provide new gradient structures for macroscopic partial differential equations arising from many-particle systems via large-deviation principles.

During the last few decades many evolutionary PDEs for models in mechanics, physics, chemistry, and biology have been studied via new approaches based on new variational techniques, combining well established tools and new ideas in clever ways. This includes classical and generalized gradient flows including rate-independent flows, geometric analysis for fluid and transport dynamics, Hamiltonian systems, and stochastic dynamics of many-particle systems. Concrete examples include the Fokker–Planck equation, porous medium equations, microfluidic systems and thin-film equations, interface evolutions, pattern formation and evolution, coarsening, micromagnetics, superconductors, materials science (crack propagation, behavior of material defects, epitaxial growth, grain boundary evolution), biological aggregation, many particle systems with interactions and randomness, and geometric flows.

The aim of this workshop was to bring together a group of experts and young researchers from calculus of variations, partial differential equations, non-smooth geometry, stochastic analysis, as well as applied and computational scientists for a stimulating interchange of ideas. This has succeeded remarkably well, and among the themes presented during the workshop, we mention here:

- optimal transport techniques and transportation distances, functional inequalities, entropic interpolation (many speakers);
- connections between gradient flows and other solution concepts, such as Brakke flow for mean curvature (Otto) and conservation laws (Brenier);
- connections between gradient flows, statistical mechanics, large deviations, and thermodynamics (Öttinger, Léonard, Maes, Zimmer);
- novel numerical methods for gradient flows and other variational evolutionary systems (Matthes, Pattacchini, Knees);
- new variational formulations (Stefanelli, Mittnenzweig, Dal Maso, Erbar, Monsaingeon), and new analysis of existing formulations (Savaré);
- reaction-diffusion problems as variational evolution (Liero);
- application of variational concepts for data analysis (García Trillos);
- entropy-entropy dissipation methods (Jüngel);
- crystallization, oscillation, and pattern formation (Niethammer, Maes);
- discrete interaction systems, evolution on graphs, and their metric and variational interpretation (Mittnenzweig, Niethammer, Erbar);
- rate-independent systems, quasi-static crack growth, and elasto-plasticity (Chambolle, Dal Maso);
- front propagation and phase-field models (Cancés, Otto).

In total, there were 17 talks of 45 minutes and 8 talks of 30 minutes leaving plenty of time for discussions, which have been greatly stimulated by the diversity

of the topics and of the contributions. As always, the friendly atmosphere and the perfect environment of Oberwolfach have also contributed to the success of the meeting in a major way.

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-1641185, "US Junior Oberwolfach Fellows". Moreover, the MFO and the workshop organizers would like to thank the Simons Foundation for supporting Jin Feng in the "Simons Visiting Professors" program at the MFO.

Workshop: Variational Methods for Evolution

Table of Contents

Abstracts

Solution by convex optimization of the Cauchy problem for hyperbolic entropic conservation laws

Yann Brenier

We want to solve by convex optimization the Cauchy problem for the class of entropic conservation laws, namely for systems of evolution PDEs of form

$$\partial_t U^\alpha + \partial_i(\mathcal{F}^{i\alpha}(U)) = 0, \quad \alpha = 1, \cdots, m, \tag{1}$$

(with implicit summation on repeated indices), where $U = U(t,x) \in \mathcal{W} \subset \mathbb{R}^m$, $t \in [0,T]$, $x \in \mathcal{D} = (\mathbb{R}/\mathbb{Z})^d$, $\partial_t = \frac{\partial}{\partial t}$, $\partial_i = \frac{\partial}{\partial x_i}$, $\mathcal{W}$ is a smooth convex open set, while the so-called "flux" $\mathcal{F} : \mathcal{W} \to \mathbb{R}^{d\times m}$ is a smooth function enjoying the symmetry property

$$\forall i \in \{1, \cdots, d\}, \ \ \forall \beta, \gamma \in \{1, \cdots, m\}, \ \ \partial^2_{\alpha\beta}\mathcal{E}\partial_\gamma \mathcal{F}^{i\alpha} = \partial^2_{\alpha\gamma}\mathcal{E}\partial_\beta \mathcal{F}^{i\alpha}, \tag{2}$$

for some smooth function, called "entropy", $\mathcal{E} : \mathcal{W} \to \mathbb{R}$, which is supposed to be convex in the strong sense that $(\partial^2_{\alpha\beta}\mathcal{E})$ is a positive definite matrix, everywhere on $\mathcal{W}$. The symmetry condition (2) enforces the conservation of entropy, in the sense that every smooth solution U to (1) must satisfy the extra-conservation law

$$\partial_t(\mathcal{E}(U)) + \partial_i(\mathcal{Q}^i(U)) = 0, \tag{3}$$

where the "entropy-flux" function $\mathcal{Q} : \mathcal{W} \to \mathbb{R}^d$ depends on $\mathcal{F}$ and $\mathcal{E}$.
This class of PDEs is of paramount importance in continuum mechanics and material sciences (Hydrodynamics, Elastodynamics, Magnetohydrodynamics, etc...[4]). The simplest example, is the so-called "inviscid" Burgers equation

$$\partial_t u + \partial_x(\frac{u^2}{2}) = 0, \quad u \in \mathbb{R}, \ \ \mathcal{E}(u) = \frac{u^2}{2}, \tag{4}$$

A richer example is the Euler equation of isothermal fluids:

$$\partial_t \rho + \nabla\cdot q = 0, \ \ \partial_t q + \nabla\cdot(\frac{q\otimes q}{\rho}) + \nabla\rho = 0, \ \ \mathcal{E} = \frac{|q|^2}{2\rho} + \rho\log\rho, \quad \rho > 0, \ \ q \in \mathbb{R}^d. \tag{5}$$

Under mild additional conditions, PDEs of that class are of hyperbolic type and are (locally) well-posed [typically in Sobolev spaces H^s for $s > d/2+1$ [4]]. [Nevertheless, in most cases, smooth solutions develop singularities, called "shock waves", in finite time, and they cease to be continuous.] Thus, our goal looks hopeless. Indeed, space-time convex minimization problems presumably lead to optimality equations of space-time *elliptic* type, in contradiction with the hyperbolic character of the equations we are interested in. However, we are able to find such a convex optimization problem with the following properties:
i) For general entropic systems of conservation laws, it is always possible to recover a smooth solution by convex optimization on a sufficiently small time interval;
ii) in the special and very simple case of the so-called "inviscid" Burgers equation, any entropy solution (à la Kruzhkov-Panov) can be recovered, at any time, by convex optimization.

Here is the way to find this convex optimization problem. We first consider the concept of weak solutions U, for which the Cauchy problem with initial condition U_0 means

$$\int_{[0,T]\times\mathcal{D}} \partial_t W_\alpha U^\alpha + \partial_i W_\alpha F^{i\alpha}(U) + \int_{\mathcal{D}} W_\alpha(0,\cdot)U_0^\alpha = 0, \tag{6}$$

for all smooth functions $W = W(t,x) \in \mathbb{R}^m$ such that $W(T,\cdot) = 0$. Weak solutions are not expected to satisfy the extra-conservation law (3). In particular

$$t \in [0,T] \to \int_{\mathcal{D}} \mathcal{E}(U(t,\cdot))$$

is not expected to be constant in t. The concept of weak solutions is quite faulty since, for a fixed initial condition U_0, weak solutions may not be unique and the conservation of entropy is generally not true. (This is now well established in the case of the Euler equations, in Hydrodynamics, through the results of Scheffer, Shnirelman, De Lellis-Székelyhidi Jr. [7, 8, 5, 10].) So, we look, given an initial condition U_0, for weak solutions that minimize over $[0,T]$ the time integral of their entropy. As just explained, this problem is not void since weak solutions may not be unique and do not conserve their entropy. Using the trial functions W in (6) as Lagrange multipliers, we get the following min-max problem

$$I = \inf_U \sup_W \int_{[0,T]\times\mathcal{D}} E(U) - \partial_t W_\alpha U^\alpha - \partial_i W_\alpha F^{\alpha i}(U) - \int_{\mathcal{D}} W_\alpha(0,\cdot)U_0^\alpha, \tag{7}$$

where $W = W(t,x) \in \mathbb{R}^m$ are smooth functions, vanishing at $t = T$. This indeed amounts to looking for a weak solution U of our system of conservation laws with initial condition U_0 that minimizes the time integral of its entropy. Let us now exchange the infimum and the supremum in the definition of I and get the lower bound

$$J = \sup_W \inf_U \int_{[0,T]\times\mathcal{D}} E(U) - \partial_t W_\alpha U^\alpha - \partial_i W_\alpha F^{i\alpha}(U) - \int_{\mathcal{D}} W_\alpha(0,\cdot)U_0^\alpha \tag{8}$$

which can be reduced to the concave maximization problem

$$J = \sup_W \int_{[0,T]\times\mathcal{D}} -K(\partial_t W, DW) - \int_{\mathcal{D}} W_\alpha(0,\cdot)U_0^\alpha \tag{9}$$

where W is still subject to $W(T,\cdot) = 0$ and K is the *convex* function defined by

$$K(A,B) = \sup_{V\in\mathcal{W}} A_\alpha V^\alpha + B_{\alpha i} F^{i\alpha}(V) - \mathcal{E}(V), \quad A \in \mathbb{R}^m, \quad B \in \mathbb{R}^{m\times d}. \tag{10}$$

This concave maximization problem is very similar to the Monge optimal mass transport problem with quadratic cost in its so-called "Benamou-Brenier" formulation [2, 1, 9]. Its numerical treatment in the style of [2] is currently under investigation.

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] J.-D. Benamou, Y. Brenier, *A Computational Fluid Mechanics solution to the Monge-Kantorovich mass transfer problem,* Numer. Math. 84 (2000) 375-393.
[3] Y. Brenier, *The initial value problem for the Euler equations of incompressible fluids viewed as a concave maximization problem ,* arXiv:1706.04180.
[4] C. Dafermos, *Hyperbolic conservation laws in continuum physics,* Springer, 2000.
[5] C. De Lellis, L. Székelyhidi Jr. *The Euler equations as a differential inclusion,* Ann. of Math. (2) 170 (2009) 1417-1436.
[6] P.-L. Lions, *Mathematical topics in fluid mechanics. Vol. 1. Incompressible models,* Oxford 1996.
[7] V. Scheffer, *An inviscid flow with compact support in space-time,* J. Geom. Anal. 3 (1993) 343-401.
[8] A. Shnirelman, *On the nonuniqueness of weak solution of the Euler equation,* Comm. Pure Appl. Math., 50 (1997) 1261-1286,
[9] C. Villani, *Topics in Optimal Transportation,* Grad. Stud. Math., vol. 58, AMS 2003.
[10] E. Wiedemann *Existence of weak solutions for the incompressible Euler equations,* Ann. Inst. H. Poincaré Anal. Non Linéaire, 28 (2011) 727-730.

A degenerate Cahn-Hilliard model as constrained gradient flow

Clément Cancès

(joint work with D. Matthes and F. Nabet)

We consider a mixture of two incompressible and immiscible fluids in a convex domain Ω of $\mathbb{R}^d$ for $d \leq 3$. The fluid composition is described by the phase concentrations $\boldsymbol{c} = (c_1, c_2)$ that are subject to the constraint

$$c_1 + c_2 = 1. \tag{1}$$

We associate an energy $\mathcal{E}(\boldsymbol{c})$ to a configuration $\boldsymbol{c}$. It is defined by

$$\mathcal{E}(\boldsymbol{c}) = \int_\Omega \left[\frac{\epsilon}{2} |\nabla c_1|^2 + \frac{1}{\epsilon} c_1 c_2 + \chi(\boldsymbol{c}) \right] \mathrm{d}\boldsymbol{x}.$$

In the above formula, χ enforces the constraint and writes

$$\chi(\boldsymbol{c}) = \begin{cases} 0 & \text{if } c_1 + c_2 = 1, \\ +\infty & \text{otherwise.} \end{cases}$$

Each phase is convected by its own speed $\boldsymbol{v}_i$, i.e.,

$$\partial_t c_i + \operatorname{div}(c_i \boldsymbol{v}_i) = 0,$$

with no-flux across $\partial\Omega$. To velocity fields $\boldsymbol{V} = (\boldsymbol{v_1}, \boldsymbol{v_2})$, we associate the dissipation

$$\mathcal{D}(\boldsymbol{c}, \boldsymbol{V}) = \sum_{i \in \{1,2\}} \frac{1}{2} \int_\Omega c_i |\boldsymbol{v}|^2 \mathrm{d}\boldsymbol{x}.$$

Then the phase speeds $\boldsymbol{v_i}$ are chosen following the following steepest descent condition [8, 10]:

$$\boldsymbol{V} \in \operatorname{Argmin}\left(\mathcal{D}(\boldsymbol{c},\boldsymbol{V}) + \max_{\boldsymbol{\mu}\in\partial\mathcal{E}(\boldsymbol{c})} \sum_{i\in\{1,2\}} \int_\Omega c_i \boldsymbol{v}_i \cdot \nabla\mu_i\right). \tag{2}$$

Since $\boldsymbol{\mu} = (\mu_1, \mu_2)$ belongs to $\partial\mathcal{E}(\boldsymbol{c})$ iff $\boldsymbol{c}$ satisfies the constraint (1) and

$$\mu_{\mathrm{g}} := \mu_1 - \mu_2 = -\epsilon\Delta c_1 + \frac{1}{\epsilon}c_1(1-c_1)$$

where μ_{g} is called the generalized chemical potential, the steepest descent condition (2) leads to the system

$$\begin{cases} \partial_t c_i - \operatorname{div}(c_i\nabla\mu_i) = 0, \\ c_1 + c_2 = 1, \\ \mu_1 - \mu_2 = -\epsilon\Delta c_1 + \frac{1}{\epsilon}c_1(1-c_1), \\ \int_\Omega (c_1\mu_1 + c_2\mu_2)\mathrm{d}\boldsymbol{x} = 0. \end{cases} \tag{3}$$

The third equation of the above system has to be complemented with homogeneous Neumann boundary conditions on $\boldsymbol{c}$, while the last equation is just there to normalize the chemical potentials that would otherwise be defined up to an additive constant.

This model coincides with the model derived by E and Palffy-Muhoray in [5], but differs from the classical degenerate Cahn-Hilliard model proposed by de Gennes in [3] that we recall now:

$$\begin{cases} \partial_t c - \operatorname{div}(c(1-c)\nabla\mu_{\mathrm{g}}) = 0, \\ \mu_{\mathrm{g}} = -\epsilon\Delta c + \frac{1}{\epsilon}c(1-c), \\ \int_\Omega \mu_{\mathrm{g}}\mathrm{d}\boldsymbol{x} = 0. \end{cases} \tag{4}$$

The system (4) can be interpreted as the gradient flow of the energy $\mathcal{E}(c, 1-c)$ for a Wasserstein like distance corresponding to the nonlinear mobility $c(1-c)$ (cf. [4, 7]). As already pointed out by Otto and E in [9], the model (3) makes the energy decrease faster than (4) does. This is confirmed by numerical simulations based on the numerical scheme proposed in [2] as shown in Figure 1.

Finally, we establish the existence of a weak solution to the problem (3) thanks to the convergence of the minimizing movement (or JKO [6]) scheme

$$\boldsymbol{c}^n \in \operatorname*{Argmin}_{\boldsymbol{c}\in\mathcal{A}} \frac{1}{2\tau}\boldsymbol{W}^2(\boldsymbol{c}, \boldsymbol{c}^{n-1}) + \mathcal{E}(\boldsymbol{c}). \tag{5}$$

In the above formula, $\tau > 0$ denotes a time step that will tend towards 0 in the convergence analysis. The set $\mathcal{A}$ is defined by

$$\mathcal{A} = \left\{ \boldsymbol{c} \in L^1(\Omega; \mathbb{R}_+^2) \;\middle|\; \int_\Omega c_i \mathrm{d}\boldsymbol{x} = \int_\Omega c_i^{\mathrm{ini}} \mathrm{d}\boldsymbol{x} \right\}$$

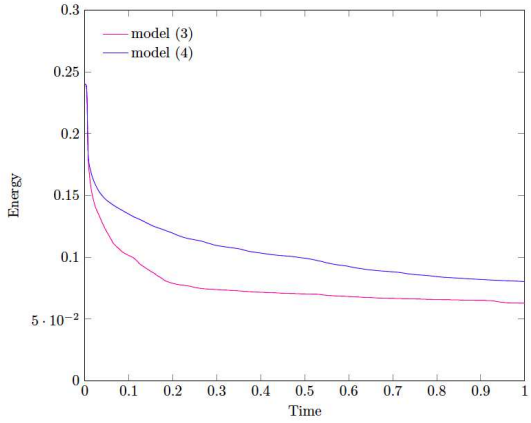

FIGURE 1. Evolution of the energy along time for a spinodal decomposition governed by both models (3) and (4).

where $\boldsymbol{c}^{\text{ini}} = (c_1^{\text{ini}}, c_2^{\text{ini}})$ is an initial data of finite energy. It is endowed with the Wasserstein distance

$$\boldsymbol{W}(\boldsymbol{c}, \check{\boldsymbol{c}}) = \left(\sum_{i \in \{1,2\}} W^2(c_i, \check{c}_i) \right)^{1/2}, \qquad \forall (\boldsymbol{c}, \check{\boldsymbol{c}}) \in \mathcal{A}^2.$$

The proof relies on tools introduced in [1] for proving the existence of a solution of multiphase porous media flows.

REFERENCES

[1] C. Cancès, T. O. Gallouët, and L. Monsaingeon, *Incompressible immiscible multiphase flows in porous media: a variational approach*, Anal. PDE **10** (2017), 1845-1876

[2] C. Cancès and F. Nabet, *Finite Volume Approximation of a Degenerate Immiscible Two-Phase Flow Model of Cahn–Hilliard Type*, in Finite Volumes for Complex Applications VIII - Methods and Theoretical Aspects, C. Cancès and P. Omnes (eds), Springer Proceedings in Mathematics and Statistics **199** (2017), 431–438.

[3] P.-G. de Gennes, *Dynamics of fluctuations and spinodal decomposition in polymer blends*, J. Chem. Phys. **72** (1980), 4756-4763.

[4] J. Dolbeault, B. Nazaret, and G. Savaré, *A new class of transport distances between measures*, Calc. Var. Partial Differential Equations **34** (2009), 193–231.

[5] W. E and P. Palffy-Muhoray, *Phase separation in incompressible systems*, Phys. Rev. E **55**, R3844.

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

[7] S. Lisini, D. Matthes, and G. Savaré, *Cahn-Hilliard and thin film equations with nonlinear mobility as gradient flows in weighted-Wasserstein metrics*, J. Differential Equations **253** (2012), 814–850.

[8] A. Mielke, *A gradient structure for reaction-diffusion systems and for energy-drift-diffusion systems*, Nonlinearity **24** (2011), 1329–1346.

[9] F. Otto and W. E, *Thermodynamically driven incompressible fluid mixtures*, J. Chem. Phys. **107** (1997), 10177-10184.
[10] M. A. Peletier, *Variational Modeling: Energies, gradient flows, and large deviations*, Lecture Notes, Würzburg (2014). Available at `http://www.win.tue.nl/∼mpeletie`

Some results on quasistatic fracture growth in linearized elasticity

Antonin Chambolle

The variational approach to fracture [13, 3] is an extension of Griffith's classical theory for the modelling and study of crack growth, where, in a quasistatic setting, a crack evolution is computed by successive (global) minimization of an energy consisting of a "bulk" (linearised elasticity) part and a "crack" term which penalises the length or surface of the discontinuity of the displacement. Namely, given a boundary datum U_0 on a part of $\Gamma^D \subseteq \partial\Omega$, $\Omega \in \mathbb{R}^d$ ($d = 2, 3$ in general), and a "time"-step $\delta t > 0$, one finds at each step $k \geq 1$ a minimiser (u, K) of

$$\min\left\{\int_\Omega \mathbb{C}e(u) : e(u)dx + \mathcal{H}^{d-1}(K) : u = (k\delta t)U_0 \text{ on } \Gamma^D, K \supset K^{k-1}\right\}. \tag{1}$$

The tensor $\mathbb{C}$ is the Hooke's law, typically $\mathbb{C}e(u) = 2\mu e(u) + \lambda \mathrm{Tr}\, e(u) Id$ and $e(u) = (Du + Du^T)/2$ is the symmetrised gradient of the displacement, here a function $u \in H^1_{loc}(\Omega \setminus K; \mathbb{R}^d)$.

In practice, one needs to relax a bit this problem (for which existence remains unknown, see [14, 9] for recent results) in the class "$[G]SBD$" [1, 2, 11] of functions u such that $Du + Du^T$ is a bounded Radon measure, which is decomposed into a part $e(u)dx$ absolutely continuous with respect to Lebesgue's measure, and a $(d-1)$-dimensional "jump" part $[u] \odot \nu_u \mathcal{H}^{d-1} \lfloor J_u$, where the countably $(d-1)$-rectifiable set J_u is defined as the set of points x where there exists (a normal unit vector) ν_u and (two different displacement values) $u^\pm$ such that, as $\rho \to 0$,

$$y \mapsto u(x + \rho y) \overset{L^1(B_1)}{\longrightarrow} u^+ \chi_{\{y\cdot\nu_u \geq 0\}} + u^- \chi_{\{y\cdot\nu_u < 0\}}$$

(the blowup of u at x converges to a function taking two values). Then, $[u] = u^+ - u^-$ and $[u] \odot \nu_u = ([u] \otimes \nu_u + \nu_u \otimes [u])/2$.

In fact, SBD is even too restrictive. The natural space where to look for a weak minimiser of (1), which corresponds to solving a problem of the form

$$\min_u \int_\Omega \mathbb{C}e(u) : e(u)dx + \mathcal{H}^{d-1}(J_u) \tag{2}$$

is the space $GSBD(\Omega)$ [11], first introduced by G. Dal Maso during the MFO Worskshop #1149 in 2011. Here, a boundary condition has to be enforced in some specific way, and J_u is the intrinsic jump set of the function u. The set $GSBD$ is defined by the properties of $1D$ slices which are required to be SBD functions. More precisely, $u \in GSBD(\Omega)$ if and only if for any $\xi \in \mathbb{S}^{d-1}$ and a.e. $x \in \xi^\perp$, the functions $s \mapsto u \cdot \xi(x + s\xi)$ are SBD_{loc} and their variations satisfy some global bound.

Such estimate give little control on the function itself (which is necessary to study problems such as (2), where the energy controls only the L^2 norm of $e(u)$ and the total surface of the jump set).

In this setting, existence of minimisers for (2) could only be proven very recently, and only in 2D, by Friedrich and Solombrino [14]. The fact that these minimisers are, in fact, strong, is shown in [9].

In this talk were reviewed and explained some recent results obtained in collaboration with S. Conti and G. Francfort [6], V. Crismale [8], S. Conti and F. Iurlano [7]. A first result is the following theorem, proved in [6] (*a priori* only for SBD displacements, however the proof is only based on integrals along lines and carries on without change to $GSBD$). Here, $GSBD_p$ is the set of functions $u \in GSBD$ such that $e(u)$ is p-integrable and J_u has finite $(d-1)$-Hausdorff measure.

Theorem 1. *Let $\theta > 0$, $\delta > 0$, $Q = (-\delta, \delta)^d$, $Q' = (1+\theta)Q$, $Q'' = (1+2\theta)Q$, $p \in (1, \infty)$. There exists $c(\theta, p, d) > 0$ such that for any $u \in GSBD_p(Q'')$:*

(1) There exists $\omega \subset Q'$ and $a = Ax + b$, $A + A^T = 0$, a linearized rigid motion, such that

$$|\omega| \leq c\delta \mathcal{H}^{d-1}(J_u)$$
$$\|u - a\|_{L^{np/(n-1)}(Q' \setminus \omega)} \leq c\delta^{1-1/n} \|e(u)\|_{L^p(Q'')}.$$

(2) Letting then $v = u\chi_{Q' \setminus \omega} + a\chi_\omega$ and given ϕ a smooth symmetric mollifier with support in $B(0, \theta/2)$, one has

$$\int_Q |e(v * \phi_\delta) - e(u) * \phi_\delta|^p dx \leq c \left(\frac{\mathcal{H}^{d-1}(J_u)}{\delta^{d-1}} \right)^s \int_{Q''} |e(u)|^p dx$$

where as usual, $\phi_\delta(x) = (1/\delta^d)\phi(x/\delta)$.

The result can be used as follows: given $u \in GSBD_p(\Omega)$ ($p \in (1, \infty)$), $\eta > 0$, $\delta > 0$ small, one covers Ω with cubes $R_z = 2\delta z + (-\delta, \delta)^d$, $z \in \mathbb{Z}^d$, of size δ (non-overlapping). One also introduces the cubes $Q_z = 2\delta z + (1+\theta)(-\delta, \delta)^d$ and their enlargements Q'_z, Q''_z as in the theorem (with sides respectively multiplied by $(1+\theta)$ and $(1+2\theta)$). Here θ is a fixed quantity (usually chosen so that Q''_z has edges of size at most $3/2$ larger than R_z). Then, one categorizes the cubes into "bad" and "good" cubes, depending whether $\mathcal{H}^{d-1}(J_u \cap Q''_z) \geq \eta\delta^{d-1}$ or $< \eta\delta^{d-1}$. In the first ones, one observes that the size of the boundary of the cube $(2d(2\delta)^{d-1})$ is controlled by $\mathcal{H}^{d-1}(J_u \cap Q''_z)/\eta$.

In the second ones, one applies the theorem, introducing sets $\omega_z \subset Q'_z$ with $|\omega_z| \leq c\delta\mathcal{H}^{d-1}(J_u \cap Q''_z)$, rigid motions a_z, and smooth approximations $v_z = \phi_\delta * ((u - a_z)\chi_{Q'_z \setminus \omega_z} + a_z)$. Then, the approximations are glued together for instance using a partition of unity of the good cubes, $\phi_z \in C_c^\infty(Q_z; [0,1])$, $\phi_z \geq \chi_{R_z}$, $\sum_z \phi_z = 1$, $|\nabla \phi_z| \leq C/\delta$ (where C obviously is like $1/\theta$).

It is then not too hard to show that if η is small enough, letting $\tilde{u} = \sum_z \psi_z v_z$ in the good cubes and $\tilde{u} = 0$ in the bad cubes, one has $\tilde{u} \in GSBD_p(\Omega)$, $\mathcal{H}^{d-1}(J_{\tilde{u}}) \leq$

$C/\eta\mathcal{H}^{d-1}(J_u)$,

$$\int |e(v_z)|^p dx \leq (1+\epsilon(\eta)) \int_\Omega |e(u)|^p dx$$

where $\epsilon(\eta)$ goes to zero as $\eta \to 0$.

Such constructions in this setting were first developed to study the convergence of phase-field approximation of fracture models with non-interpenetration, see [5] and the report of the MFO Workshop #1729.

Thanks to this construction, we could show in [8]:

Theorem 2. *Let $u \in GSBD_p(\Omega)$: then there exists $u_n \to u$ a.e. such that u_n is smooth up to its jump set J_{u_n} which is made of a finite union of (nice) pieces of C^1 hypersurfaces, with*

(1) $\lim_{n\to\infty} \mathcal{H}^{d-1}(J_{u_n}) = \mathcal{H}^{d-1}(J_u)$;
(2) $e(u_n) \to e(u)$ *strongly in* $L^p(\Omega; \mathbb{R}^{d\times d}_{sym})$.

This extends well-known results (since [4]) but removes unnecessary assumption (such as $u \in L^p$) for the orignal function u.

With a similar construction, we could also adapt the arguments in [12] for the Mumford-Shah functional and [9] (in dimension 2) to show in any dimension a result of existence of strong minimizers (i.e., with essentially closed jump set) for energies such as (2) with an additional quadratic penalization (ensuring weak compactness).

References

[1] L. Ambrosio, A. Coscia, & G. Dal Maso, *Fine properties of functions with bounded deformation*, Arch. Rational Mech. Anal. **139**-3 (1997) 201–238.

[2] G. Bellettini, A. Coscia & G. Dal Maso, *Compactness and lower semicontinuity properties in $SBD(\Omega)$*, Mathematische Zeitschrift **228** (1998) 337–351.

[3] B. Bourdin, G.A. Francfort & J.J. Marigo, *The variational approach to fracture*, J. Elasticity **91**-1,2,3 (2008) 5–148.

[4] A. Chambolle, *An approximation result for special functions with bounded variations*, J. Math Pures Appl. **83** (2004) 929–954.

[5] A. Chambolle, S. Conti & G.A. Francfort, *Approximation of a Brittle Fracture Energy with a Constraint of Non-Interpenetration*, preprint, to appear in Arch. Rational Mech. Anal., (2017).

[6] A. Chambolle, S. Conti & G.A. Francfort, *Korn-Poincaré inequalities for functions with a small jump set*, Indiana Univ. Math. J. **65**-4 (2016)1373–1399.

[7] A. Chambolle, S. Conti & F. Iurlano, *Approximation of functions with small jump sets and existence of strong minimizers of Griffith's energy*, preprint arXiv:1710.01929 (2017).

[8] A. Chambolle & V. Crismale, *A density result in $GSBD^p$ with applications to the approximation of brittle fracture energies*, preprint arXiv:1708.03281, (2017).

[9] S. Conti, M. Focardi, and F. Iurlano, *Existence of minimizers for the 2d stationary Griffith fracture model*, C. R. Math. Acad. Sci. Paris, 354 (2016), pp. 1055–1059.

[10] S. Conti, M. Focardi & F. Iurlano, *Integral representation for functionals defined on SBD^p in dimension two*, Arch. Ration. Mech. Anal., **223**-3 (2017) 1337–1374.

[11] G. Dal Maso, *Generalised functions of bounded deformation*, J. Eur. Math. Soc. **15**-5 (2013) 1943–1997.

[12] E. De Giorgi, M. Carriero, and A. Leaci, *Existence theorem for a minimum problem with free discontinuity set*, Arch. Rational Mech. Anal., 108 (1989), 195–218.

[13] G. A. Francfort and J.-J. Marigo, *Revisiting brittle fracture as an energy minimization problem*, J. Mech. Phys. Solids, 46 (1998), pp. 1319–1342.
[14] M. Friedrich & F. Solombrino, *Quasistatic crack growth in 2d-linearized elasticity*, to appear in Ann. Inst. H. Poincaré Anal. Non Linéaire (2017).

A maximal dissipation condition for dynamic fracture, with an existence result in a constrained case

Gianni Dal Maso
(joint work with Christopher J. Larsen and Rodica Toader)

We consider a model of *dynamic fracture* in dimension two. The reference configuration is a bounded open set $\Omega \subset \mathbb{R}^2$ with Lipschitz boundary $\partial\Omega = \partial_D\Omega \cup \partial_N\Omega$. On $\partial_D\Omega$ we prescribe a time dependent Dirichlet boundary condition $w(t)$, while on $\partial_N\Omega$ we impose the homogeneous Neumann boundary condition.

The *initial crack* $\Gamma_0 = \gamma_0([a_0, 0])$ is described by a curve $\gamma_0 \colon [a_0, 0] \to \overline{\Omega}$ of class $C^{3,1}$ parametrised by arc-length and such that $\gamma_0(a_0) \in \partial\Omega$, γ_0 is *transversal* to $\partial\Omega$ at $\gamma_0(a_0)$, and $\gamma_0(s) \in \Omega$ for every $s \in (a_0, 0]$. The *crack* at time $t \in [0, T]$ is described as

$$\Gamma_{s(t)} := \gamma([a_0, s(t)])\,,$$

where $\gamma \colon [a_0, b_\gamma] \to \mathbb{R}^2$ is the *arc-length parametrisation* of a $C^{3,1}$ simple curve, coinciding with γ_0 on $[a_0, 0]$, and $s \colon [0, T] \to [0, b_\gamma]$, with $s(0) = 0$, is a continuous *piecewise* $C^{3,1}$ function that, for every $t \in [0, T]$, provides the *length of the crack* produced along the curve γ.

Given $r > 0$ and $L > 0$, we assume the following properties on the unknown $C^{3,1}$ curve γ:

- (G1) *uniform tangent balls condition*: for every $s \in [a_0, b_\gamma]$ the two open *disks* of radius r *tangent* to γ at $\gamma(s)$ do *not intersect* $\gamma([a_0, b_\gamma])$;
- (G2) *uniform distance from the boundary*: $\operatorname{dist}(\gamma([0, b_\gamma]), \partial\Omega) \geq 2r$;
- (G3) *uniform bounds*: $|\gamma^{(3)}(s)| \leq L$, $|\gamma^{(3)}(s_2) - \gamma^{(3)}(s_1)| \leq L|s_2 - s_1|$, for every $s, s_1, s_2 \in [a_0, b_\gamma]$, where $\gamma^{(i)}$ denotes the i-th derivative of γ.

Given $\mu > 0$ and $M > 0$, we assume the following properties on the unknown *piecewise* $C^{3,1}$ crack length s:

- (S1) *bound on the speed*: $0 \leq \dot{s}(t) \leq \mu$;
- (S2) *uniform bounds*: $|\ddot{s}(t)| \leq M$, $|\dddot{s}(t)| \leq M$, and $|\dddot{s}(t_1) - \dddot{s}(t_2)| \leq M|t_1 - t_2|$ for every t, t_1, t_2 in an interval where s is of class $C^{3,1}$.

Given a *body force* f, a *boundary condition* w, an *initial displacement* u_0, and an *initial velocity* u_1, the *unknown displacement* $u \colon [0, T] \times \Omega \to \mathbb{R}^2$ solves the *system*

of elastodynamics in the time dependent domain $\Omega \setminus \Gamma_{s(t)}$:

$$\begin{aligned}
&\ddot{u}(t,x) - \operatorname{div}(\mathbb{C}(x)\nabla u(t,x)) = f(t,x) \quad \text{for } t \in (0,T),\ x \in \Omega \setminus \Gamma_{s(t)},\\
&u(t,x) = w(t,x) \quad \text{for } t \in (0,T),\ x \in \partial_D \Omega,\\
&(\mathbb{C}(x)\nabla u(t,x))\nu(x) = 0 \quad \text{for } t \in (0,T),\ x \in \partial_N \Omega \cup \Gamma_{s(t)},\\
&u(0,x) = u_0(x) \quad \text{for } x \in \Omega,\\
&\dot{u}(0,x) = u_1(x) \quad \text{for } x \in \Omega,
\end{aligned}$$

where $\mathbb{C}(x)$ is the *elasticity tensor*. Note that we have the homogeneous Neumann boundary condition on both sides of $\Gamma_{s(t)}$, since we assume that the crack lips are traction-free.

Our analysis is based on the following result, proved in [1].

Theorem 1. *Assume (G1), (G2), (G3), (S1), and (S2). Under suitable regularity and compatibility assumptions on f, w, u_0, and u_1, if the constant μ that bounds the crack speed is small enough, then there exists a unique weak solution of the system of elastodynamics in the time dependent domain $\Omega \setminus \Gamma_{s(t)}$. Moreover this solution depends continuously on all data, including (γ, s).*

A *localization argument* in space and time (see [3]), based on the *finite speed of propagation*, allows us to extend this result to the case $0 < \mu < \sqrt{\lambda}/2$, where λ is the *ellipticity constant* of the *elasticity tensor* (which gives an estimate of the speed of sound).

We now fix f, w, u_0, and u_1, and consider only the collection $\mathcal{C}^{piec}$ of all cracks (γ, s) satisfying (G1), (G2), (G3), (S1), (S2), and such that the corresponding displacement u satisfies the *dynamic energy-dissipation balance*:

kinetic energy at time t_2 + elastic energy at time t_2 +
energy dissipated by the crack between t_1 and t_2 =
kinetic energy at time t_1 + elastic energy at time t_1 +
work of the external forces between t_1 and t_2 .

The energy dissipated by the crack between t_1 and t_2 is assumed to be proportional to $s(t_2) - s(t_1)$, where the proportionality constant depends on the material and is interpreted as the *toughness* of the material.

The collection $\mathcal{C}^{piec}$ is *non-empty*, since every stationary crack satisfies the dynamic energy-dissipation balance. Indeed, this is a general property of the solutions of the system of elastodynamics in time-independent domains.

Therefore, we propose an additional principle, a *maximal dissipation criterion*, which forces the crack to grow when it is possible to grow in a way that preserves the dynamic energy-dissipation balance (so that stationary cracks are not always solutions). The spirit of the maximal dissipation condition is simply that the crack must *run as fast as possible*, consistent with the *dynamic energy balance*.

To be precise, we say that $(\gamma, s) \in \mathcal{C}^{piec}$ satisfies the *maximal dissipation* condition if *there exists no* $(\hat{\gamma}, \hat{s}) \in \mathcal{C}^{piec}$ such that, for some $0 \le \tau_0 < \tau_1 \le T$,

(MD1) $\operatorname{sing}(\hat{s}) \subset \operatorname{sing}(s)$,

(MD2) $\hat{s}(t) = s(t)$ and $\hat{\gamma}(\hat{s}(t)) = \gamma(s(t))$ for every $t \in [0, \tau_0]$,
(MD3) $\hat{s}(t) > s(t)$ for every $t \in (\tau_0, \tau_1]$.

Here and henceforth $\mathrm{sing}(s)$ is the set of points where s is not regular. Conditions (MD1)–(MD3) say that there is *no sufficiently regular crack* which satisfies the *dynamic energy-dissipation balance*, *coincides* with the crack described by (γ, s) up to time τ_0, and is *longer* at every time between τ_0 and τ_1.

For technical reasons, we are able to prove the existence of an admissible evolution satisfying the previous condition only in a *quantitative way*, depending on a prescribed *threshold* $\eta > 0$. This leads to the following definition: $(\gamma, s) \in \mathcal{C}^{piec}$ satisfies the *η-maximal dissipation* condition on $[0, T]$ if *there exists no* $(\hat{\gamma}, \hat{s}) \in \mathcal{C}^{piec}$ such that (MD1)–(MD3) hold for some $0 \le \tau_0 < \tau_1 \le T$ and, in addition,

(MD4) $\hat{s}(\tau_1) > s(\tau_1) + \eta$.

Our main result is the following theorem (see [3]).

Theorem 2. *Assume (G1), (G2), (G3), (S1), and (S2), with* $0 < \mu < \sqrt{\lambda}/2$. *Suppose that* f, w, u^0, u^1 *satisfy the hypotheses of Theorem 1 and let* $\eta > 0$. *Then there exists an η-maximal dissipation solution of the dynamic crack evolution problem corresponding to these data.*

The main difficulty in the definition of an η-maximal dissipation solution is the variability of the interval $(\tau_0, \tau_1]$ where the comparison crack $(\hat{\gamma}, \hat{s}) \in \mathcal{C}^{piec}$ satisfies the inequality $\hat{s}(t) > s(t)$. To overcome this problem we *discretize* time and in each time interval we use Theorem 1 to prove the existence of a *maximal pair* (γ, s) among all pairs satisfying our quantitative regularity requirements and the dynamic energy-dissipation balance. Then we prove that the function obtained by *glueing together* these maximal functions is an *η-maximal dissipation solution*, if the length of each time interval is less than η/μ, where μ is the bound for the velocity.

In the antiplane case a similar result with a prescribed crack path (i.e., γ is given and only s is unknown) is proved in [2] using the results for the scalar wave equation contained in [4].

References

[1] M. Caponi, *Linear hyperbolic systems in domains with growing cracks*, Milan J. Math. **85** (2017), 149–185.

[2] G. Dal Maso, C.J. Larsen, R. Toader, *Existence for constrained dynamic Griffith fracture with a weak maximal dissipation condition*, J. Mech. Phys. Solids **95** (2016), 697–707.

[3] G. Dal Maso, C.J. Larsen, R. Toader, *Existence for dynamic Griffith fracture with a weak maximal dissipation condition*, in preparation.

[4] G. Dal Maso, I. Lucardesi, *The wave equation on domains with cracks growing on a prescribed path: existence, uniqueness, and continuous dependence on the data*, Appl. Math. Res. Express. AMRX **1** (2017), 184–241.

A gradient flow to the homogeneous 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

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

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 $c^{-1} \leq B \leq c$ 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. $\frac{\mathrm{d}}{\mathrm{dt}}\mathcal{H}(f_t) = -D(f_t) \leq 0$, where

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

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

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

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

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

Here, we have set $\bar{\nabla}\phi = \phi' + \phi'_* - \phi - \phi_*$ and $\Lambda(f)$ is shorthand for $\Lambda\big(f f_*, f' f'_*\big)$, 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.

Cross-diffusion systems with nonstandard entropies

Ansgar Jüngel
(joint work with N. Zamponi)

Multi-species systems from physics, biology, chemistry, etc. can be modeled by reaction-diffusion equations. When the gradient of the density of one species induces a flux of another species, cross diffusion occurs. Thus, the diffusion matrix involves nonvanishing off-diagonal elements. In many applications, the evolution equations satisfy a formal gradient-flow structure, which compensates the difficulty arising from the fact that the original diffusion matrix is generally neither symmetric nor positive definite. The cross-diffusion equations have the form

$$\partial_t u_i - \sum_{j=1}^n \operatorname{div}(A_{ij}(u)\nabla u_j) = f_i(u) \quad \text{in } \Omega,\ t>0,\ i=1,\ldots,n, \tag{1}$$

where $u_i(x,t)$ is the density of the ith species, $u=(u_1,\ldots,u_n)$, $A_{ij}(u)$ are the diffusion coefficients, $f_i(u)$ is the reaction term of the ith species, and $\Omega\subset\mathbb{R}^d$ ($d\ge 1$) is a bounded domain with smooth boundary. We need to prescribe initial conditions for u_i; the boundary conditions generally depend on the considered application. Setting $A(u)=(A_{ij}(u))$ and $f(u)=(f_1(u),\ldots,f_n(u))$, we may write (1) more compactly as

$$\partial_t u - \operatorname{div}(A(u)\nabla u) = f(u) \quad \text{in } \Omega,\ t>0.$$

Generally, there do not exist maximum principles or a regularity theory for diffusion systems. Our approach is to assume that there exists a convex function h :

$\mathbb{R}^n \to \mathbb{R}$, called an entropy density, such that the matrix product $B = A(u)h''(u)^{-1}$ is positive semidefinite [2]. Then (1) writes as

$$\partial_t u - \operatorname{div}\left(B\nabla\frac{\delta H}{\delta u}\right) = f(u),$$

where $\delta H/\delta u$ is the variational derivative of the entropy $H[u] = \int_\Omega h(u)dx$. This can be recognized as a formal gradient flow with the Onsager operator $\xi \mapsto \operatorname{div}(B\nabla\xi)$. Using $h'(u)$ as a test function, this usually leads to gradient estimates needed for the global existence analysis.

In many situations from population dynamics, cell biology, and thermodynamics, the entropy is given by the sum of the Boltzmann entropies of the various species, $h(u) = \sum_{i=1}^n u_i(\log u_i - 1)$. Here, we are interested in situations, where the cross-diffusion system admits a nonstandard entropy, i.e. a function that is not the sum of Boltzmann entropies. We consider three examples.

1. Energy-transport systems for semiconductors

Under some assumptions, the transport of electrons with density $\rho(x,t)$ and electron temperature $\theta(x,t)$ can be described by the cross-diffusion system

$$\partial_t\rho = \Delta(\rho\theta^{1/2-\beta}), \quad \partial_t(\rho\theta) = \Delta(\rho\theta^{3/2-\beta}) \quad \text{in } \Omega,\ t>0, \tag{2}$$

where $\beta \in [-1/2, 1/2)$ is a parameter coming from the elastic collision rate in the Boltzmann equation from which this model is derived; see [1] for details. To simplify, we have neglected the electric field and the relaxation term. Equations (2) possess the entropy density $h(\rho,\theta) = \rho\log(\rho\theta^{-3/2})$. A formal computation shows (for suitable boundary conditions) that

$$\frac{d}{dt}\int_\Omega h(\rho,\theta)dx + C\int_\Omega \theta^{1/2-\beta}\left(\frac{|\nabla\rho|^2}{\rho} + \frac{|\nabla\theta|^2}{\theta}\right)dx \le 0,$$

where $C > 0$. Unfortunately, this estimate is not useful close to $\theta = 0$, as the positivity of θ cannot be easily proved. The idea is to use a nonphysical entropy, namely $h(\rho,\theta) = \rho^2\theta^b$ with $b \in \mathbb{R}$, which gives the estimate

$$\frac{d}{dt}\int_\Omega \rho^2\theta^b dx + C_1\int_\Omega |\nabla(\rho\theta^{(2b+1-2\beta)/4}|^2 dx \le C_2,$$

with positive constants C_1 and C_2 depending on the boundary data. Then, taking a variant of the sum $\int_\Omega \rho^2\theta^{\beta-1/2}dx + \int_\Omega \rho^2\theta^5 dx$, we are able to derive gradient estimates for the variables ρ, $\rho\theta^{1/2-\beta}$, and $\rho\theta^{3/2-\beta}$, which is necessary to define a weak solution to (2). The existence of global weak solutions was shown in [6].

2. Partial averaging in economics

Starting from the forward Kolmogorov equation for the distribution function f of an Itô process, the partial averages $u_i(x,t) = \int_{\mathbb{R}} f(x,y,t)e^{-\lambda_i y}dy$, where $x \in \mathbb{R}^d$, $y \in \mathbb{R}$, $\lambda_i > 0$, satisfy under some simplifying assumptions the equations

$$\partial_t u_i = \Delta\big(a(u_1/u_2)u_i\big) + \mu_i u_i \quad \text{in } \mathbb{T}^d,\ t>0,\ i=1,2, \tag{3}$$

where $\mathbb{T}^d$ is the d-dimensional torus and the function $a : (0,\infty) \to (0,\infty)$ and the parameters μ_i are related to the volatility appearing in the Kolmogorov equation [5]. The partial average u_i may be interpreted as an value of an economic parameter weighted by the parameter y. Surprisingly, (3) can be analyzed using entropy methods. The key idea is to employ the functional

$$H[u] = \int_{\mathbb{T}^d} h(u)dx, \quad h(u) = \left(\frac{u_1}{u_2}\right)^\alpha u_1^2 + \left(\frac{u_1}{u_2}\right)^{-\alpha} + \sum_{i=1}^2 (u_i - \log u_i),$$

where $\alpha \geq p+4$ and $u = (u_1, u_2)$, leading to the entropy inequality

$$\frac{dH}{dt} + \int_{\mathbb{T}^d} \left(\left(\frac{u_1}{u_2}\right)^{\alpha-p} + \left(\frac{u_1}{u_2}\right)^{p-\alpha}\right)\big(|\nabla u_1|^2 + |\nabla u_2|^2\big)dx \leq CH[u],$$

where $C \geq 0$ depends on μ_i and $C = 0$ if $\mu_1 = \mu_2 = 0$. This inequality gives gradient estimates for u_1 and u_2, which are the key for the global existence analysis executed in [3].

3. Flow mixture with van der Waals pressure

Consider an isothermal fluid mixture of n mass densities $\rho_i(x,t)$ in a domain $\Omega \subset \mathbb{R}^d$, whose evolution is governed by the equations

$$\partial_t \rho_i = \operatorname{div}\left(\rho_i v + \varepsilon \sum_{j=1}^n D_{ij}(\rho)\nabla\mu_j\right), \quad i = 1,\ldots,n,$$

where the velocity v is determined from Darcy's law $v = -\nabla p$, the van der Waals pressure p is defined by the Gibbs-Duhem relation $\nabla p = \sum_{i=1}^n \rho_i \nabla\mu_i$, the chemical potentials μ_i are given by $\mu_i = \partial h/\partial\rho_i$, and the free energy equals

$$h(\rho) = \sum_{i=1}^n \rho_i(\log\rho_i - 1) - \rho_{\mathrm{tot}} \log\left(1 - \sum_{j=1}^n b_j\rho_j\right) - \sum_{i,j=1}^n a_{ij}\rho_i\rho_j.$$

Here, $\varepsilon > 0$ is a small parameter, $D_{ij}(\rho)$ are the diffusion coefficients, $\rho_{\mathrm{tot}} = \sum_{i=1}^n \rho_i$ is the total mass density, $a_{ij} = a_{ji} > 0$ measures the attraction between the ith and jth species, and b_j is a measure of the size of the molecules. Putting these relations together, the evolution equations can be written as

$$\partial_t \rho_i = \operatorname{div} \sum_{j=1}^n \big((\rho_i\rho_j + \varepsilon D_{ij}(\rho))\nabla\mu_j\big), \quad i = 1,\ldots,n.$$

Since the diffusion fluxes sum up to zero, the matrix (D_{ij}) is positive definite only on the orthogonal complement of its one-dimensional kernel. The lack of full positive definiteness is compensated by the presence of the matrix $(\rho_i\rho_j)$, which is of rank one only. This property is reflected in the entropy inequality

$$\frac{d}{dt}\int_\Omega h(\rho)dx + \int_\Omega |\nabla p|^2 dx + \varepsilon \int_\Omega \nabla\mu : D(\rho)\nabla\mu dx \leq 0.$$

Both the second and third integrals lead to H^1 bounds for μ_i, which are the basis of the global existence analysis performed in [4].

References

[1] A. Jüngel, *Transport Equations for Semiconductors*, Lecture Notes in Physics **773**, Springer, Berlin, 2009.

[2] A. Jüngel, *The boundedness-by-entropy method for cross-diffusion systems*, Nonlinearity **28** (2015), 1963-2001.

[3] A. Jüngel and N. Zamponi, *A cross-diffusion system derived from a Fokker-Planck equation with partial averaging*, Z. Angew. Math. Phys. **68** (2017), 28, 15 pages.

[4] A. Jüngel, J. Mikyška, and N. Zamponi, *Existence analysis of a single-phase flow mixture model with van der Waals pressure*, to appear in SIAM J. Math. Anal., 2018. arXiv:1612.04161.

[5] P.-L. Lions, *Some new classes of nonlinear Kolmogorov equations*, talk at the 16th Pauli Colloquium, Wolfgang-Pauli Institute, Vienna, November 18, 2015.

[6] N. Zamponi and A. Jüngel, *Global existence analysis for degenerate energy-transport models for semiconductors*, J. Diff. Eqs. **258** (2015), 2339-2363.

Non-variational evolution towards a crystal-like pattern

Christian Maes
(joint work with Karel Netočný)

Nonequilibirum statistical mechanics has roughly two types of questions. One is reaching for a dynamical characterization of the equilibrium condition. Here equilibrium refers to the macroscopic state corresponding to maximal entropy given a number of constraints in terms of energy, volume, particle number and maybe also other conserved quantities. By entropy for a closed and isolated mechanical system we mean the logarithmic phase space volume of microscopic states compatible with the macroscopic condition, making thus the bridge between the microscopic description and the macroscopic looks of a system composed of many particles. The dynamical characterization of such an equilibrium condition includes the study of relaxation to equilibrium, the derivation of macroscopic evolution equation, the understanding of the emergence of dissipation from Hamiltonian dynamics, linear response theory and the theory of macroscopic dynamical fluctuations as in the theory of large deviations on path-space. That first chapter of nonequilibrium statistical mechanics thus also wants to contain derivations of variational evolutions and the understanding of their connection with the condition of detailed balance [1, 2], or the derivation of damping and friction in contact with a thermal equilibrium environment [3], etc.

The second type of question concerns the so called steady state. We are now dealing with -on the level of the total system plus reservoirs- a low entropy macroscopic condition which is assumed kept and steady over a certain time-scale. The possible variations are again enormous, but one can easily imagine systems driven by boundary conditions, where the latter make the weak contact between system and reservoirs, those being spatially separated and specified in terms of a few kinetic and thermodynamic parameters. Heat conduction is a special case. Yet, there also exists driving induced by temporal changes, like via periodic external fields, or driving via internal degrees of freedom, etc. Here again the questions

include developing a response and fluctuation theory, and to characterize possible stability and relaxation towards the nonequilibrium steady condition.
The question of the paper [4] and the present contribution is to study the emerging stability of a pattern of probes in contact with a nonequilibrium medium. That brings us towards one of the burning questions within nonequilibrium physics, to describe pattern formation from more microscopic to mesoscopic models, to get to grips with symmetry breaking and phase transitions under nonequilibrium conditions.

Morphogenesis is a traditional problem in many scientific domains. Related concepts are self-assembly, self-organization and the emergence of patterns, both spatially and temporally. It is obviously related with symmetry breaking and with issues of learning. Much of contemporary research is devoted to it for its technological and fundamental significance [5]. One of the most sensational aspects is that nature appears able to use robust mechanisms in giving a large variety of astonishing spatio-temporal patterns [6]. It is not so strange to believe that nonequilibrium conditions can be at the origin of such structures, but there are few *ab initio* studies of the mechanisms through which such patterns emerge from contact with nonequilibrium media. Indeed one would wish to avoid specially constructed or *ad hoc* modeling to deduce such structures.

In the present work [4] we present possibly the simplest scenario where calculations can be done explicitly and the emergence of a stationary spatial crystal-like pattern can be followed in great detail. The mesoscopic model consists of three layers of description. It corresponds to three types of particles, all confined on a ring of length L. First there is the viscous thermal equilibrium reservoir of fast moving small particles. It is represented in the dynamics via its temperature β^{-1} and it gives idealized (white) noise and linear friction to the other particles immersed in it. Secondly there are the driven particles forming a nonequilibrium medium. We call them colloids. They are with many, mutually independent, and are undergoing an overdamped dynamics in the thermal bath. The driving is by a constant rotational force over the ring, with amplitude ε. They are each in short range interacting with N probes. The latter have positions $x = (x(1), x(2), \ldots, x(N))$. Formally, we thus have for the position $\eta_t \in S^1$ of one colloid,

$$\zeta \frac{\mathrm{d}\eta_t}{\mathrm{d}t} = \varepsilon - \frac{\partial U(x,\eta_t)}{\partial \eta} + \Big(\frac{2\zeta}{\beta}\Big)^{1/2} \xi_t \quad (\eta \bmod L) \tag{1}$$

with $U(x,\eta) = \sum_\alpha u(x_\alpha - \eta)$ the total potential for which we suppose sufficient smoothness, the symmetry $u(z) = u(-z)$ and the range δ in that $u(z) = 0$ whenever $|z| > \delta$. The constant ε is the constant driving force. Inverse temperature β and friction ζ are fixed positive parameters, with ξ_t standard white noise on the ring. Alternatively, we can write the stationary Smoluchowski equation

$$\zeta j_x \;=\; \rho_x(\eta)\,\big[\varepsilon + \sum_\alpha u'(x(\alpha) - \eta)\big] - \frac{\rho_x'(\eta)}{\beta} \tag{2}$$

where ρ_x is the stationary density of the colloids under periodic boundary conditions and for fixed colloid mass $\int_0^L \rho_x(\eta)\,\mathrm{d}\eta = \rho^o L$; the constant j_x is the stationary current of the colloid flow and it depends on the probe positions x. Finally we imagine the probes as moving on the slowest time-scale. In that quasi-static limit we thus consider as force on the αth-probe from the driven fluid the mean force given by

$$f_\alpha(x) = -\int_0^L u'(x(\alpha)-\eta)\,\rho_x(\eta)\,\mathrm{d}\eta \tag{3}$$

In that approximation the probes then take the overdamped dynamics

$$\Gamma \dot{x}_\alpha = f_\alpha(x) + \Big(\frac{2\Gamma}{\beta}\Big)^{\frac{1}{2}} \xi_t^\alpha, \qquad \alpha = 0,\dots,N-1 \tag{4}$$

now with friction $\Gamma > 0$ and independent standard white noises ξ_t^α. The dynamics (4) is our main object of study. The main result is that it allows stable crystal-like patterns which are reached via non-variational relaxation.

There are a number of length scales in the model which are all important. There is of course the lenght L of the circle and the range δ of the probe–colloid interaction. The number of colloids being N, we want $L/N > 2\delta$ so that in equilibrium $f_\alpha(x) = 0$ whenever the probes are sufficiciently far apart. The dissipation length is $\ell_d = (\varepsilon\beta)^{-1}$ and we find that the situation is most interesting for $L \gg \ell_d$, N large and $\delta \le \ell_d$, $N\ell_d/L \ge 1$, with $L/(2\delta) \ge N$. That enters in particular in the analysis of the equidistant probe configuration.

The equidistant configuration $x = x^*$ of probes obviously has the property that f_α does not depend on α. We can write it as a rotating crystal

$$x^*(\alpha,t) = v^* t + \frac{L}{N}\,\alpha \qquad (\alpha \bmod N) \tag{5}$$

steady rotation speed v^* proportional to the stationary current j_{x^*}. In the paper [4] we prove that that crystal is stable in the following sense.

The linearized dynamics of the perturbation $x(\alpha) = x^*(\alpha) + y_\alpha$ has the standard form

$$\Gamma \dot{y}_\alpha = \sum_\gamma M_{\alpha\gamma} y_\gamma\,, \qquad M_{\alpha\gamma} = \frac{\partial f_\alpha(x^*)}{\partial x_\gamma} \tag{6}$$

with $M_{\alpha\alpha} = -\sum_{\gamma\neq\alpha} M_{\alpha\gamma}$ (zero mode) and $M_{\alpha\gamma} = m_{\gamma-\alpha}, \alpha \neq \gamma$ for certain linear elasticity coefficients m_γ. We call them effective "spring constants" but it is important to realize that they are not symmetric, $m_\gamma \neq m_{-\gamma}$, which in the end is a consequence of the driving ε in (1) (choosing a rotation direction); the induced forces between the probes are non-reactive. We show in [4] that the $m_\alpha > 0$ when the interaction u between colloids and probes is either completely attractive or completely repulsive. There is a Lyapunov function, $\Lambda = \sum_\alpha y_\alpha^2$, with $\dot{\Lambda}(t) < 0$ unless $y \equiv 0$, proving stability in (6) of the crystal pattern.

For the stability against thermal fluctuations we estimate the time of relaxation in (6) using an analogy with compound Poisson processes. The stationary distribution for the linearized (4) coincides with the Boltzmann distribution,

$$\nu(y) \propto e^{-\beta V(y)} \tag{7}$$

for the effective potential

$$V(y) = \frac{1}{2}\sum_{\alpha,\gamma} y_\alpha M_{\alpha\gamma} y_\gamma = \frac{1}{4}\sum_{\gamma>0} m_\gamma \sum_\alpha (y_{\alpha+\gamma} - y_\alpha)^2 \tag{8}$$

despite the absence of detailed balance in (4).

References

[1] A. Mielke, M.A. Peletier and D.R.M. Renger, *On the relation between gradient flows and the large-deviation principle, with applications to Markov chains and diffusion*, `arXiv:1312.7591 [math.FA]`.

[2] R. Kraaij, A. Lazarescu, C. Maes and M. Peletier, *Deriving GENERIC from a generalized fluctuation symmetry*, `https://arxiv.org/abs/1706.10115`.

[3] C. Maes and H. Tasaki, *Second law of thermodynamics for macroscopic mechanics coupled to thermodynamic degrees of freedom*, Letters in Mathematical Physics **79** (2007), 251–261.

[4] C. Maes and K. Netočný, *Non-reactive forces and pattern formation induced by a nonequilibrium medium*, `arXiv:1711.05168 [cond-mat.stat-mech]`.

[5] J. A. Pelesko, *Self Assembly: The Science of Things That Put Themselves Together*. Chapman and Hall/CRC (2007).

[6] T. Sekimura, S. Noji, N. Ueno, and P. Maini (Eds.), *Morphogenesis and Pattern Formation in Biological Systems. Experiments and Models* (Springer Japan, 2003).

A two-step time discretization of metric gradient flows

Daniel Matthes

(joint work with Simon Plazotta)

1. Variational formulation of multi-step schemes

We analyze a discrete-in-time approximation of gradient flows in metric spaces, using a variational analogue of the BDF2 formula instead of the most commonly employed BDF1 method (which is better known as minimizing movement scheme). Before stating our result in the metric context, we briefly recall the respective methods for numerical integration of ODEs: given a smooth function $F : \mathbb{R}^d \to \mathbb{R}$ and an initial datum $u^0 \in \mathbb{R}^d$, we seek an approximation of solutions $u : [0,\infty) \to \mathbb{R}^d$ of the initial value problem

$$\dot u = -\nabla F(u), \quad u(0) = u^0,$$

by sequences $(u_\tau^n)_{n=0}^\infty$ of uniform time step $\tau > 0$, i.e., such that $u_\tau^n \approx u(n\tau)$. Two canonical options are the BDF1 method,

$$u_\tau^n \quad \text{solves} \quad \frac{u - u_\tau^{n-1}}{\tau} = -\nabla F(u), \tag{1}$$

which needs one initial datum $u_\tau^0 \approx u^0$, and the BDF2 method,

$$u_\tau^n \quad \text{solves} \quad 2\left(\frac{u-u_\tau^{n-1}}{\tau}\right) - \left(\frac{u-u_\tau^{n-2}}{2\tau}\right) = -\nabla F(u), \tag{2}$$

which is initialized with two data points u_τ^0, u_τ^1. Estimates on the convergence of $(u_\tau^n)_{n=0}^\infty$ to u are classical; in the smooth setting, the BDF1 and BDF2 method are convergent, respectively, at first and second order in τ.

The essential ingredient to define analogues of these BDF methods for metric gradient flows in the sense of [1] is their variational formulation: (1) and (2) are the Euler-Lagrange equations for

$$F_\tau^{(1)}(u;v) := \frac{d(u,v)^2}{2\tau} + F(u) \quad \text{and} \quad F_\tau^{(2)}(u;v,w) := \frac{d(u,v)^2}{\tau} - \frac{d(u,w)^2}{4\tau} + F(u),$$

respectively, when $v = u_\tau^{n-1}$ and $w = u_\tau^{n-2}$, that is, the iterations (1) and (2) can be written in the form

$$u_\tau^n \in \operatorname{argmin} F_\tau^{(1)}(\,\cdot\,; u_\tau^{n-1}), \quad \text{or} \quad u_\tau^n \in \operatorname{argmin} F_\tau^{(2)}(\,\cdot\,; u_\tau^{n-1}, u_\tau^{n-2}), \tag{3}$$

respectively. Clearly, the variational functionals $F_\tau^{(1)}$ and $F_\tau^{(2)}$ can be defined also for non-smooth energies F defined on a complete metric space (X,d). And under the usual hypotheses, that F be non-negative, proper, and lower semi-continuous, the recursions (3) are well-posed in that general context.

2. Main result

So far, we have been able to prove convergence for a particularly good class of metric gradient flows, namely for flows that satisfy the evolutional variational inequalities (EVI): we say that $u : [0,\infty) \to X$ is an EVI_λ-curve for F if

$$\frac{1}{2}\frac{d}{dt} d(u(t),z)^2 + \frac{\lambda}{2} d(u(t),z)^2 \le F(z) - F(u(t)) \tag{4}$$

holds for almost all $t > 0$ and each "observer point" $z \in X$. The hypotheses that we impose on F below imply that its gradient flow consists of EVI_λ-curves. The approximation of general curves of maximal slope is left for further research.

We introduce the following notion: a functional G on (X,d) is *weirdly convex of modulus* $\mu \in \mathbb{R}$, if any γ_0, γ_1 can be joined by an absolutely continuous curve $(\gamma_s)_{s\in[0,1]}$ along which the following augmented Jensen inequality holds:

$$G(\gamma_s) \le (1-s)G(\gamma_0) + sG(\gamma_1) - \frac{\mu}{2} s(1-s) d(\gamma_0,\gamma_1)^2, \quad 0 \le s \le 1.$$

For the BDF1 method, it has been shown in [1] that if there is some $\lambda \in \mathbb{R}$ such that

$$F_\tau^{(1)}(\,\cdot\,; v) \quad \text{is weirdly convex of modulus} \quad \mu_\tau^{(1)} = \lambda + \frac{1}{\tau}$$

for each sufficiently small $\tau > 0$ and each $v \in X$, and if $d(u_\tau^0, u^0) = O(\tau)$, then the sequence of minimization problems (3) is uniquely solvable and converges with rate $\mathcal{O}(\sqrt{\tau})$ to the unique solution u of the EVI_λ-flow for F. In fact, under mild

further hypotheses, the order of convergence can be improved from one-half to one, see [1, Chapter 4]. Our own result for the BDF2 method is:

Theorem 1. *Assume that there is some $\lambda \in \mathbb{R}$ such that for each sufficiently small $\tau > 0$ and for arbitrary $v, w \in X$ that*

$$F_\tau^{(2)}(\,\cdot\,; v, w) \quad \text{is weirdly convex of modulus} \quad \mu_\tau^{(2)} = \lambda + \frac{3}{2\tau}. \tag{5}$$

Assume further that $d(u_\tau^0, u^0) = \mathcal{O}(\tau) = d(u_\tau^1, u^0)$ and that $\limsup_{\tau\to 0} F(u_\tau^0) < \infty$. Then:

- *the sequence of minimization problems* (3) *is uniquely solvable;*
- *the $(u_\tau^n)_{n=0}^\infty$ converge to an absolutely continuous limit curve $u : [0, \infty) \to X$ in the following sense: for each $T > 0$, there is a $C > 0$ such that*
$$d(u_\tau^n, u(n\tau)) \le C\sqrt{\tau} \quad \text{for all } n\tau \le T;$$
- *u is an EVI$_\lambda$-curve for F.*

We emphasize that the order of convergence that we prove is still just $O(\sqrt{\tau})$. The significant contribution is that there is *some* quantifiable rate for convergence to the — possibly very non-smooth — solution to the gradient flow in the very general context of EVI flows on metric spaces. On the other hand, *if* it is known that the limiting solution has a high degree of regularity in time when considered in another context, i.e., in a Hilbert space, then the manifold classical results on second-order convergence of BDF2 become applicable.

The connection from hypothesis (5) to the EVI formulation (4) is easily established. Indeed, from (5), one derives the following discrete EVI:

$$\begin{aligned}&\frac{3}{4}\left(\frac{d^2(u_\tau^n, z) - d^2(u_\tau^{n-1}, z)}{\tau}\right) - \frac{1}{4}\left(\frac{d^2(u_\tau^{n-1}, z) - d^2(u_\tau^{n-2}, z)}{\tau}\right) + \frac{\lambda}{2} d^2(u_\tau^n, z)\\ &\le F(z) - F(u_\tau^n) - \tau\left[\left(\frac{d(u_\tau^n, u_\tau^{n-1})}{\tau}\right)^2 - \left(\frac{d(u_\tau^n, u_\tau^{n-2})}{2\tau}\right)^2\right].\end{aligned}$$

In the course of the proof, we derive sufficiently strong a priori estimates on the u_τ^n to pass to (4) in the limit $\tau \to 0$.

3. Discussion of hypotheses

Hypothesis (5) is trivially satisfied for any λ-convex functional F on a Hilbert space X, with d being induced by the norm $\|\cdot\|$: for the connecting curve, choose the interpolating straight line, $\gamma_s := (1-s)\gamma_0 + s\gamma_1$. Then $f(s) := \|\gamma_s - a\|^2$ defines a quadratic polynomial in s for any $a \in X$ with $f'' \equiv 2\|\gamma_1 - \gamma_0\|^2$. This clearly implies that the difference of the metric terms in $F_\tau^{(2)}(\gamma_s; v, w)$ gives a contribution that is convex of modulus $\frac{3}{2\tau}\|\gamma_1 - \gamma_0\|^2$.

The main application that we have in mind is that of λ-displacement convex functionals F on the L^2-Wasserstein space $(\mathcal{P}_2(\mathbb{R}^d), W_2)$. This includes gradient flows of nonlinear drift-diffusion-aggregation type,

$$\partial_t u = \Delta(u^m) + \nabla\cdot(u\,\nabla V) + \nabla\cdot(u\,\nabla W * u),$$

where $m > (d-1)/d$, and $V, W \in C^2(\mathbb{R}^d)$ are such that their Hessians are uniformly bounded from below. Given $\gamma_0, \gamma_1, v, w \in \mathcal{P}_2(\mathbb{R}^d)$, one defines γ_s as the generalized geodesic from γ_0 to γ_1 with base point v. That is, let $\boldsymbol{t}^0, \boldsymbol{t}^1$ be optimal maps in $\gamma_0 = \boldsymbol{t}^0{}_{\#}v$ and $\gamma_1 = \boldsymbol{t}^1{}_{\#}v$, respectively, and define $\gamma_s := ((1-s)\boldsymbol{t}^0 + s\boldsymbol{t}^1)_{\#}v$. It then follows that $g(s) := W_2^2(\gamma_s, v)$ is a quadratic polynomial in s, with $g'' \equiv G := \int |\boldsymbol{t}^0(x) - \boldsymbol{t}^1(x)|^2 \, \mathrm{d}v(x) \geq W_2^2(\gamma_0, \gamma_1)$. On the other hand, $h(s) := W_2^2(\gamma_s, w)$ is not a quadratic polynomial in general, but satisfies the concavity inequality [1, Proposition 3.9.12]

$$h(s) \geq (1-s)h(0) + sh(1) - \frac{G}{2}s(1-s).$$

In summary, the metric terms in $F_\tau^{(2)}(\gamma_s; v, w)$ give a contribution that is convex of modulus $\frac{3}{2\tau}G \geq \frac{3}{2\tau}W_2^2(\gamma_1, \gamma_0)$. Finally, since we assumed that F is λ-displacement convex, $s \mapsto F(\gamma_s)$ is convex of modulus λG. This verifies (5).

References

[1] L. Ambrosio, N. Gigli, G. Savaré, *Gradient flows: in metric spaces and in the space of probability measures*, Springer Science& Business Media, 2008.

[2] D. Matthes, S. Plazotta, *A Variational Formulation of the BDF2 Method for Metric Gradient Flows*, Preprint available at `arXiv:1711.02935`.

From discrete Moran-Wright-Fisher processes to continuous Kimura equations by optimal transport

Léonard Monsaingeon

(joint work with Fabio Chalub, Ana Ribeiro & Max Souza)

In this ongoing work, we establish a new connection between discrete Moran-Wright-Fisher-type Markov processes (birth/death) and continuous Kimura diffusion equations, modeling for instance genetic drift and natural selection in population dynamics. Both models share a particular degeneracy: while in the discrete Markov setting two absorbing states exist, the continuous PDE features implicit absorbing boundary conditions. More precisely, considering a population of N individuals with probability p_i of finding i individuals with allele A (thus $N-i$ with allele B), the discrete model reads

$$\frac{d\mathbf{p}}{dt} = \mathbf{K} \cdot \mathbf{p}(t), \qquad \mathbf{p} = (p_0, \dots, p_N) \tag{MWF}$$

for some particular stochastic kernel $\mathbf{K}$ of the Moran-Wright-Fisher type. Two particular absorbing steady-states $\mathbf{e}_0, \mathbf{e}_N$ correspond to genetically homogeneous populations (only one allele A/B persists). At the continuous level, the fraction of allele A within the population is denoted by $x \in [0,1]$, and the neutral Kimura equation reads

$$\partial_t p = \partial_{xx}^2(F(x)p), \qquad t \geq 0, \, x \in [0,1]. \tag{K}$$

Here $F(x)$ is a positive function encoding the transition probabilities and vanishes on the boundaries $F(0) = F(1) = 0$, typically $F(x) = x(1-x)$. The diffusion in (MWF) therefore degenerates at the boundaries, and $x = 0, 1$ are again absorbing states corresponding to purely homogeneous population. The problem is thus implicitly supplemented with absorbing boundary conditions, one must consider measure-valued solutions with boundary Dirac singularities

$$p(t,x) = a(t)\delta_0(x) + \tilde{p}(t,x)\mathbf{d}x + b(t)\delta_1(x),$$

and two additional conservation laws must be imposed in order for the problem to be well-posed [1].

On the one hand, the convergence

$$\mathbf{p}(t) \to p(t,x) \qquad \text{weakly as } N \to \infty$$

in the limit of large population and in some appropriate sense was proved by Chalub and Souza [1]. The discrete absorbing states $i = 0, N$ naturally become absorbing boundary conditions at $x = 0, 1$ in this limit.

On the other hand, the optimal transport theory and associate discrete and continuous Wasserstein metric structures allow to interpret some diffusion problems as Wasserstein gradient flows [3, 4]. For the biologically relevant Moran-Wright-Fisher kernel $\mathbf{K} = (K_{ij})$ and Kimura diffusion $F(x)$, both (MWF) and (K) look very much like such Wasserstein gradient flows in the space of discrete/continuous probability measures, respectively. However, the absorbing states and boundary conditions prevent from applying the classical theory due to singularity issues, and neither (MWF) nor (K) can be directly interpreted as gradient flows. Instead, we show that the internal components (i-e $(p_i)_{1\dots N}$ for the discrete Markov chain or the absolutely continuous part $\tilde{p}(t,x)$ for the PDE) can be suitably rescaled in time with respect to some natural quasi-stationary distributions, and that the rescaled dynamics can be seen respectively as discrete and continuous Wasserstein gradient flows. In the rescaled q variable the driving functionals turn out to be the Boltzmann-Shannon-Gibbs entropy,

$$H_N(\mathbf{q}|\boldsymbol{\pi}_N) = \sum_{i=1}^{N} \frac{q_i}{\pi_i} \log\left(\frac{q_i}{\pi_i}\right) \pi_i$$

for the discrete case and

$$\mathcal{H}(q|\pi) = \int_0^1 \frac{\mathbf{d}q}{\mathbf{d}\pi} \log\left(\frac{\mathbf{d}q}{\mathbf{d}\pi}\right) \mathbf{d}\pi(x)$$

for the continuous case, where $\boldsymbol{\pi}_N = (\pi_i)_{i=1\dots N}$ and $\pi(x)$ are corresponding stationary measures. Furthermore, as $N \to \infty$, clearly $H_N(.|\boldsymbol{\pi}_N) \to \mathcal{H}(.|\pi)$ at least in the sense of Gamma-convergence, and the discrete Wasserstein distance $\mathcal{W}_N$ converges to the continuous Wasserstein distance $\mathcal{W}$ as $N \to \infty$ [2].

To summarize, our contribution is twofold: first, we show that the discrete and continuous dynamics enjoy similar variational structures, respectively induced by discrete and continuous optimal transport after quasi-stationary rescaling. Second,

the two variational structures are compatible with each other (in the sense of Γ-convergence of gradient-flows), since the driving functionals and metric structures also converge in the large population limit.

References

[1] F. Chalub and M. Souza. From discrete to continuous evolution models: a unifying approach to drift-diffusion and replicator dynamics. *Theor. Pop. Biol.*, 76(4):268–277, 2009.

[2] Nicola Gigli and Jan Maas. Gromov–hausdorff convergence of discrete transportation metrics. *SIAM Journal on Mathematical Analysis*, 45(2):879–899, 2013.

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

[4] Jan Maas. Gradient flows of the entropy for finite Markov chains. *J. Funct. Anal.*, 261(8):2250–2292, 2011.

Instabilities and oscillations in coagulation equations

Barbara Niethammer

(joint work with M. Bonacini, M. Herrmann, J.J.L. Velázquez)

We consider the coagulation equation

$$\partial_t f(t,\xi) = \tfrac{1}{2}\int_0^\xi K(\xi-\eta,\eta) f(t,\xi-\eta) f(t,\eta)\,d\eta - f(t,\xi)\int_0^\infty K(\xi,\eta) f(t,\eta)\,d\eta\,, \tag{1}$$

where $f(t,\xi)$ denotes the number density of clusters of size $\xi \in (0,\infty)$ at time $t > 0$. It describes the formation of larger clusters by binary coagulation of smaller ones. The microscopic details of the coagulation process are subsumed into the rate kernel K, which is typically homogeneous. It is well-known that for kernels of homogeneity larger than one gelation occurs, that is the loss of mass at finite time, while otherwise solutions conserve the mass if it is initially finite. A topic of particular interest in this latter case is whether the large-time behaviour is described by self-similar solutions. While this issue is well-understood for the constant and the additive kernel, for other kernels only few results are available. In the case of kernels with homogeneity strictly smaller than one, existence results for self-similar solutions have been established for a large class of kernels, while uniqueness and convergence to self-similar form has only recently been proved for some special cases [2].

In the case of kernels with homogeneity one, it turns out that one needs to distinguish two cases. In the first case, called class II kernels by Van Dongen and Ernst, one has $K(\xi,1) \to c_0 > 0$ as $\xi \to 0$, as for example for the solvable additive kernel $K(\xi,\eta) = \xi+\eta$. On the other hand, kernels that satisfy $\lim_{\xi\to 0} K(\xi,1) = 0$, are called class-I kernels. In order to explain more details and the difference between the two cases it is useful to rewrite (1) in conservative form, that is as

$$\partial_t\big(\xi f\big) + \partial_\xi\Big(\int_0^\xi \int_{\xi-\eta}^\infty K(\eta,\rho)\eta f(\eta) f(\rho)\,d\rho\,d\eta\Big) = 0\,. \tag{2}$$

With the change of variables $\xi = e^x$ and $u(t,x) = \xi^2 f(t,\xi)$ equation (2) becomes

$$\partial_t u(t,x) = -\partial_x\Big(\int_{-\infty}^{x} dy \int_{x+\ln(1-e^{y-x})}^{\infty} dz\, K(e^{y-z},1)u(t,y)u(t,z)\Big). \tag{3}$$

Notice also that $\int_\infty \xi f(t,\xi)\,d\xi = \int_{\mathbb{R}} u(t,x)\,dx$. Hence, mass conserving solutions to (1) correspond to nonnegative integrable solutions of (3) with conserved L^1-norm.

Equation (3) suggests to look for special solution of the form $u(t,x) = G(x-bt)$, which leads to the task of finding $b \in \mathbb{R}_+$ and a nonnegative function G satisfying

$$bG(x) = \int_{-\infty}^{0}\int_{\ln(1-e^y)}^{\infty} K(e^{y-z},1)G(x+y)G(x+z)\,dz\,dy \tag{4}$$

In fact, such special traveling wave type solutions correspond to self-similar solutions in the ξ-variable. We first notice that (4) has a nontrivial constant solution $c_0 > 0$ only if the integral $\int_{-\infty}^{0}\int_{\ln(1-e^y)}^{\infty} K(e^{y-z},1)\,dz\,dy$ is finite. This is the case for class I kernels, but not for class II kernels. In fact, for class I kernels the only consistent behaviour of solutions G of (4) is $\lim G(x) \to c_0$ as $x \to -\infty$. As a consequence, there can in this case be no nonnegative solutions to (4) with finite mass.

In the case of class II kernels we can expect that a family of solutions exist, just as for the additive kernel. There it is known that for any $b \geq 2$ there exists a solution G_b with $M = 1$, such that $G_b(x) \sim e^{-\rho x}$ with $b = \frac{\rho+1}{\rho}$ for $b > 2$ and $G_2(x) \sim e^{-\frac{1}{2}e^{-x}}$. In [1] we provide the first existence results of solutions to (4) for nonsolvable kernels. Under the assumption that K is continuous, symmetric, homogeneous of degree one and with the asymptotics $K(\xi,1) \sim 1 + \beta\xi^\alpha$ for some $\beta \in \mathbb{R}$ and $\alpha \in (0,1)$ we prove the following: for any sufficiently large b there exists a solution with decay $G_b(x) \sim e^{-\rho x}$, where also $b = \frac{\rho+1}{\rho}$. If $\beta > 0$, then there is no nonnegative solution for b sufficiently close to one. However, based on the insights gained in the proof of the latter result, we also conjecture that if $\beta < 0$, then there exists a solution for any $b \geq 1$ (with some specific asymptotics in the limit case $b = 1$).

We now return to class I kernels where we have seen that no nonnegative traveling wave solution with finite mass can exist. To get an idea what happens in the long-time limit of solutions with finite mass we consider the rescaled function $u_\varepsilon(\tau,\tilde{x}) = \frac{1}{\varepsilon}u(\frac{\tau}{\varepsilon^2},\frac{\tilde{x}}{\varepsilon})$, where $0 < \varepsilon \ll 1$, and find

$$\begin{aligned}\partial_\tau u_\varepsilon(\tau,\tilde{x}) &= -\partial_{\tilde{x}}\Big(\frac{1}{\varepsilon^2}\int_{-\infty}^{0}\int_{\varepsilon\ln\left(1-e^{\frac{y}{\varepsilon}}\right)}^{\infty} K\big(e^{\frac{y-z}{\varepsilon}},1\big)u_\varepsilon(\tau,\tilde{x}+y)u_\varepsilon(\tau,\tilde{x}+z)\,dz\,dy\Big)\\ &\approx -c_0\partial_{\tilde{x}}\big(u_\varepsilon(\tau,\tilde{x})^2\big)\end{aligned} \tag{5}$$

with $c_0 = \int_{-\infty}^{0}\int_{\ln(1-e^y)}^{\infty} K\big(e^{y-z},1\big)\,dy\,dz < \infty$. Hence, we conclude that u_ε approximately solves the Burgers equation. Recall that for integrable nonnegative data with mass M the solutions to the Burgers equation $\partial_t u + \partial_x\big(u^2\big) = 0$ converge in the long-time limit to an N-wave with the same mass, i.e. $u(t,x) \sim \frac{1}{\sqrt{t}}N(x/\sqrt{t})$

with $N(x) = \frac{x}{2}\chi_{[0,2\sqrt{M}]}$. On the other hand, if one starts with data such that $u_0(x) \to u_{-\infty} > 0$ as $x \to -\infty$, the solution converges to a traveling wave with height $u_{-\infty}$ and related speed. In order to investigate, whether the long-time behaviour of solutions to (3) is indeed similar as for the Burgers equation we performed in [3] numerical simulations of (3) for a family of kernels

$$K_\alpha(\xi,\eta) = c_\alpha(\xi+\eta)^{1-2\alpha}\xi^\alpha\eta^\alpha\,, \qquad \alpha > 0 \tag{6}$$

with a suitable normalization constant $c_\alpha > 0$. Figure 1 shows the simulations for two values of α for smooth initial data with compact support. We see that solutions

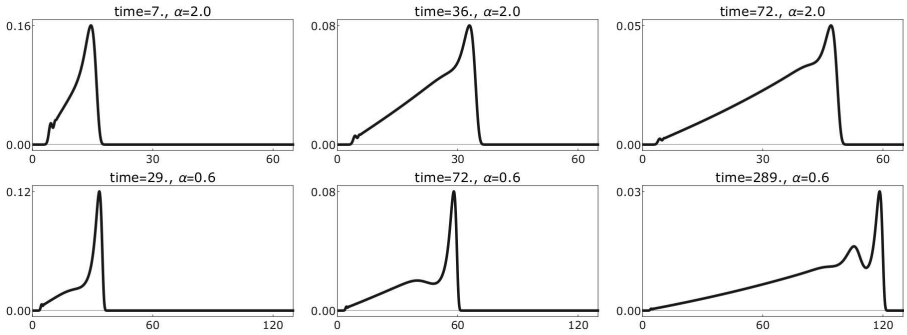

FIGURE 1. Convergence to the N-wave for $\alpha = 2$ (top), and $\alpha = 0.6$ (bottom).

indeed converge to an N-wave in the long-time limit, but that the transition at the shock front is oscillatory, with oscillations becoming stronger when α becomes smaller. One expects that the transition at the shock front is given by a rescaled traveling wave profile and simulations of (3) for Riemann data (see Figure 2) indeed confirm that the traveling wave profiles are oscillatory for small α. In [4] we construct by formal matched asymptotic expansions such traveling wave solutions for sufficiently small α.

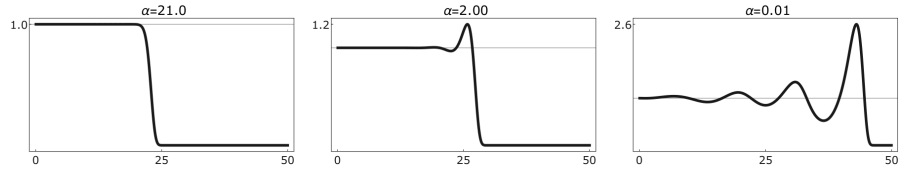

FIGURE 2. Shape of traveling wave for different values of α.

We also perform in [3] a linear stability analysis of the constant solution. It turns out that, surprisingly, for $\alpha > \alpha_{crit} \approx 35$ the constant solution is unstable, whereas for smaller α it is stable. This observation gives rise to the conjecture, that also the N-wave and the traveling wave are unstable for the same range of

α. Correspondingly we also expect that for large α, and more generally for other kernels that concentrate near the diagonal, that the long-time behaviour is not universal, and no convergence to a specific profile takes place.

References

[1] M. Bonacini, B. Niethammer, and J. J. L. Velázquez, *Self-similar solutions to coagulation equations with time-dependent tails: the case of homogeneity one*, (2016), Preprint, arxiv:1612.06610.

[2] P. Laurençot, B. Niethammer, and J. J. L. Velázquez, *Oscillatory dynamics in Smoluchowski's coagulation equation with diagonal kernel*, (2016), Preprint, arxiv:1603:02929.

[3] M. Herrmann, B. Niethammer, and J. J. L. Velázquez, *Instabilities and oscillations in coagulation equations with kernels of homogeneity one* Quart. Appl. Math. **75 (1)** (2017), 105–130.

[4] B. Niethammer, and J. J. L. Velázquez, *Oscillatory travelling wave solutions for coagulation equations*, Quart. Appl. Math, (2017) DOI: https://doi.org/10.1090/qam/1478

Is there a structure-embodying variational principle for nonequilibrium thermodynamics?

Hans Christian Öttinger

The complete structure of Hamiltonian dynamics is embodied in a variational principle, known as Hamilton's principle (principle of stationary or least action). In particular, this variational principle embodies the Poisson bracket structure and can be used to construct symplectic integrators as the stationary points of a discrete action obtained from a discrete Lagrangian [6]. The existence of this variational principle relies on variational self-adjointness, which is deeply related to symplectic structure [13]. The construction of the Lagrangian involves the inverse of the symplectic matrix so that the existence of a variational principle seems to be restricted to nondegenerate Poisson brackets.

Hamiltonian structure is the hallmark of reversible dynamics. In nonequilibrium thermodynamics, we are interested in the combination of reversible and irreversible dynamics, where a fairly abstract, geometric formulation is given by the GENERIC framework (general equation for the nonequilibrium reversible-irreversible coupling) [5, 12, 10]. Among mathematicians, this structure is also known as metriplectic [8, 9]. Important open questions are: Is there a variational principle that embodies the complete GENERIC structure? Could such a variational principle be used to construct structure-preserving GENERIC integrators [11]?

In GENERIC, the reversible contribution to dynamics possesses a Hamiltonian structure with degenerate Poisson bracket; degeneracy is essential because it allows for the existence of an entropy that is conserved under reversible dynamics. The irreversible contribution to dynamics is associated with a gradient flow, for which there exist several mathematical formulations [1]. In particular, gradient flows can be associated with dissipation potentials and a variational principle. The full dynamics can alternatively be characterized by contact Hamiltonian dynamics in

a larger space in which the physical states are characterized by a contact form [5, 3, 7, 4, 2], where the contact Hamiltonian is given in terms of the entropy production. Legendre transformations are to contact Hamiltonian dynamics what canonical transformations are to Hamiltonian dynamics.

Hamiltonian flows focus on reversible dynamics, contact Hamiltonian flows focus on irreversible dynamics. Both of them are associated with variational principles. What is missing is a single variational principle embodying all the elements of the GENERIC structure. From a more practical perspective, what is missing is a GENERIC integrator that preserves all the structural elements of combined reversible and irreversible dynamics in numerical time integration.

References

[1] S. Arnrich, A. Mielke, M. A. Peletier, G. Savaré, and M. Veneroni. Passing to the limit in a Wasserstein gradient flow: From diffusion to reaction. *Calc. Var.*, 44:419–454, 2012.

[2] A. Bravetti, H. Cruz, and D. Tapias. Contact Hamiltonian mechanics. *Ann. Phys.*, 376:17–39, 2017.

[3] D. Eberard, B. M. Maschke, and A. J. van der Schaft. An extension of Hamiltonian systems to the thermodynamic phase space: Towards a geometry of nonreversible processes. *Rep. Math. Phys.*, 60:175–198, 2007.

[4] S. Goto. Legendre submanifolds in contact manifolds as attractors and geometric nonequilibrium thermodynamics. *J. Math. Phys.*, 56:073301, 2015.

[5] M. Grmela and H. C. Öttinger. Dynamics and thermodynamics of complex fluids. I. Development of a general formalism. *Phys. Rev. E*, 56:6620–6632, 1997.

[6] J. E. Marsden and M. West. Discrete mechanics and variational integrators. *Acta Numerica*, 10:357–514, 2001.

[7] J. Merker and M. Krüger. On a variational principle in thermodynamics. *Continuum Mech. Thermodyn.*, 25:779–793, 2013.

[8] P. J. Morrison. A paradigm for joined Hamiltonian and dissipative systems. *Physica D*, 18:410–419, 1986.

[9] P. J. Morrison. Structure and structure-preserving algorithms for plasma physics. *Phys. Plasmas*, 24:055502, 2017.

[10] H. C. Öttinger. *Beyond Equilibrium Thermodynamics*. Wiley, Hoboken, 2005.

[11] H. C. Öttinger. GENERIC integrators: Structure preserving time integration for thermodynamic systems. *J. Non-Equilib. Thermodyn.*, submitted, 2017.

[12] H. C. Öttinger and M. Grmela. Dynamics and thermodynamics of complex fluids. II. Illustrations of a general formalism. *Phys. Rev. E*, 56:6633–6655, 1997.

[13] R. M. Santilli. *The Inverse Problem in Newtonian Mechanics*, volume I of *Foundations of Theoretical Mechanics*. Springer, Berlin, 1978.

Brakke - de Giorgi - Osher

Felix Otto

(joint work with Tim Laux)

Formally, the evolution of the boundary Σ of a set $\Omega \subset \mathbb{R}^d$ by its mean curvature H, meaning that we have for the normal velocity $2V = H$ (the factor of two is convenient for later), is a gradient flow of the surface energy $E = \mathcal{H}^{d-1}(\Sigma) = \int |\nabla \chi|$ (where χ is the characteristic function of Ω) wrt to the $2 \times L^2(\Sigma)$ inner

product. This follows from the fact that the first (inner) variation of the area functional wrt to a test vector field ξ is given by

$$\delta E(\chi,\xi)=\int_\Sigma H\nu\cdot\xi=\int_\Sigma \nabla^{tan}\cdot\xi=\int(\nabla\cdot\xi-\nu\cdot D\xi\nu)|\nabla\chi|, \tag{1}$$

where in the last expression, ν is the measure-theoretic normal $\frac{d\nabla\chi}{d|\nabla\chi|}$.

In view of this gradient flow structure of mean curvature flow (MCF), it is natural to study a minimizing movements (MM) scheme (with time-step size h), ie

$$\chi^n \text{ minimizes } \quad \frac{1}{2h}d^2(\chi,\chi^{n-1})+E(\chi).$$

Since the induced distance degenerates (Michor&Mumford '06, see [1] for the precise references), Almgren&Taylor&Wang introduced in '93 the proxy $d^2(\chi,\chi^0)=4\int_{\Omega\triangle\Omega^0}\mathrm{dist}(\cdot,\Sigma^0)$, and Luckhaus&Sturzenhecker established in '95 a convergence result for the piecewise constant interpolation χ_h of $\{\chi^n\}_n$ in time, under the assumptions that (for a subsequence $h\downarrow 0$)

$$\chi_h\to\chi \text{ in } L^1 \quad \text{and} \quad \int\int|\nabla\chi_h|dt\to\int\int|\nabla\chi|dt, \tag{2}$$

where the second assumption, as opposed to the first, does not follow from the natural a priori estimates and makes the convergence result a conditional one.

However, it is not known whether this limit satisfies the dissipation inequality $\frac{d}{dt}2\mathcal{H}^{d-1}(\Sigma)+\int_\Sigma H^2\le 0$, which in view of (1) can be weakly encoded by

$$\frac{d}{dt}\int|\nabla\chi|+\int(\nabla\cdot\xi-\nu\cdot D\xi\nu)|\nabla\chi|-\frac{1}{2}\int(\xi\cdot\nu)^2|\nabla\chi|\le 0 \tag{3}$$

distributionally in t for all test vector fields ξ.

In this talk based on [2], we show that a computationally highly efficient and robust scheme, the thresholding scheme introduced by Merriman&Bence&Osher in '92, produces limits that satisfy even all localized versions of (3), cf (5), which actually characterizes MCF as pointed out by Brakke in '78. In doing so, we use the tools introduced by de Giorgi for MM in metric spaces, as presented by Ambrosio-Gigli-Savaré '04.

The thresholding scheme proceeds in two steps: First, the characteristic function χ^{n-1} is convolved by the centered Gaussian of variance h, ie $G_h(z):=\frac{1}{\sqrt{2\pi h}}\exp(-\frac{|z|^2}{2h})$. Second, χ^n is the characteristic function of the set where this convolution is larger than $\frac{1}{2}$:

$$\chi^{n-1}\rightsquigarrow u^n:=G_h*\chi^{n-1}\rightsquigarrow\chi^n:=I(u^n>\frac{1}{2}). \tag{4}$$

This scheme clearly preserves the comparison principle and thus its limits have been shown to satisfy MCF in the sense of viscosity solutions (Evans '92, Barles & Georgelin '93). Its main practical interest lies in its easy extension to the multiphase case, which we can treat by the very same analysis.

The localized dissipation inequality takes the form of $\frac{d}{dt}2\int_\Sigma \zeta + \int_\Sigma(\zeta H^2 + \nabla\zeta\cdot\nu H)\le 0$ for any test function $\zeta\ge 0$. Brakke observed that this family of inequalities characterizes a smooth solution of MCF, and used it to introduce a weak notion of solution based on varifolds. On our BV-level, this takes the form of

$$\begin{aligned}\frac{d}{dt}\int\zeta|\nabla\chi| + \int(\nabla\cdot(\zeta\xi) - \nu\cdot D(\zeta\xi)\nu)|\nabla\chi| - \frac{1}{2}\int\zeta(\xi\cdot\nu)^2|\nabla\chi| \\ + \int(\triangle\zeta - \nu\cdot D^2\zeta\nu)|\nabla\chi| \le 0 \qquad (5)\end{aligned}$$

distributionally in t, for all test vector fields ξ and test functions $\zeta\ge 0$.

It was observed in Esedoglu&O. '15 that thresholding can be interpreted as a minimizing movement scheme. Indeed, (4) is equivalent to

$$\chi^n \text{ minimizes among all } v=v(x)\in[0,1] \quad \frac{1}{2h}d_h^2(v,\chi^{n-1}) + E_h(v), \qquad (6)$$

where $E_h(v) := \frac{1}{\sqrt{h}}\int(1-v)G_h * v$ Γ-converges to $c_0\int|\nabla\chi|$ (wrt to underlying L^1-convergence with the understanding that the value is $+\infty$ if the limit is not a $\{0,1\}$-valued function of bdd variation), and where the distance function is given by $\frac{1}{2}d_h^2(v,w) := \sqrt{h}\int(G_{\frac{h}{2}} * (v-w))^2$. It is this variational interpretation that allowed for an extension of thresholding to the multi-phase case with surface tensions that depend on the phases they separate. It also allowed for the first convergence result to a BV-type solution in the multi-phase case [1], however conditional like Luckhaus&Sturzenhecker, namely under condition (7). The result of this talk is

Theorem. Let $\chi^0=\chi^0(x)\in\{0,1\}$ be st $\int|\nabla\chi^0|<\infty$. Suppose that for a subsequence $h\downarrow 0$ we have the analogue of (2), that is,

$$\chi_h\to\chi \text{ in } L^1 \quad\text{and}\quad \int E_h(\chi_h)dt \to \int c_0\int|\nabla\chi|dt. \qquad (7)$$

Then χ satisfies (5).

The proof relies on the fact that next to (6), thresholding satisfies a MM principle for every localizing (test) function $\zeta\ge 0$:

$$\chi^n \text{ minimizes} \quad \frac{1}{2h}\tilde{d}_h^2(v,\chi^{n-1}) + \tilde{E}_h(v,\chi^{n-1}),$$

where $\frac{1}{2}\tilde{d}_h^2(v,w) = \sqrt{h}\int\zeta(G_{\frac{h}{2}} * (v-w))^2$ is the localization of $\frac{1}{2}d_h^2(v,w)$ but $\tilde{E}_h(\cdot,\chi^{n-1})$ is a perturbed (next to localized) version of E_h:

$$\begin{aligned}\tilde{E}_h(v,\chi) := \frac{1}{\sqrt{h}}\int\big(&\zeta(1-v)G_h * v \\ &+ (v-\chi)[\zeta, G_h*](1-\chi) + (v-\chi)[\zeta, G_{\frac{h}{2}}*]G_{\frac{h}{2}} * (v-\chi)\big).\end{aligned}$$

This MM principle allows to appeal to de Giorgi's tools for MM, yielding a discrete version of the dissipation inequality in form of

$$\tilde{E}_h(\chi^n,\chi^{n-1})-\tilde{E}_h(\chi^{n-1},\chi^{n-1})$$
$$+\frac{1}{2}|\partial\tilde{E}_h(\cdot,\chi^{n-1})|^2(\chi^n)+\frac{1}{2h}\int_{(n-1)h}^{nh}|\partial\tilde{E}_h(\cdot,\chi^{n-1})|^2(u_h(t))dt\le 0,$$

where $|\partial E|(v):=\limsup_{w\to v}\frac{(E(v)-E(w))_+}{d(v,w)}$ denotes the metric slope of a functional E and $u_h((n-1)h+\tau):=\operatorname{argmin}\big(\frac{1}{2\tau}d^2(v,\chi^{n-1})+E(v)\big)$ denotes the variational interpolation of χ^{n-1} and χ^n in a MM scheme for metric d and functional E.

Thanks to these tools, the proof is essentially soft and relies on the following ingredients. The metric slope is bounded by below by the first variation:

$$\frac{1}{2}|\partial\tilde{E}_h(\cdot,\chi)|^2(v)\ge\delta\tilde{E}_h(\cdot,\chi)(v,\xi)-\sqrt{h}\int\zeta(G_{\frac{h}{2}}*(\xi\cdot\nabla v))^2.$$

For the first variation, the relation of $\tilde{E}_h$ to E_h simplifies asymptotically

$$\big|\delta\tilde{E}_h(\cdot,\chi)(v,\xi)-\delta E_h(v,\zeta\xi)\big|\le Ch^{\frac{1}{4}}\frac{d_h(v,\chi)}{h},$$

where C is a generic constant only depending on ζ and ξ. The increments of $\tilde{E}_h(v,\chi)$ in χ are related to the first variation of the squared metric d_h^2:

$$\big|\frac{1}{h}(\tilde{E}_h(v,\chi)-\tilde{E}_h(v,v))+\frac{1}{2h}\delta d_h^2(\cdot,\chi)(v,\xi)\big|\le C\big(\sum_{p=1}^{2}(h^{\frac{1}{4}}\frac{d_h(v,\chi)}{h})^p+h^{\frac{1}{2}}E_h(\chi)\big),$$

which has to be combined with the Euler-Lagrange equation for (6). Elementary arguments show

$$\big|\delta E_h(v,\xi)-\frac{1}{\sqrt{h}}\int D\xi:(1-v)(G_h\mathrm{id}-hD^2G_h)*v\big|\le Ch^{\frac{1}{2}}E_h(v),$$
$$\big|\sqrt{h}\int\zeta(G_{\frac{h}{2}}*(\xi\cdot\nabla v))^2-\frac{1}{\sqrt{h}}\int\zeta\xi\otimes\xi:(1-v)hD^2G_h*v\big|\le Ch^{\frac{1}{2}}E_h(v),$$

so that the result relies on the assumption (7) only in the form that $v\to\chi$ in L^1 with $\int(1-v)G_h*v\to c_0\int|\nabla\chi|$ implies

$$\frac{1}{\sqrt{h}}\int\big(\zeta(1-v)G_h*v+\sigma:(1-v)h^2D^2G_h*v\big)\to c_0\int(\zeta+\nu\cdot\sigma\nu)|\nabla\chi|.$$

References

[1] T. Laux, F. Otto, *Convergence of the thresholding scheme for multi-phase mean curvature flow*, Calc. Var. **55** (2016), 1–74.

[2] T. Laux, F. Otto, *Brakke's inequality for the thresholding scheme*, arXiv:1708.03071.

A blob method for diffusion

Francesco Saverio Patacchini
(joint work with José Antonio Carrillo, Katy Craig)

For a range of partial differential equations (PDEs), from the heat and porous medium equations to the Fokker–Planck and Keller–Segel equations, solutions can be characterized as *gradient flows* with respect to the *quadratic Wasserstein distance*. In particular, solutions to the PDE

$$\partial_t \rho = \underbrace{\nabla \cdot (\rho \nabla V)}_{\text{confinement}} + \underbrace{\nabla \cdot (\rho \nabla W * \rho)}_{\text{interaction}} + \underbrace{\Delta \rho^m}_{\text{diffusion}}, \tag{1}$$

where $V : \mathbb{R}^d \to \mathbb{R}$, $W : \mathbb{R}^d \to \mathbb{R}$, $m \geq 1$, and ρ is a curve in the space of probability measures, are formally *Wasserstein gradient flows* of the energy

$$\mathcal{E}(\rho) = \int_{\mathbb{R}^d} V \, d\rho + \tfrac{1}{2} \int_{\mathbb{R}^d} W * \rho \, d\rho + \mathcal{F}^m(\rho), \ \mathcal{F}^m(\rho) = \begin{cases} \int_{\mathbb{R}^d} \rho \log \rho, & m = 1, \rho \ll \mathcal{L}^d, \\ \int_{\mathbb{R}^d} \frac{\rho^m}{m-1}, & m > 1, \rho \ll \mathcal{L}^d, \\ +\infty & \text{otherwise}, \end{cases} \tag{2}$$

where $\mathcal{L}^d$ is d-dimensional Lebesgue measure. This implies solutions ρ of (1) satisfy

$$\partial_t \rho = -\nabla_{W_2} \mathcal{E}(\rho), \tag{3}$$

for a generalized notion of gradient ∇_{W_2}, which is formally given by

$$\nabla_{W_2} E(\rho) = -\nabla \cdot \left(\rho \nabla \tfrac{\delta \mathcal{E}}{\delta \rho} \right), \tag{4}$$

where $\delta \mathcal{E} / \delta \rho$ is the first variation density of $\mathcal{E}$ at ρ. The theory of Wasserstein gradient flows has inspired new numerical methods, with a common goal of maintaining the gradient flow structure at the discrete level, albeit in different ways. One common strategy for preserving the gradient flow structure at the discrete level is to leverage the variational scheme introduced in [9] by working with different discretizations of the space of Lagrangian maps [11], by reconstructing Dirac masses over nonoverlapping balls [5, 4], or by making use of convex analysis and computational geometry [1], amongst many others.

Here we develop a deterministic particle method for Wasserstein gradient flows. The simplest particle method for (1), in the absence of diffusion, consists first in discretizing the initial datum ρ_0 as a finite sum of Dirac masses, i.e.,

$$\rho_0 \approx \rho_0^N = \textstyle\sum_{i=1}^N m_i \delta_{x_i}, \qquad x_i \in \mathbb{R}^d, \quad m_i \geq 0, \tag{5}$$

where δ_{x_i} is the Dirac mass centered at $x_i \in \mathbb{R}^d$. Without diffusion and provided sufficient regularity of V and W, the solution ρ^N of (1) with initial datum ρ_0^N remains a sum of Dirac masses at all times t, so that

$$\rho^N(t) = \textstyle\sum_{i=1}^N m_i \delta_{x_i(t)}, \tag{6}$$

and solving (1) reduces to solving the system of ordinary differential equations

$$\dot{x}_i = -\nabla V(x_i) - \sum_{j=1}^{N} m_j \nabla W(x_i - x_j), \qquad i \in \{1, \dots, N\}. \tag{7}$$

The particle solution ρ^N is the Wasserstein gradient flow of the energy (2) with initial datum ρ_0^N, so in particular the energy decreases in time along this discrete solution. This simple particle method converges to exact solutions of (1) under suitable assumptions on V and W, as has been shown in the mean-field limit sense [2, 3, 8]. In [6] the authors derive a blob method for the aggregation equation with improved rates of convergence to exact solutions for singular interaction potentials W by convolving W with a mollifier. When diffusion is present in equation (1), the fundamental assumption underlying basic particle methods breaks down: particles do not remain particles, that is, the solution to (1) is not of the form (6). One way to circumvent this is by introducing a suitable regularization of the flux of the continuity equation [7, 12, 10]; the main disadvantage of this approach is that, with the exception of the case $m = 2$, the gradient flow structure is not preserved.

The goal here is to introduce a new deterministic particle method for PDEs of the form (1) that respects the problem's underlying gradient flow structure and naturally extends to all dimensions. We follow an approach analogous to that found in [6] for the aggregation equation and regularize the associated internal energy $\mathcal{F}^m$. For a mollifier $\varphi_\varepsilon(x) = \varphi(x/\varepsilon)/\varepsilon^d$, $x \in \mathbb{R}^d$, $\varepsilon > 0$, we define

$$\mathcal{F}_\varepsilon^m(\rho) = \begin{cases} \int_{\mathbb{R}^d} \log(\varphi_\varepsilon * \rho)\, d\rho & \text{for } m = 1, \\ \int_{\mathbb{R}^d} \frac{(\varphi_\varepsilon * \rho)^{m-1}}{m-1}\, d\rho & \text{for } m > 1. \end{cases} \tag{8}$$

As $\varepsilon \to 0$, we prove that $\{\mathcal{F}_\varepsilon^m\}_\varepsilon$ Γ-converges to $\mathcal{F}^m$ for all $m \geq 1$. In the presence of a confining drift or interaction potential, so that minimizers exist, we also show that minimizers converge to minimizers. For $m \geq 2$ and semiconvex potentials $V, W \in C^2(\mathbb{R}^d)$, we show that the gradient flows of the regularized energies $\mathcal{E}_\varepsilon^m$ are well-posed and are characterized by solutions to the PDE

$$\begin{aligned} \partial_t \rho &= \nabla \cdot [(\nabla V + \nabla W * \rho)\rho] \\ &\quad + \nabla \cdot \left[\rho \left(\nabla \varphi_\varepsilon * \left((\varphi_\varepsilon * \rho)^{m-2}\rho\right) + (\varphi_\varepsilon * \rho)^{m-2}(\nabla \varphi_\varepsilon * \rho)\right)\right]. \end{aligned} \tag{9}$$

Under sufficient regularity conditions, we prove that solutions of the regularized gradient flows converge to solutions of (1) using the strategy in [13]. When $m = 2$ and the initial datum has bounded entropy, we show that these regularity conditions automatically hold, thus generalizing Lions and Mas-Gallic's result for the porous medium equation on bounded domains to (1) on all of $\mathbb{R}^d$. For this regularization particles do remain particles, and consequently our numerical method consists of taking a particle approximation for (9). We show that, under sufficient regularity conditions, our particle solutions converge to exact solutions of (1).

A key advantage of our approach is that, by regularizing the energy functional and not the flux, we preserve the problem's gradient flow structure. Further advantages include the ease with which our method may be combined with particle

methods for interaction and confinement potentials, its simplicity in any dimension, and the good numerical performance we observe for a wide choice of potentials; also, at least when $m = 2$, our regularization can be interpreted as an approximation of the porous medium equation by a nonlocal interaction potential,

We give several numerical results, in one and two dimensions, analyzing the rate of convergence with respect to the 2-Wasserstein metric, L^1-norm, and L^∞-norm and illustrating qualitative properties of the method, such as asymptotic behavior of the Fokker–Planck equation and mass criticality of the Keller–Segel equation. In particular, for the heat and porous medium equations ($V = W = 0$, $m = 1, 2, 3$), we observe that the 2-Wasserstein error depends linearly on the grid spacing $h \sim N^{-1/d}$ for $m = 1, 2, 3$, while the L^1-norm depends quadratically on the grid spacing for $m = 1, 2$ and superlinearly for $m = 3$. We conduct a detailed numerical study of equations of Keller–Segel type, including a one-dimensional variant ($V = 0, W = 2\chi \log|\cdot|, \chi > 0, m = 1, 2$) and the classical two-dimensional version ($V = 0$, $W = \Delta^{-1}$, $m = 1$); we show that the well-known mass criticality of these equations is present in our numerical solutions and demonstrate convergence of the critical mass as the grid spacing h and regularization ε are refined.

References

[1] J.-D. Benamou, G. Carlier, Q. Mérigot, and E. Oudet. *Discretization of functionals involving the Monge–Ampére operator*, Numer. Math., **134** (2016), 611–636.

[2] J. A. Carrillo, Y.-P. Choi, and M. Hauray, *The derivation of swarming models: mean-field limit and Wasserstein distances*, In Collective Dynamics from Bacteria to Crowds: An Excursion Through Modeling, Analysis and Simulation, volume 553 of CISM Courses and Lect. (2014), 1–46.

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

[4] J. A. Carrillo, Y. Huang, F. S. Patacchini, and G. Wolansky. *Numerical study of a particle method for gradient flows*, Kinet. Relat. Models, **10** (217), 613–641.

[5] J. A. Carrillo, F. S. Patacchini, P. Sternberg, and G. Wolansky, *Convergence of a particle method for diffusive gradient flows in one dimension*, SIAM J. Math. Anal., **48** (2106), 3708–3741.

[6] K. Craig and A L. Bertozzi, *A blob method for the aggregation equation*, Math. Comp., **85** (2016), 1681–1717.

[7] P. Degond and F.-J. Mustieles, *A deterministic approximation of diffusion equations using particles*, SIAM J. Sci. Statist. Comput., **11** (1990), 293–310.

[8] P.-E. Jabin, *A review of the mean field limits for Vlasov equations*, Kinet. Relat. Models, **7** (2014), 661–711.

[9] R. Jordan, D. Kinderlehrer, and F. Otto, *The variational formulation of the Fokker–Planck equation*, SIAM J. Math. Anal., **29** (9981), 1–17.

[10] P.-L. Lions and S. Mas-Gallic, *Une méthode particulaire déterministe pour des équations diffusives non linéaires*, C. R. Acad. Sci. Paris Sér. I Math., **332** (2001), 369–376.

[11] H. Osberger and D. Matthes, *Convergence of a variational Lagrangian scheme for a nonlinear drift diffusion equation*, ESAIM Math. Model. Numer. Anal., **48** (2014), 697–726.

[12] G. Russo, *Deterministic diffusion of particles*, Comm. Pure Appl. Math., **43** (1990), 697–733.

[13] S. Serfaty, *Gamma-convergence of gradient flows on Hilbert and metric spaces and applications*, Discrete Contin. Dyn. Syst., **31** (2011), 1427–1451.

A variational principle for the incompressible Navier-Stokes system

Ulisse Stefanelli
(joint work with Michael Ortiz and Bernd Schmidt)

The incompressible Navier-Stokes system

$$u_t + u \cdot \nabla u - \nu \Delta u + \nabla p = 0, \quad \operatorname{div} u = 0 \tag{1}$$

describes the flow velocity $u : \Omega \times (0,\infty) \to \Omega \in \mathbb{R}^3$ and the pressure per unit density $p : \Omega \times (0,\infty) \to \mathbb{R}$ of a viscous fluid with kinetic viscosity $\nu > 0$, and we use the notation $u \cdot \nabla u = \sum_{j=1}^{3} u_j \partial_{x_j} u$. We assume the container Ω to be Lipschitz and bounded and impose no-slip boundary conditions ($u = 0$ at $\partial\Omega$). System (1) can be classically reformulated in a weak form by letting H and V to be the closure of $\{u \in C_c^\infty(\Omega;\mathbb{R}^3) \mid \operatorname{div} u = 0\}$ in $L^2(\Omega;\mathbb{R}^3)$ and $H_0^1(\Omega;\mathbb{R}^3)$, respectively, defining the operators A, $B : V \to V'$ (dual of V) as

$$\langle Au, v\rangle = \int_\Omega \nabla u : \nabla v \, \mathrm{d}x, \quad \langle B(u), v\rangle = -\int_\Omega (u \otimes u) : \operatorname{div} v \, \mathrm{d}x \quad \forall u,\, v \in V$$

(where $\langle \cdot, \cdot \rangle$ stands for the duality pairing between V' and V and : denotes the standard contraction) and asking for

$$u_t + B(u) + \nu A u = 0 \quad \text{in } V' \text{ for a.e. } t > 0.$$

A trajectory $u \in L^2_{\mathrm{loc}}(0,\infty;V)$ with $u_t \in L^1_{\mathrm{loc}}(0,\infty;V')$ and $u(\cdot,0) = u_0 \in H$ (given) is called *Leray-Hopf* solution of (1). Such solutions exist in $L^\infty(0,T;H)$ and are weakly continuous in H, so that the initial condition makes sense [18].

We propose a new variational approach to Leray-Hopf solutions. This consists in minimizing the functionals

$$I^\varepsilon(u) = \int_0^\infty \int_\Omega \mathrm{e}^{-t/\varepsilon} \left(\frac{1}{2}|u_t + u \cdot \nabla u|^2 + \frac{\sigma}{2}|u \cdot \nabla u|^2 + \frac{\nu}{2\varepsilon}|\nabla u|^2 \right) \mathrm{d}x \, \mathrm{d}t$$

on the set of admissible trajectories

$$W^\varepsilon := \{v \in L^2_{\mathrm{loc}}(0,\infty;V) \mid v_t,\, u \cdot \nabla u \in L^2_{\mathrm{loc}}(0,\infty;H), v(0) = u_0^\varepsilon\}.$$

Here, $\varepsilon > 0$ is a small parameter, bound to converge to 0, and $\sigma > 1/8$ is a stabilization parameter. The approximating initial values $u_0^\varepsilon \in V$ are chosen in such a way that $u_0^\varepsilon \to u_0$ in H and $\|\nabla u_0^\varepsilon\|_{L^2}^2 + \varepsilon \|u_0^\varepsilon \cdot \nabla u_0^\varepsilon\|_{L^2}^2 \le C\varepsilon^{-1}$ for some $C > 0$.

The relation between the minimization of I^ε and the Navier-Stokes system (1) is revealed by formally computing the Euler-Lagrange equation for I^ε, namely

$$\begin{aligned} 0 &= u_t + u \cdot \nabla u - \nu \Delta u + \nabla p \\ &\quad - \varepsilon (u_t + u \cdot \nabla u)_t - \varepsilon \operatorname{div}((u_t + u \cdot \nabla u) \otimes u) \\ &\quad + \varepsilon \nabla u^\top (u_t + u \cdot \nabla u) - \varepsilon\sigma \operatorname{div}((u \cdot \nabla u) \otimes u) + \varepsilon\sigma \nabla u^\top (u \cdot \nabla u) \end{aligned} \tag{2}$$

so that one recovers (1) by formally taking $\varepsilon \to 0$. In fact, this program can be made rigorous and we have the following.

Variational principle [13]. *The functional I_ε admits minimizers u^ε in W^ε. Up to not relabeled subsequences, $u^\varepsilon \to u$ weakly in $L^2(0,\infty;V)$ where u is a Leray-Hopf solution of the Navier-Stokes system* (1).

Note that the Euler-Lagrange equations of I^ε correspond to an elliptic-in-time regularization of (1). An early reference in this direction is the work by Lions [10], where however no variational structure was available.

The above variational principle falls within the general class of Weighted Inertia-Dissipation-Energy (WIDE) principles. These provide a variational approach to abstract evolutive systems of the form

$$\rho\ddot{\varphi} + \partial_{\dot\varphi} D(\varphi,\dot\varphi) + \partial_\varphi E(\varphi,t) \ni 0 \tag{3}$$

where the trajectory $\varphi : (0,\infty) \to \mathcal{H}$ (Hilbert space) describes the state of the system, $\rho > 0$, $D(\varphi,\dot\varphi)$ is a nonnegative dissipation gradient, convex in the rate $\dot\varphi$ with $D(\cdot,0) = 0$, $E(\varphi,t)$ is a possibly time-dependent energy, and ∂ denotes partial (sub)differentiation. Solutions to (3) can be recovered by finding minimizers φ^ε of the functionals

$$F^\varepsilon(\varphi) = \int_0^\infty \mathrm{e}^{-t/\varepsilon} \left(\frac{\rho}{2}\|\ddot\varphi\|^2_{\mathcal{H}} + \frac{1}{\varepsilon} D(\varphi,\dot\varphi) + \frac{1}{\varepsilon^2} E(\varepsilon,t) \right) \mathrm{d}t$$

among trajectories $\varphi \in H^2_{\mathrm{loc}}(0,\infty;\mathcal{H})$ with given initial conditions $\varphi(0) = \varphi_0$ and $\rho\varphi(0) = \rho\varphi_1$ and by taking the limit $\varphi^\varepsilon \to \varphi$. This variational program has been successfully followed in a number of different parabolic ($\rho = 0$) situations including rate-independent processes ($D(\varphi,\cdot)$ positively 1-homogeneous) [11], gradient flows ($D(\varphi,\cdot)$ quadratic) [12, 14], and doubly nonlinear evolutions ($D(\varphi,\cdot)$ p-homogeneous, $1 < p \neq 2$) [1, 2]. In the dynamic case of semilinear waves ($\rho > 0$), the validity of the principle was conjectured by De Giorgi in [6] and checked in [15, 17]. See [8, 9, 16] for extensions to other situations and [5, 7, 4] for applications to mechanical problems.

Let now $\varphi : \Omega\times(0,\infty) \to \Omega$ describe the motion of the fluid. The incompressible viscous case can be modeled by letting $\mathcal{H} = L^2(\Omega;\mathbb{R}^3)$ and defining

$$D(\varphi,\dot\varphi) = \int_\Omega \rho\nu |\mathrm{dev}(\nabla\dot\varphi(\nabla\varphi)^{-1})^{\mathrm{sym}}|^2$$

('dev' and 'sym' denote respectively the deviatoric and the symmetric part of the strain rate $\nabla\dot\varphi(\nabla\varphi)^{-1}$), $E(\varphi) = 0$ if $\det\nabla\varphi = 1$ a.e. and $E(\varphi) = \infty$ otherwise. By rewriting $F^\varepsilon(\varphi)$ in Eulerian coordinates via the transformation $u = \dot\varphi \circ \varphi^{-1}$ we obtain the functional $I^\varepsilon(u)$ with $\sigma = 0$.

The σ-term in I^ε plays the role of a stabilization. It has no influence on the limiting Navier-Stokes system (see the Euler-Lagrange equations (2)) but provides enhanced coercivity, which is in turn instrumental to prove the compactness of the minimizers u^ε. Such term is reminiscent of analogous stabilizations in computational fluid mechanics [3, 19].

References

[1] G. Akagi, S. Melchionna, U. Stefanelli. Weighted Energy-Dissipation approach to doubly-nonlinear problems on the half line. *J. Evol. Equ.* (2017), to appear.

[2] G. Akagi, U. Stefanelli. Doubly nonlinear evolution equations as convex minimization problems. *SIAM J. Math. Anal.* 46 (2014), 1922–1945.

[3] A. N. Brooks, T. J. R. Hughes. Streamline upwind/Petrov-Galerkin formulations for convection dominated flows with particular emphasis on the incompressible Navier-Stokes equations. FENOMECH '81, Part I (Stuttgart, 1981). *Comput. Methods Appl. Mech. Engrg.* 32 (1982), 199–259.

[4] S. Conti, M. Ortiz. Minimum principles for the trajectories of systems governed by rate problems. *J. Mech. Phys. Solids*, 56 (2008), 1885–1904.

[5] E. Davoli, U. Stefanelli. Dynamic plasticity via convex minimization. Submitted, 2017.

[6] E. De Giorgi. Conjectures concerning some evolution problems. *Duke Math. J.* 81 (1996), 255–268.

[7] C. J. Larsen, M. Ortiz, C. L. Richardson. Fracture paths from front kinetics: relaxation and rate independence. *Arch. Ration. Mech. Anal.* 193 (2009), 539–583.

[8] M. Liero, U. Stefanelli. A new minimum principle for Lagrangian mechanics. *J. Nonlinear Sci.* 23 (2013), 179–204.

[9] M. Liero, U. Stefanelli. Weighted Inertia-Dissipation-Energy functionals for semilinear equations. *Boll. Unione Mat. Ital. (9)*, 6 (2013), 1–27.

[10] J. L. Lions. Sur certaines équations paraboliques non linéaires. (French) *Bull. Soc. Math. France*, 93 (1965), 155–175.

[11] A. Mielke, M. Ortiz. A class of minimum principles for characterizing the trajectories and the relaxation of dissipative systems. *ESAIM Control Optim. Calc. Var.* 14 (2008), 494–516.

[12] A. Mielke, U. Stefanelli. Weighted energy-dissipation functionals for gradient flows. *ESAIM Control Optim. Calc. Var.* 17 (2011), 52–85.

[13] M. Ortiz, B. Schmidt, U. Stefanelli. A variational approach to Navier-Stokes. In preparation, 2017.

[14] R. Rossi, G. Savaré, A. Segatti, U. Stefanelli. A variational principle for gradient flows in metric spaces. *C. R. Math. Acad. Sci. Paris*, 349 (2011), 1224–1228.

[15] E. Serra, P. Tilli. Nonlinear wave equations as limits of convex minimization problems: proof of a conjecture by De Giorgi. *Ann. of Math. (2)*, 175 (2012), 1551–1574.

[16] E. Serra, P. Tilli. A minimization approach to hyperbolic Cauchy problems. *J. Eur. Math. Soc. (JEMS)*, 18 (2016), 2019–2044.

[17] U. Stefanelli. The De Giorgi conjecture on elliptic regularization. *Math. Models Meth. Appl. Sci.*, 21 (2011), 1377–1394.

[18] R. Temam. *Navier-Stokes equations. Theory and numerical analysis.* Revised edition. With an appendix by F. Thomasset. Studies in Mathematics and its Applications, 2. North–Holland Publishing Co., Amsterdam–New York, 1979.

[19] L. Tobiska, R. Verfürth. Analysis of a streamline-diffusion finite element method for the Stokes and Navier-Stokes equations. *SIAM J. Numer. Anal.* 33 (1996), 107–127.

Asymptotical analysis of a weighted very fast diffusion equation arising in quantization of measures via the JKO scheme

Filippo Santambrogio
(joint work with Mikaela Iacobelli, Francesco Saverio Patacchini)

The problem of quantization of a measure consists in finding the atomic measure ν, with fixed number N of atoms, best approximating in the sense of the Wasserstein distance, a given density f in a compact of $\mathbb{R}^d$: $\min\{W_p(\nu, f) : \#\mathrm{spt}(\nu) = N\}$. This is equivalent to

$$\min\{\int d(x,\Sigma)^p f(x)dx : \#\Sigma = N\}.$$

The only unknown are the points of the support Σ of the optimal measure, and a classical question concerns the asymptotics when $N \to \infty$. In this case it is known that the limit density of these points is a minimizer of the energy functional

$$F(\rho) = \int \frac{f}{\rho^{p/d}} dx.$$

This is a classical fact (see [7, 6, 13]), then translated into a Γ-convergence statement (see [4]) in [1] (a more general proof, has been established in a similar case in [14]). See also [11] and [8].

A natural problem is to study the evolution of the points of Σ when they follow the steepest descent curves of the functional they minimize, and to compare it to the gradient flow (in the Wasserstein sense) of F. Notice that the Lagrangian evolution of Σ exactly translates into a W_2-gradient flow for ρ. In [2, 3] a comparison between the two is studied, essentially in 1D or in a very particular 2D case.

In this work we only concentrate on the limit equation, the gradient flow of F. This equation reads

$$\partial_t \rho - \nabla \cdot (\rho \nabla U'(\frac{\rho}{m})) = 0,$$

where $m = f^{1/(\alpha+1)}$ and $U(s) = s^{-\alpha}$ with $\alpha = p/d$. The function m is usually supposed to be bounded from below and from above, and, when needed, smooth. This can be considered as a very fast diffusion equation (because of the negative exponent in U, while fast diffusion corresponds to exponents between 0 and 1), and we call it weighted because of the presence of the non-constant weight factor m. It was already studied, with particular attention to the asymptotical behavior when $t \to \infty$, in [9], but only in 1D.

The approach that we use is to use the so-called JKO scheme (see [10]), which requires to iteratively solve

$$\min F(\rho) + \frac{W_2^2(\rho, \rho_k)}{2\tau}.$$

We prove some properties of the minimizers of this problem, which can be iterated in k and, when $\tau \to 0$, provide properties of the solution.

In particular we are able to prove

- Upper and lower bounds. If $\rho_k \le a^+$, then the minimizer ρ also satisfies $\rho_k \le a^+$ and, if $\rho_k \ge a^-$, we also have $\rho \ge a^-$. This works independently of the regularity of m.
- BV bounds. If $\rho_k/m \in BV$ and satsifies upper and lower bounds, then we the same is true for ρ/m, provided $\log m$ satisfies $D^2(\log m) \le \lambda I$. More precisely, we have

$$(1-\lambda C\tau)\int |\nabla(\frac{\rho}{m})|m \le \int |\nabla(\frac{\rho_k}{m})|m,$$

  where the constant $C>0$ depends on the upper and lower bound of ρ/m (which are the same as ρ_k/M and hence as ρ_0/m). In particular, when $\lambda<0$, this implies exponential convergence in BV to the steady state and, when $\lambda>0$, implies controlled BV bounds. This result is proven via the technique developed in [5].
- Instantaneous regularization. The third point evoked during the talk concerns the fact that, even if the initial density ρ_0 is such that ρ_0/m is neither L^∞ nor BV (just an inital summability $\rho_0 \in L^q$ for a suitable q is required), these properties appear instantaneously, in the sense that, in the time-continuous equation, for every $t>0$ the L^∞ and the BV norm of ρ_t/m are bounded by a negative power of t. For the BV norm, this is obtained by differentiating in time $G(\rho) := \int (\rho/m)^q m$ and obtaining

$$\frac{d}{dt}\int \left(\frac{\rho_t}{m}\right)^q m = -c\int \left|\nabla\left((\frac{\rho_t}{m})^{(q-\alpha-2)/2}\right)\right|^2 m.$$

  Integrating in time, this proves that we have $L^2(H^1)$ regularity for the function $(\rho/m)^{(q-\alpha-2)/2}$. This implies that, for $t>0$, ρ_t/m is H^1 and hence BV, and the BV regularity is then preserved in time. For the L^∞ regularity, this has to be coupled with a Moser-like iteration (see [15]) estimate, which allows to improve the summability exponent. These computations can be done in the discrete JKO scheme, using the so-called flow interchange technique (taken from [12]), but require the geodesic convexity of G (which is true for m log-concave). Adapatations via approximation exist anyway, even without this assumption. Another difference is that we do not obtain instantaneous L^∞ regularization (because we cannot perform Moser's iteration infinitely many times), but only on estimate on the $L^{p(\tau)}$ norm, where $\lim_{\tau\to 0} p(\tau) = +\infty$.

References

[1] G. Bouchitté, C. Jimenez, M. Rajesh: Asymptotique d'un problme de positionnement optimal, *C. R. Acad. Sci. Paris* Ser. I, **335** (2002) 1–6.

[2] E. Caglioti, F. Golse and M. Iacobelli, *A gradient flow approach to quantization of measures* Math. Models Methods Appl. Sci., 25 (2015), no. 10, 1845-1885.

[3] E. Caglioti, F. Golse and M. Iacobelli, *Quantization of measures and gradient flows: a perturbative approach in the 2-dimensional case.* To appear in Annales de l'Institut Henri Poincaré C, Analyse Non Linéaire

[4] G. Dal Maso: *An Introduction to* $\Gamma-$*convergence.* Birkhauser, Basel, 1992.

[5] G. De Philippis, A. R. Mészáros, F. Santambrogio, B. Velichkov, BV estimates in optimal transportation and applications, *Arch. Rat. Mech. An.* 219 (2), 829-860, 2016.
[6] L. Fejes Tóth: *Lagerungen in der Ebene auf der Kugel und im Raum*, Die Grundlehren der Math. Wiss., Vol. 65, Springer-Verlag, Berlin, 1953.
[7] S. Graf and H. Luschgy. *Foundations of quantization for probability distributions*, volume 1730 of *Lecture Notes in Mathematics*. Springer-Verlag, Berlin, 2000.
[8] M. Iacobelli, *Asymptotic quantization for probability measures on Riemannian manifolds.* ESAIM Control Optim. Calc. Var., 22 (2016), no. 3, 770 ? 785.
[9] M. Iacobelli, *Asymptotic analysis for a very fast diffusion equation arising from the 1D quantization problem.* To appear in Discrete Contin. Dyn. Syst.
[10] R. Jordan, D. Kinderlehrer and F. Otto, *The variational formulation of the Fokker-Planck equation*, SIAM J. Math. Anal. **29** (1998), 1–17.
[11] B. Kloeckner, *Approximation by finitely supported measures*, ESAIM: Control, Optimisation and Calculus of Variations, 18(2), 2012.
[12] 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.
[13] F. Morgan, R. Bolton: Hexagonal Economic Regions Solve the Location Problem, *Amer. Math. Monthly* **109**, no. 2 (2002), 165–172.
[14] S. Mosconi and P. Tilli: $\Gamma-$Convergence for the Irrigation Problem, 2003. *J. of Conv. Anal.* **12**, no.1 (2005), 145–158.
[15] J. Moser, *A new proof of De Giorgi's theorem concerning the regularity problem for elliptic differential equations*, Communications on Pure and Applied Mathematics 13: 457?468, 1980.

The Hellinger-Kantorovich distance as a generalization of optimal-transport distances to scalar reaction-diffusion problems

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

1. Entropy-transport problems

Let $\mu_0 \in \mathcal{M}_{\geq 0}(X)$ and $\mu_1 \in \mathcal{M}_{\geq 0}(Y)$ be nonnegative, finite Borel measures on Polish spaces X and Y, respectively. In particular, we consider the case that μ_0 and μ_1 have different total masses $\mu_0(X) \neq \mu_1(Y)$, and the theory of optimal transport is not meaningful. However, by relaxing the marginal constraint for transport plans and introducing a penalization in the form of entropy functionals we arrive at the new formulation of entropy-transport problems:

$$(1)\quad \mathsf{ET}(\mu_0,\mu_1) = \inf\Big\{ \int_X F_0\Big(\frac{\mathrm{d}\eta_0}{\mathrm{d}\mu_0}\Big)\mathrm{d}\mu_0 + \int_Y F_1\Big(\frac{\mathrm{d}\eta_1}{\mathrm{d}\mu_1}\Big)\mathrm{d}\mu_1 + \int_{X\times Y} \mathsf{c}(x,y)\mathrm{d}\boldsymbol{\eta}\Big\},$$

where $F_i : \mathbb{R}_{\geq 0} \to \mathbb{R}_\infty$ are lower semicontinuous and convex functions with $\mathrm{dom}(F_i)\cap]0,\infty[\neq \emptyset$ to avoid trivial cases and $\mathsf{c} : X\times Y \to [0,\infty]$ is a lower semicontinuous cost function. The infimum runs over all transport plans $\boldsymbol{\eta} \in \mathcal{M}_{\geq 0}(X\times Y)$ with marginals $\eta_i = \pi^i_\sharp \boldsymbol{\eta}$. Under reasonable assumptions, minimizers for (1) exist.

The concave transform $F_i^\circ(\phi) = \inf_{s>0}\{F_i(s) + s\phi\}$ leads to the equivalent dual problem

$$\mathsf{ET}(\mu_0,\mu_1) = \sup\Big\{ \int_X F_0^\circ(\phi_0(x))\mathrm{d}\mu_0 + \int_Y F_1^\circ(\phi_1(y))\mathrm{d}\mu_1 \,\Big|\, \phi_0(x)+\phi_1(y) \le \mathsf{c}(x,y)\Big\}. \tag{2}$$

Using the transformation $\psi_i = F_i^\circ(\phi_i)$ the dual problem can be written in terms of a linear objective functional in ψ_i but with the nonlinear constraint $G_0(\psi_0(x)) + G_1(\psi_1(y)) \le \mathsf{c}(x,y)$, where $G_i(\psi) = \sup_{r>0}\{r\psi - rF_i(1/r)\}$.

We can find a third equivalent formulation in terms of the marginal perspective function, which allows us to formulate the entropy-transport problem as an optimal transport problem on an extended space, where a marginal constraint now has to hold for barycentric projections.

For this, we use the Lebesgue decomposition $\mu_i = \rho_i\eta_i + \mu_i^\perp$ that allows us to rewrite the integrals in (1) as a single integral with respect to η, namely

$$\int_{X\times Y} \Big[\rho_0 F_0\Big(\frac{1}{\rho_0}\Big) \oplus \rho_1 F_1\Big(\frac{1}{\rho_1}\Big) + \mathsf{c}\Big]\mathrm{d}\boldsymbol{\eta} + F_0(0)\mu_0^\perp(X) + F_1(0)\mu_1^\perp(Y),$$

where $(f\oplus g)(x,y) := f(x)+g(x)$. Since we are dealing with a convex minimization problem we can perform a Young measure relaxation by considering the extended spaces $\mathbf{X} = X\times[0,\infty[$ or $\mathbf{Y} = Y\times[0,\infty[$ such that a pair (ρ_i,η_i) corresponds to the measure $\delta_{\rho_i}\otimes\eta_i$. In particular, measures λ on $\mathbf{X}$ and $\mathbf{Y}$ are related to measures μ on X and Y via the barycentric projection $\mu = \mathfrak{P}\lambda = \int_{[0,\infty[} r\mathrm{d}\lambda(r)$. Note that $\mathfrak{P}$ satisfies the scaling invariance $\mathfrak{P}\lambda_\theta = \mathfrak{P}\lambda$, where the rescaled measure is given via $\lambda_\theta = (\mathrm{prd}_\theta)_\sharp(\theta\lambda)$ for $\mathrm{prd}_\theta(x,r) = (x, r/\theta(x))$ and a positive function $\theta : X \to [0,\infty[$. After in inner minimization with respect to positive functions $(x,y)\mapsto\theta(x,y)$, we arrive at the equivalent formulation

$$\mathsf{ET}(\mu_0\mu_1) = \min\Big\{ \int_{\mathbf{X}\times\mathbf{Y}} H(x,r,y,s)\mathrm{d}\boldsymbol{\lambda} \,\Big|\, \mathfrak{P}_i\boldsymbol{\lambda} = \mu_i\Big\} \tag{3}$$

for plans $\boldsymbol{\lambda}\in\mathcal{M}_{\ge0}(\mathbf{X}\times\mathbf{Y})$ and with the marginal perspective function

$$H(x,r,y,s) = \inf_{\theta>0}\Big\{ rF_0\Big(\frac{\theta}{r}\Big) + sF_1\Big(\frac{\theta}{s}\Big) + \theta\mathsf{c}(x,y)\Big\}.$$

Note that $H(x,r,y,s) \to F_0(0)r + F_1(0)s$ for $\mathsf{c}(x,y)\to\infty$ and $H(x,r,y,s) \le F_0(0)r + F_1(0)s$ for every $(x,r)\in\mathbf{X}, (y,s)\in\mathbf{Y}$.

2. The Hellinger–Kantorovich distance

For $X = Y$, a special entropy-transport problem is given by the choice $F_0(z) = F_1(z) = z\log z - z + 1$ and $\mathsf{c}(x,y) = -2\log(\cos(\mathsf{d}_X(x,y)\wedge\pi/2))$ as it induces a distance on the space of nonnegative, finite Borel measures, which can be seen as an interpolation between the Hellinger–Kakutani distance and the Kantorovich–Wasserstein distance and hence shall be called Hellinger–Kantorovich distance HK.

Indeed, the homogeneous marginal perspective function in this case reads

$$H(x_0, r_0^2, x_1, r_1^2) = r_0^2 + r_1^2 - 2r_0 r_1 \cos(\mathsf{d}_X(x_0, x_1) \wedge \pi) =: \mathsf{d}_{\mathfrak{C}}([x_0, r_0], [x_1, r_1])^2. \tag{4}$$

The function $\mathsf{d}_{\mathfrak{C}}$ defines a distance on the cone space over the metric space X. The latter is given by the quotient space $\mathfrak{C}_X := \big(X \times [0, \infty[\big)/_\sim$ with

$$(x_0, r_0) \sim (x_1, r_1) \quad \text{iff } x_0 = x_1,\ r_0 = r_1 > 0, \text{ or } r_0 = r_1 = 0. \tag{5}$$

In particular, the associated entropy-transport problem for μ_0, μ_1 consists of computing the minimal Kantorovich-Wasserstein distance $\mathsf{W}_{\mathfrak{C}}$ among all measures λ_0, λ_1 on the cone space having μ_0 and μ_1 as barycenters. The HK distance inherits many properties from the Kantorovich-Wasserstein distance on the cone.

In the case $X = \mathbb{R}^d$, the Hellinger–Kantorovich distance has an equivalent dynamical formulation in the sense of Benamou–Brenier

$$\mathsf{HK}(\mu_0, \mu_1)^2 = \min \Big\{ \int_0^1 \int_{\mathbb{R}^d} \big[|\mathbf{v}_s|^2 + 4w_s^2\big] \mathrm{d}\mu_s \mathrm{d}s \,\Big|\, \mu_{s=0} = \mu_0,\ \mu_{s=1} = \mu_1 \\ \frac{\mathrm{d}}{\mathrm{d}s}\mu_s + \nabla \cdot (\mathbf{v}_s \mu_s) = 4 w_s \mu_s \Big\}, \tag{6}$$

where the generalized continuity equation for the velocity field $\mathbf{v}_s(x) \in \mathbb{R}^d$ and the scalar field $w_s(x) \in \mathbb{R}$ holds in the distributional sense. In particular, the distance is induced by the Onsager operator $\mathbb{K}_{\mathsf{HK}}(\rho)\xi = -\nabla \cdot (\rho \nabla \xi) + 4\rho\xi$, which gives rise to gradient-flow formulations of scalar reaction-diffusion equations of type

$$\dot\rho_t = -\mathbb{K}_{\mathsf{HK}}(\rho_t) \mathrm{D}\mathcal{F}(\rho_t) \qquad \Longleftrightarrow \qquad \dot\rho_t - \nabla \cdot (\rho_t F''(\rho_t) \nabla \rho_t) + 4\rho_t F'(\rho_t) = 0,$$

where $\mathcal{F}(\rho) = \int F(\rho) \mathrm{d}x$ is the driving functional for the evolution.

3. Convex functionals along Hellinger–Kantorovich geodesics

Similar to the case of the Kantorovich–Wasserstein distance, where geodesic curves are induced by the geodesic curves in the underlying space X, we can find an explicit characterization of geodesics in the Hellinger–Kantorovich space for $X = \mathbb{R}^d$: If $s \mapsto \mu_s$, for $s \in [0,1]$, satisfies $\mathsf{HK}(\mu_{s_0}, \mu_{s_1}) = |s_1 - s_0| \mathsf{HK}(\mu_0, \mu_1)$ then there exists a dynamical plan $\pi \in \mathrm{Prob}(\mathrm{C}([0,1]; \mathfrak{C}_X))$ such that

(i) π is concentrated on geodesic curves in $\mathfrak{C}_X$,
(ii) $\lambda_s := (\mathsf{e}_s)_\sharp \pi$ is a geodesic curve in $(\mathrm{Prob}(\mathfrak{C}_X), \mathsf{W}_{\mathfrak{C}})$, where $\mathsf{e}_s(\mathsf{z}) = \mathsf{z}(s)$,
(iii) $\mu_s = \mathfrak{P}\lambda_s$ and $(\mathsf{e}_{s_0}, \mathsf{e}_{s_1})_\sharp \pi \in \mathcal{M}(\mathfrak{C}_X \times \mathfrak{C}_X)$ is optimal in (3) for μ_{s_0} and μ_{s_1}.

We consider functionals on the space of nonnegative, finite Borel measures consisting of an internal and potential part, i.e.

$$\mathcal{F}(\mu) = \begin{cases} \int F(\rho)\mathrm{d}x + \int V(x)\mathrm{d}\mu & \text{if } \mu = \rho(x)\mathrm{d}x \\ +\infty & \text{otherwise.} \end{cases}$$

Considering only the potential energy given by $V : \mathbb{R}^d \to \mathbb{R}_\infty$, we find that it is geodesically λ-convex iff the function $(r, x) \mapsto r^2 V(x)$ is geodesically convex with respect to the cone metric. For the internal energy part, we use that for absolutely

continuous measures $\mu_i = \rho_i \mathrm{d}x$ a connecting Hellinger–Kantorovich geodesic curve also has a density which is given via

$$\rho_s(x) = \rho_0(x) \frac{R_s(\phi_0(x), \nabla\phi_0(x))^2}{\det\left(\mathrm{D}\mathbf{T}_s(\phi_0(x), \nabla\phi_0(x), \mathrm{D}^2\phi_0(x))\right)}$$

where R_s and the transport map $\mathbf{T}_s$ are in turn given explicit in terms of an optimal potential ϕ_0 coming from the dual formulation in (2).

Plugging the formula for ρ_s into the functional $\mathcal{F}$ and computing the second derivative with respect to s, we find that $\mathcal{F}$ is geodesically 0-convex iff the function $G(r,u) = ur^d F(1/(u^2 r^d))$ satisfies

$$\forall u: \ r \mapsto G(r,u) \text{ is non-increasing}, \quad \text{and} \quad (r,u) \mapsto G(r,u) \text{ is convex.} \tag{7}$$

Note that for fixed u these are McCann's conditions for geodesically convex functionals with respect to the Kantorovich–Wasserstein distance, while for fixed r the conditions imply convexity with respect to the Hellinger–Kakutani distance. We emphasize, that the joint convexity in (r,u) is an even stronger requirement.

In case of $F(\rho) = \rho\log\rho$ the conditions in (7) are not satisfied, hence the associated functional is not convex with respect to the HK distance. In contrast, $F(\rho) = \rho^p$ yields convex functionals for $p \geq 1$.

References

[1] M. Liero, A. Mielke, G. Savaré, *Optimal transport in competition with reaction: The Hellinger–Kantorovich distance and geodesic curves.*, SIAM J. Math Analysis, 18: 2107-2165, 2016.

[2] M. Liero, A. Mielke, G. Savaré, *Optimal entropy-transport problems and a new Hellinger–Kantorovich distance between positive measures*, to appear in Inventiones Mathematicae.

[3] M. Liero, A. Mielke, G. Savaré, *On geodesic λ-convexity with respect to the Hellinger–Kantorovich distance.*, in preparation.

Entropic transport, deterministic transport and large deviations

Christian Léonard

No effort of rigor is done in this abstract, in particular the full underlying assumptions of some results will be left unprecise.
Consider two spaces $\mathcal{X}$ and $\mathcal{Y}$, a cost function $c : \mathcal{X}\times\mathcal{Y} \to [0,\infty)$ such that $\inf c = 0$, two probability measures $\mu \in \mathrm{P}(\mathcal{X})$ on $\mathcal{X}$ and $\nu \in \mathrm{P}(\mathcal{Y})$ on $\mathcal{Y}$.

Deterministic optimal transport. We call deterministic the standard optimal transport in order to distinguish it from the entropic optimal transport to be introduced later. Deterministic optimal transport (DOT) consists in solving the Monge-Kantorovich problem:

$$\text{(MK)} \qquad \int_{\mathcal{X}\times\mathcal{Y}} c(x,y)\, \mathsf{p}(dxdy) \to \min; \qquad \mathsf{p} \in \Pi(\mu,\nu)$$

where $\Pi(\mu,\nu) := \{\mathsf{p} \in \mathrm{P}(\mathcal{X}\times\mathcal{Y}) : \mathsf{p}_X = \mu,\ \mathsf{p}_Y = \nu\}$ is the set of all couplings of μ and ν. As a result, it is known [7] that any coupling $\mathsf{p} \in \Pi(\mu,\nu)$ solves (MK)

if and only if there exist two functions $\varphi : \mathcal{X} \to [-\infty, \infty)$ and $\psi : \mathcal{Y} \to [-\infty, \infty)$ such that

$$(1) \qquad \begin{cases} \varphi \oplus \psi & \leq & c, & \text{everywhere,} \\ \varphi \oplus \psi & = & c, & \mathsf{p}\text{-almost everywhere.} \end{cases}$$

These functions are called the Kantorovich potentials and they can be chosen such that

$$(2) \qquad \begin{cases} \varphi(x) & = & \inf_{b \in \mathcal{Y}} \{c(x,b) - \psi(b)\}, & x \in \mathcal{X}, \\ \psi(y) & = & \inf_{a \in \mathcal{X}} \{c(a,y) - \varphi(a)\}, & y \in \mathcal{Y}. \end{cases}$$

Entropic optimal transport. The relative entropy of the probability measure p with respect to the reference probability measure r is defined by

$$H(\mathsf{p}|\mathsf{r}) = \int \log(d\mathsf{p}/d\mathsf{r})\, d\mathsf{p} \in [0, \infty]$$

Take $\mathsf{r}^\infty \in \mathrm{P}(\mathcal{X} \times \mathcal{Y})$ and for any $\epsilon > 0$, define the probability measure

$$\mathsf{r}^\epsilon = \exp(-\epsilon^{-1}\lambda^\epsilon) \exp(-\epsilon^{-1} c)\, \mathsf{r}^\infty \in \mathrm{P}(\mathcal{X} \times \mathcal{Y})$$

where $\lambda^\epsilon \in \mathbb{R}$ is a normalizing constant. Entropic optimal transport (EOT) consists in solving the Schrödinger problem of order ϵ:

$$(\mathrm{S}^\epsilon) \qquad \epsilon H(\mathsf{p}|\mathsf{r}^\epsilon) \to \min; \qquad \mathsf{p} \in \Pi(\mu, \nu).$$

By a direct computation, $\epsilon H(\mathsf{p}|\mathsf{r}^\epsilon) = \int c\, d\mathsf{p} + \epsilon H(\mathsf{p}|\mathsf{r}^\infty) + \lambda^\epsilon \underset{\epsilon \to 0}{\longrightarrow} \int c\, d\mathsf{p}$, since by the Laplace principle: $\lambda^\epsilon := \epsilon \log \int e^{-c/\epsilon}\, d\mathsf{r}^\infty \underset{\epsilon \to 0}{\to} \inf c = 0$. This suggests that $(\mathrm{S}^\epsilon) \underset{\epsilon \to 0}{\to}$ (MK) in some sense.

Convergence of (EOT) to (DOT). Indeed, as a general result, one can prove that [4, 2]: $\Gamma\text{-}\lim_{\epsilon \downarrow 0} \left(\epsilon H(\cdot|\mathsf{r}^\epsilon) + \iota_{\Pi(\mu,\nu)}\right) = \langle c, \cdot \rangle + \iota_{\Pi(\mu,\nu)}$, where $\iota_{\Pi(\mu,\nu)}(\mathsf{p}) = 0$ if $\mathsf{p} \in \Pi(\mu, \nu)$ and $+\infty$ otherwise. Consequently, $\lim_{\epsilon \downarrow 0} \inf (\mathrm{S}^\epsilon) = \inf (\mathrm{MK})$ and any limit point of the sequence (p^ϵ) of solutions to (S^ϵ) solves (MK).

Large deviation principle. A sequence (q^ϵ) of Borel probability measures on some topological space $\mathcal{Z}$ is said to obey the large deviation principle (LDP) with rate function $J : \mathcal{Z} \to [0, \infty]$, if for any Borel subset $A \subset \mathcal{Z}$,

$$-\inf_{\operatorname{int} A} J \leq \liminf_{\epsilon \downarrow 0} \epsilon \log \mathsf{q}^\epsilon(A) \leq \limsup_{\epsilon \downarrow 0} \epsilon \log \mathsf{q}^\epsilon(A) \leq -\inf_{\operatorname{cl} A} J,$$

where $\operatorname{int} A$ and $\operatorname{cl} A$ are the topological interior and closure of A, see [3]. A shorthand for this LDP is

$$\mathsf{q}^\epsilon \underset{\epsilon \downarrow 0}{\asymp} \exp(-\epsilon^{-1} J).$$

Large deviations of (p^ϵ)**.** The sequence (p^ϵ) of solutions to (S^ϵ) satisfies the following large deviation principle. There exist two functions $\varphi : \mathcal{X} \to [-\infty, \infty)$ and $\psi : \mathcal{Y} \to [-\infty, \infty)$ such that

$$\mathsf{p}^\epsilon \underset{\epsilon\downarrow 0}{\asymp} \exp(-\epsilon^{-1} I),$$

where the rate function I is given by:

$$I(x,y) = c(x,y) - \varphi(x) - \psi(y). \tag{3}$$

As a direct corollary, we recover the standard result (1) about DOT by expressing that, as a large deviation rate function, I is nonnegative and the support of any limit point of (p^ϵ) is included in $\{I = 0\}$.
The proof of this LDP is based on the representation of the unique solution p^ϵ of the Schrödinger problem (S^ϵ) as

$$\mathsf{p}^\epsilon(dxdy) = f^\epsilon(x) g^\epsilon(y)\, \mathsf{r}^\epsilon(dxdy)$$

for some measurable nonnegative functions f^ϵ on $\mathcal{X}$ and g^ϵ on $\mathcal{Y}$, see [5] for instance. Some manipulations based on this product form formula, inspired by Mikami's article [6], allow to obtain the equation

$$\mu(dx)\nu(dy) = \Big[\int_{\mathcal{X}\times\mathcal{Y}} \frac{d(\mathsf{r}_X^{\epsilon,y} \otimes \mathsf{r}_Y^{\epsilon,x})}{d\mathsf{r}^\epsilon}\; d\mathsf{p}^\epsilon \Big]\; \frac{d(\mathsf{r}_X^\epsilon \otimes \mathsf{r}_Y^\epsilon)}{d\mathsf{r}^\epsilon}(x,y)\; \mathsf{p}^\epsilon(dxdy), \tag{4}$$

where the unknown is p^ϵ, the data are μ, ν and r^ϵ, and we denote: $\mathsf{r}_X^\epsilon(dx) := \mathsf{r}^\epsilon(X \in dx)$, $\mathsf{r}_Y^\epsilon(dy) := \mathsf{r}^\epsilon(Y \in dy)$, $\mathsf{r}_X^{\epsilon,y}(dx) := \mathsf{r}^\epsilon(dx \mid Y = y)$ and $\mathsf{r}_Y^{\epsilon,x}(dy) := \mathsf{r}^\epsilon(dy \mid X = x)$, with $X(x,y) = x$ and $Y(x,y) = y$ the marginal projections. As the sequence (r^ϵ) satisfies the LDP

$$\mathsf{r}^\epsilon \underset{\epsilon\downarrow 0}{\asymp} \exp(-\epsilon^{-1} c),$$

it follows that $(\mathsf{r}_X^\epsilon), \ldots$ also satisfy LDPs with rate functions expressed in terms of c. For instance, $\mathsf{r}_Y^\epsilon \underset{\epsilon\downarrow 0}{\asymp} \exp(-\epsilon^{-1} c_Y)$ with $c_Y(y) = \inf_a c(a,y)$, and $\mathsf{r}_X^{\epsilon,y} \underset{\epsilon\downarrow 0}{\asymp} \exp(-\epsilon^{-1} c_X^y)$ with $c_X^y(x) = c(x,y) - c_Y(y)$, under the assumption that $\mathsf{r}^\infty \ll \mathsf{r}_X^\infty \otimes \mathsf{r}_Y^\infty$. On the other hand, since $\Pi(\mu,\nu)$ is compact, from any subsequence of (p^ϵ) one can extract a subsubsequence (still denoted (p^ϵ)) such that

$$\mathsf{p}^\epsilon \underset{\epsilon\downarrow 0}{\asymp} \exp(-\epsilon^{-1} I),$$

for some rate function I. Plugging all these exponential estimates in the right-hand side of (4) provides us with an LDP. It remains to identify the resulting rate function with the rate function $\iota_{\mathrm{supp}\,\mu \times \mathrm{supp}\,\nu}$ of the left-hand side to obtain an equation leading to (2) and (3).

JKO revisited with large deviations. This last section is based on a joint work in progress with Johannes Zimmer. It is known since [1] that the JKO scheme of convergence for the heat equation is related to the large deviations of the empirical

measure of a large system of independent Brownian particles. This LDP (Sanov's theorem) leads to the entropy minimization problem

$$H(P|R) \to \min; \qquad P \in \mathrm{P}(\Omega) : P_0 = \mu_0, \tag{5}$$

where R is the path measure describing a reversible Brownian motion, $P \in \mathrm{P}(\Omega)$ is a probability measure on the path space Ω with initial marginal P_0. For any time step $\tau > 0$, let R^τ be the τ-discretization of R and P^τ the solution of (5) with R replaced by R^τ. The Γ-convergence of $H(\cdot|R^\tau)$ to $H(\cdot|R)$ as τ tends down to zero, together with a rewriting of $H(P|R)$ by means of current and osmotic actions: A^{cu} and A^{os}, allows to prove that the solution of the De Giorgi scheme based on

$$J^\tau(\mu,\nu) = H(\mu|\mathrm{vol})/2 - H(\nu|\mathrm{vol})/2 + \tau^{-1} C^\tau(\mu,\nu)$$

with $C^\tau(\mu,\nu) := \inf\left\{A^{cu}(P) + \tau^2 A^{os}(P); P \in \mathrm{P}(\Omega) : P_0 = \mu, P_1 = \nu\right\}$, converges to the solution of the heat equation. Note that by the Benamou-Brenier formula, we have: $\lim_{\tau\downarrow 0} C^\tau = W_2^2$, the squared Wasserstein distance.

References

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

[2] G. Carlier, V. Duval, G. Peyré, and B. Schmitzer. Convergence of entropic schemes for optimal transport and gradient flows. *SIAM Journal on Mathematical Analysis*, 49(2):1385–1418, 2017.

[3] A. Dembo and O. Zeitouni. *Large deviations techniques and applications. Second edition.* Applications of Mathematics 38. Springer Verlag, 1998.

[4] C. Léonard. From the Schrödinger problem to the Monge-Kantorovich problem. *J. Funct. Anal.*, 262(1879-1920), 2012.

[5] C. Léonard. A survey of the Schrödinger problem and some of its connections with optimal transport. *Discrete Contin. Dyn. Syst. A*, 34(4):1533–1574, 2014.

[6] T. Mikami. Monge's problem with a quadratic cost by the zero-noise limit of h-path processes. *Probab. Theory Relat. Fields*, 129:245–260, 2004.

[7] C. Villani. *Optimal Transport. Old and New*, volume 338 of *Grundlehren der mathematischen Wissenschaften*. Springer, 2009.

Variational convergence of EVI-flows

Giuseppe Savaré

(joint work with Matteo Muratori)

Consider a complete metric space (X, d) and a lower semicontinuous functional $\phi : X \to (-\infty, +\infty]$, with proper domain $D(\phi) := \{x \in X : \phi(x) < +\infty\}$. The slope of ϕ [1] is defined as

$$\begin{cases} \limsup\limits_{y\to x} \dfrac{(\phi(x)-\phi(y))_+}{\mathsf{d}(x,y)} & \text{if } x \in D(\phi) \text{ and not isolated,} \\ 0 & \text{if } x \in D(\phi) \text{ is an isolated point,} \\ +\infty & \text{if } x \notin D(\phi). \end{cases}$$

ϕ is geodesically λ-convex, $\lambda \in \mathbb{R}$, if for every $x_0, x_1 \in D(\phi)$ there exists a (minimal, constant speed) geodesic $\mathsf{x} : [0,1] \to D(\phi)$ such that $\mathsf{x}(0) = x_0,\ \mathsf{x}(1) = x_1$,

$$\mathsf{d}(\mathsf{x}(\vartheta_0), \mathsf{x}(\vartheta_1)) = |\vartheta_1 - \vartheta_0| \mathsf{d}(\mathsf{x}(0), \mathsf{x}(1)) \quad \text{for every } \vartheta_i \in [0,1],$$

and

$$\phi(\mathsf{x}(\vartheta)) \le (1-\vartheta)\phi(x_0) + \vartheta\phi(x_1) - \frac{\lambda}{2}\vartheta(1-\vartheta)\mathsf{d}^2(x_0,x_1) \quad \text{for every } \vartheta \in [0,1].$$

A continuous curve $u : [0,\infty) \to D(\phi)$ satisfies the Evolution Variational Inequalities $\mathrm{EVI}_\lambda(\phi)$ associated to ϕ if

$$\frac{1}{2}\frac{\mathrm{d}_+}{\mathrm{d}t}\mathsf{d}^2(u(t),y) + \frac{\lambda}{2}\mathsf{d}^2(u(t),y) \le \phi(y) - \phi(u(t)) \quad \text{for every } t \ge 0,\ y \in D(\phi),$$

where $\frac{\mathrm{d}+}{\mathrm{d}t}\zeta(t) := \limsup_{h\downarrow 0}(\zeta(t+h)-\zeta(t))/h$ is the right upper Dini derivative of a real function ζ.

Definition [1, 7, 5]. *The functional ϕ admits a λ-gradient flow in (X,d) if there exists a continuous semigroup $\mathsf{S}_t : D(\phi) \to D(\phi)$, $t \ge 0$, satisfying the properties*

$$\mathsf{S}_{s+t}(x) = \mathsf{S}_s(\mathsf{S}_t(x)), \quad \lim_{t\downarrow 0}\mathsf{S}_t(x) = \mathsf{S}_0(x) = x \quad \textit{for every } x \in D(\phi),$$

such that for every $x \in D(\phi)$ the curve $t \mapsto \mathsf{S}_t(x)$ solves $\mathrm{EVI}_\lambda(\phi)$.

Remark. When X is an Hilbert space and d is the distance induced by the scalar product, a functional ϕ is geodesically λ-convex if and only if $x \mapsto \phi_\lambda(x) := \phi(x) - \frac{\lambda}{2}|x|^2$ is convex. In this case, a curve $u : [0,\infty) \to D(\phi)$ is a solution to $\mathrm{EVI}_\lambda(\phi)$ if and only if u is locally Lipschitz and solves the differential inclusion

$$\frac{\mathrm{d}}{\mathrm{d}t}u(t) \in A(u(t)) \quad \text{a.e. in } (0,\infty), \tag{1}$$

where A is the multivalued λ-convex subdifferential of ϕ characterized by

$$\xi \in A(x) \quad \Leftrightarrow \quad \langle \xi, y-x\rangle + \frac{\lambda}{2}|y-x|^2 \le \phi(y) - \phi(x),$$

or, equivalently,

$$A - \lambda I = \partial\phi_\lambda.$$

Since $A - \lambda I$ is a maximal monotone operator, Crandall-Liggett generation theorem and Brézis regularization result [3] guarantee that for every $u_0 \in D(\phi)$ there exists a unique solution of (1) and ϕ admits a λ-gradient flow. □

Remark. Let $(X,\mathsf{d},\mathfrak{m})$ be a complete and separable metric-measure space satisfying the exponential volume growth condition $\mathfrak{m}(B_\rho(x_0)) \le A\mathrm{e}^{B\rho}$ for suitable constants $A, B \ge 0$ and $x_0 \in X$. $(X,\mathsf{d},\mathfrak{m})$ satisfies the *Riemannian* Ricci lower curvature condition $\mathrm{RCD}(K,\infty)$, $K \in \mathbb{R}$, [2] if the relative entropy functional $\mu \mapsto \phi(\mu) := \int \varrho\log\varrho\,\mathrm{d}\mathfrak{m}$, $\varrho = \mathrm{d}\mu/\mathrm{d}\mathfrak{m}$, admits a K gradient flow in the Wasserstein space $(\mathcal{P}_2(X), W_\mathsf{d})$.

We are interested to study the stability of the notion of λ-gradient flow with respect to perturbations of the functional ϕ in the variational framework of Γ-convergence.

Let us consider a sequence of extended-real functionals $(\phi^h)_{h\in\mathbb{N}}$ defined in the metric space (X,d). The inferior and superior Γ-limits [4] are defined by

$$\Gamma\text{-}\liminf_{h\to\infty}\phi^h(x)=\inf\Big\{\liminf_{h\to\infty}\phi^h(x^h):x^h\to x\Big\}=\lim_{\rho\downarrow 0}\liminf_{h\to\infty}\inf_{B_\rho(x)}\phi^h,$$

$$\Gamma\text{-}\limsup_{h\to\infty}\phi^h(x)=\inf\Big\{\limsup_{h\to\infty}\phi^h(x^h):x^h\to x\Big\}=\lim_{\rho\downarrow 0}\limsup_{h\to\infty}\inf_{B_\rho(x)}\phi^h,$$

where $B_\rho(x):=\{y\in X:\mathsf{d}(y,x)<\rho\}$.

We say that $\phi=\Gamma\text{-}\lim_{h\to\infty}\phi^h$ if $\phi=\Gamma\text{-}\liminf_{h\to\infty}\phi^h=\Gamma\text{-}\limsup_{h\to\infty}\phi^h$. It is well known that the Γ-limit ϕ can also be characterized by the following properties

$$x^h\to x\quad\Rightarrow\quad\liminf_{h\to\infty}\phi^h(x^h)\ge\phi(x),$$
$$\forall\,x\in X\quad\exists\,(x^h)_h:\quad x^h\to x,\quad \phi^h(x^h)\to\phi(x).$$

We prove the following stability result:

THEOREM. *Let ϕ^h,ϕ, $h\in\mathbb{N}$ be geodesically λ-convex functionals in the metric space (X,d). We suppose that for every $h\in\mathbb{N}$ ϕ^h admits a λ-gradient flow $(\mathsf{S}^h_t)_{t\ge0}$ and that* at least one *of the following properties holds*

(a) ***Recovery sequence for ϕ and $|\partial\phi|$:***

$$\Gamma\text{-}\limsup_{h\to\infty}\phi^h=\phi,\quad \Gamma\text{-}\limsup_{h\to\infty}|\partial\phi^h|\le|\partial\phi|\qquad\text{on } D(|\partial\phi|),$$

i.e. for every $u\in D(|\partial\phi|)$ there exist $u^h\in D(|\partial\phi^h|)$ such that

$$u^h\to u,\quad \phi^h(u^h)\to\phi(u),\quad \limsup_{h\to\infty}|\partial\phi^h|(u^h)\le|\partial\phi|(u).$$

(b) ***Qualified Γ-convergence:*** *$\Gamma\text{-}\limsup_{h\to\infty}\phi^h\le\phi$ and for every $u\in D(\phi)$, $\varepsilon>0$ and $\bar\tau>0$, there exists $\tau\in(0,\bar\tau)$ such that*

$$\liminf_{h\to\infty}\inf_{B_\tau(u)}\phi^h\ge\inf_{B_\tau(u)}\phi-\varepsilon\tau.$$

(c) ***Local Yosida regularization:*** *$\Gamma\text{-}\limsup_{h\to\infty}\phi^h\le\phi$ and for every $u\in D(\phi)$, $\varepsilon>0$ and $\bar\tau>0$, there exists $\tau\in(0,\bar\tau)$ such that*

$$\liminf_{h\to\infty}\inf_v\phi^h(v)+\frac{1}{2\tau}\mathsf{d}^2(u,v)\ge\inf_v\phi(v)+\frac{1}{2\tau}\mathsf{d}^2(u,v)-\varepsilon\tau.$$

Then we have

- ***Existence of the limit λ-flow:*** *the functional ϕ admits a λ-gradient flow $(\mathsf{S}_t)_{t\ge0}$.*
- ***Convergence of the flows:*** *whenever $x^h\to x\in D(\phi)$, $x^h\in D(\phi^h)$,*

$$\lim_{h\to\infty}\mathsf{S}^h_t(x^h)=\mathsf{S}_t(x)\quad\text{for every } t\ge0.$$

- ***Γ-convergence for ϕ^h and $|\partial\phi^h|$:*** $\Gamma\text{-}\lim_{h\to\infty}\phi^h=\phi$, $\Gamma\text{-}\lim_{h\to\infty}|\partial\phi^h|=|\partial\phi|$.

The above theorem is a metric version of the corresponding result in Hilbert spaces, due to Bénilan, Crandall, Pazy [3]. We can also prove a slightly different statement, involving an auxiliary topology and modeled on Mosco convergence in Hilbert spaces [6]. To this aim, let us suppose that X is endowed with a weaker Hausdorff topology σ such that

- $x_n \xrightarrow{\sigma} x$ yields $\liminf_{n\to\infty} \mathsf{d}(x_n, y) \geq \mathsf{d}(x, y)$ for every $y \in X$;
- every d-bounded sequence $(x^h)_h$ with $\sup_h \phi^h(x^h) < +\infty$ admits a σ-converging subsequence.

We can say that ϕ^h converges to ϕ in the *metric*-Mosco sense if [1, 7]

$$\begin{aligned} x^h \xrightarrow{\sigma} x \quad &\Rightarrow \quad \liminf_{h\to\infty} \phi^h(x^h) \geq \phi(x), \\ \forall\, x \in X \quad \exists\, (x^h)_h : \quad & x^h \to x, \quad \phi^h(x^h) \to \phi(x). \end{aligned}$$

THEOREM *If ϕ^h admits a λ-flow S^h and converges to ϕ in the metric-Mosco sense as $h \to \infty$, then also ϕ admits a λ-flow $(\mathsf{S}_t)_{t\geq 0}$ and the same conclusions of the previous Theorem hold.*

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 Verlag, Basel, 2008.

[2] L. Ambrosio, N. Gigli, G. Savaré, *Metric measure spaces with Riemannian Ricci curvature bounded from below*, Duke Math. J., **163** (2014) 1405–1490.

[3] H. Brézis, *Opérateurs maximaux monotones et semi-groupes de contractions dans les espaces de Hilbert*, North-Holland Mathematics Studies, No. 5. Notas de Matemática (50), 1973.

[4] G. Dal Maso, *An introduction to Γ-convergence*, Progress in Nonlinear Differential Equations and their Applications, **8**, Birkhäuser Boston, 1993.

[5] S. Daneri, G. Savaré, *Lecture notes on gradient flows and optimal transport*, in Optimal transportation, London Math. Soc. Lecture Note Ser., **413** (2014), 100–144.

[6] U. Mosco, *Convergence of convex sets and of solutions of variational inequalities*, Advances in Math., **3** (1969), 510–585.

[7] G. Savaré, *Gradient flows and diffusion semigroups in metric spaces under lower curvature bounds*, C. R. Math. Acad. Sci. Paris, **345**, (2007) 151–154.

Gromov-Hausdorff limit of Wasserstein spaces on point clouds

Nicolás García Trillos

In this talk we suggest how variational techniques for evolution can be used to study the stability of certain data analysis procedures in the large data limit. In particular, we present a first theoretical result that provides the basis for the study of statistical consistency of such procedures.

Mathematically speaking, the type of evolutions we are interested in are gradient flows in the Wasserstein space associated to a graph. The nodes of the graph are random data points $x_1, \ldots, x_n$ sampled from some distribution on $\mathbb{R}^d$; two points are connected with an edge if their distance is smaller than a given length-scale

$\varepsilon > 0$ controlling data connectivity. The main question we are interested in is the following:

(1) How and when can one establish the stability of evolution equations defined on geometric graphs that can be described as gradient flows in discrete Wasserstein space, as the number of points $n \to \infty$ and $\varepsilon \to 0$?

Discrete Wasserstein spaces on graphs have been introduced independently in [1, 2, 3] as follows. Let us denote by $\mathcal{P}(X_n)$ the set of probability measures supported on $X_n := \{x_1, \dots, x_n\}$. For any $\rho_0, \rho_1 \in \mathcal{P}(X_n)$ their discrete Wasserstein distance is defined by

$$W_n(\rho_0, \rho_1) := \inf \int_0^1 \sum_{x_i \sim x_j} V_t(x_i, x_j)^2 \hat{\rho}_t(x_i, x_j) dt$$

where the inf is taken over all curves $t \in [0, 1] \mapsto (\rho_t, V_t)$, starting at ρ_0 and ending at ρ_1, that solve the (discrete) continuity equation

$$\frac{d}{dt}\rho_t + div_n(\hat{\rho}_t V_t) = 0.$$

Here div_n stands for the divergence operator acting on discrete vector fields (functions defined on the edges of the graph) and $\hat{\rho}_t$ is an interpolation of ρ_t that define a "density" on edges of the graph (e.g. interpolation with the logarithmic mean). The above definition is motivated by the Benamu-Brenier characterization of the Wasserstein space on Euclidean space. In [1], it is shown that the heat equation on the graph can be described as the gradient flow of the relative entropy (with respect to the uniform measure on the graph) with respect to the discrete Wassesrtein distance.

Our approach to answer (1) is to use the notion of Γ-convergence for sequences of gradient flows that was introduced in [4]. In our setting, the evolutions are all defined in different metric spaces, and hence, the variational stability of the sequence of gradient flows depends on the stability of the metric spaces in the Gromov-Hausdorff sense. The main theoretical result that we discuss in this talk is the following Gromov-Hausdorff convergence of Wasserstein spaces on point clouds in a localized setting.

Theorem:(Main Theorem in [5]) Suppose $X_n := \{x_1, \dots, x_n\}$ are i.i.d. samples from the uniform distribution on the d-dimensional flat torus. Suppose the parameter $\varepsilon := \varepsilon_n$ used to construct the geometric graph is such that

$$\frac{\log(n)^{p_d}}{\varepsilon_n n^{1/d}} \to 0, \quad \varepsilon_n \sqrt{\log(n\varepsilon_n^d)} \to 0,$$

where $p_2 = 3/4$ and $p_d = 1/d$ for $d \geq 3$. Then, with probability one, the discrete Wasserstein space $(\mathcal{P}(X_n), W_n)$ (appropriately rescaled) converges in the Gromov-Hausdorff sense towards the Wasserstein space $(\mathcal{P}(T^d), W)$, as $n \to \infty$.

This result is analogue to the work [6] where the Gromov-Hausdorff convergence of Wasserstein spaces defined on regular grids towards the Wasserstein space at the continuum level is established. As suggested earlier, the main motivation for our mathematical considerations, is the analysis of data analysis procedures of

evolution type (like for example the mean shift algorithm [7]) in the large sample limit. To connect the motivating applications with the mathematical theory described above we need to answer the following questions:

- Are there any existing machine learning algorithm for clustering or classification of evolution type, that can be directly described as gradient flows in Wasserstein space or that can be associated to one such flow?
- Can one design relevant machine learning algorithms with a gradient flow structure that by their very nature would inherit desirable properties and in particular exhibit stability properties as $n \to \infty$?

Many conceptual, theoretical, and computational aspects of these questions remain to be developed. At the mathematical level one immediate task to complete is to extend the results in the **Theorem** to more general geometric settings (e.g. data sampled from a measure supported on a manifold) and more general graph constructions on random data (e.g. k-NN graphs). By formulating all these questions we hope to open and explore new research directions in mathematical analysis, with the intention of establishing rigorous analytical results supporting the use of algorithms for the analysis of data.

References

[1] J. Maas, *Gradient flows of the entropy for finite Markov chains*, Journal of Functional Analysis (2011),2250–2292

[2] S. Chow, W. Wen, Y. Li and H. Zhou, *Fokker-Planck equations for a free energy functional or Markov process on a graph*, Arch. Ration. Mech. Anal. (2012), 969–1008

[3] A. Mielke, *Geodesic convexity of the relative entropy in reversible Markov chains* , Calc. Var. (2013) 1–31

[4] S. Serfaty, *Gamma-convergence of gradient flows on Hilbert and metric spaces and applications*, Discrete and Continuous Dynamical Systems (2011), 1427–1451

[5] N. Garcia Trillos, *Gromov-Hausdorff convergence of Wasserstein spaces on point clouds*, Preprint (2017)

[6] N. Gigli and J. Maas, *Gromov-Hausdorff convergence of discrete transportation metrics*, SIAM J. Math. Anal. (2013), 879–899

[7] M. Perpiñán, *A review of mean-shift algorithms for clustering*, Preprint (2015)

On fluctuations in Markov chains

Johannes Zimmer

(joint work with Robert L. Jack, Marcus Kaiser)

In recent years, new connections between the theory of gradient flows on the macroscopic scale and underlying reversible Markov chains on a mesoscopic scale have been discovered; we refer for example to the work of Maas [4]. These links shed new light on the famous work by Onsager; in particular, the near-equilibrium theory of Onsager has recently been extended to a more general setting [6], and it is shown there that this leads to nonlinear evolution equations, so-called generalised gradient flows. We argue that this setting, when cast in terms of forces and currents (fluxes) as the classic Onsager theory, can also be extended to irreversible

Markov chains. We show this setting is natural in the sense that classic physical quantities appear. Furthermore, an associated variational principle appears as rate functional of a large deviation principle.

Before explaining the setting and results, we first recall elements of Onsager theory and Macroscopic Fluctuation Theory, as the setting for Markov chains is a natural generalisation. Let ρ denote state of the system; then Onsager considers n currents $j = (j^\alpha)_{\alpha=1}^n$ and conjugate forces $F(\rho) = (F^\alpha)_{\alpha=1}^n$ (such that $F(\rho)$ is independent of the state ρ). He establishes a variational setting, which says that the most likely current minimises

$$\Phi(\rho, j, f) = \|j - Lf\|^2_{L^{-1}} = \|j\|^2_{L^{-1}} - \langle j, f\rangle + \|f\|^2_L \tag{1}$$

with norm $\|j\|^2_{L^{-1}} := \frac{1}{2} j^T L^{-1} j$. Note that in Onsager theory, the relation between forces and fluxes is linear; for a given force F, the most likely current J, say, satisfies $J = LF$. Moreover, Φ is independent of ρ, even if we include ρ as argument for later convenience. Also, without going into details, we remark that (1) is the integrand in a so-called rate functional of a large deviation principle, which describes the probability of (rare) dynamical trajectories, in other words fluctuations of macroscopic properties of physical systems. Namely, with the choices made above, we can cast the Onsager-Machlup theory for dynamical fluctuations [7] as follows. For some large parameter $\mathcal{N}$ (for example the size of an ensemble),

$$\mathrm{Prob}\left((\hat\rho_t^{\mathcal{N}}, \hat\jmath_t^{\mathcal{N}})_{t\in[0,T]} \approx (\rho_t, j_t)_{t\in[0,T]}\right) \asymp \exp\left\{-\mathcal{N} I_{[0,T]}\left((\rho_t, j_t)_{t\in[0,T]}\right)\right\} \tag{2}$$

with the *rate functional*

$$I_{[0,T]}\big((\rho_t, j_t)_{t\in[0,T]}\big) = I_0(\rho_0) + \frac{1}{2}\int_0^T \Phi(\rho_t, j_t, F(\rho_t))\,\mathrm{d}t; \tag{3}$$

here I_0 is a functional related to a given initial condition.

Macroscopic Fluctuation Theory (MFT) [1] can be seen as extension of Onsager theory. In this setting, one considers the macroscopic description of $\mathcal{N}$ indistinguishable particles moving on a lattice Λ_L, in the hydrodynamic limit, i.e., in the limit $L \to \infty$ such that density converges to a number $\bar\rho$, say. The system is described by a local (mass) density $\rho\colon \Lambda \to [0,\infty)$ and local current $J\colon \Lambda \to \mathbb{R}^d$. The time evolution depends on the applied force $F(\rho)\colon \Lambda \to \mathbb{R}^d$, and one can see that as, in Onsager theory, there is a relation between applied forces and resulting fluxes. As for the Onsager setting, this can be described by a large deviation principle of the form (2)–(3), however with integrand

$$\Phi_{\mathrm{MFT}}(\rho, j, f) := \frac{1}{2}\int_\Lambda \big(j - \chi(\rho)f\big)\cdot\chi(\rho)^{-1}\big(j - \chi(\rho)f\big)\,\mathrm{d}x. \tag{4}$$

A natural questions is: is there a similar structure on the level of Markov chains? We answer this for Markov chains on a finite state space V, though many results carry over to countable state spaces [3]. We denote the transition rates by r_{xy} and assume the stationary distribution π is unique, with $\pi(x) > 0$ for all $x \in V$. We recall that a Markov chain is *reversible* if it satisfies the detailed balance condition $\pi(x)r_{xy} = \pi(y)r_{yx}$, and *non-reversible* otherwise.

Reversible and non-reversible Markov chains exhibit distinctly different behaviour. For example, reversible Markov chains are typically models for equilibrium systems, are convenient to use in Monte Carlo algorithms, and have steady states without persistent currents. In contrast, non-reversible Markov chains are typically models for non-equilibrium systems, are less often used in Monte Carlo algorithms, and have steady states with persistent currents.

Given these significant differences, it is natural to compare these two classes of Markov chains. This might enable us to understand better how to model nonequilibrium systems and how to design new Monte-Carlo algorithms, for example. It is known that breaking detailed balance can accelerate convergence to equilibrium, see for example [2] and the references therein. Yet, the structure analysed in [3] sketched below applied to both reversible and non-reversible chains.

We recall that the free energy (*Kullback-Leibler divergence*) with respect to π is $\mathcal{F}(\rho) := \sum_x \rho(x) \log\Big(\frac{\rho(x)}{\pi(x)}\Big)$. As in Onsager theory and MFT, we consider forces and currents. Namely, the *probability current* is

$$J_{xy}(\rho) := \rho(x) r_{xy} - \rho(y) r_{yx}. \tag{5}$$

Further, we assign to each edge in E a force F (*affinity* [9]) and a *mobility* a,

$$F_{xy}(\rho) := \log \frac{\rho(x) r_{xy}}{\rho(y) r_{yx}} \quad \text{and} \quad a_{xy}(\rho) := 2\sqrt{\rho(x) r_{xy} \rho(y) r_{yx}}. \tag{6}$$

Then it follows directly from (5) and (6) that force-flux relation is nonlinear,

$$J_{xy}(\rho) = a_{xy}(\rho) \sinh\big(\tfrac{1}{2} F_{xy}(\rho)\big). \tag{7}$$

Furthermore, a large deviation principle of the form (2) holds, where the rate functional is finite for paths satisfying $\dot\rho_t + \operatorname{div} j_t = 0$, and then takes the form

$$I_{[0,T]}\big((\rho_t, j_t)_{t\in[0,T]}\big) = \mathcal{F}(\rho_0) + \int_0^T \frac{1}{2} \left[\Psi(\rho_t, j_t) - \langle j_t, F(\rho_t)\rangle + \Psi^\star(\rho_t, F(\rho_t))\right] \mathrm{d}t, \tag{8}$$

where Ψ and $\Psi^\star$ are Legendre duals, with

$$\Psi^\star(\rho, F) := \sum_{xy} a_{xy}(\rho) \big(\cosh\big(\tfrac{1}{2} F_{xy}(\rho)\big) - 1\big), \tag{9}$$

see [5, 8, 3].

We remark that the large deviation principle (2) with a rate functional (8) is general in the sense that it covers the Markov chain setting as well as Onsager theory and MFT, with the only difference being that for the latter two settings the Legendre duals Ψ and $\Psi^\star$ are quadratic functions, while there are genuinely nonquadratic in the case of Markov chains, (9). From this observation, it immediately follows that the force-flux relation is linear for Onsager theory and MFT, but leads for Markov chains necessarily to the nonlinear relation (7).

This "thermodynamic" structure for Markov chains can provide insight in several ways. For example, for a natural decomposition of forces in "symmetric" and "antisymmetric" parts, $F^S_{xy}(\rho) := \log \frac{\rho(x)\pi(y)}{\rho(y)\pi(x)}$ and $F^A_{xy} := \log \frac{\pi(x) r_{xy}}{\pi(y) r_{yx}}$ (F^S

is the same for the adjoint process, while F^A changes sign for the adjoint, so $F^* = F^S - F^A$) the following generalised orthogonality holds

$$\Psi^\star\big(\rho, F^S(\rho) + F^A\big) = \Psi^\star\big(\rho, F^S(\rho) - F^A\big).$$

Also, it is natural to conjecture that the macroscopic quadratic functionals of Onsager theory and MFT, (1) and (4), appear in a scaling limit, where the non-quadratic structure (9) is in the limit approximated by its quadratic Taylor expansion; this is a question for future investigation.

References

[1] Lorenzo Bertini, Alberto De Sole, Davide Gabrielli, Giovanni Jona-Lasinio, and Claudio Landim. Macroscopic fluctuation theory. *Rev. Modern Phys.*, 87(2):593–636, 2015.

[2] Marcus Kaiser, Robert L. Jack, and Johannes Zimmer. Acceleration of convergence to equilibrium in Markov chains by breaking detailed balance. *J. Stat. Phys.*, 168:259–287, 2017.

[3] Marcus Kaiser, Robert L. Jack, and Johannes Zimmer. Canonical structure and orthogonality of forces and currents in irreversible Markov chains. `http://arxiv.org/abs/1708.01453`, 2017.

[4] Jan Maas. Gradient flows of the entropy for finite Markov chains. *J. Funct. Anal.*, 261(8):2250–2292, 2011.

[5] C. Maes and K. Netočný. Canonical structure of dynamical fluctuations in mesoscopic nonequilibrium steady states. *Europhys. Lett. EPL*, 82(3):Art. 30003, 6, 2008.

[6] A. Mielke, M. A. Peletier, and D. R. M. Renger. A generalization of Onsager's reciprocity relations to gradient flows with nonlinear mobility. *J. Non-Equilib. Thermodyn.*, 41(2):141–149, 2016.

[7] L. Onsager and S. Machlup. Fluctuations and irreversible processes. *Physical Rev. (2)*, 91:1505–1512, 1953.

[8] D. R. Michiel Renger. Large deviations of specific empirical fluxes of independent Markov chains, with implications for Macroscopic Fluctuation Theory. Weierstrass Institute, Preprint 2375, 2017.

[9] J. Schnakenberg. Network theory of microscopic and macroscopic behavior of master equation systems. *Rev. Modern Phys.*, 48(4):571–585, 1976.

Convergence analysis of time-discretization schemes for rate-independent systems

Dorothee Knees

It is well known that rate-independent systems involving nonconvex energy functionals in general do not allow for time-continuous solutions even if the given data are smooth in time. Several solution concepts are proposed to deal with these discontinuities, among them the meanwhile classical global energetic approach and the more recent vanishing viscosity approach. Both approaches generate solutions with a well characterized jump behavior. However, the solution concepts are not equivalent. Numerical discretization schemes are needed that efficiently and reliably approximate directly that type of solution that one is interested in. For instance, in the vanishing viscosity context it is reasonable to couple the viscosity parameter with the time-step size. Other approaches rely on different versions of local minimization or on alternate minimization strategies. The aim of this short

note is to discuss different time-discretization schemes proposed in literature with respect to their convergence properties.

The general setting is as follows, [3]: Let $\mathcal{Z}, \mathcal{V}$ be Hilbert spaces, $\mathcal{X}$ a Banach space with dense and compact or continuous embeddings, i.e. $\mathcal{Z} \Subset \mathcal{V} \subset \mathcal{X}$. The energy functional $\mathcal{I} : [0,T] \times \mathcal{Z} \to \mathbb{R}$ is of the type $\mathcal{I}(t,z) = \frac{1}{2}\langle Az, z\rangle + \mathcal{F}(z) - \langle \ell(t), z\rangle$, where $A \in \mathrm{Lin}(\mathcal{Z}, \mathcal{Z}^*)$ is symmetric and $\mathcal{Z}$-elliptic, $\mathcal{F} \in C^2(\mathcal{Z}; [0,\infty))$ plays the role of a nonconvex term of lower order, is weakly continuous and satisfies some further growth assumptions and $\ell \in C^1([0,T]; \mathcal{V}^*)$. The dissipation potential $\mathcal{R} : \mathcal{X} \to [0,\infty)$ is assumed to be convex, lower semicontinuous, positively homogeneous of degree one and bounded, i.e. there exist $c, C > 0$ such that for all $x \in \mathcal{X}$ we have $c\,\|x\|_{\mathcal{X}} \le \mathcal{R}(x) \le C\,\|x\|_{\mathcal{X}}$.

The aim is to approximate solutions of the following doubly nonlinear inclusion: Given $z_0 \in \mathcal{Z}$ determine $z : [0,T] \to \mathcal{Z}$ with $z(0) = z_0$ and

$$0 \in \partial\mathcal{R}(\dot z(t)) + \mathrm{D}_z\mathcal{I}(t, z(t))\,. \tag{1}$$

A function $z : [0,T] \to \mathcal{Z}$ is a *local solution* to (1) if $0 \in \partial\mathcal{R}(0) + \mathrm{D}_z\mathcal{I}(t, z(t))$ for almost all t together with the energy-dissipation estimate

$$\mathcal{I}(t_1, z(t_1)) + \mathrm{diss}_{\mathcal{R}}(z; [t_0, t_1]) \le \mathcal{I}(t_0, z(t_0)) + \int_{t_0}^{t_1} \partial_t\mathcal{I}(r, z(r))\mathrm{d}r \tag{E,$\le$}$$

that is valid for every $0 \le t_0 < t_1 \le T$, [7]. The dissipation along the curve z is given by $\mathrm{diss}_{\mathcal{R}}(z, [0,t]) := \sup_{\text{partitions } (t_i)_i \text{ of } [0,t]} \sum_{i=1}^N \mathcal{R}(z(t_i) - z(t_{i-1}))$. All solutions discussed in the next sections belong to this class. For simplifying the presentation we assume an equidistant partition of $[0,T]$ with $0 = t_0^N < \ldots < t_N^N$, $N \in \mathbb{N}$, $\tau^N = T/N$, $t_k^N = k\tau^N$.

Global energetic solutions. Global energetic solutions are characterized by the global stability condition $\mathcal{I}(t, z(t)) \le \mathcal{I}(t, v) + \mathcal{R}(v - z(t))$ that is fulfilled for all $v \in \mathcal{Z}$ and $t \in [0,T]$, together with an energy-dissipation identity, i.e. (E,$\le$) with equality instead of an inequality, [7]. They can be approximated by a time-incremental (global) minimization scheme via

$$z(0) = z_0, \quad z_k \in \mathrm{Argmin}\{\, \mathcal{I}(t_k, v) + \mathcal{R}(v - z_{k-1})\,;\, v \in \mathcal{Z}\,\}, \quad 1 \le k \le N\,.$$

BV-solutions. Given $\varepsilon > 0$ and starting from the minimization scheme

$$z_k \in \mathrm{Argmin}\{\, \mathcal{I}(t_k, z) + \frac{\varepsilon}{2\tau}\,\|z - z_{k-1}\|_{\mathcal{V}}^2 + \mathcal{R}(z - z_{k-1})\,;\, z \in \mathcal{Z}\,\} \tag{2}$$

in the limit $\tau \to 0$, $\varepsilon \to 0$ and $\tau/\varepsilon \to 0$ one obtains vanishing viscosity solutions, [6]. These are a special subclass of balanced viscosity solutions (BV-solutions). The crucial estimate for the limit analysis is a BV-estimate of the type $\sup_N \sum_{k=1}^N \|z_k - z_{k-1}\|_{\mathcal{Z}} < \infty$. This estimate provides a uniform bound on the length of the discrete curves $\{\,(t, z_N(t))\,;\, t \in [0,T]\,\}$ and allows to reformulate the system with respect to an arc-length parameter, i.e. one considers curves $\{\,(\hat t_N(\varrho), \hat z_N(\varrho))\,;\, \varrho \in [0, S_N]\,\}$ and studies the limit in this new setting.

Definition 1. *A pair $(\hat t, \hat z)$ with $\hat t \in W^{1,\infty}((0,S),\mathbb{R})$, $\hat z \in W^{1,\infty}((0,S);\mathcal{V}) \cap L^\infty((0,S);\mathcal{Z})$ is a parametrized BV-solution with initial value z_0, if $\hat t(0) = 0$,*

$\hat{t}(S) = T$, $\hat{z}(0) = z_0$, $\hat{t}'(s) \geq 0$, $\hat{t}'(s) + \|\hat{z}'(s)\|_{\mathcal{V}} \leq 1$, *if for all* $\sigma \in [0, S]$

$$\mathcal{I}(\hat{t}(\sigma), \hat{z}(\sigma)) + \int_0^\sigma \mathcal{R}(\hat{z}'(s)) + \|\hat{z}'(s)\|_{\mathcal{V}} \operatorname{dist}_{\mathcal{V}}(-\mathrm{D}_z\mathcal{I}(\hat{t}(s), \hat{z}(s)), \partial\mathcal{R}(0))\mathrm{d}s$$
$$= \mathcal{I}(0, \hat{z}(0)) + \int_0^\sigma \partial_t\mathcal{I}(\hat{t}(s), \hat{z}(s))\mathrm{d}s$$

and if $\hat{t}'(s) \operatorname{dist}_{\mathcal{V}}(-\mathrm{D}_z\mathcal{I}(\hat{t}(s), \hat{z}(s)), \partial\mathcal{R}(0)) = 0$ *for a.a.* $s \in (0, S)$.

In general it is difficult to choose the discretization parameters ε_N, τ_N in such a way that discrete solutions with small N already show approximately the correct jump behavior, see the discussion in [5, 3].

Incremental local minimization. A different way to approximate BV-solutions was proposed in [2]. Instead of introducing a viscosity parameter, this scheme relies on a local minimization procedure: Given $h > 0$, $t_0 = 0$, $z_0 \in \mathcal{Z}$, for $k \geq 1$

$$z_k^h \in \operatorname{Argmin}\{\, \mathcal{I}(t_{k-1}^h, z) + \mathcal{R}(z - z_{k-1}^h)\,;\, z \in \mathcal{Z},\, \left\|z - z_{k-1}^h\right\|_{\mathcal{V}} \leq h\,\}$$
$$t_k^h = \min\left\{t_{k-1}^h + h - \left\|z_k^h - z_{k-1}^h\right\|_{\mathcal{V}},\, T\right\}.$$

While the limit curve, a BV-solution, was characterized already in [2] it is shown recently in [3] that the proposed scheme reaches the final time T after a finite number of time-steps and that again a uniform BV-estimate is valid that provides a uniform upper bound on the length (with respect to $\|\cdot\|_{\mathcal{Z}}$) of the curves generated by the pairs (t_k^h, z_k^h).

Relaxed local minimization. Instead of considering strict local minimization the authors of [1] analyze a relaxed version and obtain local solutions in the limit. In [3] the following slightly modified version was investigated: For $1 \leq k \leq N$, $i \in \mathbb{N}$ and $t_k = k\tau$, $z_{k,0} := z_{k-1}$ find

$$z_{k,i} \in \operatorname{Argmin}\{\, \mathcal{I}(t_k, v) + \frac{\eta}{2}\|v - z_{k,i-1}\|_{\mathcal{V}}^2 + \mathcal{R}(v - z_{k,i-1})\,;\, v \in \mathcal{Z}\,\},$$

and define $z_k := z_{k,\infty} := \lim_{i\to\infty} z_{k,i}$ (weak limit in $\mathcal{Z}$). In [1], the term $\mathcal{R}(v - z_k)$ was used instead of our term $\mathcal{R}(v - z_{k,i-1})$. It is shown in [3] that the above scheme is well defined and that suitable interpolants of the incremental solutions again converge (for $\tau \to 0$ and $\eta \to \infty$) to BV-solutions. A discussion of the case with a fixed $\eta > 0$ can be found in [3], as well.

Alternate minimization with a penalty term. Rate-independent systems typically play a role in the modeling of complex material behavior like plasticity or damage propagation. In the small strain regime and neglecting inertia and visco-elastic effects one obtains models where the equation of linear elasticity for determining the displacements u is coupled to a rate independent system describing the change of the inner variable z. Alternate minimization schemes or staggered schemes are frequently used in the computations for such systems. Let the energy functional $\mathcal{E} : [0,T] \times \mathcal{U} \times \mathcal{Z} \to \mathbb{R}$ be given by $\mathcal{E}(t, u, z) = \frac{1}{2}\langle \mathcal{A}\binom{u}{z}, \binom{u}{z}\rangle + \mathcal{F}(z) - \langle \ell(t), z\rangle$ and choose $\mathcal{R} : \mathcal{Z} \to [0, \infty)$ as before. Find $(u, t) : [0, T] \to \mathcal{U} \times \mathcal{Z}$ with $z(0) = z_0$ and

$$0 = \mathrm{D}_u\mathcal{E}(t, u(t), z(t)), \quad 0 \in \partial_{\dot{z}}\mathcal{R}(\dot{z}(t)) + \mathrm{D}_z\mathcal{E}(t, u(t), z(t)).$$

The scheme from [3] reads with $(u_{k,0}, z_{k,0}) := (u_{k-1}, z_{k-1})$: For $i \geq 1$ determine

$$u_{k,i} \in \operatorname{Argmin}\{\, \mathcal{E}(t_k, v, z_{k,i-1})\,;\, v \in \mathcal{U}\,\},$$
$$z_{k,i} \in \operatorname{Argmin}\{\, \mathcal{E}(t_k, u_{k,i}, \zeta) + \mathcal{R}(\zeta - z_{k,i-1}) + \frac{\eta}{2}\,\|\zeta - z_{k,i-1}\|_{\mathcal{V}}^2\,;\, \zeta \in \mathcal{Z}\,\},$$
$$\text{stop if } \|z_{k,i} - z_{k,i-1}\|_{\mathcal{V}} \leq \delta;\ (u_k, z_k) := (u_{k,i}, z_{k,i})\,.$$

For $\tau, \delta \to 0$, $\eta \to \infty$ and $\eta\delta \to 0$ suitable interpolating curves converge to a limit that is a BV-solution with respect to z and that satisfies in addition $\mathrm{D}_u\mathcal{E}(\hat{t}(\varrho), \hat{u}(\varrho), \hat{z}(\varrho)) = 0$. This scheme was studied for the Ambrosio-Tortorelli model with $\eta = 0$ in [4]. Visco-elastic and visco-damage terms appear in the limit although viscous terms were not introduced into the time incremental problems.

References

[1] M. Artina, F. Cagnetti, M. Fornasier, and F. Solombrino. Linearly constrained evolutions of critical points and an application to cohesive fractures. *Math. Models Methods Appl. Sci.*, 27(2):231–290, 2017.

[2] M. A. Efendiev and A. Mielke. On the rate-independent limit of systems with dry friction and small viscosity. *J. Convex Anal.*, 13(1):151–167, 2006.

[3] D. Knees. Convergence of local minimization, relaxed local minimization and alternate minimization algorithms for rate-independent systems. Technical report, University of Kassel, 2017. submitted.

[4] D. Knees and M. Negri. Convergence of alternate minimization schemes for phase-field fracture and damage. *Math. Models Methods Appl. Sci.*, 27(9):1743–1794, 2017.

[5] D. Knees and A. Schröder. Computational aspects of quasi-static crack propagation. *Discrete Contin. Dyn. Syst., Ser. S*, 6(1):63–99, 2013.

[6] A. Mielke, R. Rossi, and G. Savaré. Balanced viscosity (BV) solutions to infinite-dimensional rate-independent systems. *J. Eur. Math. Soc. (JEMS)*, 18(9):2107–2165, 2016.

[7] A. Mielke and T. Roubíček. *Rate-independent systems. Theory and application.* New York, NY: Springer, 2015.

A variational approach to quantum master equations coupled to a semiconductor PDE

Markus Mittnenzweig

(joint work with Markus Kantner, Thomas Koprucki, Alexander Mielke)

Quantum master equations of Lindblad type are widely used in applications to model the dynamics of open quantum systems interacting with the environment. They are of the form

$$\dot{\rho} = \frac{\mathrm{i}}{\hbar}\left[\rho, H\right] + \mathcal{L}(\rho) \tag{1}$$

where $\mathcal{L}(\rho)$ is

$$\mathcal{L}(\rho) = \sum_{\alpha} \gamma_\alpha \big(A_\alpha \rho A_\alpha^\dagger - \frac{1}{2}\big\{ A_\alpha^\dagger A_\alpha, \rho \big\} \big)$$

with γ_α positive. Recently it has been shown [3, 2], that Lindblad operators $\mathcal{L}$ satisfying a quantum detailed balance condition also possess a gradient flow

structure with respect to the quantum free energy. More precisely, the following holds true.

Proposition 1 (Carlen-Maas 2017, Mittnenzweig-Mielke 2017)**.** *For every finite-dimensional Lindblad operator satisfying the GNS quantum detailed balance condition, there exists an Onsager operator* $\mathbb{K}(\rho)$ *such that*

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

where $\mathcal{F}(\rho) = \mathrm{tr}(\rho \log \rho + \beta H \rho)$ *is the quantum free energy.*

Here, we want to focus on the application of this gradient structure to the formulation of hybrid quantum-classical models for semiconductor quantum dots (QD) [1]. Modern optoelectronic devices increasingly use the quantum-mechanical properties of few quantum dots such as entanglement or non-classical correlations which makes it necessary to use density-matrix based descriptions for the dynamics of the QDs. On the other hand, the simulation of whole semiconductor devices needs a more coarse-grained description of the electrons and holes, such as the well-established reaction-drift-diffusion transport models for instance. Thus, the necessity of a hybrid model arises, where parts of the system are modeled quantum-mechanically and other parts classically by a semiconductor PDE. Both parts need to interact such that a general coupled quantum-classical model reads

$$-\nabla \cdot \varepsilon \nabla \psi = q\left(p - n + C + Q\left(\rho\right)\right), \tag{2}$$

$$\partial_t n - \frac{1}{q}\nabla \cdot \mathbf{j}_n = -R - S_n\left(\rho; \psi, n, p\right), \tag{3}$$

$$\partial_t p + \frac{1}{q}\nabla \cdot \mathbf{j}_p = -R - S_p\left(\rho; \psi, n, p\right), \tag{4}$$

$$\frac{\mathrm{d}}{\mathrm{d}t}\rho = -\frac{\mathrm{i}}{\hbar}\left[H, \rho\right] + \mathcal{L}\left(\rho; \psi, n, p\right). \tag{5}$$

The first equation is the Poisson equation for the electrostatic potential ψ. Equations (3) and (4) model the transport of the electrons and holes, where R contains the creation and recombination of electrons and holes. The last equation gives the dynamics of the quantum system. Note that n, p and ρ are coupled by the terms $S_{n/p}$ and $\mathcal{L}$ which model the scattering of electrons and holes from the semiconductor bulk into the quantum system. In the following we want to focus on this coupling part and we want to show that the coupling terms used in [1] are of gradient type, i.e. they are generated by a gradient structure

$$\begin{pmatrix} -S_n(\rho; \psi, n, p) \\ -S_p(\rho; \psi, n, p) \\ \mathcal{L}(\rho; \psi, n, p) \end{pmatrix} = -\mathbb{K}(\rho, n, p)\mathrm{D}\mathcal{F}(\rho, n, p)$$

where $\mathcal{F}$ is the free energy of the coupled system. The advantage of such a variational modeling approach is that it automatically gives thermodynamically consistent equations. It ensures positivity of the entropy production and preserves the global charge of the coupled system as we shall see. Conservation of charge is a

necessary condition on every coupling from a physical point of view. The charge of the quantum system is given by $\mathrm{Tr}(\rho N)$, where N is the charge operator.

Using Proposition 1 one can show that there exists Onsager operators such that the following $\mathcal{L}$ is a Lindblad operator satisfying detailed balance:

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

For every scattering process, one can now construct a corresponding Onsager operator that generates it. In the case of electron scattering, it has the form

$$\begin{pmatrix} -S_n(\rho;\mu_e,\psi) \\ \mathcal{L}_n(\rho) \end{pmatrix} = -\mathbb{K}_e \begin{pmatrix} \mu_e + \psi \\ \log\rho + \beta(H-\mu N) \end{pmatrix}$$

with μ being the chemical potential and

$$\mathbb{K}_e = \begin{pmatrix} \mathbb{K}(\rho,\mu)\square & -\langle\square\rangle_w \mathbb{K}(\rho,\mu)N \\ -w(x)\mathrm{tr}\big(N\mathbb{K}(\rho,\mu)\square\big) & w(x)\mathrm{tr}\big(N\langle\square\rangle_w\,\mathbb{K}(\rho,\mu)N\big). \end{pmatrix}$$

The notation $\square$ refers to the position, where the argument is to be inserted and

$$\langle f\rangle_w = \int_\Omega w(x)f(x)\,\mathrm{d}x$$

is the averaging of the macroscopic quantity f with respect to the coupling function $w(\cdot)$. This weight $w(x)$ is the central function that connects the density matrix with the PDE system. The weight w is non-negative and integrates to 1. Moreover, $\mathbb{K}$ is a symmetric and positive operator, so it is indeed an Onsager operator and thus ensures positivity of the entropy production. The conservation of charge of the coupled system also follows from the special structure of the Onsager operator. Quantities of the PDE system such as electrostatic or chemical potential couple to the quantum system by averaging with w. So the quantum system experiences only a local average of the macroscopic system. Quantum-mechanical quantities embed into the PDE by taking the trace with ρ and multiplication with $w(x)$. The quantum-mechnanical contribution $Q(\rho)$ to the Poisson equation for instance is given by

$$Q(\rho) = \mathrm{tr}\big(N\rho\big)\cdot w(x).$$

The central point of the above variational coupling is that it suggests to average the thermodynamic potentials and not the electron or hole densities instead, thereby creating thermodynamically consistent equations. Please see [1] for more details on the model as well as its application to the simulation of an electrically driven single-photon sources.

References

[1] M. Kantner, M. Mittnenzweig, T. Koprucki, *Hybrid quantum-classical modeling of quantum dot devices*, Phys. Rev. B, **96** (2017), 205301.

[2] M. Mittnenzweig, A. Mielke, *An entropic gradient structure for Lindblad equations and couplings of quantum systems to macroscopic models*, J. Stat. Phys. **167** (2017), 205–233.

[3] E. A. Carlen, J. Maas, *Gradient flow and entropy inequalities for quantum Markov semigroups with detailed balance*, J. Funct. Anal. **273** (2017), 1810–1869.

In between Energetic and Balanced Viscosity solutions to rate-independent systems: the Visco-Energetic concept

Riccarda Rossi
(joint work with Giuseppe Savaré)

A broad class of *rate-independent systems* (hereafter abbreviated as RIS) is mathematically modeled in terms of

- a time-dependent, *driving* energy functional $\mathcal{E} : [0,T] \times X \to (-\infty, \infty]$, with $[0,T]$ the time span during which the system is observed, and X the space of the states of the system,
- a dissipation functional $\mathcal{D} : X \times X \to [0,\infty]$, keeping track of the energy dissipated by the curve $u : [0,T] \to X$ describing the system evolution.

The aim of this note is to illustrate the novel approach of *Visco-Energetic* (VE) solutions to the weak solvability of RIS, and to compare it with the existing notions of *Energetic* (E) and *Balanced Viscosity* (BV) solutions. We will confine the discussion to the simple, yet significant case in which (X, d) is a (complete) metric space and the dissipation $\mathcal{D}$ is induced by the distance d via $\mathcal{D}(u_1, u_2) = \mathsf{d}(u_1, u_2)$.

First, let us give a quick overview of the E and BV concepts, referring to [1] and the bibliography therein for all precise statements. E solutions can be constructed by recursively solving the time-incremental minimization scheme

$$(\mathrm{IM}_\tau) \qquad U_\tau^n \in \mathrm{Argmin}_{U \in X} \left(\mathsf{d}(U_\tau^{n-1}, U) + \mathcal{E}(t_\tau^n, U) \right), \quad n = 1, \dots, N_\tau \,,$$

where $\{t_\tau^n\}_{n=0}^{N_\tau}$ is a partition of $[0,T]$ with (fixed) time step $\tau > 0$ and we set $U_\tau^0 := u_0$, the initial datum. Under suitable conditions on $\mathcal{E}$, the piecewise constant interpolants $(\overline{U_\tau})_\tau$ of the discrete solutions $(U_\tau^n)_{n=1}^{N_\tau}$ converge as $\tau \downarrow 0$ to an E solution of the RIS $(X, \mathcal{E}, \mathsf{d})$, namely a curve $u \in \mathrm{BV}_{\mathsf{d}}([0,T]; X)$ fulfilling

- the *global* stability condition

$$(\mathrm{S}_{\mathsf{d}}) \qquad \mathcal{E}(t, u(t)) \le \mathcal{E}(t, v) + \mathsf{d}(u(t), v) \quad \text{for every } v \in X,$$

- the energy-dissipation balance for all $t \in [0,T]$

$$(\mathrm{E}_{\mathsf{d}}) \qquad \mathcal{E}(t, u(t)) + \mathrm{Var}_{\mathsf{d}}(u, [0,t]) = \mathcal{E}(0, u(0)) + \int_0^t \partial_t \mathcal{E}(s, u(s)) \, \mathrm{d}s \,.$$

It can be shown that, at every point t in the jump set J_u of u, the left and right limits $u(t-)$ and $u(t+)$ fulfill the *jump conditions*

$$(1) \qquad \begin{cases} \mathcal{E}(t, u(t-)) - \mathcal{E}(t, u(t)) = \mathsf{d}(u(t-), u(t)), \\ \mathcal{E}(t, u(t)) - \mathcal{E}(t, u(t+)) = \mathsf{d}(u(t), u(t+)), \end{cases}$$

which are strongly influenced by the *global* energy landscape. It is well known that, because of this, E solutions driven by nonconvex energies may have to jump 'too early' and 'too long'. This fact has motivated the introduction of an alternative weak solvability concept for the RIS $(X, \mathcal{E}, \mathsf{d})$, which can be obtained starting from this modified time-incremental scheme, for $n = 1, \dots, N_\tau$,

$$(\mathrm{IM}_{\varepsilon,\tau}) \qquad U_\tau^n \in \mathrm{Argmin}_{U \in X} \left(\mathsf{d}(U_\tau^{n-1}, U) + \frac{\varepsilon}{2\tau} \mathsf{d}^2(U_\tau^{n-1}, U) + \mathcal{E}(t_\tau^n, U) \right).$$

Here, the *viscous correction* $\mathsf{d}^2(U_\tau^{n-1}, U)$ is modulated by a parameter ε such that $\frac{\varepsilon}{\tau} \uparrow \infty$. Under appropriate conditions on $\mathcal{E}$, $(\overline{U}_\tau)_\tau$ originating from $(\mathrm{IM}_{\varepsilon,\tau})$ converge as $\tau \downarrow 0$ to a BV solution of the RIS, i.e. $u \in \mathrm{BV}_{\mathsf{d}}([0,T];X)$ fulfilling

- the *local* stability condition

$$|\mathrm{D}_u\mathcal{E}|(t,u(t)) \le 1 \quad \text{for every } t \in [0,T] \setminus \mathrm{J}_u, \tag{$\mathrm{S}_{\mathrm{loc}}$}$$

- the 'BV energy-dissipation' balance for all $t \in [0,T]$

$$\mathcal{E}(t,u(t)) + \mathrm{Var}_{\mathsf{d},\mathsf{v}}(u,[0,t]) = \mathcal{E}(0,u(0)) + \int_0^t \partial_t\mathcal{E}(s,u(s))\,\mathrm{d}s\,. \tag{E_{BV}}$$

In $(\mathrm{E}_{\mathrm{BV}})$, the total variation functional $\mathrm{Var}_{\mathsf{d},\mathsf{v}}$ measures the energy dissipated at a jump point $t \in \mathrm{J}_u$ in terms of a Finsler-type cost $\mathsf{v}(t,\cdot,\cdot)$, defined by

$$\begin{aligned}\mathsf{v}(t,u(t-),u(t+)) := \inf\Big\{ \int_0^1 |\vartheta'|(r)\ (|\mathrm{D}_u\mathcal{E}|(t,\vartheta(r))\vee 1)\ \mathrm{d}r\ :\\ \vartheta \in \mathrm{AC}([0,1];X),\ \vartheta(0)=u(t-),\ \vartheta(1)=u(t+)\Big\}\end{aligned} \tag{2}$$

with $|\vartheta'|$ and $|\mathrm{D}_u\mathcal{E}|$ the metric derivative of ϑ and local slope of $\mathcal{E}(t,\cdot)$, respectively. Accordingly, at every $t \in \mathrm{J}_u$ there hold the jump conditions

$$\begin{cases} \mathcal{E}(t,u(t-)) - \mathcal{E}(t,u(t)) = \mathsf{v}(t,u(t-),u(t)), \\ \mathcal{E}(t,u(t)) - \mathcal{E}(t,u(t+)) = \mathsf{v}(t,u(t),u(t+))\,. \end{cases} \tag{3}$$

The BV cost $\mathsf{v}(t,\cdot,\cdot)$ indeed encodes the onset of *viscosity*, hence of rate-dependence, into the description of the system behavior at jumps. Furthermore, because of the *local* character of the stability condition $(\mathrm{S}_{\mathrm{loc}})$, BV solutions driven by nonconvex energies have mechanically feasible jumps. Nonetheless, a crucial requirement underlying the whole Balanced Viscosity theory is that the energy $\mathcal{E}$ complies with a chain-rule type condition, which is ultimately related to regularity properties and unavoidably restricts the range of applicability of BV solutions.

That is why, VE solutions have been proposed in the recent [2] as an alternative solution notion. The key idea is to broaden the class of admissible viscous corrections of the time-incremental scheme: VE solutions indeed originate in the time-continuous limit of discrete solutions to

$$U_\tau^n \in \mathrm{Argmin}_{U\in X}\left(\mathsf{d}(U_\tau^{n-1},U) + \delta(U_\tau^{n-1},U) + \mathcal{E}(t_\tau^n,U)\right), \quad n = 1,\dots,N_\tau,$$

where $\delta : X \times X \to [0,\infty]$ is a general, lower semicontinuous functional. However, we shall confine the present exposition to the simpler, but still significant, case $\delta(u,v) = \frac{\mu}{2}\mathsf{d}^2(u,v)$, with $\mu > 0$ a *fixed* parameter, and refer to [2] for the analysis of the general case. The time-incremental minimization scheme thus becomes

$$(\mathrm{IM}_\mu)\quad U_\tau^n \in \mathrm{Argmin}_{U\in X}\left(\mathsf{d}(U_\tau^{n-1},U) + \frac{\mu}{2}\mathsf{d}^2(U_\tau^{n-1},U) + \mathcal{E}(t_\tau^n,U)\right) \text{ with } \mu > 0.$$

In [2, Thm. 3.9] the discrete solutions $(\overline{U}_\tau)_\tau$ of (IM_μ) were shown to converge, as $\tau \downarrow 0$, to a VE solution of the RIS, i.e. $u \in \mathrm{BV}_{\mathsf{d}}([0,T];X)$ fulfilling

- the viscously perturbed, but still global, stability condition

$$(\mathrm{S_{VE}})\quad \mathcal{E}(t,u(t)) \le \mathcal{E}(t,v)+\mathsf{d}(u(t),v)+\frac{\mu}{2}\mathsf{d}^2(u(t),v) \quad \forall v\in X \text{ and } \forall t\in[0,T]\setminus \mathrm{J}_u,$$

- the energy-dissipation balance for all $t\in[0,T]$

$$(\mathrm{E_{VE}})\qquad \mathcal{E}(t,u(t)) + \mathrm{Var}_{\mathsf{d},\mathsf{c}}(u,[0,t]) = \mathcal{E}(0,u(0)) + \int_0^t \partial_t\mathcal{E}(s,u(s))\,\mathrm{d}s\,.$$

In ($\mathrm{E_{VE}}$), the contribution to the total variation functional $\mathrm{Var}_{\mathsf{d},\mathsf{c}}$ at a jump point $t\in \mathrm{J}_u$ now features the 'VE cost', which also enters into the jump conditions,

$$\mathsf{c}(t,u(t-),u(t+)) := \inf\Big\{\mathrm{Trc_{VE}}(t;\vartheta,E)\ :E\Subset\mathbb{R},\vartheta\in \mathrm{C}(E;X),\\ \vartheta(\inf E)=u(t-),\ \vartheta(\sup E)=u(t+)\Big\},$$

where the transition cost $\mathrm{Trc_{VE}}(t;\vartheta,E)$ is given by $\mathrm{Trc_{VE}}(t;\vartheta,E)=\mathrm{Var}_{\mathsf{d}}(\vartheta,E)+\mathrm{GapVar}_{\mathsf{d}}(\vartheta,E)+\sum_{s\in E\setminus\sup E}\mathcal{R}(t,\vartheta(s))$ and

- since E need not be interval, it may have 'holes' (i.e., connected components of $[\inf E,\sup E]\setminus E$) and

$$\mathrm{GapVar}_{\mathsf{d}}(\vartheta,E)=\sum_{I \text{ hole of } E}\frac{\mu}{2}\mathsf{d}^2(\vartheta(\inf I),\vartheta(\sup I)),$$

- the residual function $\mathcal{R}:[0,T]\times X$ measures the failure of the VE-stability:

$$\mathcal{R}(t,u):=\sup_{v\in X}\Big(\mathcal{E}(t,u)-\mathcal{E}(t,v)-\mathsf{d}(u,v)-\frac{\mu}{2}\mathsf{d}^2(u,v)\Big)\,.$$

In fact, VE solutions are in between E and BV solutions in these respects:

- the assumptions for the existence theory, weaker than for E and stronger than for BV solutions (cf. [2, Thm. 3.9]), and the modeling of jumps;
- the singular limits $\mu\downarrow 0$ and $\mu\uparrow\infty$: VE solutions converge to an E and to a BV solution, respectively, as shown in [3];
- the range of applicability of the theory to rate-independent systems in solid mechanics, cf. [4].

References

[1] A. Mielke and T. Roubíček, Rate-independent systems. Theory and application. Applied Mathematical Sciences, 193. Springer, New York, 2015.

[2] L. Minotti and G. Savaré, *Viscous corrections of the Time Incremental Minimization Scheme and Visco-Energetic Solutions to Rate-Independent Evolution Problems*, Arch. Rational Mech. Anal. (2017). https://doi.org/10.1007/s00205-017-1165-5.

[3] R. Rossi and G. Savaré, *From Visco-Energetic to Energetic and Balanced Viscosity Solutions of Rate-Independent Systems*, pp. 489–531 in Colli P., Favini A., Rocca E., Schimperna G., Sprekels J. (eds), "Solvability, Regularity, and Optimal Control of Boundary Value Problems for PDEs". Springer INdAM Series, vol 22. Springer, Cham.

[4] R. Rossi, *Visco-energetic solutions of some rate-independent systems in delamination, damage, and plasticity*, arXiv: 1803.03768, 2018..

Generalised Gradient Systems and GENERIC in Flux Space

Michiel Renger

This work is largely motivated by the quest for thermodynamically consistent structures that characterise an evolution. Since many different structures could describe the same evolution equation, we restrict to structures that can be derived naturally by statistical mechanical arguments. This fits in the philosophy of Boltzmann, who related free energy to large deviations of a microscopic system in equilibrium, and Onsager, who related a gradient flow of free energy to dynamic large deviations of a system close to equilibrium. In our previous work [1], we related large deviations of systems satisfying detailed balance to generalised gradient systems. Deriving thermodynamic structures for systems that do not satisfy detailed balance remains the central challenge of non-equilibrium thermodynamics.

Breaking of detailed balance typically occurs due to the occurence of non-trivial but divergence-free fluxes. It is therefore widely believed that a good understanding of these fluxes are one of the missing pieces of the puzzle. This sparks the question whether flux large deviations could induce a structure like a gradient system or Generic [3], where the state large deviations fail to induce something alike. This turns out to be impossible.

In a general setting, we are given a Lagrangian $\mathcal{L} : \mathcal{W} \to \mathbb{R}_+$ on a manifold $\mathcal{W}$ of (integrated) fluxes, a manifold of states $\mathcal{X}$ and a mapping $\phi : \mathcal{W} \to \mathcal{X}$. The corresponding Lagrangian $\hat{\mathcal{L}} : \mathcal{X} \to \mathbb{R}_+$ is then obtained by

$$\hat{\mathcal{L}}(\rho, \dot\rho) := \inf_{w:\phi(w)=\rho} \inf_{\dot w : d\phi_w \dot w = \dot\rho} \mathcal{L}(w, \dot w), \tag{1}$$

where the relation $d\phi_w \dot w = \dot\rho$ is simply a continuity equation.

We will say that the Lagrangian $\mathcal{L}$ induces a generalised gradient system (GGS) $(\mathcal{W}, \Psi, \mathcal{F})$ whenever one can write

$$\mathcal{L}(w, \dot w) = \Psi(w, \dot w) + \Psi^*\big(w, -d_w\mathcal{F}(w)\big) + \langle d_w\mathcal{F}(w), \dot w\rangle,$$

where $\mathcal{F}$ is a free energy and Ψ, Ψ^* is a convex dual pair such that $\Psi(w, 0) = 0 = \Psi^*(w, 0)$. This generalises a linear gradient system, where $\Psi(w, \dot w) = \frac{1}{2}\langle \dot w, K(w)\dot w\rangle$ for some metric tensor K.

The first abstract result states that if the flux Lagrangian $\mathcal{L}$ induces a GGS $(\mathcal{W}, \Psi, \mathcal{F})$ and $\mathcal{F}(w) = \hat{\mathcal{F}}(\phi(w))$ for some $\hat{\mathcal{F}}$, that is, the free energy depends on the state only, then the state Lagrangian $\hat{\mathcal{L}}$ also induces a GGS $(\mathcal{X}, \hat\Psi, \hat{\mathcal{F}})$. Moreover, if for the corresponding Hamiltonian $\mathcal{H}(w, d\phi_w^{\mathsf{T}}\xi)$ also depends on w through $\phi(w)$ only, then we can calculate $\hat\Psi, \hat\Psi^*$ explicitly through

$$\hat\Psi\big(\phi(w), \dot\rho\big) = \inf_{d\phi_w \dot w = \dot\rho} \Psi(w, \dot w) \qquad \text{and} \qquad \hat\Psi^*\big(\phi(w), \xi\big) = \Psi^*(w, d\phi_w^{\mathsf{T}}\xi). \tag{2}$$

On the other hand, if a state Lagrangian $\hat{\mathcal{L}}$ induces a GGS $(\mathcal{X}, \hat\Psi, \hat{\mathcal{F}})$, then we can always modify the flux Lagrangian by

$$\tilde{\mathcal{L}}(w, \dot w) := \mathcal{L}\Big(w, \dot w + d_\zeta \mathcal{H}\big(w, d\phi_w^{\mathsf{T}} d_\rho \hat{\mathcal{F}}(\phi(w))\big)\Big),$$

such that $\tilde{\mathcal{L}}$ induces a GGS $(\mathcal{W}, \Psi, \mathcal{F})$ where $\mathcal{F}(w) = \hat{\mathcal{F}}(\phi(w))$, and we still have $\hat{\mathcal{L}}(\rho, s) = \inf_{\phi(w)=\rho} \inf_{d\phi_w j = s} \tilde{\mathcal{L}}(w, \dot{w})$.

In a similar fashion, we say that the Lagrangian $\mathcal{L}$ induces a generalised Generic system (GGEN) $(\mathcal{W}, \Psi, \mathcal{F}, L, \mathcal{E})$ whenever one can write

$$\mathcal{L}(w, \dot{w}) = \Psi\big(w, \dot{w} - L(w) d_w \mathcal{E}(w)\big) + \Psi^*\big(w, -d_w \mathcal{F}(w)\big) + \langle d_w \mathcal{F}(w), \dot{w} \rangle,$$

where $\mathcal{F}, \mathcal{E}$ are energies, Ψ, Ψ^* is again a convex dual pair such that $\Psi(w, 0) = 0 = \Psi^*(w, 0)$, L is a so-called Poisson-structure, and the following two non-interaction conditions are satisfied (see [2]):

$$\Psi^*\big(w, \zeta + \lambda d_w \mathcal{E}(w)\big) = \Psi^*(w, \zeta) \text{ for all } (w, \zeta) \in T^*\mathcal{W} \text{ and } \lambda \in \mathbb{R},$$
$$L(w) d_w \mathcal{F}(w) = 0 \text{ for all } w \in \mathcal{W}.$$

The second abstract result states that if the flux Lagrangian $\mathcal{L}$ induces a GGEN $(\mathcal{W}, \Psi, \mathcal{F}, L, \mathcal{E})$ and $\mathcal{F}(w) = \hat{\mathcal{F}}(\phi(w)), \mathcal{E}(w) = \hat{\mathcal{E}}(\phi(w))$ for some $\hat{\mathcal{F}}, \hat{\mathcal{E}}$, and

$$\hat{L}\big(\phi(w)\big) := d\phi_w L(w) d\phi_w^{\mathsf{T}} \tag{3}$$

depends on the state $\phi(w)$ only, then the state Lagrangian $\hat{\mathcal{L}}$ also induces a GGEN $(\mathcal{X}, \hat{\Psi}, \hat{\mathcal{F}}, \hat{L}, \hat{\mathcal{E}})$. In particular, the (3) leaves the 'Poissonian' structure of L intact.

A typical example is that of a system of independent Markov chains $X_k(t)$ with generator matrix Q on a finite state space I. The state variable is $\rho_i^{(n)}(t) := \frac{1}{n}\#\{X_k(t) = i\}$ and $W_{ij}^{(n)}(t), i < j$ measures the total net number of jumps from i to j in time $(0, t]$, also scaled with $1/n$. Now $(\phi(w))_i = \rho_i(0) - div_i\, w = \rho_i(0) - \sum_j w_{ij}$. The large-deviation Lagrangian is given by

$$\mathcal{L}(w, \dot{w}) = \inf_{\dot{w} = u - u^{\mathcal{T}}} \sum_{i,j \in I} u_{ij} \log \frac{u_{ij}}{\phi_i(w) Q_{ij}} - u_{ij} + \phi_i(w) Q_{ij}, \tag{4}$$

where the infimum comes from a contraction over one-way fluxes u. The Lagrangian $\mathcal{L}$ induces a GGS $(\mathbb{R}^{\{i<j\}}, \Psi, \mathcal{F})$ if and only Q satisfies detailed balance w.r.t. invariant measure π, in which case

$$\mathcal{F}(w) = \frac{1}{2} \sum_{i \in I} \phi_i(w) \log \frac{\phi_i(w)}{\pi_i}, \qquad \text{and}$$
$$\Psi^*(w, \zeta) = 2 \sum_{i<j} \sqrt{\phi_i(w)\phi_j(w) Q_{ij} Q_{ji}}\big(\cosh(\zeta_{ij}) - 1\big).$$

By result stated above, the state Lagrangian $\hat{\mathcal{L}}$ now also induces a GGS $(\mathcal{X}, \hat{\Psi}, \hat{\mathcal{F}})$. From (2) and $(d\phi_w^{\mathsf{T}} \xi)_{ij} = \nabla_{ij} \xi = \xi_j - \xi_i$, we retrieve the GGS as found in [1]:

$$\hat{\Psi}^*(\rho, \xi) = 2 \sum_{i<j} \sqrt{\rho_i \rho_j Q_{ij} Q_{ji}}\big(\cosh(\xi_j - \xi_i) - 1\big).$$

By the above arguments one should not expect to find a gradient or Generic system in flux space if the evolution in state space does not have such structure. Nevertheless, there are *good reasons to study flux Lagrangians*. The Lagrangian in flux space is often a simple expression, whereas the state Lagrangian may not even have an explicit formula. More precisely, if the dynamics consists of multiple

mechanisms that drive the system, then the flux Lagrangian is a sum over these mechanisms, whereas the state Lagrangian (1) has the form of an inf-convolution over fluxes. In the example above, we needed to study net fluxes in order to derive a GGS, but the corresponding Lagrangian on one-way fluxes would be simply (4) without the infimum. Naturally, having such an explicit Lagrangian could be very practical, e.g. when studying multiscale expansions or phase transitions. In a sense, the different mechanisms driving a system become decoupled in flux space. This principle could also be useful to derive operator splitting techniques for systems with interacting mechanisms, like for example reaction-diffusion equations.

References

[1] A. Mielke, M.A. Peletier, and D.R.M. Renger, *On the relation between gradient flows and the large-deviation principle, with applications to Markov chains and diffusion*, Potential Analysis **41** (2014), 1293–1327.

[2] A. Mielke *Formulation of thermoelastic dissipative material behavior using GENERIC*, Continuum Mech. Thermodyn. **23** (2011), 233–256.

[3] H.C. Öttinger, *Beyond equilibrium thermodynamics*, Wiley-Interscience, Hoboken, NJ, USA (2005).

Reporter: Artur Stephan

Participants

Dr. Jean-David Benamou
INRIA Paris
2, Rue Simone Iff
75012 Paris Cedex
FRANCE

Prof. Dr. Yann Brenier
DMA
Ecole Normale Supérieure
45, rue d'Ulm
75005 Paris Cedex
FRANCE

Dr. Clément Cancès
Equipe RAPSODI
INRIA Lille-Nord Europe
40, Avenue Halley
59650 Villeneuve-d'Ascq
FRANCE

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

Prof. Dr. Antonin Chambolle
Centre de Mathématiques Appliquées
École Polytechnique
Plateau de Palaiseau
91128 Palaiseau Cedex
FRANCE

Prof. Dr. Gianni Dal Maso
SISSA
Via Bonomea 265
34136 Trieste
ITALY

Thibaut Demaerel
Departement Wiskunde
Faculteit der Wetenschappen
Katholieke Universiteit Leuven
Celestijnenlaan 200B
3001 Leuven
BELGIUM

Prof. Dr. Marco Di Francesco
Dipartimento di Ingegneria e Scienze
dell'Informazione e Matematica (DISIM)
Università degli Studi dell'Aquila
Via Vetoio 1, Coppito
67010 L'Aquila (AQ)
ITALY

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

Prof. Dr. Jin Feng
Department of Mathematics
University of Kansas
405 Snow Hall
Lawrence, KS 66045-7567
UNITED STATES

Dr. Julian Fischer
Institute of Science and
Technology Austria (IST Austria)
Am Campus 1
3400 Klosterneuburg
AUSTRIA

Thomas Frenzel
Weierstraß-Institut für
Angewandte Analysis und Stochastik
Mohrenstrasse 39
10117 Berlin
GERMANY

Dr. Nicolas Garcia-Trillos
Department of Mathematics
Brown University
Box F
182 George Street
Providence, RI 02912
UNITED STATES

Prof. Dr. Adriana Garroni
Dipartimento di Matematica
"Guido Castelnuovo"
Universita di Roma "La Sapienza"
Piazzale Aldo Moro, 2
00185 Roma
ITALY

Prof. Dr. Nicola Gigli
SISSA
Via Bonomea 265
34136 Trieste
ITALY

Prof. Dr. Michael Herrmann
Institut Computational Mathematics
Technische Universität Braunschweig
Universitätsplatz 2
38106 Braunschweig
GERMANY

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

Prof. Dr. Dorothee Knees
FB 10 - Mathematik und Naturwissenschaften
Arbeitsgruppe AAM
Universität Kassel
Heinrich-Plett-Strasse 40
34132 Kassel
GERMANY

Dr. Vaios Laschos
Weierstraß-Institut für Angewandte Analysis und Stochastik
Mohrenstrasse 39
10117 Berlin
GERMANY

Prof. Dr. Christian Léonard
Département de Mathématiques
Université de Paris X
200, Ave. de la Republique
92001 Nanterre Cedex
FRANCE

Matthias Liero
Weierstraß-Institut für Angewandte Analysis und Stochastik
Mohrenstrasse 39
10117 Berlin
GERMANY

Dr. Jan Maas
IST Austria
Am Campus 1
3400 Klosterneuburg
AUSTRIA

Prof. Dr. Christian Maes
Instituut voor Theoretische Fysika
Katholieke Universiteit Leuven
Celestijnenlaan 200 D
3001 Leuven
BELGIUM

Prof. Dr. Daniel Matthes
Zentrum Mathematik
Technische Universität München
Boltzmannstrasse 3
85747 Garching bei München
GERMANY

Stefano Melchionna
Fakultät für Mathematik
Universität Wien
Oskar-Morgenstern-Platz 1
1090 Wien
AUSTRIA

Prof. Dr. Alexander Mielke
Weierstraß-Institut für Angewandte
Analysis und Stochastik
Mohrenstrasse 39
10117 Berlin
GERMANY

Markus Mittnenzweig
Weierstraß-Institut für
Angewandte Analysis und Stochastik
Mohrenstrasse 39
10117 Berlin
GERMANY

Dr. Léonard Monsaingeon
Institut Élie Cartan
UMR 7502 du CNRS
Université de Lorraine
BP 70239
54506 Vandoeuvre-lès-Nancy
FRANCE

Prof. Dr. Delio Mugnolo
Fakultät für Mathematik und Informatik
Lehrgebiet Analysis
Fernuniversität zu Hagen
58084 Hagen
GERMANY

Prof. Dr. Barbara Niethammer
Institut für Angewandte Mathematik
Universität Bonn
Endenicher Allee 60
53115 Bonn
GERMANY

Prof. Dr. Hans Christian Öttinger
Institut für Polymere
ETH Zürich
HCP F47.2
Leopold-Ruzicka-Weg 4
8093 Zürich
SWITZERLAND

Prof. Dr. Felix Otto
Max-Planck-Institut für Mathematik
in den Naturwissenschaften
Inselstrasse 22 - 26
04103 Leipzig
GERMANY

Dr. Francesco S. Patacchini
Department of Mathematical Sciences
Carnegie Mellon University
Pittsburgh, PA 15213-3890
UNITED STATES

Prof. Dr. Mark A. Peletier
Department of Mathematics
Eindhoven University of Technology
P.O.Box 513
5600 MB Eindhoven
NETHERLANDS

Dr. Michiel Renger
Weierstraß-Institut für
Angewandte Analysis und Stochastik
Mohrenstrasse 39
10117 Berlin
GERMANY

Dr. Riccarda Rossi
Dipartimento di Matematica
Universita di Brescia
Via Valotti, 9
25133 Brescia
ITALY

Dr. Tomas Roubicek
Mathematical Institute
Charles University
Sokolovska 83
186 75 Praha 8
CZECH REPUBLIC

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
Centre d'Orsay
Université Paris-Sud
Batiment 425
91405 Orsay Cedex
FRANCE

Prof. Giuseppe Savare
Dipartimento di Matematica
"F. Casorati"
Universita di Pavia
Via Ferrata, 1
27100 Pavia
ITALY

Dr. Upanshu Sharma
CERMICS - ENPC
Cite Descartes, Champs-sur-Marne
6 et 8 Ave. Blaise Pascal
77455 Marne-la-Vallée Cedex 2
FRANCE

Prof. Dr. Dejan Slepcev
Department of Mathematical Sciences
Carnegie Mellon University
5000 Forbes Avenue
Pittsburgh, PA 15213-3890
UNITED STATES

Prof. Dr. Ulisse Stefanelli
Fakultät für Mathematik
Universität Wien
Oskar-Morgenstern-Platz 1
1090 Wien
AUSTRIA

Artur Stephan
Fachbereich Mathematik
Humboldt Universität Berlin
Unter den Linden 6
10099 Berlin
GERMANY

Dr. Marita Thomas
Weierstraß-Institut für
Angewandte Analysis und Stochastik
Mohrenstrasse 39
10117 Berlin
GERMANY

Dr. Oliver T. C. Tse
Department of Mathematics
Eindhoven University of Technology
P.O.Box 513
5600 MB Eindhoven
NETHERLANDS

Prof. Dr. Max von Renesse
Mathematisches Institut
Universität Leipzig
Augustusplatz 10
04109 Leipzig
GERMANY

Prof. Dr. Stephen J. Watson
Department of Mathematics
University of Glasgow
University Place
Glasgow G12 8QW
UNITED KINGDOM

Prof. Dr. Michael Westdickenberg
Institut für Mathematik
RWTH Aachen
Templergraben 55
52062 Aachen
GERMANY

Prof. Dr. Barbara Zwicknagl
Lehrstuhl für Mathematik XI
Universität Würzburg
Emil-Fischer-Strasse 40
97074 Würzburg
GERMANY

Prof. Dr. Johannes Zimmer
Department of Mathematical Sciences
University of Bath
Bath BA2 7AY
UNITED KINGDOM