Mathematisches Forschungsinstitut Oberwolfach

Report No. 29/2020

DOI: 10.4171/OWR/2020/29

Variational Methods for Evolution (hybrid meeting)

Organized by
Alexander Mielke, Berlin
Mark Peletier, Eindhoven
Dejan Slepčev, Pittsburgh

13 September – 19 September 2020

Abstract. Variational principles for evolutionary systems take advantage of the rich toolbox provided by the theory of the calculus of variations. Such principles are available for Hamiltonian systems in classical mechanics, gradient flows for dissipative systems, but also time-incremental minimization techniques for more general evolutionary problems. The new challenges arise via the interplay of two or more functionals (e.g. a free energy and a dissipation potential), new structures (systems with nonlocal transport, gradient flows on graphs, kinetic equations, systems of equations) thus encompassing a large variety of applications in the modeling of materials and fluids, in biology, in multi-agent systems, and 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, optimal transport, gradient flows, and large-deviation principles for time-continuous Markov processes, Γ-convergence and homogenization.

Mathematics Subject Classification (2020): 49-06, 35-06, 70-06, 58E30, 60F10, 82C05.

Introduction by the Organizers

Variational approaches to evolution systems are relevant to a number of seemingly distant branches of mathematics and physical sciences, as well as, more recently, data science. The variational structure provides a unifying theme that creates

connections and a powerful viewpoint to study a large family of problems. Classically, this includes studies of geometric aspects like metric structures in infinite-dimensional function spaces and of the geometry of relevant energy landscapes. 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. New directions include related, geometric, variational descriptions to evolution in rate-independent systems, kinetic systems (e.g. Boltzmann equation), fluid dynamics, systems of reaction-diffusion equations, as well as equations on discrete graphs. Connection to stochastic analysis (in particular large-deviation principles) and statistical mechanics provide a rigorous connection between microscopic, particle-based, descriptions of the systems and their macroscopic dissipative structure.

The tools that were developed during the last few decades have shed new light on many evolutionary equations for models in mechanics, physics, chemistry, and biology. 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 researchers with a variety of backgrounds and perspectives ranging from calculus of variations, partial differential equations, stochastic analysis, to physics and numerical analysis for a stimulating interchange of ideas. Further goal was to provide a group of promising young researchers with opportunity to interact and have in extended discussions with senior researchers in the relaxed setting that MFO provides. A challenge that arose was the COVID-19 crisis that forced the workshop to run in a hybrid mode and prevented more than half of the participants from being physically present. To stimulate the exchange of ideas, seven discussion sessions were included in the program. Some of them ran in parallel and enabled in-depth discussions of pertinent problems in smaller groups. Nearly all of these separate discussion sessions were planned on-the-fly in response to questions that arose in the regular presentations.

The facilities that MFO provided enabled a surprisingly effective interaction between the in-person and online participants throughout the workshop. While the essential aspect of direct human interaction was not available with online participation, the efforts of many participants made the interactions effective and created many insightful discussions.

There was a total of 26 talks, which were, due to the hybrid format of the workshop, 30 minutes long. The talks discussed the following topics:

- gradient flows in systems with nonlocal transport (Tse, Schlichting)
- connection between (interacting) particle systems and the energy driven systems that arise as their continuum limits (Daneri, Feng, Júngel, Hraivoronska, Esposito, Jungel, Wolfram)

- large deviation principles and microscopic description of energy driven systems (Patterson, Renger, Zimmer)
- kinetic systems – gradient flow structure and asymptotic (Carrillo, Daneri)
- discrete interaction systems, evolution on graphs, and their metric and variational interpretation (Erbar, Maas, Schlichting)
- optimal transport techniques and transportation distances (Benamou, Brenier, Maas)
- energy-reaction-diffusion systems (Hopf, Stephan)
- exponential gradient structures and applications (Burger, Craig)
- gradient flows for measure sampling and quantization (Hoffmann, Iacobelli)
- waiting-time phenomena in dissipative systems (Matthes)
- regularity of solutions of free boundary problems (Figalli)
- rate-independent energy-driven systems (Knees)
- asymptotics of gradient flows (Westdickenberg)

Workshop (hybrid meeting): Variational Methods for Evolution

Table of Contents

Abstracts

Exponential gradient structures

Martin Burger

Gradient structures such as gradient flows in metric spaces or generalizations thereof have developed to an important tool in the analysis of nonlinear partial differential equations and Markov processes recently. A particularly celebrated example is the Fokker-Planck equation, which can be derived from discrete Markov processes with jump rates depending exponentially on the potential difference between jump and target location. Though widely used, there are application cases where it is not clear whether a potential difference is really the appropriate model, instead a simple direct dependence on the potential may be more suitable. An example is cell migration by chemotaxis, where the approach based on potential differences is still under debate, since it is not clear that all cells can sense a gradient of an external field and the alternative approach is understood as *local sensing* (cf. [8]).

The mathematical formulation of such local sensing problems leads to equations of the form

$$\partial_t u = -L(e^{E'(u)} - 1) \tag{1}$$

with a positive semidefinite operator L. We call equation (1) an *exponential gradient structure*, if there exists an operator D and a positive semidefinite operator $M(v)$ such that

$$\langle v, Le^v \rangle = \langle Dv, M(v)Dv \rangle.$$

The latter property guarantees the energy dissipation property

$$\partial_t E(u) = -\langle E'(u), D(e^{E'(u)})E'(u) \rangle \leq 0.$$

Such models recently arise in a variety of applications, ranging from cell motion (cf. [8, 1]) over collective animal behaviour (cf. [2]) to crystal growth (cf. [4, 7]).

The canonical examples are the negative Laplacian $L = -\Delta$ with D the usual gradient and $M(v)$ the multiplication with e^v, or a (negative) graph Laplacian

$$(Lv)_i = \sum_j w_{ij}(v_i - v_j).$$

The simplest model of this kind might be a linear reaction equation

$$\partial_t u = -\alpha u + \beta,$$

with the linear operator $L = \alpha I$ and the energy $E(u) = u \log u + \frac{\beta}{\alpha} u$, which however allows for several different gradient formulations due to the simplicity of L. In this case we can choose $D = I$ and $M(v)$ is the multiplication with the logarithmic mean $\frac{v-1}{\log v}$. In the general case D is the graph gradient and $M(v)$ is a diagonal operator on edge functions multiplying by $\frac{v_i - v_j}{\log v_i - \log v_j}$.

A paradigmatic case arises is the original model in jump processes (cf. [8]), where jump rates are due to local distances (described by some kernel K) and exponential statistics related to some potential V. This leads to

$$\partial_t u(x_i,t) = -\sum_j K(x_i - x_j)(e^{V(x_i)}u(x_i,t) - e^{V(x_j)}u(x_j,t)),$$

giving an exponential structure as in (1) with L being the graph Laplacian arising from kernel K and $E(u) = \sum_i u(x_i)(\log u(x_i) + V(x_i))$. In a localized continuum limit we obtain again the structure with the local Laplacian

$$\partial_t u = \Delta(e^V u) = \nabla \cdot (e^V u \nabla(\log u + V), \tag{2}$$

which displays the similarity to the classical Fokker-Planck equation

$$\partial_t u = \nabla \cdot (\nabla u + u\nabla V) = \nabla \cdot (u\nabla(\log u + V). \tag{3}$$

1. Analogue of Fokker-Planck Equation

In the following we further discuss (2) and its relation to the standard Fokker-Planck equation (3). The latter has been analyzed as a gradient flow in the Wasserstein-metric (cf. [5]) and in a similar way we can treat (2) as a gradient flow in a weighted metric

$$d(\rho_0,\rho_1)^2 = \inf_{\rho_t + \nabla\cdot j_t = 0} \int a(x) \frac{|j_t|^2}{\rho_t}\, dx$$

as carried out in [6]. In the Wasserstein case we have $a \equiv 1$, while in the exponential case we find $a = e^{-V}$, hence it is apparent that such a metric is topologically equivalent to the Wasserstein one if V is bounded. The geodesics may change however, which has an impact on the geodesic convexity of the energy functional E. However, due to the direct relation between mobility and energy, the analysis in [6] simplifies and indeed convexity of V still implies geodesic convexity and the Bakry-Emery condition is directly the positivity of $e^V \Delta V$ (cf. [6, Thm. 1.5]).

The long-time behaviour can also be understood from the relative entropy method. In particular for any Csiszar f-Divergence with a smooth convex function Φ we have

$$\frac{d}{dt}\int \Phi(\frac{u}{u_\infty})\, dx = -c\int \Phi''(\frac{u}{u_\infty}) \left|\nabla \frac{u}{u_\infty}\right|^2\, dx, \tag{4}$$

where $u_\infty = ce^{-V}$ is the unique equilibrium. In the case $\Phi(z) = (z-1)^2$ the Poincare inequality directly implies exponential convergence to equilibrium, interestingly with a rate that depends on the scaling constant c, i.e. on the mass of the solution.

2. Nonlocal Energies

In the case of nonlocal energies, the exponential gradient structure seems to induce much stronger changes compared to the Wasserstein gradient structure . Apart from the fact that an analogue of the Wasserstein metric with a mobility $ue^{E'(u)}$ is difficult to analyze per se, the properties of the metric and equation can be very

different if E' is unbounded. Such a case was analyzed in [1], where the gradient flow for the same energy as in the well-kown Keller-Segel model of chemotaxis (cf. [9]) was analyzed. It was shown that in the supercritical case, when the energy is not bounded from below there is a changed blow-up phenomenon. In the Keller-Segel model a celebrated result shows the finite-time blow up, while in the case of the exponential gradient structure the blow-up is delayed to the inifinite time limit. The key change in this case is a H^{-1} bound in finite time that naturally arises from the structure $\partial_t u = \Delta v$, also related to duality estimates for cross-diffusion systems (cf. [3]). On bounded sets in H^{-1} the energy is indeed bounded below, which is the main reason for global existence of solutions.

Another example that was analyzed so far is the structure for the heat equation, leading to the so-called exponential thin-film equation

$$\partial_t u = \Delta e^{-\Delta u}, \qquad E(u) = \frac{1}{2}\int |\nabla u|^2 \, dx, \tag{5}$$

but the analysis is based on the fact that the equation is also an L^2 gradient flow for the energy

$$F(u) = \int e^{-\Delta u} \, dx.$$

A very relevant and still completely open case is the analogous equation for E being the total variation instead of the Dirichlet energy, where one has to deal with the nonsmoothness of the functional and needs a variational argument to select an appropriate subgradient. Let us finally mention that similar statements may hold for structures with the graph Laplacian, but there are no results on this case yet.

References

[1] M. Burger, P. Laurençot, A. Trescases, *Delayed blow-up for chemotaxis models with local sensing*, Preprint arXiv:2005.02734 (2020).

[2] M. Burger, J. Haskovec, M.T. Wolfram, *Individual based and mean-field modeling of direct aggregation*, Physica D: Nonlinear Phenomena **260** (2013) , 145-158.

[3] L. Desvillettes, T. Lepoutre, A. Moussa, *Entropy, duality, and cross diffusion*, SIAM J. Math. Analysis, **46** (2014), 820–853.

[4] Y. Gao, J.G. Liu, X.Y. Lu, *Gradient flow approach to an exponential thin film equation: global existence and latent singularity*, ESAIM: Control, Optimisation and Calculus of Variations **25** (2019), 49.

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

[6] S.Lisini, *Nonlinear diffusion equations with variable coefficients as gradient flows in Wasserstein spaces*, ESAIM: Control, Optimisation and Calculus of Variations, **15** (2009), 712-740.

[7] J.G. Liu, J. Lu, D. Margetis, J.L. Marzuola, *Asymmetry in crystal facet dynamics of homoepitaxy by a continuum model*, Physica D: Nonlinear Phenomena **393** 2019, 54-67.

[8] H.G. Othmer, A. Stevens, *Aggregation, blowup, and collapse: the ABC's of taxis in reinforced random walks. SIAM Journal on Applied Mathematics*, **57** (1997), 1044-1081.

[9] B. Perthame, *Transport Equations in Biology*, Birkhäuser, Basel, 2006.

Two approaches to optimal convergence rates for (mildly) nonconvex gradient flows

Maria G. Westdickenberg
(joint work with Olga Chugreeva, Felix Otto, and Sebastian Scholtes)

We are interested in obtaining optimal, algebraic-in-time convergence rates for gradient flows with respect to nonconvex energies. We present two methods, the so-called HED method based on the natural norm associated to the gradient flow, and an L^1 method developed more recently.

To explain the idea behind the first method, we begin with an abstract gradient flow

$$u_t = -\nabla E(u) \qquad \text{with dissipation} \qquad D := ||\nabla E||^2.$$

Suppose that u converges to u_* for $t \to \infty$ and define the square-distance $H := ||u-u_*||^2$. An observation of Brezis reveals that a convex (but not strictly convex) gradient flow satisfies the algebraic relationship

$$E \leq \sqrt{HD}$$

as well as the differential relationships $\frac{d}{dt}H \leq 0$, $\frac{d}{dt}D \leq 0$, from which it is easy to deduce algebraic decay of the energy and the dissipation according to

$$E \leq \frac{H_0}{t}, \qquad D \leq \frac{H_0}{t^2}.$$

We are interested in adapting the observation of Brezis to handle nonconvex gradient flows. In [2], we observe that the one-dimensional Cahn-Hilliard equation on the line satisfies $E \leq C\sqrt{HD}$ and that $\frac{d}{dt}H$, $\frac{d}{dt}D$, while not nonpositive, are "small." Here "small" means small enough so that we are able to close a simple ODE argument to deduce algebraic decay of the energy. Rather than explaining the details of this work, we explain how the HED method is applied in [1] to analyze relaxation of a curved interface to the axis for the Mullins-Sekerka evolution in the plane. Again the ingredients are natural algebraic and differential relationships among the quantities H, E, and D and the result is

$$E \leq C\frac{H_0}{t}, \qquad D \leq C\frac{H_0}{t^2}.$$

In the second part of the talk, we return to the one-dimensional Cahn-Hilliard equation and ask whether we can obtain optimal decay rates under the assumption of a bounded L^1 disturbance. Explicitly: Our bound should not depend on the distance of the disturbance from the origin. In [3] we study disturbances to the centered "kink" profile v, i.e., the minimizer of the underlying energy subject to ± 1 boundary conditions at infinity, centered so that $v(0) = 0$. Since any shift $v_a := v(\cdot - a)$ of v is still a minimizer, energetic convergence alone will not determine the final state of the solution u of the Cahn-Hilliard equation, however

the conserved quantity

$$\int_{\mathbb{R}} u - v \, dt = \int_{\mathbb{R}} u_0 - v \, dx$$

can be used to identify to which kink the solution converges in the long-time limit; our goal in [3] is to determine the rate.

We consider initial data normalized so that

$$\int_{\mathbb{R}} u_0 - v \, dx = 0$$

and a bounded distance in L^1 away:

$$\int_{\mathbb{R}} |u_0 - v| \, dx < \infty.$$

At any time t, we associate to the solution u a shifted kink v_c that minimizes the L^2 distance:

$$v_c = \operatorname{argmin} ||u - v_c||^2_{L^2(\mathbb{R})}.$$

We use the shifted kink to define the excess mass

$$V := \int_{\mathbb{R}} |u - v_c| \, dx.$$

The main idea is to establish and make use of the Nash-type inequality

$$\mathcal{E} \leq C D^{1/3} (V + 1)^{4/3},$$

where the energy gap $\mathcal{E} := E(u) - E(v)$. Indeed, letting $V_T := \sup_{t \leq T} V + 1$, it is easy to show that

$$\mathcal{E} \leq C \, \frac{V_T^2}{t^{1/2}}.$$

Hence the main and nontrivial challenge is to show

$$V_T \leq C(V_0 + 1).$$

This is accomplished using a duality argument and decay estimates for the adjoint of the linearized equation.

Work in progress with Sarah Biesenbach, Felix Otto, and Richard Schubert analyzes relaxation rates and metastability of the one-dimensional Cahn-Hilliard equation in a more subtle setting, in which no simple conserved quantity identifies the limit among the collection of energy minimizers.

References

[1] O. Chugreeva, F. Otto, and M. G. Westdickenberg, *Relaxation to a planar interface in the Mullins-Sekerka problem*, Interfaces and Free Boundaries **21** (2019), 21–40.

[2] *Relaxation to equilibrium in the one-dimensional Cahn-Hilliard equation*, SIAM J. Math. Anal. **46** (2014), no. 1, 720–756.

[3] F. Otto, S. Scholtes, and M. G. Westdickenberg, *Optimal L^1-type relaxation rates for the Cahn-Hilliard equation on the line*, SIAM J. Math. Anal. **51** (2019), no. 6, 4645–4682.

The Landau equation as a Gradient Flow

José A. Carrillo
(joint work with Matias G. Delgadino, Laurent Desvillettes, Jeremy Wu)

The Landau equation is an important partial differential equation in kinetic theory. It gives a description of colliding particles in plasma physics [8], and it can be formally derived as a limit of the Boltzmann equation where grazing collisions are dominant [9]. Similar to the Boltzmann equation, the rigorous derivation of the Landau equation from particle dynamics is still a huge challenge. For a spatially homogeneous density of particles $f = f_t(v)$ for $t \in (0, \infty), v \in \mathbb{R}^d$ the homogeneous Landau equation reads

(1)
$$\partial_t f(v) = \nabla_v \cdot \left(f(v) \int_{\mathbb{R}^d} |v - v_*|^{2+\gamma} \Pi[v - v_*](\nabla_v \log f(v) - \nabla_{v_*} \log f(v_*)) f(v_*) dv_* \right).$$

For notational convenience, we sometimes abbreviate $f = f_t(v)$ and $f_* = f_t(v_*)$. We also denote the differentiations by $\nabla = \nabla_v$ and $\nabla_* = \nabla_{v_*}$. The physically relevant parameters are usually $d = 2, 3$ and $\gamma \geq -d - 1$ with $\Pi[z] = I - \frac{z \otimes z}{|z|^2}$ being the projection matrix onto $\{z\}^\perp$. In this paper, for simplicity we will focus in the case $d = 3$ and vary the weight parameter γ, although most of our results are valid in arbitrary dimension. The regime $0 < \gamma < 1$ corresponds to the so-called *hard potentials* while $\gamma < 0$ corresponds to the *soft potentials* with a further classification of $-2 \leq \gamma < 0$ as the moderately soft potentials and $-4 \leq \gamma < -2$ as the very soft potentials. The particular instances of $\gamma = 0$ and $\gamma = -d$ are known as the Maxwellian and Coulomb cases respectively.

The purpose of this work is to propose a new perspective inspired from gradient flows for weak solutions to (1), which is in analogy with the relationship of the heat equation and the 2-Wasserstein metric, see [7, 2]. One of the fundamental steps is to symmetrize the right hand of (1). More specifically, if we consider a test function $\phi \in C_c^\infty(\mathbb{R}^d)$ we can formally characterize the equation by

(2)
$$\frac{d}{dt} \int_{\mathbb{R}^d} \phi f dv = -\frac{1}{2} \iint_{\mathbb{R}^{2d}} f f_* |v - v_*|^{2+\gamma} (\nabla \phi - \nabla_* \phi_*) \cdot \Pi[v - v_*](\nabla \log f - \nabla_* \log f_*) dv_* dv,$$

where the change of variables $v \leftrightarrow v_*$ has been exploited. Building our analogy with the heat equation and the 2-Wasserstein distance, we define an appropriate gradient

$$\tilde{\nabla} \phi := |v - v_*|^{1+\frac{\gamma}{2}} \Pi[v - v_*](\nabla \phi - \nabla_* \phi_*),$$

so that equation (2) now looks like

$$\frac{d}{dt} \int_{\mathbb{R}^d} \phi f dv = -\frac{1}{2} \iint_{\mathbb{R}^{2d}} f f_* \tilde{\nabla} \phi \cdot \tilde{\nabla} \log f dv_* dv,$$

noting that $\Pi^2 = \Pi$. To highlight the use of this interpretation, we notice that $\tilde{\nabla} \phi = 0$, when we choose as test functions $\phi = 1,\ v_i,\ |v|^2$ for $i = 1, \ldots, d$ which immediately shows that formally the equation conserves mass, momentum and energy. The action functional defining the Landau metric mimics the Benamou-Brenier formula [3] for the 2-Wasserstein distance. In fact, the Landau metric is

built by considering a minimizing action principle over curves that are solutions to the appropriate continuity equation, that is

$$d_L(f,g) := \min_{\substack{\partial_t \mu + \frac{1}{2}\tilde{\nabla}\cdot(V\mu\mu_*)=0 \\ \mu_0 = f,\ \mu_1 = g}} \left\{ \frac{1}{2} \int_0^1 \iint_{\mathbb{R}^{2d}} |V|^2 \, d\mu(v) d\mu(v_*) dt \right\}, \tag{3}$$

where the $\tilde{\nabla}\cdot$ is the appropriate divergence; the formal adjoint to the appropriate gradient. Also, we notice that analogously to the heat equation, written as the continuity equation $\partial_t f = \nabla \cdot (f \nabla \log f)$, the Landau equation can be formally re-written as

$$\partial_t f = \frac{1}{2} \tilde{\nabla} \cdot (f f_* \tilde{\nabla} \log f),$$

equivalent to the continuity equation with non-local velocity field given by

$$\begin{cases} \partial_t f + \nabla \cdot (U(f) f) = 0 \\ U(f) := -\displaystyle\int_{\mathbb{R}^d} |v - v_*|^{2+\gamma} \Pi[v - v_*] \left(\nabla \log f - \nabla_* \log f_* \right) f_* dv_* \, . \end{cases} \tag{4}$$

Considering the evolution of Boltzmann entropy we formally obtain

$$\frac{d}{dt} \int_{\mathbb{R}^d} f \log f dv =: -D(f_t) = -\frac{1}{2} \iint_{\mathbb{R}^{2d}} |\tilde{\nabla} \log f|^2 f f_* dv_* dv \leq 0. \tag{5}$$

In physical terms this is referred to as the *entropy dissipation* or *entropy production* for it formally shows that the entropy functional

$$\mathcal{H}[f] := \int_{\mathbb{R}^d} f \log f dv$$

is non-increasing along the dynamics of the Landau equation. Moreover, by integrating equation (5) in time one formally obtains

$$\mathcal{H}[f_t] + \int_0^t D(f_s) ds = \mathcal{H}[f_0]. \tag{6}$$

Similar to H-solutions our approach will also be based on the entropy dissipation (6). Following De Giorgi's minimizing movement ideas [1, 2], we characterize the Landau equation by its associated Energy Dissipation Inequality. More specifically, we show that weak solutions to (1) with initial data f_0 are completely determined by the following functional inequality:

$$\mathcal{H}[f_t] + \frac{1}{2} \int_0^t |\dot{f}|^2_{d_L}(s) \, ds + \frac{1}{2} \int_0^t D(f_s) \, ds \leq \mathcal{H}[f_0] \qquad \text{for a.e. every } t > 0,$$

where $|\dot{f}|^2_{d_L}(s)$ stands for the metric derivative associated to the Landau metric defined above. Our analysis is also largely inspired by Erbar's approach in viewing the Boltzmann equation as a gradient flow [6] and recent numerical simulations of the homogeneous Landau equation in [4] based on a regularized version of (4). In contrast with the classical 2-Wasserstein metric, one of the main features of the Landau equation (1) and metric (3) is that they are non-local. Hence, the

convergence analysis usually relying on convexity and lower-semi continuity needs to be adapted to deal with the non-locality of this equation.

In order to deal with the Landau equation using the theory of metric measure spaces, we carefully study the Landau distance d_L. Moreover, we showed for a regularized version of the Landau equation that we can construct gradient flow solutions, curves of maximal slope, via the corresponding variational scheme. The main result obtained for the Landau equation shows that the chain rule can be rigorously proved for the grazing continuity equation, this implies that H-solutions with certain apriori estimates on moments and entropy dissipation are equivalent to gradient flow solutions of the Landau equation. We crucially make use of estimates on Fisher-like quantities in terms of the Landau entropy dissipation developed in [5].

References

[1] L. Ambrosio. Minimizing movements. *Rend. Accad. Naz. Sci. XL Mem. Mat. Appl. (5)*, 19:191–246, 1995.

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

[3] J.-D. Benamou and Y. Brenier. A computational fluid mechanics solution to the Monge-Kantorovich mass transfer problem. *Numer. Math.*, 84(3):375–393, 2000.

[4] J. A. Carrillo, J. Hu, L. Wang, and J. Wu. A particle method for the homogeneous Landau equation. *J. Comput. Phys. X*, 7:100066, 2020.

[5] L. Desvillettes. Entropy dissipation estimates for the Landau equation in the Coulomb case and applications. *J. Funct. Anal.*, 269(5):1359–1403, 2015.

[6] M. Erbar. A gradient flow approach to the Boltzmann equation. *arXiv preprint arXiv:1603.00540*, 2016.

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

[8] E. M. Lifshitz and L. P. Pitaevskiĭ. *Course of theoretical physics ["Landau-Lifshits"]. Vol. 10*. Pergamon International Library of Science, Technology, Engineering and Social Studies. Pergamon Press, Oxford-Elmsford, N.Y., 1981. Translated from the Russian by J. B. Sykes and R. N. Franklin.

[9] C. Villani. On a new class of weak solutions to the spatially homogeneous Boltzmann and Landau equations. *Arch. Rational Mech. Anal.*, 143(3):273–307, 1998.

On hydrodynamic limit for non-interacting deterministic particles, A Hamilton-Jacobi PDE in space of probability measure approach

Jin Feng

(joint work with Toshio Mikami)

We propose a method for understanding hydrodynamic limit of N-particle Hamiltonian dynamics by proving limit theorems for Hamilton-Jacobi equations in space of probability measures. We start with finite particle dynamic which corresponds to minimizer of action. The action is further characterized by Hamilton-Jacobi PDEs. Convergence of the PDEs imply convergence of the actions, hence the minimizers. In the infinite particle continuum limit, action becomes defined over

probability-measure-valued curves and Hamilton-Jacobi is defined in space of probability measures. Although these PDEs are defined in abstract spaces and the method is indirect, such approach has potential practical advantages over usual method of directly verifying convergence of Hamiltonian ODEs to system of conservation law PDEs – the abstract Hamilton-Jacobi is a scalar equation with maximum principle, the above steps can be made rigorous.

We illustrate the above scheme using models of non-interacting deterministic Hamiltonian particles with a two-scale hydrodynamic limit. We justify the Hamilton-Jacobi convergence through a singular perturbation/homogenization argument generalized to a setting of PDEs in space of probability measures. A "cell equation" arising in such context coincides with a finite dimensional deterministic averaging problem considered in weak KAM (Kolmogorov-Arnold-Moser) theory. As a consequence, we derive a continuum level effective-Hamiltonian as scaling limit from the particle ones. The whole derivation avoids direct involvement of micro-canonical and canonical ensembles through deterministic ergodic arguments in any strong sense. It also avoids the even more difficult task of relating these ensembles.

1. Microscopic N-particle dynamics and embedding into a continuum

Let $\mathsf{H} := \mathsf{H}(q,p) : \mathbb{T}^d \times \mathbb{R}^d \mapsto \mathbb{R}$. We consider N-particle non-interacting Hamiltonian dynamics given by

$$\dot{q}_i = \nabla_{p_i}\mathsf{H}(q_i,p_i), \quad \dot{p}_i = -\nabla_{q_i}\mathsf{H}(q_i,p_i).$$

With a hydrodynamical scaling $\mathbf{x}(t) := \mathbf{x}_\epsilon(t) := \epsilon\mathbf{q}(\frac{t}{\epsilon}), \mathbf{P}(t) := \mathbf{P}_\epsilon(t) := \mathbf{p}(\frac{t}{\epsilon})$, we arrive at $\dot{x}_i(t) = \nabla_2\mathsf{H}\big(\frac{x_i}{\epsilon}, P_i\big)$, $\dot{P}_i(t) = -\frac{1}{\epsilon}\nabla_1\mathsf{H}\big(\frac{x_i}{\epsilon}, P_i\big)$, where the $\nabla_1\mathsf{H}(q,p) := \nabla_q\mathsf{H}(q,p)$ and $\nabla_2\mathsf{H}(q,p) := \nabla_p\mathsf{H}(q,p)$. Considering symmetries in particle labels, we model micro-state evolution processes using $\rho_{N,\epsilon}(t;dx) := \frac{1}{N}\sum_{i=1}^N \delta_{x_i(t)}(dx)$ and $\mu_{N,\epsilon}(t;dx,dq) := \frac{1}{N}\sum_{i=1}^N \delta_{x_i(t),q_i(\frac{t}{\epsilon})}(dx,dq) = \rho_{N,\epsilon}(t,dx)\delta_{\frac{x}{\epsilon}}(dq)$. We will take $N \to \infty$ and $\epsilon \to 0$.

We look at Hamilton-Jacobi PDEs associated with the above ODE flows. These PDEs should be defined in space of the micro- and macro-states. First, we clarify on notion of *smooth* functions for studying these problems. We denote $(\mathbb{R}^d, |\cdot|)$ the usual d-dimensional Euclidean space, and consider $\mathsf{X}_N := (\mathbb{T}^d)^N$ as a metric space with tangent at any given point identified with $\big((\mathbb{R}^d)^N, |\cdot|_N\big)$, where the weighted norm $|\mathbf{x}|_N^2 := \frac{1}{N}\sum_{i=1}^N |x_i|^2$. Let $\nabla_{N,\mathbf{x}}$ denote the associated notion of gradients for functions in X_N. Note that we already used $\nabla_{\mathbf{x}}$ to denote the usual gradients for functions when tangent is identified with the usual un-weighted Euclidean space. It follows that $\nabla_N = N\nabla$. Let G_N be permutation group on N-particle labels. It introduces an isometric group action on X_N through $\big(\pi\mathbf{x}\big)_i := \mathbf{x}_{\pi(i)}$ for every $\pi \in \mathsf{G}_N$. The quotient metric space $\mathsf{X}_N^* := \mathsf{X}_N/\mathsf{G}_N$ has a natural metric $\mathsf{d}_N^2(\mathbf{x}^*,\mathbf{y}^*) = \inf_{\pi\in\mathsf{G}_N}\frac{1}{N}\sum_{i=1}^N \mathsf{d}^2(x_i,y_{\pi(i)})$. We realize that X_N^* can be isometrically identified with the space of N-particle empirical probability measures: $\mathbf{x}^* \sim \rho := \frac{1}{N}\sum_{i=1}^N \delta_{x_i}$ and $\mathbf{y}^* \sim \gamma := \frac{1}{N}\sum_{i=1}^N \delta_{y_i}$. Moreover, $\mathsf{d}_N^2(\mathbf{x}^*,\mathbf{y}^*) = W^2(\rho,\gamma)$ is the order-2

Wasserstein metric, and $\mathsf{X}_N^* \subset \mathsf{X}_{N+1}^* \subset \ldots \subset \mathsf{X}^*$ with X^* the order-2 Wasserstein space. This reveals that the X_N^* should be considered as non-smooth space, with singularities arise whenever particles collide (i.e. $x_i = x_j$ for some $i \neq j$). These X_N^*s are Alexandrov spaces with nonnegative curvature. Therefore, a rigorous PDE theory can be built upon using test functions composed of combinations of distance square functions. In the following, in order to be intuitive, we pursue a different class of G_N-invariant smooth test functions:

$$\begin{aligned} f(\rho) &:= \psi(\langle\vec{\varphi},\rho\rangle) := \psi(\langle\varphi_1,\rho\rangle,\ldots,\langle\varphi_K,\rho\rangle) \\ &= \psi\Big(\frac{1}{N}\sum_{i=1}^N \varphi_1(x_i),\ldots,\frac{1}{N}\sum_{i=1}^N \varphi_K(x_i)\Big) =: f(\mathbf{x}), \end{aligned}$$

where $\varphi_k \in C_c^1(\mathbb{T}^d\times\mathbb{R}^d)$ and $\psi \in C^1(\mathbb{R}^K)$. We also write $\frac{\delta f}{\delta\rho} := \sum_{k=1}^K \partial_k\psi(\langle\vec{\varphi},\rho\rangle)\varphi_k$.

Introducing $H_{N,\epsilon}(\mathbf{x},\mathbf{P}) := \frac{1}{N}\sum_{i=1}^N \mathsf{H}(\frac{x_i}{\epsilon}, P_i)$, then the dynamics of $(\mathbf{x},\mathbf{P})$ can be rewritten as a Hamiltonian flow $\dot{x}_i = \nabla_{N,P_i}H_N$ and $\dot{P}_i = -\nabla_{N,x_i}H_N$. This observation motivates us to define

$$H_{N,\epsilon}f(\rho) := H_{N,\epsilon}\big(\mathbf{x},\nabla_{N,\mathbf{x}}f\big) = \int_{\mathbb{T}^d} \mathsf{H}\big(\frac{x}{\epsilon},\nabla\frac{\delta f}{\delta\rho}(x)\big)\rho(dx).$$

2. Separation of scales and homogenization of Hamilton-Jacobi PDE in Wasserstein space

We show that the $H_{N,\epsilon}$ converges to a limiting *effective Hamiltonian* H in operator graph sense (e.g. Kurtz [5]), and identify this limit.

Given a test function of macro-state variable $f = f(\rho)$, we look for perturbed test functions $f_{N,\epsilon}(\rho) := f(\rho) + \epsilon g(\mu)$ with $g(\mu) := \langle\phi,\mu\rangle$, $\phi := \phi(x,P;q) \in C_c^2(\mathbb{T}^d\times\mathbb{R}^d\times\mathbb{T}^d)$ and $\mu(dx,dP;dq) := \rho(dx)\delta_{(\nabla\frac{\delta f}{\delta\rho}(x),\frac{x}{\epsilon})}(dP,dq)$. The choice of ϕ should depend on f which is to be determined soon. With a slight rewriting, we can also view $g(\mu) = g(\rho) := \langle\varphi,\rho\rangle$ with $\varphi := \varphi_\epsilon(x) := \phi\big(x,\nabla\frac{\delta f}{\delta\rho}(x);\frac{x}{\epsilon}\big)$. Therefore,

$$\begin{aligned} H_{N,\epsilon}f_{N,\epsilon}(\rho) &= \int_{\mathbb{T}^d\times\mathbb{T}^d} \mathsf{H}\Big(q,\nabla_x\frac{\delta f}{\delta\rho}(x) + \nabla_q\phi\big(x,\nabla\frac{\delta f}{\delta\rho}(x);q\big)\Big)\mu(dx,dq) + o(1) \\ &= \int_{\mathbb{T}^d\times\mathbb{R}^d\times\mathbb{T}^d} \mathsf{H}\big(q,P+\nabla_q\phi(x,P;q)\big)\mu(dx,dP;dq) + o(1). \end{aligned}$$

Suppose that we can solve an auxiliary PDE problem

$$\mathsf{H}\big(q,P+\nabla_q\phi(q)\big) = \bar{\mathsf{H}}(P), \quad \forall q\in\mathbb{T}^d.$$

By solution, we mean that for each P fixed, there exists a pair $(\bar{\mathsf{H}}(P),\phi(P;\cdot)) \in \mathbb{R}\times C(\mathbb{T}^d)$ satisfying the above identity. Then we conclude that,

$$f_{N,\epsilon}\to f, \quad H_{N,\epsilon}f_{N,\epsilon}\to Hf$$

with $Hf(\rho) := \int_{\mathbb{T}^d \times \mathbb{R}^d} \bar{\mathsf{H}}(P)\mu(dx, dP)$, where the $\mu(dx, dP) := \rho(dx)\delta_{\nabla \frac{\delta f}{\delta \rho}(x)}(dP)$ is precisely gradient of f in Alexandrov analysis sense. The auxiliary PDE problem was considered in homogenization theory (Lions, Papanicolaou and Varadhan [6]). It is also extensively studied in weak KAM theory for finite dimensional Hamiltonian dynamical systems (Fathi [4], E [2], Evans and Gomez [3]). In general, we know that there is no $\phi \in C^1$, but a generalized sense solution (i.e. viscosity) can be found. More importantly, the single particle level *effective Hamiltonian* can always be identified implicitly through variational formula

$$\bar{\mathsf{H}}(P) := \inf_{\phi} \sup_{q} \mathsf{H}(q, P + \nabla_q \phi).$$

Resolvent and Cauchy problems for H have been studied in Ambrosio and Feng [1].

3. An example

We consider the case $\mathbb{T} := [0, 1]$ with $\mathsf{H}(q, p) = |p|^2 - V(q)$ and $\min V = 0$. Let $c_V := \int_{\mathbb{T}} \sqrt{V} dq$. The paper [6] identified that $\bar{\mathsf{H}}(P) = 0$ when $|P| \leq c_V$, and $\bar{\mathsf{H}}(P) = \lambda$ solving $|P| = \int_{\mathbb{T}} dq\sqrt{V + \lambda}$, when $|P| > c_V$. If $V \equiv 0$, then $\bar{\mathsf{H}}(P) = |P|^2$. In this case, minimal action path in the continuum satisfies compressible Euler equation modeling ideal gas.

References

[1] L. Ambrosio, J. Feng, *On a class of first order Hamilton-Jacobi equations in metric spaces*, J. Diff. Eqn., **256(7)**, (2014), 2194–2245.

[2] Weinan E, *Aubry-Mather theory and periodic solutions of the forced Burgers equation*, Comm. Pure Appl. Math., **52**, (1999), 811–828.

[3] L.C. Evans, D. Gomes, *Effective Hamiltonians and averaging for Hamiltonian dynamics. I, II*, Arch. Ration. Mech. Anal., **157**, (2001),1–33. **161**, (2002), 271–305.

[4] A. Fathi, *The Weak KAM Theorem in Lagrangian Dynamics*, Book manuscript, (English Version 10 circa 2008).

[5] Thomas Kurtz, *Convergence of sequences of semigroups of nonlinear operators with an application to gas kinetics*, Trans. Amer. Math. Soc., **186**, (1973), 259–272.

[6] P.L. Lions, G. Papanicolaou, S.R.S. Varadhan, *Homogenization of Hamilton-Jacobi equations*, Unpublished Preprint, circa 1987.

Interpreting LDPs without detailed balance

Robert I. A. Patterson
(joint work with D. R. Michiel Renger, Upanshu Sharma)

We present in outline a rather general connection between flux space large deviations, free energy dissipation bounds and macroscopic fluctuation theory (MFT) [1].

1. Detailed balance and gradient flows

It is by now fairly classical that, for systems of interacting particles $X_1, \dots X_n$ with a finite time horizon T, satisfying a detailed balance condition and some further technical assumptions, there is an LDP for the empirical measures $\rho^n(t) := \frac{1}{n}\sum_{i=1}^n \delta_{X_i(t)}$, which informally says

$$\mathbb{P}\left(\rho^n \approx \rho\right) \sim \exp\left(-n\int_0^T \widehat{\mathcal{L}}(\rho(t), \dot\rho(t)\mathrm{d}t\right)$$

for some $\widehat{\mathcal{L}} \geq 0$ and one has a decomposition [5]

$$\widehat{\mathcal{L}}(x,u) = \Psi(x,u) + \Psi^*\left(x, -\frac{1}{2}\mathrm{D}I_0(x)\right) + \left\langle \frac{1}{2}\mathrm{D}I_0(x), u\right\rangle.$$

Here Ψ and Ψ^* are a convex dual pair of dissipation potentials, I_0 is the large deviations rate function for the invariant distributions of the ρ^n and D indicates a Gateaux derivative. This characterises the zero cost flows ($\widehat{\mathcal{L}}(\rho(t), \dot\rho(t)) = 0$ for all t) by

$$\dot\rho(t) = \partial\Psi^*\left(\rho(t), -\frac{1}{2}\mathrm{D}I_0(\rho(t))\right),$$

where the differential of Ψ^* is with respect to its second argument. This provides a full description of changes in I_0, which can be interpreted as a free energy. However, detailed balance plays a fundamental role, even if one includes GENERIC and related extensions [6, 4].

2. Fluxes

In order to make progress beyond detailed balance we describe a particle process not just by an empirical measure ρ, but also using a flux w so that

$$\rho(t) - \rho(0) = -\nabla\cdot w(t), \tag{1}$$

where $\nabla\cdot$ and its adjoint $-\nabla$ are adapted to the problem (e.g. graph divergence and gradient for particles hopping on a graph). We assume an LDP of the form

$$\mathbb{P}\left((\rho^n, w^n) \approx (\rho, w)\right) \sim \exp\left(-n\int_0^T \mathcal{L}(\rho(t), \dot w(t)\mathrm{d}t\right),$$

where we restrict to pairs (ρ, w) satisfying (1) for all $t \in [0,T]$, $\mathcal{L}$ is its own convex bidual and $\mathcal{L}^* =: \mathcal{H}$. (Throughout this abstract convex duality refers to the final argument of a function.)

We weaken the assumption of detailed balance for the dynamics of the empirical measure to simply assuming that there is an invariant distribution for ρ and a large deviations rate function I_0 for samples from this invariant distribution so that for all ρ in a large class

$$\mathcal{H}(\rho, \nabla \mathrm{D} I_0) = 0. \tag{2}$$

Provided this equation makes sense for a sufficiently large class of ρ and some mild reversibility conditions, weaker than detailed balance, hold, we show that

$$\begin{aligned}\mathcal{L}(\rho(t), \dot{w}(t)) = \mathcal{L}_{F^{\mathrm{asym}}}(\rho(t), \dot{w}(t)) \\ - \mathcal{H}\left(\rho(t), -\frac{1}{2}\nabla \mathrm{D} I_0(\rho(t))\right) + \left\langle \frac{1}{2}\nabla \mathrm{D} I_0(\rho(t)), \dot{w}(t)\right\rangle.\end{aligned} \tag{3}$$

$\mathcal{L}_{F^{\mathrm{asym}}}$ can be interpreted as the large deviations rate function for paths under (precisely defined) time-antisymmetrised dynamics, see [3], so that $\mathcal{L}_{F^{\mathrm{asym}}} \geq 0$ and in some sense

$$\mathbb{P}_{\mathrm{asym}}\left((\rho^n, w^n) \approx (\rho, w)\right) \sim \exp\left(-n \int_0^T \mathcal{L}_{F^{\mathrm{asym}}}(\rho(t), \dot{w}(t)\mathrm{d}t\right).$$

Additionally we are able to check that $\mathcal{H}\left(\rho(t), -\frac{1}{2}\nabla \mathrm{D} I_0(\rho(t))\right) \leq 0$. All these results follow by fairly simple convexity arguments. They can be seen on the one hand as a generalisation of Fisher Information inequalities for free energy dissipation [2] and on the other hand as an abstract approach to MFT.

In principle our results are purely analytic in character and would be valid for any $\mathcal{L}$ that is its own convex bidual and for any I_0 satisfying (2). We find I_0 solving (2) as the large deviations rate function of the invariant measures of an underlying family of stochastic particle models. However, other strategies for finding solutions to this Hamilton–Jacobi equation could also be used in order to develop the theory.

3. The purely asymmetric dynamics

We conjecture that the dynamics given by $\mathcal{L}_{F^{\mathrm{asym}}}$ should be interpreted as the counterpart of a gradient flow, and that it will describe a purely non-dissipative dynamics.

In the special case of independent particles each performing a (non-degenerate) random walk on a discrete set $\mathcal{X}$ with generator matrix Q and stationary distribution π such that $\pi^T Q = 0$, the zero cost flow of $\mathcal{L}_{F^{\mathrm{asym}}}$ is given by the system of ODEs

$$\dot{\rho}_x(t) = \sum_{y \neq x}\left(\sqrt{\rho_y(t)\rho_x(t)\pi_y/\pi_x}Q_{yx} - \sqrt{\rho_x(t)\rho_y(t)\pi_x/\pi_y}Q_{xy}\right) \qquad \forall x \in \mathcal{X}. \tag{4}$$

With the help of other participants it was possible to identify this as a Hamiltonian flow where the conserved energy is $\sum_{x\in\mathcal{X}}\sqrt{\rho_x(t)}$. It was also noted that (4) does not have a Lipschitz right hand side close to points at which for one or more $x \in \mathcal{X}$ one has $\rho_x(t) = 0$.

4. Acknowledgement

This research has been partially funded by the Deutsche Forschungsgemeinschaft (DFG) through grant CRC 1114 "Scaling Cascades in Complex Systems", Project C08

References

[1] L. Bertini, A. De Sole, D. Gabrielli, G. Jona-Lasino, and C. Landim. Macroscopic fluctuation theory. *Reviews of Modern Physics*, 87(2), 2015.
[2] B. Hilder, M. A. Peletier, U. Sharma, and O. Tse. An inequality connecting entropy distance, fisher information and large deviations. *Stochastic Processes and their Applications*, 130(5), 2020.
[3] M. Kaiser, R. Jack, and J. Zimmer. Canonical structure and orthogonality of forces and currents in irreversible markov chains. *Journal of Statistical Physics*, 170(6):1019–1050, 2018.
[4] R. Kraaij, A. Lazarescu, C. Maes, and M.A. Peletier. Deriving GENERIC from a generalized fluctuation symmetry. *Journal of Statistical Physics*, 170(3):492–508, 2018.
[5] 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(4):1293–1327, 2014.
[6] H.-C. Öttinger. Beyond equilibrium thermodynamics. John Wiley & Sons, 2005.

Homogenization of discrete optimal transport

Jan Maas
(joint work with Peter Gladbach, Eva Kopfer, Lorenzo Portinale)

In the past decades there has been intense research activity in the area of optimal transport. In continuous settings, a key result in the field is the Benamou–Brenier formula [1], which provides a dynamical formulation of the classical Mong–Kantorovich problem of optimal transport. In discrete settings, the equivalence between static and dynamical optimal transport breaks down, and it turns out that the dynamical formulation (introduced in [5, 6]) is essential in applications to evolution equations, discrete Ricci curvature, and functional inequalities.

The limit passage from discrete dynamical transport to continuous optimal transport turns out to be nontrivial. In fact, it has been shown that seemingly natural discretizations of the Benamou–Brenier formula do not necessarily converge to the Kantorovich distance $\mathbb{W}_2$, even in one-dimensional settings. The main result in [2] asserts that, for a sequence of meshes on a bounded convex domain in $\mathbb{R}^d$, an isotropy condition on the meshes is required to obtain the convergence of the discrete dynamical transport distances to $\mathbb{W}_2$.

In this talk we present recent work in which we identify the limiting behaviour of the discrete metrics on $\mathbb{Z}^d$-periodic graphs where the isotropy condition fails to hold.

For an (undirected) graph $(\mathcal{X},\mathcal{E})$ and probability measure $m^0, m^1 \in \mathcal{P}(\mathcal{X})$, we consider the dynamical transport cost given by

$$\mathcal{C}(m^0,m^1) = \inf_{(m,J)\in \mathrm{CE}(m^0,m^1)} \Big\{ \int_0^1 \sum_{(x,y)\in\mathcal{E}} c_{xy}\big(m_t(x), m_t(y), J_t(x,y)\big)\,\mathrm{d}t \Big\},$$

where $\mathrm{CE}(m^0,m^1)$ denotes the class of all solutions to the discrete continuity equation connecting m^0 and m^1, i.e., all curves of probability measures $m:[0,1]\to\mathcal{P}(\mathcal{X})$ and all time-dependent discrete vector fields (i.e., anti-symmetric functions) $J:[0,1]\to\mathbb{R}^{\mathcal{E}}$ satisfying the discrete continuity equation

$$\partial_t m_t(x) + \sum_{y:y\sim x} J_t(x,y) = 0, \qquad m_0 = m^0,\ m_1 = m^1.$$

The cost functions $c_{xy}:\mathbb{R}_+\times\mathbb{R}_+\times\mathbb{R}\to\mathbb{R}_+$ are assumed to be lower-semicontinuous and convex.

In [3, 4] we consider a $\mathbb{Z}^d$-periodic graph $(\mathcal{X},\mathcal{E})$ embedded in $\mathbb{R}^d$. For any $\varepsilon>0$ with $1/\varepsilon\in\mathbb{N}$, we let $(\mathcal{X}_\varepsilon,\mathcal{E}_\varepsilon)$ be the rescaled graph wrapped around the torus $\mathbb{T}^d$, i.e.,

$$\mathcal{X}_\varepsilon = \varepsilon\mathcal{X}/Z^d \quad\text{and}\quad \mathcal{E}_\varepsilon = \varepsilon\mathcal{X}/\{(z,z): z\in\mathbb{Z}^d\}.$$

The rescaled transport cost is given by

$$\mathcal{C}_\varepsilon(m^0,m^1) = \inf_{(m,J)\in\mathrm{CE}(m^0,m^1)} \Big\{ \int_0^1 \sum_{(x,y)\in\mathcal{E}_\varepsilon} \varepsilon^d c_{xy}\Big(\frac{m_t(x)}{\varepsilon^d}, \frac{m_t(y)}{\varepsilon^d}, \frac{J_t(x,y)}{\varepsilon^{d-1}}\Big)\,\mathrm{d}t \Big\}.$$

Loosely speaking, our main result asserts that these discrete transport problems converge to a continuous transport problem with a homogenized cost function, as $\varepsilon\to 0$. Namely, for any weakly converging sequences of probability measures $m^0_\varepsilon \rightharpoonup \mu^0$ and $m^1_\varepsilon\rightharpoonup\mu^1$, we have convergence $\mathcal{C}_\varepsilon(m^0_\varepsilon,m^1_\varepsilon)\to\mathbb{C}(\mu^0,\mu^1)$, where the limiting transport problem is given by

$$\mathbb{C}(\mu^0,\mu^1) = \inf_{\mu,j}\Big\{ \int_0^1\int_{\mathbb{T}^d} c_{\mathrm{hom}}\Big(\frac{\mathrm{d}\mu_t}{\mathrm{d}\mathcal{L}^d}, \frac{\mathrm{d}j_t}{\mathrm{d}\mathcal{L}^d}\Big)\,\mathrm{d}x\,\mathrm{d}t \ :\ \partial_t\mu + \nabla\cdot j = 0,\ \mu_{0,1}=\mu^{0,1}\Big\}.$$

The homogenized cost $c_{\mathrm{hom}}:\mathbb{R}_+\times\mathbb{R}^d\to\mathbb{R}$ is given by the cell formula

$$c_{\mathrm{hom}}(\mu,j) = \inf_{m,J}\Big\{ \sum_{x\in\mathcal{X}\cap[0,1)^d}\ \sum_{y\sim x} c_{xy}\big(m(x),m(y),J(x,y)\big)\Big\},$$

where the infimum runs over all $\mathbb{Z}^d$-periodic functions $m:\mathcal{X}\to\mathbb{R}_+$ and all $\mathbb{Z}^d$-periodic discrete vector fields (i.e., all anti-symmetric functions $J:\mathcal{E}\to\mathbb{R}$ with $\sum_{y:y\sim x}J(x,y)=0$ for all $x\in\mathcal{X}$) satisfying

$$\sum_{x\in\mathcal{X}\cap[0,1)^d} m(x) = \mu \quad\text{and}\quad \frac{1}{2}\sum_{x\in\mathcal{X}\cap[0,1)^d}\ \sum_{y\sim x} J(x,y)(y-x) = j.$$

This cell formula describes how the limiting transport cost is affected by the geometry of the periodic graph. The rigorous formulation of our main result is given in terms of Γ-convergence for curves in the space of probability measures.

References

[1] J.-D. Benamou and Y. Brenier. *A computational fluid mechanics solution to the Monge-Kantorovich mass transfer problem*, Numer. Math. **84** (2000), 375–393.

[2] P. Gladbach, E. Kopfer, and J. Maas. *Scaling limits of discrete optimal transport*, *SIAM J. Math. Anal.* **52** (2020), 2759–2802.

[3] P. Gladbach, E. Kopfer, J. Maas, and L. Portinale. *Homogenisation of one-dimensional discrete optimal transport*, *J. Math. Pures Appl.* **139** (2020), 204–234.

[4] P. Gladbach, E. Kopfer, J. Maas, and L. Portinale. *Homogenisation of functionals on measures-valued curves with dynamical constraints*, in preparation.

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

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

Many-particle limit for a system of interaction equations driven by Newtonian potentials

Antonio Esposito

(joint work with Marco Di Francesco, Markus Schmidtchen)

The problem of approximating transport PDEs by the empirical measure associated to moving particles is quite classical in many contexts such as particle physics and gravitation. A prototype model and variation of the pure transport PDE which gained great attention in the last decades is the nonlocal transport-diffusion equation

$$\partial_t \rho = \operatorname{div}(\nabla a(\rho) + \rho \nabla G * \rho), \tag{1}$$

where $a = a(\rho)$ is a nonlinear diffusion function and G is a space dependent kernel modelling nonlocal interaction. In the aforementioned contexts in particle physics and gravitation, G is typically a singular kernel, which makes the analysis of (1) quite challenging. Starting from the early 2000 years, the theory of gradient flows in Wasserstein spaces developed in [8, 10, 1] became an important tool to provide well-posedness results for the class of models (1). The result in [4] is also relevant in this context in that it allowed to extend the theory to kernels G displaying a discontinuity of the gradient at the origin in the re-solution of the JKO-scheme and in the proof of λ-convexity [9] of the related functional. The role of λ-convexity of the functional is essential in order to prove a stability result for two solution curves, which often implies as a byproduct a *many-particle approximation* result for the target equation (1).

Recently, there has been an increasing interest in *systems* of gradient flows, i.e. systems of more than one transport equations of the form (1), modelling the mutual interplay of more than one species of individuals. The case with diffusion has a very rich literature in that it is quite challenging at the level of well-posedness due to the possibility of cross-diffusion effects, see for instance the recent [5] and the references therein. In the following we consider a system of nonlocal interaction PDEs with no (cross-)diffusion part, which is not included in the work [7], and

whose analysis has lead to an interesting connection with a nonliner hyperbolic system, see [3].

In particular, in [6], we deal with a deterministic particle approximation for a system of continuity equations driven by Newtonian nonlocal interactions in one dimension. More precisely, setting $N(x) := |x|$, we prove that the PDE system

$$\begin{cases} \partial_t \rho = -\partial_x(\rho \partial_x N * \rho) + \partial_x(\rho \partial_x N * \eta), \\ \partial_t \eta = -\partial_x(\eta \partial_x N * \eta) + \partial_x(\eta \partial_x N * \rho), \end{cases} \tag{2}$$

can be obtained as the *many-particle limit* of the deterministic ODE system

$$\begin{cases} \dot{x}_i(t) = \displaystyle\sum_{x_k(t)\neq x_i(t)} m_k \mathrm{sign}(x_i(t) - x_k(t)) - \sum_{y_k(t)\neq x_i(t)} n_k \mathrm{sign}(x_i(t) - y_k(t)), \\ \dot{y}_j(t) = \displaystyle\sum_{y_k(t)\neq y_j(t)} n_k \mathrm{sign}(y_j(t) - y_k(t)) - \sum_{x_k(t)\neq y_j(t)} m_k \mathrm{sign}(y_j(t) - x_k(t)), \end{cases} \tag{3}$$

with $i = 1, .., N$, and $j = 1, ..., N$. System (3) models the movement of N particles for each species, with masses $m_1, \ldots, m_N$ for the x-species and $n_1, \ldots, n_N$ for the y-species, under the effect of repulsive Newtonian potentials for same-species interactions and attractive Newtonian potentials for cross-species interactions. We stress that particles in the ODE system (3) *may overlap*. When this happens, the right-hand side of (3) features a jump discontinuity, which brings additional difficulties. To bypass this problem and to better understand the dynamics of (3), we frame it rigorously as the (finite dimensional) gradient flow of the (convex, in a suitable metric sense) functional

$$-\frac{1}{2}\sum_{i,j} m_i m_j |x_i - x_j| - \frac{1}{2}\sum_{i,j} n_i n_j |y_i - y_j| + \sum_{i,j} m_i n_j |x_i - y_j|,$$

in the convex cone $\mathcal{C}^N \times \mathcal{C}^N$ of ordered configurations

$$x_1 \leq x_2 \leq \ldots \leq x_N\,, \qquad y_1 \leq y_2 \leq \ldots \leq y_N.$$

. More precisely, among other issues:

- We prove that the sub-differential of this functional is always non-empty for any given configuration in $\mathcal{C}^N \times \mathcal{C}^N$ (including overlapping of particles of opposite species).
- We analyse *collisions* among particles (which are possible because particles do not "slow down" when they get very close due to the lack of regularity of the interaction potential) and prove that *particles of the same species never collide*. Moreover, we provide explicit necessary and sufficient conditions for particles of opposite species to cross each other.
- We explore the case of initial overlapping of particles and provide the explicit solution to the corresponding particle system.

These properties are preparatory to prove the rigorous derivation of solutions to (2) with L^1 initial data as many-particle limits of the empirical measures of the particle system (3).

Such a result heavily relies on uniform estimates at the discrete level. In particular, we observe that in the 2-species case weak compactness in the measure sense by itself is insufficient to obtain consistency in the limit due to the cross-interaction terms. Indeed, since the cross-interaction terms cannot be symmetrised (unlike, for instance, the Keller-Segel one-species model), *weak L^1 compactness* is needed in this case.

As already mentioned, system (2) is not included in the theory of [7] since the interaction potential in the (repulsive) intraspecific part is neither convex nor λ-convex, i.e., convex up to a quadratic perturbation. In one dimension this problem can be overcome as shown in [3]. In particular, in case of absolutely continuous initial data ρ_0, η_0 [3] proves global-in-time existence and uniqueness of solutions by posing system (2) as gradient flow of the *interaction energy functional*

$$\mathcal{F}(\rho,\eta) = -\frac{1}{2}\int_{\mathbb{R}} N*\rho\, d\rho - \frac{1}{2}\int_{\mathbb{R}} N*\eta\, d\eta + \int_{\mathbb{R}} N*\eta\, d\rho, \tag{4}$$

where $N(x) := |x|$, $x \in \mathbb{R}$. When dealing with general measures as initial data, in particular Dirac deltas, the sub-differential of $\mathcal{F}$ may be empty. Hence, in [3], global-in-time existence and uniqueness of solutions to system (2) is proven by (formally) re-writing the system in the pseudo-inverse formalism and by using the concept of gradient flows in Hilbert spaces *à la Brézis*, cf. [2]. Another difference with [7] is that the analysis in [3] implies that *particle solutions* are not *gradient flow solutions* to system (2). Thus, the mean-field limit cannot be treated via the stability result mentioned previously since the atoms of the empirical measure may diffuse instantaneously.

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] H. Brézis, *Opérateurs maximaux monotones et semi-groupes de contractions dans les espaces de Hilbert*, North-Holland Publishing Co., Amsterdam, 1973.

[3] J. A. Carrillo, M. Di Francesco, A. Esposito, S. Fagioli, M. Schmidtchen, *Measure solutions to a system of continuity equations driven by Newtonian nonlocal interactions*, Discrete & Continuous Dynamical Systems - A **40, no. 2** (2020), 1191–1231.

[4] J. A. Carrillo, M. Di Francesco, A. Figalli, T. Laurent, D. Slepčev, *Global-in-time weak measure solutions and finite-time aggregation for nonlocal interaction equations*, Duke Math. J. **156 no. 2** (2011), 229–271.

[5] M. Di Francesco, A. Esposito, S. Fagioli, *Nonlinear degenerate cross-diffusion systems with nonlocal interaction*, Nonlinear Analysis **169** (2018), 94–117.

[6] M. Di Francesco, A. Esposito, M. Schmidtchen, *Many-particle limit for a system of interaction equations driven by Newtonian potentials*, preprint arXiv:2008.11106 (2020).

[7] M. Di Francesco, S. Fagioli, *Measure solutions for nonlocal interaction PDEs with two species*, Nonlinearity, **26** (2013), 2777–2808.

[8] R. Jordan, D. Kinderlehrer, F. Otto, *The variational formulation of the Fokker–Planck equation*, SIAM journal on mathematical analysis **29, no. 1** (1998), 1–17.

[9] R. J. McCann, *A convexity principle for interacting gases*, Advances in mathematics, **128 no. 1** (1997), 153–179.

[10] F. Otto, *The geometry of dissipative evolution equations: the porous medium equation*, Comm. Partial Differential Equations **26, no. 1–2** (2001), 101–174.

Rigorous derivation of cross-diffusion equations from interacting particle systems

Ansgar Jüngel
(joint work with L. Chen, E. Daus, and A. Holzinger)

Cross-diffusion models describe the evolution of multicomponent systems arising in, for instance, cell biology, gas mixture theory, and population dynamics. Their derivation from microscopic models is important to determine the range of validity of the diffusive equations and to understand their possible formal gradient-flow or entropy structure. We review in this note two many-particle limits from stochastic interacting particle systems, based on the works [2, 3].

The aim is the rigorous derivation of quasilinear parabolic systems of the form

$$\partial_t u_i = \operatorname{div}\left(\sum_{j=1}^n A_{ij}(u)\nabla u_j\right) + \operatorname{div}(u_i\nabla U_i), \quad u_i(0) = u_{0,i} \quad \text{in } \mathbb{R}^d,\ t>0, \tag{1}$$

where $i = 1,\ldots,n$ is the species index, $u = (u_1,\ldots,u_n)$ is the vector of (particle) densities, $A_{ij}(u)$ are the diffusion coefficients, and $U_i(x)$ are environmental potentials, from stochastic interacting particle systems of the type

$$\mathrm{d}X_{k,i} = -a^{N,\eta}(X)\mathrm{d}t + b^{N,\eta}(X)\mathrm{d}W_{k,i}(t), \quad X_{k,i}(0) = \xi_{k,i}, \quad i = 1,\ldots,n, \tag{2}$$

where $k = 1,\ldots,N$ is the particle number, $X = (X_{k,i})$ is the vector of random positions of the particles, the parameter $\eta > 0$ models the interaction radius, $(W_{k,i})$ are d-dimensional Brownian motions, and $\xi_{1,i},\ldots,\xi_{N,i}$ are independent and identically distributed random variables. We wish to prove the limit $N\to\infty$ and $\eta\to 0$ (in a certain sense) in (2) leading to (1). For this, we consider two examples for $a^{N,\eta}(X)$ and $b^{N,\eta}(X)$.

1. First model: interactions in the drift term

The first model is given by

$$\mathrm{d}X_{k,i}^{N,\eta}(t) = -\sum_{j=1}^n \frac{1}{N}\sum_{\ell=1}^N \nabla B_{ij}^\eta\big(X_{k,i}^{N,\eta} - X_{\ell,j}^{N,\eta}\big)\mathrm{d}t + \sqrt{2\sigma_i}\mathrm{d}W_{k,i}(t), \tag{3}$$

where $\sigma_i > 0$ are constant diffusion coefficients and the smooth interaction potentials B_{ij} satisfy $B_{ij}^\eta(x) = \eta^{-d}B_{ij}(|x|/\eta)$ for $x\in\mathbb{R}^d$ with $\int_{\mathbb{R}^d} B_{ij}(|x|)dx =: a_{ij}$ and $B_{ij}^\eta \to a_{ij}\delta_0$ in the sense of distributions as $\eta\to 0$.

The limit $N\to\infty$, $\eta\to 0$ has to be understood in the following sense (see [6, 8]). For fixed $\eta > 0$, system (3) is approximated for $N\to\infty$ by the intermediate system

$$\mathrm{d}\bar{X}_{k,i}^\eta(t) = -\sum_{j=1}^n (\nabla B_{ij}^\eta * u_{\eta,j})(\bar{X}_{k,i}^\eta(t),t)\mathrm{d}t + \sqrt{2\sigma_i}\mathrm{d}W_{k,i}(t), \tag{4}$$

where $u_{\eta,j} = u_{\eta,j}(x,t)$ satisfies the nonlocal cross-diffusion system

$$\partial_t u_{\eta,i} = \sigma_i\Delta u_{\eta,i} + \operatorname{div}\left(\sum_{j=1}^n u_{\eta,i}\nabla B_{ij}^\eta * u_{\eta,j}\right) \quad \text{in } \mathbb{R}^d,\ t>0.$$

System (4) depends on the particle index $k = 1, \ldots, N$ only via the initial data $\bar{X}^k_{\eta,i}(0) = \xi_{k,i}$, i.e., $\bar{X}^\eta_{k,i}(t)$ are N independent copies of the solution to (4). Since $\nabla B^\eta_{ij} * u_{\eta,j} \to a_{ij}\nabla u_j$ in L^2, the limit $\eta \to 0$ in (4) leads to the limiting system

$$\mathrm{d}\widehat{X}_{k,i}(t) = -\sum_{j=1}^n a_{ij}\nabla u_j(\widehat{X}_{k,i}(t))\mathrm{d}t + \sqrt{2\sigma_i}\mathrm{d}W_{k,i}(t),$$

where the law of $\widehat{X}_{k,i}$, $u_i = \mathrm{law}(\widehat{X}_{k,i})$, is a solution to

$$\partial_t u_i = \sigma_i\Delta u_i + \mathrm{div}\left(\sum_{j=1}^n a_{ij}u_i\nabla u_j\right) \quad \text{in } \mathbb{R}^d,\ t > 0, \tag{5}$$

and $u_i(0) = u_{i,0}$ is the common probability density function of $\xi_{k,i}$. This model describes segregation effects in multi-species populations [1, 5].

The main result of [2] is the proof of the estimate

$$\sup_{k=1,\ldots,N} \mathrm{E}\left(\sum_{i=1}^n \sup_{0<s<t} \left|X^{N,\eta}_{k,i}(s) - \widehat{X}_{k,i}(s)\right|\right) \le C(t)\eta,$$

under the condition that $\eta^{-(2d+4)} \le \varepsilon \log N$, where $\varepsilon > 0$ is sufficiently small. This estimate implies propagation of chaos [9]. The idea of the proof is to estimate the differences $|X^{N,\eta}_{k,i} - \bar{X}^\eta_{k,i}|$ and $|\bar{X}^\eta_{k,i} - \widehat{X}_{k,i}|$. The first difference is of order $N^{-1}\eta^{-d-2}$, coming from the properties of ∇B^η_{ij}, while the second difference is of order η, which comes from estimating $|\nabla B^\eta_{ij} * \nabla u_j - a_{ij}\nabla u_j|$ in terms of $\eta|D^2 u_j|$.

2. Second model: interactions in the diffusion term

The second model is given by

$$\mathrm{d}X^{N,\eta}_{k,i} = -\nabla U_i(X^{N,\eta}_{k,i})\mathrm{d}t + \left(2\sigma_i + 2\sum_{j=1}^n \frac{1}{N}\sum_{\ell=1}^N B^\eta_{ij}(X^{N,\eta}_{k,i} - X^{N,\eta}_{\ell,j})\right)^{1/2}\mathrm{d}W_{k,i}(t), \tag{6}$$

where we exclude $(\ell, j) \ne (k, i)$ in the sum over ℓ. The idea of the many-particle limit is as before. We first pass to the limit $N \to \infty$, leading to an intermediate nonlocal system, and then perform the limit $\eta \to 0$, giving

$$\mathrm{d}\widehat{X}_{k,i} = -\nabla U_i(\widehat{X}_{k,i})\mathrm{d}t + \left(2\sigma_i + 2\sum_{j=1}^n a_{ij}u_j(\widehat{X}_{k,i})\right)^{1/2}\mathrm{d}W_{k,i}(t),$$

and the function $u_i = \mathrm{law}(\widehat{X}_{k,i})$ satisfies

$$\partial_t u_i = \mathrm{div}(u_i\nabla U_i) + \Delta\left(\sigma_i u_i + u_i\sum_{j=1}^n a_{ij}u_j\right), \quad u_i(0) = u_{i,0}, \tag{7}$$

where $i = 1, \ldots, n$. Model (7) corresponds to the population system suggested by Shigesada, Kawasaki, and Teramoto [7]. It distinguishes from the first model

(5) by the additional diffusion $\operatorname{div}(\sum_{j=1}^n a_{ij}u_j\nabla u_i)$. It is shown in [3] that if $U_i(x) = -\frac{1}{2}|x|^2$ and $\eta^{-(2d+2)} \le \varepsilon \log N$ for some sufficiently small $\varepsilon > 0$ then

$$\sup_{k=1,\ldots,N} \mathrm{E}\bigg(\sum_{i=1}^n \sup_{0<s<t} \big|X_{k,i}^{N,\eta}(s) - \widehat{X}_{k,i}(s)\big|\bigg) \le C(t)\eta.$$

This result can be extended in various directions. First, we may choose general smooth potentials U_i such that ∇U_i is globally Lipschitz continuous, D^2U_i is positive semidefinite, and D^kU_i is sufficiently small for $kg_\epsilon 3$. Second, the diffusion coefficient in (6) can be replaced by

$$\bigg(2\sigma_i + 2\sum_{j=1}^n f_\eta\bigg(\frac{1}{N}\sum_{\ell=1}^N B_{ij}^\eta(X_{k,i}^{N,\eta} - X_{\ell,j}^{N,\eta})\bigg)\bigg)^{1/2},$$

where f_η is a globally Lipschitz continuous approximation of a function f that may be only locally Lipschitz continuous (for instance, $f(z) = z^p$ for $p > 1$). Then the sum $\sum_{j=1}^n a_{ij}u_j$ in (7) has to be replaced by $\sum_{j=1}^n f(a_{ij}u_j)$. This generalization provides a derivation of the porous-medium equation from interacting particle systems. Indeed, let $n = 1$, $\sigma_1 = 0$, $U_1 = 0$, and $a_{11} = 1$. Then (7) can be written as $\partial_t u = \Delta(uf(u))$. We remark that another derivation was published in [4] assuming a double-convolution potential in the drift term.

References

[1] M. Bertsch, M. Gurtin, D. Hilhorst, L. Peletier, *On interacting populations that disperse to avoid crowding: preservation of segregation*, J. Math. Biol. **23** (1985), 1–13.

[2] L. Chen, E. S. Daus, A. Jüngel, *Rigorous mean-field limit and cross-diffusion*, Z. Angew. Math. Phys. **70** (2019), no. 122, 21 pages.

[3] L. Chen, E. S. Daus, A. Holzinger, A. Jüngel, *Rigorous derivation of population cross-diffusion systems from moderately interacting particle systems*, work in preparation, 2020.

[4] A. Figalli, R. Philipowski, *Convergence to the viscous porous medium equation and propagation of chaos*, Alea **4** (2008), 185–203.

[5] G. Galiano, V. Selgas, *On a cross-diffusion segregation problem arising from a model of interacting particles*, Nonlin. Anal. Real World Appl. **18** (2014), 34–49.

[6] K. Oelschläger, *On the derivation of reaction-diffusion equations as limit dynamics of systems of moderately interacting stochastic processes*, Prob. Theory Rel. Fields **82** (1989), 565–586.

[7] N. Shigesada, K. Kawasaki, and E. Teramoto, *Spatial segregation of interacting species*, J. Theor. Biol. **79** (1979), 83–99.

[8] A. Stevens, *The derivation of chemotaxis equations as limit dynamics of moderately interacting stochastic many-particle systems*, SIAM J. Appl. Math. **61** (2000), 183–212. (Erratum: **61** (2000), 2200–2200.)

[9] A. S. Sznitman, *Topics in propagation of chaos*, in: P. L. Hennequin (ed.), École d'Été de Probabilités de Saint-Flour, XIX-1989, Lect. Notes Math. **1464**. Berlin, Springer, 1991.

Global existence analysis of energy-reaction-diffusion systems

Katharina Hopf
(joint work with Julian Fischer, Michael Kniely, Alexander Mielke)

Reaction-diffusion systems (RDS) arise in the modelling of various physical phenomena and often have an underlying entropy structure linked to the thermodynamic origin of the problem. In this talk, we propose thermodynamically consistent extensions of RDS to the non-isothermal setting taking into account variations in temperature, and discuss issues concerning the global-in-time existence of solutions.

As a starting point, consider an isothermal system of a number of I chemical species diffusing and reacting in a bounded Lipschitz domain $\Omega \subset \mathbb{R}^d$, described by a vector of non-negative concentrations $c = (c_1, \dots, c_I)$ obeying the evolution law

$$\dot{c}_i = \operatorname{div}(a_i \nabla c_i) + R_i(c), \qquad i = 1, \dots, I. \tag{1}$$

Here, $c_i = c_i(t, x)$, where $t > 0$ denotes the time variable, $x \in \Omega$ the space variable, $\dot{c} = \frac{\mathrm{d}}{\mathrm{d}t} c$, and $\nabla = \nabla_x$ the spatial gradient. The constant $a_i > 0$ is the diffusion coefficient of the i-th species, and $R_i \in C([0,\infty)^I)$ describe the reactions. In many applications (e.g. for a single reversible reaction following mass action kinetics), this problem has an underlying Lyapunov functional that essentially consists of a sum of Boltzmann entropies.

In the non-isothermal case, the thermodynamic equilibria $w_i = w_i(u)$ of the concentrations c_i depend on the distribution of the thermal variable $u = u(t, x)$, which is itself an unknown of the problem. Following [7, 5, 8] we model changes in temperature using the *internal energy* density u, the advantage being that typical driving functionals (= the negative of the physical entropy) are convex in $z = (c, u)$ (but fail to be convex when considered as a function of c and the temperature θ). Moreover, since chemical reactions do not change the internal energy, the evolution law for u is source-free. Our models are based on Lyapunov functions of the form (cf. [7, 8, 5])

$$h(c, u) = b(c|w(u)) - \sigma(u) = \sum_{i=1}^{I} \Big(\lambda(c_i) - c_i \log w_i(u) \Big) - \hat{\sigma}(u), \tag{2}$$

where $\lambda(s) = s \log s - s + 1$ is the Boltzmann function, $b(c|w) := \sum_{i=1}^{I} w_i \lambda(c_i/w_i)$ the relative Boltzmann entropy and $\sigma(u)$ the thermal part of the entropy density when the concentrations $c = (c_i)_{i=1}^I$ are in equilibrium. We generally assume that $w_i \in C([0,\infty)) \cap C^2((0,\infty))$ are positive, non-decreasing and concave, while the C^2 function $\hat{\sigma}(s) := \sigma(s) - \sum_{i=1}^{I} w_i(s) + I$ is supposed to be strictly concave and increasing. These properties ensure that h is strictly convex and that $u \mapsto h(c, u)$ is decreasing. The temperature θ can then be recovered via $\theta = \frac{1}{-\partial_u h(c,u)}$ and is non-negative.

Letting $A(z) := \mathbb{M}(z)D^2h(z)$, $z = (c,u)$, and $R^\circ(z) := (R(z),0)$, the *energy-reaction-diffusion systems* (ERDS) we consider take the form

$$\begin{aligned} \dot z &= \operatorname{div}\big(A(z)\nabla z\big) + R^\circ(z), && t>0,\ x\in\Omega,\\ 0 &= \nu\cdot A(z)\nabla z, && t>0,\ x\in\partial\Omega. \end{aligned}$$

Under the basic hypotheses that $\mathbb{M}$ be positive semi-definite and R° satisfy the pointwise inequality $Dh(z)\cdot R^\circ(z)\le 0$, solutions z of the above system formally enjoy the entropy estimate

$$\frac{\mathrm{d}}{\mathrm{d}t}\int_\Omega h(z) + \int_\Omega \nabla z : D^2h(z)\mathbb{M}(z)D^2h(z)\nabla z\,\mathrm{d}x \le 0. \tag{3}$$

Due to the entropic coupling between c_i and u, the Hessian of h is non-diagonal. It takes the form

$$D^2h(c,u) = \begin{pmatrix} \ddots & & & \vdots \\ & \frac{1}{c_i} & & -\frac{w_i'}{w_i} \\ & & \ddots & \vdots \\ \dots & -\frac{w_i'}{w_i} & \dots & \partial_u^2 h \end{pmatrix},$$

where $w_i = w_i(u)$. Hence, in general the matrix $A(z) = \mathbb{M}(z)D^2h(z)$ is not diagonal, usually not symmetric, and positive semi-definiteness cannot be expected.

Our existence analysis focuses on (symmetric) mobility matrices of the form

$$\mathbb{M}(z) := \operatorname{diag}\big(m_1,\dots,m_I,0\big) + \pi_1(z)\mu\otimes\mu \tag{4}$$

for certain continuous functions $m_i(c,u)\ge 0$ (e.g. $m_i = a_ic_i$ for some $a_i\in\mathbb{R}^+$) and $\pi_1(c,u)\ge 0$, and where $\mu = (\mu_1,\dots,\mu_I,1)$ with $\mu_i(c,u) := \frac{w_i'(u)}{w_i(u)}c_i$. This ansatz is motivated by the structure of the inverse Hessian of h, and allows in particular to model thermodynamically consistent energy-dependent extensions of the basic isothermal problem (1).

In case that the flux term $A(z)\nabla z$ and the reaction rates $R(z)$ can be suitably controlled in terms of the formal a priori bounds obtained from the entropy structure, entropy methods for reaction-cross-diffusion systems [6] can be adapted to our ERDS and allow us to construct global-in-time weak solutions. The key here is to exploit the availability of a strictly convex entropy function $h(z)$ to define a transformation $w = Dh(z)$ upon which the system $\dot z = \operatorname{div}(A(z)\nabla z)$, with $A(z) = \mathbb{M}(z)D^2h(z)$, becomes $\dot z = \operatorname{div}(\mathbb{M}(z)\nabla w)$, where now $z = (Dh)^{-1}(w)$ is a function of the new unknown variables w. Letting g denote the Legendre transform of h, we have $z = Dg(w)$ (and recall that g is convex). Thus, this change of variables reveals the parabolic structure of the PDE system. A crucial technical advantage of this transformation lies in the fact that in the new variables the system can be regularised using standard vanishing viscosity approximations while preserving the main a priori estimates obtained from the entropy-entropy dissipation inequality (3). Issues with low time regularity can be taken care of by first considering a suitable time-discrete approximation that preserves the entropy

estimate (e.g. an implicit Euler scheme). Once, approximate solutions have been constructed, one returns to the original variables in order to exploit (3).

In contrast to many of the models in [6], which are of volume filling type and provide a priori L^∞ bounds on the concentrations, our structural assumptions do not imply pointwise upper bounds. Instead, integrability of the concentrations solely relies on the entropy estimate (3): it is based on Gagliardo–Nirenberg inequalities interpolating between the uniform-in-time upper bound on $\int_\Omega h(z)\,\mathrm{d}x$ and the gradient control obtained from entropy-dissipation coercivity bounds, which in our applications typically take the form

$$\nabla z : D^2h(z)\mathbb{M}(z)D^2h(z)\nabla z \gtrsim \sum_{i=1}^{I} |\nabla\sqrt{c_i}|^2 + \alpha(z)|\nabla u|^2.$$

(Here, due to the sublinearity of $\hat\sigma$ the function α may be very degenerate as $u \to \infty$. Thanks to the absence of reactions in the energy equation, the u-component has, however, an additional quasilinear heat equation-like structure, which gives rise to separate a priori bounds.)

The lack of good integrability bounds for the concentrations can lead to serious issues concerning the global-in-time existence analysis. In fact, we encounter situations, where the a priori bounds obtained from entropy estimates are not sufficient to give a meaning to the flux term $A(z)\nabla z$ in the distributional sense. At the same time, these formal bounds do prevent global blow-up of the concentrations and provide some control of the gradient ∇z, which calls for an adaptation of the classical concept of weak solutions when faced with the question of global existence. Related issues are well-known for Boltzmann equations, where a concept of renormalised solutions has been introduced [3, 2]. This concept has recently been adapted to (isothermal) reaction-diffusion systems such as (1) to deal with continuous reaction rates not obeying any growth hypothesis [4]. This notion of solution requires compositions $\xi \circ z$ for $\xi \in C^\infty([0,\infty)^{I+1})$ with compactly supported derivative $D\xi$, to satisfy in the distributional sense an evolution equation obtained by a formal application of the chain rule. In the context of our ERDS, this equation states

(5)
$$\tfrac{\mathrm{d}}{\mathrm{d}t}\xi(z) = \sum_{i,j} \operatorname{div}\big(\partial_i\xi(z)A_{ij}(z)\nabla z_j\big) - \sum_{i,j} A_{ij}(z)\nabla z_j\cdot\nabla\partial_i\xi(z) + \sum_i \partial_i\xi(z)R_i^\circ(z).$$

This definition is meaningful if, e.g., $\sqrt{z_i} \in L^2_tH^1_x$ and $\chi_{\{|z|\le E\}}A(z)\nabla z \in L^2_{t,x}$ for all $E \ge 1$. Using this solution concept, we have developed an existence analysis for the above-mentioned problem with uncontrolled flux and reactions without growth bounds. In the construction of renormalised solutions we face two new challenges (w.r.t. existing works [4, 1]). The first difficulty is to find a suitable regularisation giving rise to global-in-time regular (or weak) approximate solutions since this requires a modification at the level of the diffusive operator, which features cross effects and is defined upon a matrix $A(z)$ lacking positive definiteness. Here, we

exploit the entropy form $A(z) = \mathbb{M}(z)D^2h(z)$, where the crucial point is a 'regularisation' (adaptation) of the mobility matrix $\mathbb{M}(z)$. A second complication arises due to the presence of the squared gradients in the renormalised formulation (5), leading to compactness issues when passing to the limit in the approximate problem. This is particularly delicate in the uncontrolled cross terms induced by energy gradients (those are the reason for the lack of $L^1_{t,x}$ a priori bounds on $A(z)\nabla z$). The key to resolve this issue lies in the heat equation-like structure of the energy component, which allows to show strong convergence in $L^2_{t,x}$ at the level of ∇u.

References

[1] X. Chen and A. Jüngel. Global renormalized solutions to reaction-cross-diffusion systems with self-diffusion. *J. Differential Equations*, 267(10):5901–5937, 2019.

[2] R. J. DiPerna and P.-L. Lions. On the Fokker-Planck-Boltzmann equation. *Comm. Math. Phys.*, 120(1):1–23, 1988.

[3] R. J. DiPerna and P.-L. Lions. On the Cauchy problem for Boltzmann equations: global existence and weak stability. *Ann. of Math. (2)*, 130(2):321–366, 1989.

[4] J. Fischer. Global existence of renormalized solutions to entropy-dissipating reaction-diffusion systems. *Arch. Ration. Mech. Anal.*, 218(1):553–587, 2015.

[5] J. Haskovec, S. Hittmeir, P. Markowich, and A. Mielke. Decay to equilibrium for energy-reaction-diffusion systems. *SIAM J. Math. Anal.*, 50(1):1037–1075, 2018.

[6] A. Jüngel. The boundedness-by-entropy method for cross-diffusion systems. *Nonlinearity*, 28(6):1963–2001, 2015.

[7] A. Mielke. Thermomechanical modeling of energy-reaction-diffusion systems, including bulk-interface interactions. *Discrete Contin. Dyn. Syst. Ser. S*, 6(2):479–499, 2013.

[8] A. Mielke and M. Mittnenzweig. Convergence to equilibrium in energy-reaction-diffusion systems using vector-valued functional inequalities. *J. Nonlinear Sci.*, 28(2):765–806, 2018.

On fluctuations in particle systems and their links to macroscopic models

Johannes Zimmer

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

There are two views on diffusion, say for example

$$\dot\rho_t = \Delta\rho_t + \nabla\cdot(\rho_t\nabla V), \tag{1}$$

with a (smooth and confining) external potential V. Namely, we can interpret (1) as gradient flow for a functional $\mathcal{F}(\rho)$ and a metric $M(\rho)$:

$$\dot\rho_t = -M(\rho_t)\frac{\delta\mathcal{F}}{\delta\rho_t},$$

for example with $\mathcal{F}$ the Boltzmann entropy and M the Wasserstein metric. On the other hand, (1) can be interpreted as a continuity equation $\dot\rho_t = -\nabla\cdot J(\rho_t)$, with *current* or *flux* $J(\rho) = -\nabla\rho - \rho\nabla V$. We can define a *force* $F(\rho) = -\nabla\log\rho - \nabla V = -\nabla\frac{\delta\mathcal{F}}{\delta\rho}$, such that

$$J(\rho) = \rho F(\rho).$$

These structures exist on the continuum. One can think of this macroscopic scale as scaling limit of small-scale processes. In particular, Markov processes

can lead to macroscopic diffusive effects. Are there analogous structures on the Markov chain level?

For reversible Markov chains, Maas and Mielke have independently discovered a gradient flow structure [5, 7]. The irreversible case corresponds to

$$\dot\rho_t = \Delta\rho_t + \nabla\cdot(\rho_t E),$$

with a non-gradient field E. The non-gradient structure destroys the gradient flow structure we know for (1). However, the force-flux relation makes sense even in this non-gradient setting. It is therefore natural to seek a force-flux structure for Markov chains. It turns out that it is possible to define a force j and current J for finite state Markov chains, see [3, 6] and references therein. One sees that the force-flux relation is nonlinear,

$$J_{xy}(\rho) = a_{xy}(\rho)\sinh\big(\tfrac12 F_{xy}(\rho)\big);$$

here J_{xy} is the current between nodes x and y and F_{xy} is the force.

Furthermore, one can show that a large deviation principle holds, where the rate functional is finite for paths satisfying a continuity equation, and then takes the form

(2)
$$I_{[0,T]}\big((\rho_t,j_t)_{t\in[0,T]}\big) = \mathcal F(\rho_0) + \int_0^T \frac12\left[\Psi(\rho_t,j_t) - \langle j_t, F(\rho_t)\rangle + \Psi^\star(\rho_t,F(\rho_t))\right]\,\mathrm dt;$$

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

(3)
$$\Psi^\star(\rho,F) := \sum_{xy} a_{xy}(\rho)\big(\cosh\big(\tfrac12 F_{xy}(\rho)\big) - 1\big),$$

see [6, 8, 3].

The proof in [3] is an application of Sanov's theorem, giving the rate functional as relative entropy of a path with respect to a reference path. A similar argument can be developed for a class of particle processes, including (non-condensing) zero-range processes. Again a large deviation principle is can be established, giving the relative entropy of a process with respect to a reference process. This rate functional is again of $\Psi-\Psi^\star$ form as (2) and (3) above. The hydrodynamic limit is a nonlinear diffusion process, with quadratic rate functional. We discuss here a simple case without potential, where the rate functional can be written as

(4)
$$\int_0^T \|\dot\rho_t - \Delta\phi(\rho)\|^2_{1,\chi(\rho_t)}\,dt = S(\rho(0)) - S(\rho(T)) + \frac12\int_0^T\Big[\|\dot\rho_t\rho\|^2_{1,\chi(\rho_t)} + \|\Delta\rho\|^2_{1,\chi(\rho_t)}\Big]\,dt.$$

One can show that this quadratic rate functional emerges as limit of the non-quadratic $(\Psi-\Psi^\star)$ rate functional (2) discussed before. Namely, under strong assumptions, notably concentration of limit paths, one can show that, among others, the following convergences hold [4] (the subscript L indicates that the

functional is defined on the lattice level, with $L \to \infty$ being the continuum limit in a suitable scaling):

(1) We have

$$\lim_{L\to\infty} \frac{1}{L^d} \int_0^T \Psi_L\big(\rho_t^L, \jmath_t^L\big)\,\mathrm{d}t = \frac{1}{2}\int_0^T \|\dot\rho_t\|^2_{-1,\chi(\rho_t)}\,\mathrm{d}t.$$

(2) Also, as $L \to \infty$,

$$\frac{1}{L^d} \int_0^T \Psi_L^\star\big(\rho_t^L, F_\alpha^V(\rho_t^L)\big)\,\mathrm{d}t \to \frac{1}{2}\int_0^T \|\Delta\phi(\rho_t)\|^2_{-1,\chi(\rho_t)}\,\mathrm{d}t.$$

Furthermore, the mixed term obtained in an expansion of the left-hand side in (4), which gives rise to the entropy difference on the right, can be obtained in a limit from the relevant expression for the particle functional. This is a subtle point, however, as the expansion requires suitable regularity. In particular, we need to know that for the analogous term on the particle level, it holds that

$$\mathcal{F}_\alpha^V(\rho_{t_2}) - \mathcal{F}_\alpha^V(\rho_{t_1}) = \int_{t_1}^{t_2} \big\langle \dot\rho_t, \frac{\delta \mathcal{F}_\alpha^V}{\delta \rho_t}\big\rangle\,\mathrm{d}t = -\int_{t_1}^{t_2} \langle \dot\rho_t, \Delta\phi(\rho_t)\rangle_{-1,\chi(\rho_t)}\,\mathrm{d}t,$$

This requires a Wasserstein-type calculus for the particle process under consideration. In [4], we use equivalence of the topology to the classic Wasserstein one. It would be worthwhile to build the limit passage relying entirely on the geometry associated with the particle process in question; this can lead to interesting connections between particle dynamics and Wasserstein-type geometries. For example, the result [4] is restricted to one space dimension, as it uses the McCann condition for geodesic convexity, linked to the pressure in the particle system.

The aim of [4] was to establish the passage from the $\Psi - \Psi^\star$ structure on the particle level to the macroscopic structure (4), thereby giving a rigorous derivation and interpretation of the quadratic structure of the latter functional. The classic hydrodynamic limit passage, however, requires significantly weaker assumptions than those made in [4]. The underlying idea of deriving macroscopic gradient flows follows that of [2]. It is possible that ideas from [1] can be used profitably to weaken the assumptions.

References

[1] Matthias Erbar, Max Fathi, Vaios Laschos, André Schlichting. Gradient flow structure for McKean-Vlasov equations on discrete spaces. *Disc. Contin. Dyn. Syst.*, 36:12, 6799–6833, 2016.

[2] Max Fathi and Marelle Simon. The gradient flow approach to hydrodynamic limits for the simple exclusion process. In: *From particle systems to partial differential equations. III*, 167–184, 2016.

[3] Marcus Kaiser, Robert L. Jack, and Johannes Zimmer. Canonical structure and orthogonality of forces and currents in irreversible Markov chains. *J. Stat. Phys.*, 170(6):1019–1050, 2018.

[4] Marcus Kaiser, Robert L. Jack, and Johannes Zimmer. A Variational Structure for Interacting Particle Systems and their Hydrodynamic Scaling Limits. *Commun. Math. Sci.*, 17(3):739–780, 2019.

[5] Jan Maas. Gradient flows of the entropy for finite Markov chains. *J. Funct. Anal.*, 261(8):2250–2292, 2011.
[6] C. Maes and K. Netočný. Canonical structure of dynamical fluctuations in mesoscopic nonequilibrium steady states. *Europhys. Lett. EPL*, 82(3):Art. 30003, 6, 2008.
[7] Alexander Mielke. A gradient structure for reaction-diffusion systems and for energy-drift-diffusion systems. *Nonlinearity*, 24(4):1329–1346, 2011.
[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.

Nonlocal equations on discrete spaces inspired by numerical schemes

André Schlichting
(joint work with Antonio Esposito, Francesco Patacchini, Dejan Slepčev, Christian Seis)

The nonlocal-nonlocal interaction equation. We introduce a general framework for studying interaction equations on families of graphs and their limits as the number of vertices n goes to ∞ based on the preprint [4]. In particular, in the applications to machine learning, the graphs considered are random samples of some underlying measure in Euclidean space, and the edge weights, as well as the interaction energy, depend on the positions of the vertices.

The vertices are points in $\mathbb{R}^d$. The edges are given in terms of a non-negative symmetric weight function $\eta\colon \{(x,y)\in\mathbb{R}^d\times\mathbb{R}^d : x\neq y\}\to[0,\infty)$, which defines the set of edges as $G=\{(x,y)\in\mathbb{R}^d\times\mathbb{R}^d : x\neq y,\, \eta(x,y)>0\}$. This setting also covers discrete graphs with vertices $X=\{x_1,\dots,x_n\}\subset\mathbb{R}^d$ by considering μ as the empirical measure of the set of points, $\mu=\frac{1}{n}\sum_{i=1}^n \delta_{x_i}$. The distribution of mass over the vertices is described by a measure $\rho\in\mathcal{P}(\mathbb{R}^d)$.

The evolution of interest is the gradient descent of the energy $\mathcal{E}:\mathcal{P}(\mathbb{R}^d)\to\mathbb{R}$

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

where $K\colon\mathbb{R}^d\times\mathbb{R}^d\to\mathbb{R}$ is symmetric. The gradient flow equation takes the form

$$\begin{aligned}
&\partial_t\rho_t(x) = -\int_{\mathbb{R}^d} j_t(x,y)\eta(x,y)\,d\mu(y) =: -(\overline{\nabla}\cdot j_t)(x),\\
&j_t(x,y) = \rho_t(x)v_t(x,y)_+ - \rho_t(y)v_t(x,y)_-,\\
&v_t(x,y) = -\left(K*\rho_t(y) - K*\rho_t(x) + P(y) - P(x)\right).
\end{aligned} \tag{NL^{2}IE}$$

Firstly, the system (NL^{2}IE) consists of a nonlocal continuity equation, where the divergence $\overline{\nabla}\cdot$ is encoded by the graph structure described through μ and η. Secondly, it involves a mapping from velocity to flux, which here is the upwind flux and encodes the geometry of the gradient structure. Finally, the third equation identifies the driving velocity as the nonlocal gradient of the variation of the energy (1). Overall, we obtain that (NL^{2}IE) is the gradient flow of the energy $\mathcal{E}$ with respect to a generalization of the *Upwind transportation metric*.

Upwind nonlocal transportation metric. Let us set

$$\mathcal{V}^{\mathrm{as}}(G) := \{v\colon G \to \mathbb{R} : v(x,y) = -v(y,x) \text{ for all } (x,y) \in G\} \tag{2}$$

and call its elements *nonlocal (antisymmetric) vector fields* on G; for any pair $(x,y) \in G$ the value $v(x,y)$ can be regarded as a jump rate from x to y. Let us fix a final time $T > 0$ and let a family $\{v_t\}_{t\in[0,T]} \subset \mathcal{V}^{\mathrm{as}}(G)$ be given. In the case $\rho_t \ll \mu$ for all $t \in [0,T]$, it is possible to combine the first two equations in (NL^{2}IE) in order to arrive at the nonlocal continuity equation for μ-a.e. $x \in \mathbb{R}^d$

$$\partial_t \rho_t(x) + \int_{\mathbb{R}^d} \left(\rho_t(x) v_t(x,y)_+ - \rho_t(y) v_t(x,y)_-\right) \eta(x,y)\, d\mu(y) = 0. \tag{3}$$

For general curves $\rho\colon [0,T] \to \mathcal{P}(\mathbb{R}^d)$, it is possible to consider the weak form of (3), which is discussed in [4].

We use the nonlocal continuity equation (3) to define a *nonlocal Wasserstein quasi-distance* in analogy to the Benamou–Brenier formulation [1] for the Kantorovich–Wasserstein distances. That is, for two probability measures $\rho_0, \rho_1 \in \mathcal{P}_2(\mathbb{R}^d)$, let

$$\mathcal{T}_\mu(\rho_0,\rho_1)^2 := \inf_{(\rho,v)\in \mathrm{CE}(\rho_0,\rho_1)} \left\{ \int_0^1 \iint_G |v_t(x,y)_+|^2 \eta(x,y)\, d\rho_t(x)\, d\mu(y)\, dt \right\}, \tag{4}$$

where $\mathrm{CE}(\rho_0,\rho_1)$ is the set of weak solutions ρ to the nonlocal continuity equation on $[0,1]$ with $\rho(0) = \rho_0$ and $\rho(1) = \rho_1$. We note that the notion of the nonlocal Wasserstein distance for measures on $\mathbb{R}^d$ was introduced by Erbar in [3], who used it to study the fractional heat equation. One difference is that the interpolation we consider is beyond the scope of [3]. Very recently, [5] has extended the gradient flow viewpoint of the jump processes to generalized gradient structures driven by a broad class of internal energies. However, the particular choice of the upwind interpolation seems not to be covered.

Another difference is that here the measure μ plays an important role in how the action is measured and allows to incorporate seamlessly both the continuum case (e.g., μ is the Lebesgue measure on $\mathbb{R}^d$) and the graph case (μ is the empirical measure of the set of vertices).

Relation to the numerical finite-volume scheme. Equation (3) is very closely connected to the numerical Upwind scheme, the workhorse of finite-volume methods. To draw the connection, let $\{x_1,\dots,x_n\}$ be a suitable representative of a tessellation $\{K_1,\dots,K_n\}$, for instance a Voronoi tessellation, of some bounded domain $\Omega \subset \mathbb{R}^d$. Let μ be the Lebesgue measure on Ω and take η to be the transmission coefficient common in finite-volume schemes: $\eta(x_i,x_j) = \mathcal{H}^{d-1}(\overline{K_i} \cap \overline{K_j})/\operatorname{Leb}(K_i)$, for $i,j \in \{1,\dots,n\}$, where $\mathcal{H}^{d-1}(\overline{K_i} \cap \overline{K_j})$ is the $d-1$ dimensional Hausdorff measure of the common face between K_i and K_j. With this choice the equation (3) becomes the (continuous-time) discretization of the classical continuity equation

$$\partial_t \rho_t + \nabla \cdot \left(\mathbf{v}_t\, \rho_t\right) = 0,$$

for some vector field $\mathbf{v}_t\colon \Omega \to \mathbb{R}^d$. Hereby, the discretized vector field v_t is obtained from $\mathbf{v}_t$ by taking the average over common interfaces:

$$v_t(x_i, x_j) = \frac{1}{\mathcal{H}^{d-1}(\overline{K_i} \cap \overline{K_j})} \int_{\overline{K_i}\cap\overline{K_j}} \mathbf{v}_t \cdot \nu_{K_i,K_j}\, d\mathcal{H}^{d-1},$$

where ν_{K_i,K_j} is the unit normal to K_i pointing from K_i to K_j. We refer to the recent work [2] for a variational interpretation of the upwind scheme, which is close to that we propose for the more general equation (3).

The Scharfetter–Gummel scheme for aggregation-diffusion equations. In the preprint [7], a numerical scheme based on the Scharfetter–Gummel flux interpolation [6] is suggested to obtain a structure preserving scheme for the Wasserstein gradient flow associated to a free energy, which besides the interaction energy from (1), contains an entropy at some temperature $\kappa > 0$

$$\mathcal{F}_\kappa(\rho) = \kappa \int_{\mathbb{R}^d} \log \rho(x)\, d\rho(x) + \mathcal{E}(\rho), \tag{5}$$

The scheme can be rewritten as (NL^{2}IE), once the Upwind-interpolation for the flux is replaced by the Scharfetter–Gummel interpolation

$$j_t(x,y) = \theta_\kappa(\rho_t(x), \rho_t(y), v_t(x,y)) := v_t(x,y) \frac{\rho_t(x) e^{\frac{v_t(x,y)}{2\kappa}} - \rho_t(y) e^{-\frac{v_t(x,y)}{2\kappa}}}{e^{\frac{v_t(x,y)}{2\kappa}} - e^{-\frac{v_t(x,y)}{2\kappa}}}, \tag{6}$$

with $v_t(x,y)$ given as in (NL^{2}IE). Let us note, that $\theta_\kappa(a,b,v) \to a(v_+)^2 + b(v_-)^2$ for $\kappa \to 0$, which is exactly the Upwind interpolation. The velocity-flux relation in (6) encodes through the function θ_κ is derived by solving a suitable cell problem, which is given as the following boundary value problem

$$\begin{aligned} \theta_\kappa(\rho(x), \rho(y), v(x,y)) = -\kappa \partial_\xi \overline{\rho}(\xi) + v(x,y)\overline{\rho}(\xi) \quad &\text{for } \xi \in (0,1), \\ \overline{\rho}(0) = \rho(x) \qquad \text{and} \qquad \overline{\rho}(1) &= \rho(y). \end{aligned} \tag{7}$$

Hence besides θ_κ, the function $[0,1] \ni \xi \mapsto \overline{\rho}(\xi)$ is part of the unknown in (7).

The numerical scheme is obtained by a further implicit discretization of time. With the choice (6) of the flux-interpolation, it is shown in [7] that a discretized version of the free energy (5) is a Lyapunov function for the scheme. This allows to characterize stationary states of the scheme and show also its longtime behavior. In addition, the free-energy–dissipation gives suitable a prior bounds on discrete gradients of solutions to the scheme, which provide enough compactness for passing to the limit in the mesh size for obtaining weak solutions to the aggregation-diffusion equation

$$\partial_t \rho_t = \nabla \cdot \big(\kappa \nabla \rho_t + \rho_t \nabla W * \rho_t\big). \tag{8}$$

References

[1] J.-D. Benamou, Y. Brenier, *A computational fluid mechanics solution to the Monge-Kantorovich mass transfer problem*, Numer. Math. **84** (2000) 375–393.

[2] C. Cancès, T.O. Gallouët, G. Todeschi, *A variational finite volume scheme for Wasserstein gradient flows*, arXiv:1907.08305.

[3] M. Erbar, *Gradient flows of the entropy for jump processes*, Ann. l'Institut Henri Poincaré, Probab. Stat. **50** (2014) 920–945.
[4] A. Esposito, F.S. Patacchini, A. Schlichting, D. Slepčev, *Nonlocal-interaction equation on graphs: gradient flow structure and continuum limit*, arXiv: 1912.09834.
[5] M.A. Peletier, R. Rossi, G. Savaré, O. Tse, *Jump processes as Generalized Gradient Flows*, arXiv:2006.10624.
[6] D.L. Scharfetter, H.K. Gummel, *Large-signal analysis of a silicon Read diode oscillator*, IEEE Trans. Electron Devices. 16 (1969) 64–77.
[7] A. Schlichting, C. Seis, *The Scharfetter–Gummel scheme for aggregation-diffusion equations*, arXiv:2004.13981.

Mean-field models for segregation dynamics

Marie-Therese Wolfram
(joint work with Martin Burger, Jan Haskovec, Jan-Frederik Pietschmann, Helene Ranetbauer, Christian Schmeiser)

Aggregation and segregation dynamics are often observed in the collective motion of large interacting particle systems in the life and social sciences. Hereby the formation of clusters or aggregates serve diverse functions such as reproduction or reducing the risk by predators. There has been an increased interest in understanding the underlying mechanisms of aggregation and segregation dynamics in the last decades. There is an extensive literature on mathematical models for collective dynamics on the micro- as well as macroscopic level. In all these models the qualitative behaviour of solutions - such as equilibration and steady states - is of particular interest.

Aggregation dynamics can be linked to different underlying individual dynamics. The first popular class of models relate to *self-organisation* in large animal groups, like fish schools or bird flocks. Hereby individuals align characteristic features, such as their velocity. The respective continuum models for the individual density $\rho = \rho(x,t)$ are of the form

$$\partial_t \rho(x,t) = \nabla \cdot (\rho \nabla K * \rho),$$

where K is a radially symmetric pairwise interaction kernel. Attraction towards a chemical gradient (produced by the individuals themselves) is another popular aggregation mechanism. The most prominent example is the Patlak-Keller-Segel model

$$\partial_t \rho(x,t) = \nabla \cdot (\nabla \rho^m + \rho \nabla (\rho * V)),$$

where $m \geq 1$ and V is a given attractive potential. In *density dependent random walks* individuals decrease the amplitude of their random motion if others are close. The respective continuum models are generally of the form

$$\partial_t \rho(x,t) = \nabla \cdot (\Phi(\rho) \nabla \rho),$$

where the diffusion coefficient might be degenerate, that is $\Phi(\rho) \geq 0$.

In the following we will discuss two different microscopic models, which describe segregation dynamics and their respective mean-field limits. Both approaches will lead to parabolic degenerate diffusion equations or systems thereof.

In the first approach, see [1], individual dynamics are based on first or second order models for the position $X_i^t \in \mathbb{R}^d$ and the velocity $V_i^t \in \mathbb{R}^d$ of the i-th individual, $i = 1, \dots N$. We assume that individuals undergo a density dependent random walk, in which the amplitude of the random walk depends on the sensed average density

$$\eta_i^t = \frac{1}{N} \sum_{i=1}^{N} W(X - X_i^t),$$

where $W(x) = w(|x|)$ with $w : \mathbb{R}^+ \to \mathbb{R}^+$ being a bounded, non-negative and non-increasing weight function (such as the indicator function on the interval $[0, R]$, $R > 0$). Then the first order dynamics are given by

$$dX_i^t = G(\eta_i^t) dB_i^t, \qquad i = 1, \dots N, \tag{1}$$

where B_i^t are independent d-dimensional Brownian motions and $G : \mathbb{R}^+ \to \mathbb{R}^+$ is a bounded, non-negative, non-increasing function. Possible choices are $G(s) = \exp(-s)$ or $G(s) \equiv 0$ for $s \geq s_0 > 0$.

In case of second order dynamics the individually sensed density also depends on the direction of motion, that is

$$\eta_i^t = \frac{1}{N} \sum_{i=0}^{N} W(X - X_i^t, V_i^t),$$

where $W(x, v) = w(|x|, \frac{x \cdot v}{|x||v|})$. Then

$$dX_i^t = V_i^t dt, \tag{2a}$$

$$dV_i^t = -H(\eta_i^t) V_i^t dt + G(\eta_i^t) dB_i^t, \tag{2b}$$

where G is of the same form as in (1) and $H : \mathbb{R}^+ \to \mathbb{R}^+$ is a non-negative and non-decreasing damping term. In the large particle In [3] we derive the respective mean field models. In case of first order dynamics, we obtain the following continuum equation for the individual density $\rho = \rho(x, t)$:

$$\partial_t \rho(x, t) = \Delta_x (G(W * \rho)^2 \rho), \tag{3}$$

where $W * \rho = \int_{\mathbb{R}^d} W(x - y) \rho(y, t) dy$. For the second order dynamics we derive a Fokker-Planck equation for the particle density with respect to their position x and velocity v:

$$\partial_t f(x, v, t) + v \cdot \nabla_x f(x, v, t) = \nabla_v \cdot \left(H(W * f) v f + \frac{1}{2} \nabla_v (G(W * f)^2 f) \right) \tag{4}$$

where we define the convolution operator as $W * f = \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} W(x - y, v) f(y, v, t) dy dv$.

Stationary states and linear stability analysis: Equation (3) and (4) exhibit complex stationary states, which can be analysed by linear stability analysis. To do so we consider (3) on a bounded domain with periodic boundary conditions.

Then $\rho \equiv const$ is a trivial stationary state, but more complex stationary states can be easily constructed. If $G(s) = 0$ for $s > s_0$ then one can construct atomic measures

$$\rho(x) = \sum_{i=1}^{N} c_i \delta(x - x_i),$$

for some $n \in \mathbb{N}$, $x_i \in \Omega$ and c_i such that the Delta Diracs don't see each other (due to the cut off for values larger than s_0). Plain vanilla linear stability analysis around the constant steady state shows that certain wavenumbers are unstable, leading to the formation of aggregates and clusters. Furthermore if the diffusion depends on the local density only, that is $W = \delta_0$, then equation (3) can be written as

$$\partial_t \rho(x,t) = \frac{1}{2} \nabla \cdot \left[G(\rho) \left(2G'(\rho)\rho + G(\rho) \right) \nabla \rho \right]. \tag{5}$$

Note that equation (5) is parabolic (and therefore well-posed) if $2G(\rho)G'(\rho)\rho + G^2(\rho)$ is strictly positive (which is the same condition that can be found in the linear stability analysis). The non-locality does however introduce a stabilising effect as a formal expansion of W shows. Let $W_\epsilon := \epsilon^{-d} W(\frac{s}{\epsilon})$ such that $W_\epsilon \to \delta_0$ as $\epsilon \to 0$. Then this formal expansion yields

$$\partial_t \rho = \frac{1}{2} \Delta \left[G(\rho)^2 \rho + \epsilon^2 \beta G(\rho) G'(\rho) \rho \Delta \rho \right] + \mathcal{O}(\epsilon^4).$$

This yields a Cahn Hilliard type equation, which is well-posed for every $\epsilon > 0$. However, the regularising effect of the non-locality is lost for $\epsilon = 0$.

In the second model, see [2], we focus on segregation dynamics for two species, called red and blue in the following. Let $r = r(x,t)$ and $b = b(x,t)$ denote the probability to find a red or blue particle at location x at time t. Then the individual dynamics are given by

$$\partial_t c(x,t) = \int_{\mathbb{R}^d} K_1(x - x') \left[(1-\rho) D_c' c' - (1-\rho') D_c c \right] dx', \tag{6}$$

where $c = r, b$, $\rho = r + b$ is the total density (normalised to one) and the dash indicates the evaluation at x'. The diffusion coefficient are of the form

$$D_c(x,t) = D_c((K_2 * r(\cdot,t))(x), (K_2 * b(\cdot,t))(x)),$$

with convolution $(K_2 * u(\cdot,t))(x) = \int_{\mathbb{R}^d} K_2(x - x') u(x',t) dx'$. The factor $(1-\rho)$ in (6) includes finite size effects - particles can only move to a site x if it is not already occupied. The kernel K_1 corresponds to the jump rate, while K_2 to the sensing area of the density. The first term on the right hand of (6) is a gain term since individuals move from x' into x, while the second is a loss term as individuals move from x to x'.

By scaling the kernels K_1 and K_2 appropriately we can either localise one or the other. If individuals only move in their local neighbourhood (which corresponds

to re-scaling K_1), we obtain the following partial differential equation system with cross diffusion terms

$$\partial_t r(x,t) = \nabla \cdot [(1-\rho)\nabla D_r(K_2 * r, K_2 * b)r) + rD_r(K_2 * r, K2 * b)\nabla\rho] \tag{7a}$$

$$\partial_t b(x,t) = \nabla \cdot [(1-\rho)\nabla D_b(K_2 * r, K_2 * b)b) + bD_b(K_2 * r, K_2 * b)\nabla\rho]\,. \tag{7b}$$

If we localise the sensing area (hence re-scaling K_2) we obtain an integro-differential equation instead

$$\partial_t r(x,t) = \int_{\mathbb{R}^d} K_1(x-x')[(1-\rho)D_r' r' - (1-\rho')D_r r]dx' \tag{8a}$$

$$\partial_t b(x,t) = \int_{\mathbb{R}^d} K_1(x-x')\left[(1-\rho)D_b' b' - (1-\rho')D_b b\right]dx'. \tag{8b}$$

Since we assume that individuals like to stick to their own group, the diffusion coefficients are non-increasing functions with respect to the own species. In particular we assume that they take the form

$$D_r(p,q) = C_r e^{-C_{rr}p + C_{rb}q} \text{ and } D_b(p,q) = C_b e^{C_{br}p - C_{bb}q}, \tag{9}$$

with positive constants $C_{ij},\ i,j \in \{r,b\}$.

Existence and uniqueness of solutions to (7) follows from the boundedness by entropy principle and from Picard Lindelöf theorem for (8).

Stationary states and linear stability: The PDE as well as the integro-differential equation system exhibit the trivial stationary state $\rho \equiv const$ on the bounded domain with periodic boundary conditions. The linear stability analysis for both systems shows that there exist unstable modes, which lead to the formation of clusters and segregated states. The inherent separation property of the PDE system (3) can be seen from its Cahn-Hilliard structure. For diffusion coefficients of the form (9) we can rewrite system as (7) as

$$\begin{aligned}\partial_t r(x,t) &= \nabla \cdot [D_r k r(1-\rho)\nabla\left(-C_{rr}\Delta_K r + C_{rb}\Delta_K b + \partial_r W\right)]\\ \partial_t b(x,t) &= \nabla \cdot [D_b k b(1-\rho)\nabla\left(C_{br}\Delta_K r - C_{bb}\Delta_K b + \partial_b W\right)]\,,\end{aligned}$$

where $k = \int_{\mathbb{R}^d} K_2(x)dx$ and the non-local Laplacian is defined via

$$-\Delta_k u = u - \frac{1}{k}K * u.$$

The potential W is given by

$$\begin{aligned}W(r,b) = {} & r\log r + b\log b + (1-\rho)\log(1-\rho)\\ & - \frac{C_{rr}}{2}r^2 + C_{rb}rb + C_{br}rb - \frac{C_{bb}}{2}b^2 + \frac{C_{rr}}{2}r + \frac{C_{bb}}{2}b.\end{aligned}$$

If we ignore the diffusion terms the potential has three global minimisers (r,b) at the corners of the unit triangle, that is $(1,0)$, $(0,1)$ and $(0,0)$. The first two correspond to segregated states, the last one to void. These energy minimisers play a main role in the long time asymptotics. In case of small diffusion the minimisers are shifted to the interior and the segregation will be less pronounced. We refer to [3] for a more detailed discussion in case of similar model.

These first insights into the qualitative behaviour of solutions to nonlinear degenerate diffusion models for aggregation and segregation dynamics show that already very minimalistic and basic interaction rules lead to the formation of complex stationary states.

References

[1] M. Burger, J. Haskovec, M.-T. Wolfram, *Individual based and mean-field modeling of direct aggregation*, Physica D **260** (2013).
[2] M. Burger, J.-F. Pietschmann, H. Ranetbauer, C. Schmeiser, M.T. Wolfram, *Mean field models for segregation dynamics*, accepted for publication at EJAM (2020).
[3] J. Berendsen, M. Burger, J.-F. Pietschmann, *On a cross-diffusion model for multiple species with nonlocal interaction and size exclusion* Nonlinear Analysis 159 (2017)

Survival of the waiting time phenomenon under Lagrangian discretization

Daniel Matthes
(joint work with Julian Fischer)

For definiteness, we consider initial value problems for the one-dimensional porous medium equation (PME)

$$\partial_t u = (u^m)_{xx}$$

with exponent $m > 1$ and the thin film equation (TFE)

$$\partial_t u = -(u u_{xxx})_x,$$

on the real line, with a compactly supported initial datum $\bar{u} \geq 0$. For the TFE, we consider strong solutions, see e.g. [2], so in particular the contact angle is zero. Without loss of generality, we further assume that the (conserved) mass is $\int_{\mathbb{R}} u(t,x)\,dx = 1$ and that the initial support is $\operatorname{supp} \bar{u} = [0,b]$ for some $b > 0$.

For both PME and TFE, it is well-known that *for large times*, the size $R(t)$ of the support of the solution expands at least at an algebraic rate: $R(t) > t^{1/(m+1)}$ for the PME, and $R(t) > t^{1/5}$ for the TFE. This follows by comparison with self-similar solutions [2]. What is probably less known is that *for short times*, the waiting time phenomenon may occur: if $\bar{u}(x)$ is "very flat" for small positive x, then the left edge of support remains at $x = 0$ for a certain waiting time $T_* > 0$. More specifically, if $\bar{u}(x) \leq K x_+^p$ for some $p > \frac{2}{m-1}$ in case of the PME, or $p > 4$ in case of the TFE, then the solution satisfies $\inf \operatorname{supp} u(t) = 0$ for all $0 \leq t \leq T_*$. These criteria are essentially optimal, see e.g. [1, 4].

Intuitively, waiting times account for a slow transfer of energy from the bulk through the thin tail to the interface. The by-now-standard method for proving the occurence of the phenomenon rigorously and also for estimation of waiting times from below builds on this intuition by considering localized entropy estimates: the core idea is to compare the values and time derivatives of the functionals

$$H_h(t) = \int_{\mathbb{R}} (h-x)_+^q\, u(t,x)^p\,dx \tag{1}$$

with suitable powers $p, q > 0$ and varying penetration depth $h \geq 0$. A Stampacchia estimate then allows to conclude from flatness of $\bar{u}(x)$ near $x = 0$ that $H_0(t) = 0$ for $t \in [0, T_*]$, which proves the one-sided waiting time phenomenon, i.e., that there is no expansion of support before time T_*. This method has been developed in [3] in the context of more general thin film equations, unifying and simplifying various earlier approaches.

Our own contribution is two-fold. First, we develop a similar method in the Lagrangian picture of the evolution; it is even slightly stronger since it directly proves the double sided waiting time phenomenon. And second, we identify a mesh-independent discrete waiting time phenomenon for the Lagrangian discretizations of PME and TFE that had been proposed in [6, 7].

The Lagrangian picture of PME and TFE is obtained by passing from the spatial coordinate $x \in \mathbb{R}$ to the mass coordinate $\xi \in [0, 1]$. The conversion is given via the time-dependent Lagrangian map $X(t; \cdot) : [0, 1] \to \mathbb{R}$, characterized by

$$\int_{-\infty}^{X(t;\xi)} u(t; x')\, dx' = \xi \quad \text{for all } 0 < \xi < 1.$$

That is, $t \mapsto X(t; \xi)$ is the trajectory of the mass particle that has been tagged by ξ initially. In particular, $t \mapsto X(t; 0)$ is the curve of the left edge of support. In terms of X and $Z := u \circ X = \frac{1}{\partial_\xi X}$, the PME and TFE, respectively, become

$$\partial_t X = -\partial_\xi (Z^m), \quad \text{and} \quad \partial_t X = \partial_\xi \big(2Z^3 \partial_{\xi\xi} Z + Z^2 (\partial_\xi Z)^2\big). \tag{2}$$

Theorem 1. *Let X be the Lagrangian map for the solution to the PME with an initial datum $\bar{u}$ such that*

$$\bar{B} := \sup_{h>0} \frac{\int_0^h \bar{u}(x)\, dx}{\left(\int_0^h \bar{u}(x)\right)^{\frac{3m-1}{m+1}}} < \infty. \tag{3}$$

Then $X(t, 0) = 0$ at least for all $t \leq T$, where

$$T := K \bar{B}^{-\frac{m+1}{m-1}} \tag{4}$$

with a universal constant K.

An analogous, technically more complicated result holds for the TFE [5].

Idea of proof: Lacking a notion of "outside of u's support" in the Lagrangian picture, we cannot directly use the functionals from (1) with a fixed localization in x-space. Instead, we perform a localization — moving in time with the Lagrangian map — from inside u's support. Specifically, we consider for $\rho \in [0, 1]$:

$$H_\rho(t) = \int_0^\rho Z(t)^{m-1}\, d\xi, \quad D_\rho(t) = \int_0^\rho \big(\partial_\xi \big[Z(t)^m\big]\big)^2\, d\xi, \quad G_\rho(t) = \int_0^\rho Z(t)^{2m}\, d\xi,$$

and obtain (up to irrelevant constants) the localized entropy dissipation estimate

$$\sup_{0\leq t\leq T} H_\rho(t) + \int_0^T D_\rho(t)\, dt \leq H_{2\rho}(0) + \rho^{-2} \int_0^T G_{2\rho}(t)\, dt. \tag{5}$$

By means of a functional inequality as simple as $G_\rho \le \rho^{\frac{m+1}{2m}} D_\rho^{\frac{m+1}{2m}} H_\rho$, one arrives at the following family of estimates, parametrized by ρ:

$$A_\rho(T) \le T^{\frac{m-1}{2m}} \big(A_\rho(T) + \bar{B}\big)^{1+\frac{m+1}{2m}} \quad \text{for} \quad A_\rho(T) := \rho^{-1-\frac{4m}{m+1}} \int_0^T G_\rho(t)\, dt.$$

This yields the recursive relation for $a_n := A_{2^{-n}}(T)$,

$$a_{n+1} \le c(a_n + \bar{B})^{1+\frac{m+1}{2m}}, \tag{6}$$

with $c := T^{\frac{m-1}{2m}}$. By an elementary induction argument, one concludes that the a_n are n-uniformly bounded since $\bar{B} < \infty$ and T has been chosen as in (4). From that bound, one easily deduces directly from the evolution equations that $\int_0^T \big(\partial_t X(t;0)\big)^2\, dt = 0$, which is the claim.

For discretization of X, we use finite differences in mass space. u is then approximated by piecewise constant functions with respect to a moving mesh on $\mathbb{R}$. The discrete evolution equations are formally identical to (2), with derivatives replaced by finite differences. As shown in [6, 7], these discretizations are convergent. And moreover, they inherit the dissipative (in fact: even the gradient flow) structure of the original PME and TFE. Thus, for appropriately defined discretized versions of H_ρ, D_ρ and G_ρ, one proves a dissipation estimate in full analogy to (5). The main difference is that the cascade of estimates (6) is now finite, since there is a left-most cell of finite mass $\delta_1 > 0$.

Theorem 2. *Let $(x_0(t), x_1(t), \dots, x_M(t))$ be the moving Lagrangian mesh corresponding to a solution of the discretized PME with a $\bar{u}$ satisfying* (3)*. Then $|x_0(t)| \le c\delta_1^{\frac{m-1}{m+1}} \bar{B}^{-\frac{2}{m-1}}$ at least for all $t \le T$, with the $T > 0$ from* (4)*.*

Again, we derive a similar result for the discretized TFE [5].

Open problem: In our numerical simulations, the waiting time phenomenon ist captured atonishingly well, particularly for adapted meshes that become finer and finer closer to the interfaces. For a mesh-independent waiting time, the interface appears to be not moving at all, and then changes to positive speed quite abruptly. The deviation of $x_0(t)$ from zero during the inital phase appears to be much smaller than predicted by the theorem above. We would like to understand if our approach can be modified to derive sharper estimates and to capture the quality of the numerical observation.

References

[1] D.G. Aronson, L.A. Caffarelli, S. Kamin, *How an initially stationary interface begins to move in porous medium flow*, SIAM J. Math. Anal. **14** (1983), 639–658.

[2] J.A. Carrillo, G. Toscani, *Long-time asymptotics for strong solutions of the thin film equation*, Comm. Math. Phys. **225** (2002) 551–571.

[3] R. Dal Passo, L. Giacomelli, G. Grün, *A waiting time phenomenon for thin film equations*, Ann. Scuola Norm. Sup. Pisa Cl. Sci. **30** (2001), 437–463.

[4] J. Fischer, *Behaviour of free boundaries in thin-film flow: the regime of strong slippage and the regime of very weak slippage*, Ann. Inst. H. Poincaré Anal. Non Linéaire **33** (2016), 1301–1327.

[5] J. Fischer, D. Matthes, *The waiting time phenomenon in spatially discretized porous medium and thin film equations*, arXiv:1911.04185 (2019).
[6] D. Matthes, H. Osberger, *Convergence of a variational Lagrangian scheme for a nonlinear drift diffusion equation*, ESAIM Math. Model. Numer. Anal. **48** (2014), 697–726.
[7] H. Osberger, D. Matthes, *Convergence of a fully discrete variational scheme for a thin-film equation*, Radon Ser. Comput. Appl. Math. **17**, De Gruyter, Berlin (2017), 356–399.

Optimal control of rate-independent systems

Dorothee Knees
(joint work with Stephanie Thomas)

In this short note, we focus on the optimal control of a *rate-independent system*, for details we refer to [2]. The system is given in terms of a *state variable* $z : [0,T] \to \mathcal{Z}$, a time-dependent *external load* ℓ, a *stored energy functional* $\mathcal{E}$ depending on ℓ and z, and a *dissipation potential* $\mathcal{R} : \mathcal{Z} \to [0,\infty)$, which captures the dissipation due to internal friction. To be more precise, we assume that the state space $\mathcal{Z}$ is a separable Hilbert space which fulfills the embedding $\mathcal{Z} \Subset \mathcal{V} \subset \mathcal{X}$ for another separable Hilbert space $\mathcal{V}$ and a Banach space $\mathcal{X}$ and choose $\ell \in H^1(0,T;\mathcal{V}^*)$. We are working with a *semilinear model*, i.e., we assume that there are a linear operator $A \in \mathrm{Lin}(\mathcal{Z},\mathcal{Z}^*)$ and a nonlinearity $\mathcal{F} : \mathcal{Z} \to [0,\infty)$ such that $\mathcal{E} : [0,T] \times \mathcal{Z} \to \mathbb{R}$ is given by

$$\mathcal{E}(t,z) := \tfrac{1}{2}\langle Az, z\rangle_{\mathcal{Z}^*,\mathcal{Z}} + \mathcal{F}(z) - \langle \ell(t), z\rangle_{\mathcal{V}^*,\mathcal{V}} = \mathcal{I}(z) - \langle \ell(t), z\rangle_{\mathcal{V}^*,\mathcal{V}}.$$

$\mathcal{F}$ is supposed to be *nonconvex* and of lower order with respect to A. The precise assumptions on A, $\mathcal{F}$ and ℓ can be found in [2]. In order to obtain rate-independence, the dissipation potential $\mathcal{R}$ is assumed not only to be continuous and convex, but also *positively homogeneous of degree one*. We are dealing with a *bounded dissipation potential*, i.e. there are constants $c, C > 0$ such that

$$\text{for all } z \in \mathcal{X}: \quad c\|z\|_{\mathcal{X}} \le \mathcal{R}(z) \le C\|z\|_{\mathcal{X}}. \tag{1}$$

With these ingredients, the evolution of the state variable z can be described by means of the *doubly nonlinear equation*

$$0 \in \partial\mathcal{R}(\dot z(t)) + \mathrm{D}_z\mathcal{E}(t, z(t)) \quad \text{for a.a. } t \in [0,T], \tag{2}$$

where $\mathrm{D}_z\mathcal{E}$ is the Gâteaux derivative of $\mathcal{E}$ w.r.t. z and $\partial\mathcal{R} : \mathcal{Z} \to \mathcal{P}(\mathcal{Z}^*)$ denotes the convex subdifferential of $\mathcal{R}$. The main result in [2] is the existence of a globally optimal solution of an optimal control problem of the type

$$\left.\begin{array}{ll} \min & \|\hat z - z_{\mathrm{des}}\| + \alpha\|\ell\|_{H^1(0,T;\mathcal{V}^*)} \\ \text{s.t.} & \hat z \in M_{ad}, \end{array}\right\} \tag{3}$$

where the external load ℓ is the control variable, $\alpha > 0$ is a fixed Tikhonov parameter, and z_{des} is a given desired state. The problem is restrained to an admissible set M_{ad} consisting of all solutions of (2) in the sense of *parametrized BV solutions*.

It is well known that rate-independent systems with nonconvex energy $\mathcal{E}$ in general do not admit solutions that are continuous in time. Several solution concepts are available in the literature that allow for discontinuous solutions and we refer

to [5] for more details and an overview on solution concepts. Independently of the chosen solution concept, solutions of rate-independent systems with nonconvex energies in general are not unique. This is a major challenge when it comes to optimal control of such systems.

The literature concerning the optimal control of rate-independent systems with nonconvex energies formulated on infinite dimensional spaces is rather scant. We mention here [6, 7], where the existence of optimal solutions to a variant of the problem (3) constrained to global energetic solutions and a reverse approximation property are shown. To the best of our knowledge, no existence results are available in the literature for optimal control problems constrained to BV solutions.

The solution concept in this short note are normalized, p-parametrized balanced viscosity solutions (BV solutions). Thereby, the solutions are represented with respect to an artificial arc length parameter as in [1] or in [4, Definition 4.2]. The existence of BV solutions can be shown via a vanishing viscosity approach, where the equation (2) is approximated by a sequence of equations

$$0 \in \partial\mathcal{R}(\dot{z}_\varepsilon(t)) + \varepsilon\mathbb{V}\dot{z}_\varepsilon(t) + \mathrm{D}_z\mathcal{E}(t, z_\varepsilon(t)) \quad \text{for a.a. } t \in [0,T], \tag{4}$$

where $\mathbb{V} \in \mathrm{Lin}(\mathcal{V}, \mathcal{V}^*)$ is an elliptic and symmetric operator. These types of *viscous* systems have been analyzed in the past (see, e.g., [3]) and are known to have absolutely continuous solutions $z_\varepsilon \in W^{1,1}(0,T;\mathcal{V})$. In order to identify the limit as the viscosity ε tends to zero, one option is to reformulate the viscous system with respect to an artificial arc length parameter so that the trajectory $t \mapsto (t, z_\varepsilon(t))$ is rewritten as $s \mapsto (\hat{t}_\varepsilon(s), \hat{z}_\varepsilon(s))$. There are several possibilities for choosing the reparametrization. For our purpose, the reparametrization based on the vanishing viscosity contact potential $\mathrm{p}(\cdot,\cdot)$ is the most appropriate one, see [4]. Here, one defines

$$s_\varepsilon(t) := t + \int_0^t \mathrm{p}(\dot{z}_\varepsilon(\tau), -\mathrm{D}\mathcal{E}(\tau, z_\varepsilon(\tau)))\mathrm{d}\tau \tag{5}$$

with $\mathrm{p}(v,\xi) := \mathcal{R}(v) + \|v\|_{\mathbb{V}}\,\mathrm{dist}_{\mathbb{V}}(\xi, \partial\mathcal{R}(0))$ and chooses $\hat{t}_\varepsilon$ as the inverse function of s_ε. Defining $\hat{z}_\varepsilon = z_\varepsilon \circ \hat{t}_\varepsilon : [0, S_\varepsilon] \to \mathcal{Z}$, it is then possible to pass to the limit for vanishing viscosity (i.e., for $\varepsilon \to 0$) and obtain limits $S \in [0,\infty)$ of S_ε, $\hat{z} \in \mathrm{AC}(0,S;\mathcal{X})$ of $\hat{z}_\varepsilon$ and $\hat{t} \in W^{1,\infty}(0,S;\mathbb{R})$ of $\hat{t}_\varepsilon$. Simultaneously passing to the limit in the reparametrized energy dissipation balance associated with (4), one also obtains the energy dissipation balance fulfilled by $(\hat{t}, \hat{z})$, which reads

$$\begin{aligned}\mathcal{E}(\hat{t}(s), \hat{z}(s)) + \int_0^s \mathcal{R}[\hat{z}'](r)\mathrm{d}r + \int_{(0,s)\cap G} \|\hat{z}'(r)\|_{\mathbb{V}}\mathrm{dist}_{\mathbb{V}}(-\mathrm{D}\mathcal{E}(\hat{t}(r), \hat{z}(s)), \partial\mathcal{R}(0))\mathrm{d}r \\ = \mathcal{E}(0, z_0) - \int_0^s \langle \ell'(\hat{t}(r))\hat{t}'(r), \hat{z}(r)\rangle \mathrm{d}r\,.\end{aligned}$$

It is possible to show that $\hat{z} \in \mathrm{AC}_{\mathrm{loc}}(G;\mathcal{V})$ is differentiable almost everywhere on

$$G = \{s \in [0,S] \,|\, \mathrm{dist}_{\mathbb{V}}(-\mathrm{D}\mathcal{E}(\hat{t}(s), \hat{z}(s)), \partial\mathcal{R}(0)) > 0\},$$

so that the second integrand is defined almost everywhere, [4]. Normalized, p-parametrized BV solutions then are defined as triples $(S, \hat{t}, \hat{z})$ with certain regularities that satisfy the energy dissipation identity and that are normalized in the sense of [2, Definition 3.1]. One advantage of using the parametrization (5) is that limits of solutions $(\hat{t}_\varepsilon, \hat{z}_\varepsilon)_\varepsilon$ are automatically normalized, a property that is essential in the analysis for the optimal control problem.

As already mentioned, BV solutions typically are not unique. Hence, for the purpose of optimal control one needs to show the sequential closedness of the graph of the set-valued solution operator and a compactness property, [2, Theorem 3.12]. For the proof, the main challenge is to derive a priori estimates for the driving forces $\mathrm{D}\mathcal{E}(\hat{t}, \hat{z})$ on the set G. In order to obtain these, we use the fact that for each parametrized BV solution, there exists a Lagrange parameter $\lambda : (0, S) \to [0, \infty)$ with $\lambda(s) = 0$ on $(0, S) \setminus G$ such that the inclusion

$$0 \in \partial\mathcal{R}(\hat{z}'(s)) + \lambda(s)\mathbb{V}\hat{z}'(s) + \mathrm{D}\mathcal{I}(\hat{z}(s)) - \ell(\hat{t}(s)) \tag{6}$$

is fulfilled almost everywhere on G, [4]. For each connected component of G, one then chooses a reparametrization in such a way that the transformed functions are solutions of the autonomous system

$$0 \in \partial\mathcal{R}(\dot{z}(t)) + \mathbb{V}\dot{z}(t) + \mathrm{D}\mathcal{I}(z(t)) - \ell_* \quad \text{for } t > 0\,. \tag{7}$$

Analyzing this system yields the desired estimates and in particular guarantees the compactness of solution sets, see [2, Theorem 3.12]. With this, the following existence result for the *optimal control problem governed by* (2), which is constrained to the admissible set

$$M_{ad} := \big\{(S, \hat{t}, \hat{z}, \ell) \,|\, (S, \hat{t}, \hat{z}) \text{ is a parametrized BV solution for } (z_0, \ell)\big\} \tag{8}$$

is immediate ([2, Theorem 4.1]:

Theorem. Let $\alpha > 0$ be a fixed Tikhonov parameter, $z_0 \in \mathcal{Z}$ be chosen such that there exists $\ell \in H^1(0, T; \mathcal{V}^*)$ such that $\mathrm{D}_z\mathcal{E}(0, z_0) \in \mathcal{V}^*$ and let $j : \mathcal{V} \to \mathbb{R}$ be bounded from below and continuous. Then, the optimal control problem (3) has a globally optimal solution.

References

[1] Messoud A. Efendiev and Alexander Mielke, *On the rate-independent limit of systems with dry friction and small viscosity.*, J. Convex Anal. **13** (2006), no. 1, 151–167.

[2] Dorothee Knees and Stephanie Thomas, *Optimal control for a rate-independent systems constrained to parametrized balanced viscosity solutions*, Tech. report, University of Kassel, 2018, http://arxiv.org/abs/1810.12572.

[3] A. Mielke, R. Rossi, and G. Savaré, *BV solutions and viscosity approximations of rate-independent systems*, ESAIM Control Optim. Calc. Var. **18** (2012), no. 1, 36–80.

[4] Alexander Mielke, Riccarda Rossi, and Giuseppe Savaré, *Balanced viscosity (BV) solutions to infinite-dimensional rate-independent systems.*, J. Eur. Math. Soc. (JEMS) **18** (2016), no. 9, 2107–2165 (English).

[5] Alexander Mielke and Tomáš Roubíček. *Rate-independent systems. Theory and application.* New York, NY: Springer, 2015.

[6] Filip Rindler, *Optimal control for nonconvex rate-independent evolution processes*, SIAM J. Control Optim. **47** (2008), no. 6, 2773–2794. MR 2466091
[7] Filip Rindler, *Approximation of rate-independent optimal control problems*, SIAM J. Numer. Anal. **47** (2009), no. 5, 3884–3909. MR 2576524

On the sticky particle solutions to the multi-dimensional pressureless Euler system

Sara Daneri
(joint work with Stefano Bianchini)

We consider the pressureless Euler system in $[0,T]\times\mathbb{R}^d$

$$\begin{cases}\partial_t\rho+\operatorname{div}(\rho v)=0\\ \partial_t(\rho v)+\operatorname{div}(\rho v\otimes v)=0,\end{cases}\tag{1}$$

where ρ is the distribution of particles and v is their velocity.

Such a model has been proposed by Zeldovich [9] as a simplified model for the early stages of the formation of galaxies, when a dust of particles moving without pressure should start to collide and aggregate into bigger and bigger clusters.

Since then, several authors devoted attention to the search of sticky particle solutions, namely solutions to (1) which satisfy the following adhesion principle: if two particles of fluid do not interact, then they move freely keeping constant velocity, otherwise they join with velocity given by the balance of momentum.

The great majority of the results in the literature are concerned with the one-dimensional pressureless dynamics (see e.g. [2, 3, 5, 6, 7, 8]) In this case, exploiting the density of finite particle solutions, one can obtain from quite general initial data a global measure solution of (1) satisfying an adhesion principle.

In general dimension, much less is known. For initial data given by a finite number of particle pointing each in a given direction, it is easy to show that a global sticky particle solution always exists and is unique. However, in dimension $d\geq 2$, one sees immediately already from a finite number of particles that the sticky particle solutions do not depend continuously on the initial data.

In [4] it is shown that, in general, both existence and uniqueness might fail: it is indeed possible to build initial data of non-existence or non-uniqueness for the sticky particle solutions. In particular, in dimension $d\geq 2$ one cannot hope for a well-posedness of the Cauchy problem in the set of sticky particle solutions for all measure-type initial data as in the one-dimensional case.

Thus the natural question of whether one can still find particle solutions for a large class of data (hence excluding the counterexamples in [4]) remained unanswered. In [1] we give a positive answer to this question.

In order to state our main result, define

$$\mathcal{P}_{2,1}(\mathbb{R}^d\times\mathbb{R}^d):=\Big\{\nu_0\in\mathcal{P}(\mathbb{R}^d\times\mathbb{R}^d):\int|x|^2\mathtt{p}_{x\#}\nu_0\leq 1,\int|v|^2\mathtt{p}_{v\#}\nu_0\leq 1\Big\},$$

where $\mathcal{P}(\mathbb{R}^d \times \mathbb{R}^d)$ are the probability measures on $\mathbb{R}^d \times \mathbb{R}^d$, $(x, v) \in \mathbb{R}^d \times \mathbb{R}^d$ the position-velocity coordinates and $\mathtt{p}_x$ ($\mathtt{p}_v$) the projection operators on the first (last) d coordinates. Moreover, we consider the problem of existence and uniqueness in a larger class of solutions which we call *dissipative* since in particular their kinetic energy is decreasing but their trajectories might cross without joining at later times. By *free flow* we mean a flow in which trajectories are disjoint straight lines which never intersect.

Our main result is the following: there is a set $D_0 \subset \mathcal{P}_{2,1}(\mathbb{R}^d \times \mathbb{R}^d)$ such that, for any $\nu_0 \in D_0$ there exists a unique dissipative solution η with initial data ν_0 and it is given by a free flow. Such a set is a dense G_δ set (i.e. of second category) in the weak topology on $\mathcal{P}_{2,1}(\mathbb{R}^d \times \mathbb{R}^d)$.

Since our notion of dissipative solution includes the classical sticky particle solutions, the above theorem implies that, even though the sticky particle solutions are not well-posed for every measure-type initial data, there exists a comeager set of initial data in the weak topology giving rise to a unique sticky particle solution. Moreover, for any of these initial data the sticky particle solution is unique also in the larger class of dissipative solutions (where trajectories are allowed to cross) and is given by a trivial free flow concentrated on trajectories which do not intersect. In particular for such initial data there is only one dissipative solution and its dissipation is equal to zero. Thus, for a comeager set of initial data the problem of finding sticky particle solutions is well-posed, but the dynamics that one sees is trivial.

Both the concepts of dissipative and classical sticky particle solutions are defined at a Lagrangian level as measures on the space of curves with finite energy. The class of dissipative solutions turns out to be the compact weak closure of the set of classical sticky particle solutions. Then we show that dissipative solutions can be approximated in the weak topology by classical sticky particle solutions. Once one has the density of finite particle solutions one can use the fact that the dimension is greater than or equal to 2 to modify the initial data of such finite particle solutions in order to have the trivial free flow as unique solution, while staying close in the weak topology w.r.t. the initial data. Such data in particular have the property that every dissipating solution starting from them has zero dissipation. The fact that such initial data are a G_δ set follows from the compactness of the set of dissipative solutions and the upper semicontinuity of the dissipation. Hence, for a dense G_δ-set of initial data the weak solutions constructed by any reasonable approximation scheme coincide with our dissipative solutions, i.e. the free flow.

References

[1] S. Bianchini and S. Daneri, *On the sticky-particle solutions to the multi-dimensional Euler equations*, Preprint arXiv:2004.06557 (2020).

[2] Y. Brenier, W. Gangbo, G. Savaré and M. Westdickenberg, *Sticky particle dynamics with interactions*, J. Math. Pures Appl. **99** (5) (2013), 577–617.

[3] Y. Brenier and E. Grenier, *Sticky particles and scalar conservation laws*, SIAM Jour. Num. Anal. **35** (1998), 2317–2328.

[4] A. Bressan and T. Nguyen, *Non-existence and non-uniqueness for multidimensional sticky particle systems*, Kinetic and related models **7** (2) (2014), 205–218.
[5] F. Cavalletti, M. Sedjro and M. Westdickenberg, *A simple proof of global existence for the 1D pressureless gas dynamics equations*, SIAM J. Math. Anal. **47** (2015), 66–79.
[6] W. E, Y.G. Rykov and Y.G. Sinai, *Generalized variational principles, global weak solutions and behaviour with random initial data for systems of conservation laws arising in adhesion particle dynamics*, Comm. Math. Phys. **177** (1996), 349–380.
[7] L. Natile and G. Savaré, *A Wasserstein approach to the one-dimensional sticky particle system*, SIAM J. Math. Anal. **41** (2009), 1340–1365.
[8] T. Nguyen and A. Tudorascu, *Pressureless Euler/Euler-Poisson system via adhesion dynamics and scalar conservation laws*, SIAM J. Math. Anal. **40** (2) (2008), 754–775.
[9] Ya. B. Zeldovich, *Gravitational instability: an approximate theory for large density perturbations*, Astron. & Astrophys.**5** (1970), 84–89.

From quantization of measures to ultrafast diffusion equations.

Mikaela Iacobelli
(joint work with Emanuele Caglioti, François Golse, Francesco Saverio Patacchini, Filippo Santambrogio)

The problem of quantization of measures consists of finding an atomic measure, with fixed number N of atoms, that best approximates in the Wasserstein distance a given density ρ in a compact subset Ω of $\mathbb{R}^d$. This problem can be proved to be equivalent to minimizing the functional

$$F_{N,r}(x_1,\dots,x_N) := \int_\Omega \Big(\min_{i=1}^N |x-x_i|\Big)^r \rho(x)\,dx.$$

A classical question concerns the asymptotic of the location of the optimal points $\{\hat{x}_i\}_{i=1}^N$ when $N\to\infty$, and it is by now well-known that $\frac{1}{N}\sum_{i=1}^N \delta_{\hat{x}_i}$ converge to a minimizer of the functional

$$f \mapsto \mathcal{F}[f] := \int_\Omega \frac{\rho}{f^{r/d}}\,dx.$$

Part 1: A gradient flow approach to quantization of measures.
In joint papers with Caglioti and Golse, we introduced a natural way to constructively find the optimal locations $\hat{x}_i$ by studying the evolution of an arbitrary configuration of points $(x_1^0,\dots,x_N^0)$ as they follow the steepest descent curves of the functional $F_{N,r}$. In other words, one solves the ODE system in $(\mathbb{R}^d)^N$ given by

$$\dot{X}(t) = -\nabla F_{N,r}(X(t)), \qquad X(0) = (x_1^0,\dots,x_N^0). \tag{1}$$

Unfortunately, this gradient flow is highly non-convex, so a strategy introduced in [1] has been to compare this evolution to a Lagrangian version of the Wasserstein gradient flow of $\mathcal{F}$. This corresponds to considering a suitable limit of (1) as $N\to\infty$, and the advantage of this approach is that this limit is much more regular. In particular, as shown in [1, 2], one can use this idea (at least in 1D, or in some specific cases in 2D) to extract information on the above ODE and prove quantitative convergence results.

Part 2: Weighted ultrafast diffusion equations.
In a sequent work with Patacchini and Santambrogio, we investigated the Wasserstein gradient flow of $\mathcal{F}$. This reads as

$$\partial_t f - \operatorname{div}\Big(f\,\nabla\Big[U'\Big(\frac{f}{m}\Big)\Big]\Big) = 0, \tag{2}$$

where $m = \rho^{\frac{1}{\alpha+1}}$, $U(s) = s^{-\alpha}$, and $\alpha = r/d$. This PDE is coupled with periodic or Neumann boundary conditions, and the function m is supposed to be bounded away from zero and infinity inside Ω.

The equation (2) is of very fast diffusion-type, and we call it weighted because of the presence of the non-constant weight factor m. The asymptotic behavior of this PDE was already studied, in the smooth case, in [3]. Then, in [4], we exploited the gradient-flow structure of this equation, in particular through the so-called JKO scheme, to introduce a suitable notion of weak solutions for which we can:
1) prove existence and uniqueness, with BV and H^1 regularity estimates;
2) prove L^1 weighted contractivity, Harnack inequalities, and exponential convergence to the steady-state.

References

[1] E. Caglioti, F.Golse, M. Iacobelli, *A gradient flow approach to quantization of measures*, Math. Models Methods Appl. Sci., **25** (2015), no. 10, 1845–1885.
[2] E. Caglioti, F.Golse, M. Iacobelli, *A gradient flow approach to quantization of measures*, Ann. Inst. H. Poincaré Anal. Non Linéaire, **35** (2018), no. 6, 1531–1555.
[3] M. Iacobelli, *Asymptotic analysis for a very fast diffusion equation arising from the 1D quantization problem.* Discrete Contin. Dyn. Syst., **39** (2019), no. 9, 4929–4943.
[4] M. Iacobelli, F. S. Patacchini, F. Santambrogio, *Weighted ultrafast diffusion equations: from well-posedness to long-time behaviour.* Arch. Ration. Mech. Anal., **232** (2019), no. 3, 1165–1206.

Jump Processes and Generalized Gradient Flows

Oliver Tse

(joint work with Mark Peletier, Giuseppe Savaré, Riccarda Rossi)

The theory for variational evolutions—evolutions driven by one or more energiesor entropies—in spaces of measures has seen tremendous growth in the last decades, of which resulted in a rich framework for classical gradient systems in general metric spaces by Ambrosio, Gigli and Savaré [2], where the Wasserstein metric of optimal transport theory plays a fundamental role; and a theory for rate-independent systems [4]. While these theories have allowed massive development of variational evolutions in a certain direction—gradient flows with homogeneous dissipation—physics and large-deviation theory suggest the study of generalized gradient flows—gradient flows with non-homogeneous dissipation—which are not covered in either theories [1, 3].

To remedy this deficiency, we have created, in [5], a functional framework for a class of generalized gradient systems in the space of nonnegative measures, which includes the Forward Kolmogorov equations for the laws of Markov jump processes on Polish spaces, that can formally be written as

$$\begin{aligned}
&(\textit{Continuity equation}) && \partial_t \rho_t + \overline{\mathrm{div}} j_t = 0,\\
&(\textit{Force-Flux relation}) && j_t = \mathrm{D}_\zeta \mathsf{R}^*(\rho, \zeta_t),\\
&(\textit{Force-Energy relation}) && \zeta_t = -\overline{\nabla}\, \mathrm{D}_\rho \mathsf{E}(\rho),
\end{aligned}$$

in terms of a *driving functional* E and a *dual dissipation potential* $\mathsf{R}^* = \mathsf{R}^*(\rho, \zeta)$ (and its associated Legendre dual R w.r.t. the ζ variable). Here, D_ζ and D_ρ denote derivatives with respect to ζ and ρ, and $(\overline{\nabla}, \overline{\mathrm{div}})$ denotes the discrete gradient-divergence pair, where $(\overline{\nabla}\varphi)(x, y) = \varphi(y) - \varphi(x)$, $x, y \in V$.

To illustrate the theory, we consider a Markov jump process $(X_t)_{t\geq 0}$ on a standard Borel space V, whose law at each $t \geq 0$ are time-dependent probability measures $t \mapsto \rho_t \in \mathcal{P}(V)$, satisfying the Kolmogorov Forward equation

$$\partial_t \rho_t = Q^* \rho_t,$$

where Q^* is the dual of the infinitesimal generator $Q : \mathrm{B}_b(V) \to \mathrm{B}_b(V)$ given by

$$(Q\varphi)(x) = \int_V [\varphi(y) - \varphi(x)]\, \kappa(x, dy),$$

for any arbitrary bounded Borel function $\varphi \in \mathrm{B}_b(V)$. The jump kernel κ in Q characterizes the process: $\kappa(x, \cdot) \in \mathcal{M}^+(V)$ (with $\mathcal{M}^+(V)$ the space of positive finite Borel measures) is the infinitesimal rate of jumps of a particle from the point $x \in V$ to Borel sets in V, and is assumed to satisfy $\sup_{x\in V} \kappa(x, V) < \infty$, and be in detailed balance with respect to an invariant measure $\pi \in \mathcal{M}^+(V)$, i.e. the joint measure $\kappa(x, dy)\pi(dx)$ is symmetric.

The functional that drives the evolution is the relative entropy with respect to the invariant measure π,

$$\mathsf{E}(\rho) = \begin{cases} \int_V \Phi(u(x))\, \pi(dx) & \text{if } \rho \ll \pi, \text{ with } u = d\rho/d\pi,\\ +\infty & \text{otherwise,} \end{cases}$$

where for the Markov jump process above, Φ is the Boltzmann–Shannon entropy $\Phi(s) = s \log s - s + 1$. The dual dissipation functional R^* is defined by

$$\begin{aligned}
\mathsf{R}^*(\rho, \zeta) &:= \frac{1}{2} \iint_{V\times V} \Psi^*(\zeta(x, y))\, \nu_\rho(dxdy), \qquad \zeta \in \mathrm{B}_b(V)\\
\nu_\rho(dxdy) &:= \alpha(u(x), u(y))\kappa(x, dy)\pi(dx), \qquad u = d\rho/d\pi,
\end{aligned}$$

for suitable functions $\Psi^* : \mathbb{R} \to [0, \infty)$ and $\alpha : [0, \infty) \times [0, \infty) \to [0, \infty)$. This class of dual dissipation potentials gives rise to a large family of Dynamical-Variational transport costs that is then used to generalize away from metric-based theories.

This new framework comprises of

(1) a definition of a notion of solutions on the basis of the *Energy-Dissipation balance*

$$\mathcal{L}(\rho, j) := \int_0^T \mathsf{R}(\rho_t, j_t) + \mathsf{R}^*(\rho_t, -\overline{\nabla}\, \mathrm{D}_\rho \mathsf{E}(\rho_t))\, dt + \mathsf{E}(\rho_T) - \mathsf{E}(\rho_0) \equiv 0,$$

for a pair (ρ, j) satisfying the continuity equation;

(2) a method to prove existence via a *chain rule inequality* for pairs (ρ, j) satisfying the continuity equation having finite action $\int_0^T \mathsf{R}(\rho_t, j_t)\, dt < \infty$, and a generalization of the *Minimizing-Movement scheme* for Dynamical-Variational Transport costs;

(3) an archetype uniqueness result based on the linear convexity of the Energy-Dissipation functional $\mathcal{L} = \mathcal{L}(\rho, j)$.

All these steps are done using only the structure that is provided directly by the driving functional E, and dissipation/dual dissipation pair $(\mathsf{R}, \mathsf{R}^*)$, which need not be homogeneous, and we do not appeal to any metric structure.

References

[1] S. Adams, N. Dirr, M. Peletier, J. Zimmer, *Large deviations and gradient flows*, Phil. Trans. R. Soc. A, **371**(2005):20120341, 2013.

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

[3] A. Mielke, M. Peletier, M. Renger, *On the relation between gradient flows and the large-deviation principle, with applications to Markov chains and diffusion*, Potential Analysis, **41**(4):1293–1327, 2014.

[4] A. Mielke, T. Roubíček. *Rate-Independent Systems: Theory and Application*, Applied Mathematical Sciences. Springer New York, 2015.

[5] M. Peletier, G. Savaré, R. Rossi, O. Tse, *Jump processes as generalized gradient flows*, arXiv:2006.10624.

Sampling inspired by gradient flows

Franca Hoffmann

(joint work with Alfredo Garbuno-Iñigo, Wuchen Li, Andrew M. Stuart)

Solving inverse problems without the use of derivatives or adjoints of the forward model is highly desirable in many applications arising in science and engineering. In this talk we propose a new version of such a methodology, a framework for its analysis, and numerical evidence of the practicality of the method proposed, see [1]. Consider the inverse problem of finding $u \in \mathbb{R}^d$ from $y \in \mathbb{R}^K$ where

$$y = \mathcal{G}(u) + \eta, \tag{1}$$

$\mathcal{G} : \mathbb{R}^d \to \mathbb{R}^K$ is a known nonlinear forward operator and η is the unknown observational noise drawn from $\eta \sim \mathsf{N}(0, \Gamma)$ for a known covariance matrix $\Gamma \in \mathbb{R}^{K \times K}$. Our objective is to find information about the truth $u^\dagger$ when the forward map $\mathcal{G}$, the covariance Γ and the data y are all viewed as given.

Using a Bayesian approach to inversion we place a prior distribution $\mathsf{N}(0,\Gamma_0)$ on the unknown u, with Lebesgue density $\pi_0(u)$, then the posterior density on $u|y$, denoted $\pi(u)$, is given by

$$\pi(u) \propto \exp\bigl(-\Phi(u)\bigr)\pi_0(u) \propto \exp\bigl(-\Phi_R(u)\bigr)\,,$$
$$\Phi_R(u) = \frac{1}{2}\|y-\mathcal{G}(u)\|_\Gamma^2 + \frac{1}{2}\|u\|_{\Gamma_0}^2\,.$$

We seek to understand the shape of the posterior distribution π. However, for computationally expensive nonlinear forward maps $\mathcal{G}$, it is a known difficult problem to sample from the target distribution π. Our starting point is an ensemble of over-damped Langevin diffusions which interact through a single preconditioner computed as the empirical ensemble covariance:

$$\dot{u}^{(j)} = -\mathsf{C}(U)\nabla\Phi_R(u^{(j)}) + \sqrt{2\,\mathsf{C}(U)}\,\dot{\mathbf{W}}^{(j)}\,, \tag{2}$$
$$\mathsf{C}(U) = \frac{1}{J}\sum_{k=1}^{J}(u^{(k)}-\bar{u})\otimes(u^{(k)}-\bar{u}) \in \mathbb{R}^{d\times d}\,, \qquad \bar{u} = \frac{1}{J}\sum_{j=1}^{J}u^{(j)}\,.$$

Here, the $\{\mathbf{W}^{(j)}\}$ are a collection of i.i.d. standard Brownian motions in the space $\mathbb{R}^d$. The empirical covariance $\mathsf{C}(U)$ acts as a preconditioner which is known to be able to accelerate the dynamics [9], a choice which is motivated by an underlying gradient flow structure. System (2) can be re-written as

$$\dot{u}^{(j)} = -\frac{1}{J}\sum_{k=1}^{J}\langle D\mathcal{G}(u^{(j)})\bigl(u^{(k)}-\bar{u}\bigr), \mathcal{G}(u^{(j)})-y\rangle_\Gamma\, u^{(k)} \;-\; \mathsf{C}(U)\Gamma_0^{-1}u^{(j)} + \sqrt{2\mathsf{C}(U)}\,\dot{\mathbf{W}}^{(j)}. \tag{3}$$

In many applications, derivatives of the forward map $\mathcal{G}$ are either not available, or extremely costly to obtain. A common technique used in ensemble Kalman methods is to approximate the gradient $\nabla\Phi_R$ by differences in order to obtain a derivative-free algorithm for inverting $\mathcal{G}$. To this end, consider the dynamical system (3) and invoke the approximation $D\mathcal{G}(u^{(j)})\bigl(u^{(k)}-\bar{u}\bigr) \approx \bigl(\mathcal{G}(u^{(k)})-\bar{\mathcal{G}}\bigr)$. This leads us to introduce the following derivative-free algorithm to generate approximate samples from the posterior distribution, which we call *ensemble Kalman sampling* (EKS):

$$\dot{u}^{(j)} = -\frac{1}{J}\sum_{k=1}^{J}\langle \mathcal{G}(u^{(k)})-\bar{\mathcal{G}}, \mathcal{G}(u^{(j)})-y\rangle_\Gamma\, u^{(k)} \;-\; \mathsf{C}(U)\Gamma_0^{-1}u^{(j)} + \sqrt{2\mathsf{C}(U)}\,\dot{\mathbf{W}}^{(j)}. \tag{4}$$

This dynamical system is similar to the noisy Ensemble Kalman Inversion (EKI), but has a different noise structure (noise in parameter space not data space) and explicitly accounts for the prior on the right hand side (rather than having it enter through initialization).

In order to write down the mean field limit of (3), we define the macroscopic mean and covariance:

$$m(\rho) := \int v\rho\, dv\,, \qquad \mathcal{C}(\rho) := \int \big(v - m(\rho)\big) \otimes \big(v - m(\rho)\big)\, \rho(v)\, dv\,.$$

Taking the large particle limit leads to the mean field equation

$$\dot{u} = -\,\mathcal{C}(\rho)\nabla\Phi_R(u) + \sqrt{2\,\mathcal{C}(\rho)}\,\dot{W}, \tag{5}$$

with corresponding nonlinear Fokker-Planck equation

$$\partial_t\rho = \nabla\cdot\big(\rho\,\mathcal{C}(\rho)\nabla\left(\Phi_R(u) + \log\rho\right)\big)\,. \tag{6}$$

The rigorous derivation of the mean-field equations (5) and (6) has recently been accomplished in [2], and an important correction term for the finite particle description was pointed out in [3]. Since the only probability densities for which $\mathcal{C}(\rho)$ vanishes are Diracs, we obtain a manifold of stationary states in addition to the equilibrium given by the posterior distribution π solving the inverse problem (1).

The nonlinear Fokker-Planck equation (6) has a generalized Wasserstein-2 gradient flow structure,

$$\partial_t\rho = \nabla\cdot\left(\rho\,\mathcal{C}(\rho)\nabla\frac{\delta E}{\delta\rho}\right)\,, \quad E(\rho) = \int\Big(\Phi_R(u) + \log\rho(u)\Big)\rho(u)\, du\,, \tag{7}$$

for the weighted Wasserstein-2 metric

$$W_{\mathcal{C}}(\rho^0,\rho^1)^2 := \inf\int_0^1\int_\Omega \langle\nabla\phi_t\,,\,\mathcal{C}(\rho_t)\nabla\phi_t\rangle\ \rho_t\, du$$
$$\text{subject to}\quad \partial_t\rho_t + \nabla\cdot(\rho_t\mathcal{C}(\rho_t)\nabla\phi_t) = 0,\ \rho_0 = \rho^0,\ \rho_1 = \rho^1,$$

where $\rho^0,\,\rho^1 \in \big\{\rho\in L^1(\Omega)\,:\,\rho>0 \text{ a.e. }\,,\int\rho(u)\,du = 1\big\}\cap C^\infty(\Omega)$. The weighted Wasserstein-2 metric above was also identified in [4, 5]. In that work a nonlocal transport equation, mapping one Gaussian into another in unit time, is studied and the transport equation is shown to have a gradient structure closely related to the one we study here for the nonlinear nonlocal Fokker-Planck equation (7). The work in [4, 5] emerges from understanding ensemble Kalman filtering from the perspective of mean field limits and may be viewed as a progenitor of our work.

Using this gradient structure, we investigate large time properties of the Fokker-Planck equation in [1], showing that its invariant measure coincides with that of a single Langevin diffusion, and demonstrating exponential convergence to the invariant measure in a number of settings. In particular, for a linear forward map $\mathcal{G}(u) = Au$, we obtain closed equations for the mean and covariance of ρ, and therefore Gaussians remain Gaussians along the flow of (7). Recently, precise convergence rates have been obtained in [7], answering the question of equilibration posed in [10] for the linear setting: If ρ_t^1, ρ_t^2 are solutions to (7) with initial conditions ρ_0^1, ρ_0^2 respectively, then under suitable conditions on $\rho_0^1, \rho_0^2, A, \Gamma_0, \Gamma$,

$$W_2(\rho_t^1,\rho_t^2) \le Ce^{-t}W_2(\rho_0^1,\rho_0^2)\,,$$

where C only depends on the first two moments of ρ_0^1, ρ_0^2, and on A, Γ_0, Γ. In addition, the authors showed that these rates are optimal.

Several interesting open questions come out of this work. On the application side, implementation and performance of the EKS are to be investigated, for example as part of an algorithmic framework such as the Calibrate-Emulate-Sample approach proposed in [8], or improvements such as the ALDI algorithm proposed in [6] satisfying affine invariance whilst avoiding the computation of the term $\sqrt{\mathsf{C}(U)}$ in front of the noise in (4) - a computational bottleneck for implementation of the EKS in higher dimensions. On the theoretical side, it is of independent interest to obtain a better understanding of the properties of such generalized Wasserstein-2 metric structures. More precisely, one could seek conditions on a general matrix $K(\rho, u, t) \in \mathbb{R}^{d\times d}$ such that the following dynamical description provides a well-defined metric structure and geodesic equations:

$$W_K(\mu,\nu)^2 := \inf \int_0^1 \int_\Omega \langle \nabla\phi_t \, , \, K(\rho_t, u, t)\nabla\phi_t \rangle \; \rho_t \, du$$
$$\text{subject to} \quad \partial_t \rho_t + \nabla \cdot (\rho_t K(\rho_t, u, t)\nabla\phi_t) = 0, \; \rho_0 = \mu, \; \rho_1 = \nu \, .$$

Some results are known for the scalar case $K(\rho) \in \mathbb{R}$ (see eg [11] and references therein), or for the constant matrix $K \in \mathbb{R}^{d\times d}$ (see eg [4] in the context of data assimilation); a more general understanding however is lacking.

References

[1] A. Garbuno-Inigo, F. Hoffmann, W. Li, A.W. Stuart, *Interacting Langevin diffusions: gradient structure and ensemble Kalman sampler*, SIAM Journal on Applied Dynamical Systems **19** (2020), 1:412–441.

[2] Z. Ding, Q. Li, *Ensemble Kalman Sampler: mean-field limit and convergence analysis*, (2019), arXiv:1910.12923[math].

[3] N. Nüsken, S. Reich, *Note on Interacting Langevin Diffusions: Gradient Structure and Ensemble Kalman Smoother by Garbuno-Inigo, Hoffmann, Li and Stuart*, (2019), arXiv:1908.10890.

[4] S. Reich,C.J. Cotter, *Ensemble filter techniques for intermittent data assimilation*, Large scale inverse problems **13** (2013), 91–134.

[5] S. Reich,C.J. Cotter, *Probabilistic forecasting and Bayesian data assimilation*, Cambridge University Press, New York (2015), x+297.

[6] A. Garbuno-Inigo, N. Nüsken, S. Reich, *Affine invariant interacting Langevin dynamics for Bayesian inference*, SIAM Journal on Applied Dynamical Systems **19** (2020), 3:1633–1658.

[7] J. A. Carrillo, U. Vaes, *Wasserstein stability estimates for covariance-preconditioned Fokker-Planck equations*, (2019), arXiv:1910.07555[math].

[8] E. Cleary, A. Garbuno-Inigo, S. Lan, T. Schneider, A. M. Stuart, *Calibrate, Emulate, Sample*, Journal of Computational Physics (2020), 109716.

[9] B. Leimkuhler, C. Matthews, J. Weare, *Ensemble preconditioning for Markov chain Monte Carlo simulation*, Statistics and Computing **28** (2018), 2:277–290.

[10] M. Herty, G. Visconti, *Kinetic Methods for Inverse Problems*, (2018), arXiv:1811.09387.

[11] J.A. Carrillo, S. Lisini, G. Savaré, D. Slepčev, *Nonlinear mobility continuity equations and generalized displacement convexity*, Journal of Functional Analysis **258** (2010), 4:1273–1309.

Capacity constrained Entropic Optimal Transport, Sinkhorn Saturated Domain Out-Summation and Vanishing Temperature

Jean-David Benamou
(joint work with Mélanie Martinet)

1. Context

Entropic Optimal Transport and Sinkhorn Algorithm, its companion numerical method is used routinely to compute the Wasserstein 2 distance between probability measures. A comprehensive description of the method and its applications can be found in [3]. It is based on a regularization of the classical Optimal Transport cost with an Entropy functional. The strength of the regularization is given by a positive parameter ϵ. Following [2] it will be called the *temperature*.

Sinkhorn Algorithm computes iteratively $(f_\epsilon^m, g_\epsilon^m) \in \mathcal{C}(X) \times \mathcal{C}(Y)$, m is the iteration index, such that in the limit the measure on the product space $X \otimes Y$

$$\gamma_\epsilon^m(x,y) = \exp(\frac{1}{\epsilon}(f_\epsilon^m(x) \oplus g_\epsilon^m(y) - c(x,y))\, \mu(x) \otimes \nu(y), \tag{1}$$

has marginals μ and ν, X and Y are two compact subsets of a metric space (M, c). In this presentation we will restrict to the torus $M = \mathbb{R}^d/\mathbb{Z}^d$ and $c(x,y) = \inf_{k\in\mathbb{Z}^d} \frac{1}{2}\|x + k - y\|^2$, the periodic Eudlidean distance squared.

When μ and ν have smooth continuous densities (still denoted μ and ν by abuse of notation), it is well known that the classical ($\epsilon = 0$) Optimal Transport problem between μ and ν has a unique solution $\gamma_0^* := (Id \otimes T)\#\mu \in \mathcal{P}(X \otimes Y)$ where T is the *optimal transport map*. We will use the notation : $x \in X \mapsto T(x) := y_x^\star \in Y$. The optimal transport plan γ_0^* is sparse and supported only on the graph $\{(x, y_x^*),\, x \in X\}$.

In practice Sinkhorn is applied on discrete sets $X_N = \{x_i\}_{i=1..N}$, $Y_N = \{y_j\}_{j=1..N}$ and μ and ν need to be discretized accordingly

$$\mu_N = \sum_{i=1}^N \frac{\mu(x_i)}{\sum_j \mu(x_j)} \delta_{x_i}, \qquad \nu_N = \sum_{j=1}^N \frac{\nu(y_j)}{\sum_j \nu(y_j)} \delta_{y_j} \tag{2}$$

We are interested in using Sinkhorn to approximate the Optimal Transport problem at 0 temperature trough the above N discretization. As can be observed on (1), ϵ acts as *bandwith* and the discretization must be scaled accordingly. A convergence result linking m the Sinkhorn iteration index, ϵ and N has recently been given in [1]. We assume to simplifiy the formulation that X_N and Y_N are cartesian grids with edgelength $h = \frac{1}{N^{1/d}}$:

Theorem 1 (Berman joint convergence for the plan - part of corollary 1.3 [1])**.** *We assume μ and ν are in $\mathcal{C}^{2,\alpha}$, N and ϵ are dependent parameters : $N = (1/\epsilon)^d$ (or $N = C_\delta/\epsilon^{d/(2(1+\delta))}$ for C^∞ marginals). We further assume that the discretisation (X_N, Y_N, μ_N, ν_N) satisfies $(\mu_N, \nu_N) \rightharpoonup (\mu, \nu)$. Then there exists a positive constant A_0 such that for any $A > A_0$ the folowing holds : Setting $m_\epsilon = [-A\log(\epsilon)/\epsilon]$ as the final number of Sinkhorn iterations, the discrete probability measures $\gamma_\epsilon^{m_\epsilon}$*

on $X_N \times Y_N$ (defined in (1)) converges weakly to the optimal transport plan $\gamma_0^\star$ in the $\epsilon \to 0$ temperature limit and satisfies the the following estimate

$$\gamma_\epsilon^{m_\epsilon} \leq \mathcal{B}_\epsilon \, \mu_N \otimes \nu_N, \tag{3}$$

where

$$(x, y) \mapsto \mathcal{B}_\epsilon(x, y) := \frac{p}{\epsilon^p} \exp(-\frac{c(y, y_x^\star)}{\epsilon \, p}) \tag{4}$$

The parameter p is positive, depends on the marginals $\mu \; \nu$ but is unfortunately not explicit. Estimate (4) is nevertheless important for two reasons : First, it shows that *the mass of the entropic transport plan concentrates exponentially fast in a neighbourhood of the transport map.* So even though $\gamma_\epsilon^{m_\epsilon}$ has support on the full $N \times N$ grid, it will be negligible , and even null in finite precision, for an increasing number of points as ϵ decreases. Secondly, as the support of $\{\mathcal{B}_\epsilon > \lambda\}$, $\lambda << 1$ converges monotonically to the graph of $\{(x, y_x^*), \, x \in X\}$ and $\gamma_\epsilon^{m_\epsilon}$ is dominated by $\mathcal{B}_\epsilon$, it suggests that *for $\epsilon' < \epsilon$ support information for $\{\gamma_\epsilon^{m_{\epsilon'}} > \lambda\}$ can be recovered from $\gamma_\epsilon^{m_\epsilon}$.*

This, of course has been observed numerically long before theorem 1. A detailed description of these phenomea and a proposed heuristic multiscale method in ϵ can be found in [5]. A GPU implementation can also been found in [6]. They both observe numerically the hoped for $\mathcal{O}(N \, log(N)$ complexity arising from the sparsity in the limit of the transport plan instead of the naive more than quadratic complexity, Just sketching the idea : the method uses the criterium $f_\epsilon^m(x) \oplus g_\epsilon^m(y) - c(x, y) < -M$ after a few iterations m for some large M to discard points for the next $\epsilon' < \epsilon$. This is done for a decreasing sequence ϵ (thus increasing N) but there is no proof of convergence. Also some mass transport may be missed if the truncation is too strong (or M too small). We show in this presentation that theorem 1 can be used to provide guidance both in the design and also the complexity analysis of a truncation/multi-scale method in ϵ. The details cannot be given in full in this abstract but can be found in [4], the main ideas follow.

We work on a vertical slice of the plan $\gamma_\epsilon^{m_\epsilon}$, i.e. x is fixed, let us call it x_0. Introducing a new parameter λ and assuming that $y_{x_0}^*$ is known, then the sub-level set inclusion $S_{\epsilon,\lambda}^{x_0} := \{\gamma_\epsilon^{m_\epsilon}(x_0, y) < \lambda\} \subset \{\mathcal{B}_\epsilon(x_0, y) < \lambda\}$ could be used as a rule to truncate the $X_N \otimes Y_N$ domain. for some small λ. The number of points N depends on ϵ and so is the number of iterations m_ϵ. They are all prescribed in theorem 1. The memory requirements and number of operations in Sinkhorn can be computed explicitely. Like in the [5] strategy, we want to choose λ in order to maximize the number of discarded point without destabilizing the Entropic OT problem at scale ϵ (the problem can even become infeasible if the truncated coupling cannot satisfy the marginal constraints). There are in summary two difficulties. First, of course estimating $y_{x_0}^*$ which itself is the solution at $\epsilon = 0$. Secondly we need to build a ϵ, λ sequence of problems on truncated domains which converge to the $\epsilon = 0$ temperature Optimal Transport problem.

We notice that the set $\{\mathcal{B}_\epsilon(x_0, y) < \lambda\}$ is a ball of center $y_{x_0}^*$ and its radius $d\mathcal{B}_{\epsilon,\lambda}$ is known as well as its size in term of N as a function of ϵ, λ. We use it to obtain

the super inclusion $\{\mathcal{B}_\epsilon(x_0,y) < \lambda\} \subset \S^{x_0}_{\epsilon,\lambda} := \cup_{y\in\{\gamma^{m_\epsilon}_\epsilon(x_0,y)<\lambda\}} B(y, d\mathcal{B}_{\epsilon,\lambda})$. The set on the right can be be computed without knowing $y^*_{x_0}$. Then, the simple choice $\lambda \equiv \epsilon$ is sufficient to guarantuee $S^{x_0}_{\epsilon',\lambda'} \subset \{\mathcal{B}_{\epsilon'}(x_0,y) < \lambda'\} \subset \{\mathcal{B}_\epsilon(x_0,y) < \lambda\}\S^{x_0}_{\epsilon,\lambda_\epsilon}$ and thus control the truncation at $\epsilon' < \epsilon$ using the solution at ϵ.

The second difficutly is tackled through the introduction of a *capacity constraint* $\gamma \geq \lambda$ to the Entropic Optimal Transport problem. Sinkhorn algorithm can modified accordingly and the truncation domain becomes a saturation domain $\gamma = \lambda$ wich can be *out-summed* analytically from Sinkhorn. The full story is more technical as one has to ensure that the estimate (4) still holds for the capacity constrained Sinkhorn solutions, this can again be found in [4].

We are finally able to give a *rigorous* estimate of the complexity of this new ϵ multicale method. Assuming the decreasing ϵ sequence is finite and that the smaller ϵ is the sought for precision, we find a computational cost of $\mathcal{O}(\log(N)^{1+d/2}\, N^{3/2+1/d})$ if $N = (1/\epsilon)^d$ and $\mathcal{O}(\log(N)^{1+d/2}\, N^{1+1/d})$ if $N = (1/\sqrt{\epsilon})^d$ (this is if μ and ν are $\mathcal{C}^\infty$), N is of course the largest reached discretization. This is not as good as the observed $\mathcal{O}(N\,\log(N)$ in [5] but is certainly not optimal as (4) is not sharp. The dependance in the dimension d is linked to the prescribed number of iterations m_ϵ in theorem 1.

References

[1] Robert J Berman. The sinkhorn algorithm, parabolic optimal transport and geometric monge-amp\ere equations. *Numerische Mathematics*, 145, June 2020.

[2] Christian Léonard. A survey of the schrödinger problem and some of its connections with optimal transport, 2013.

[3] G. Peyré and M. Cuturi. Computational Optimal Transport. *ArXiv e-prints*, March 2018.

[4] J.-D. Benamou and M. Martinet Capacity Constrained Entropic Optimal Transport, Sinkhorn Saturated Domain Out-Summation and Vanishing Temperature. *HAL preprint*, May 2020.

[5] Bernhard Schmitzer. Stabilized Sparse Scaling Algorithms for Entropy Regularized Transport Problems. *SIAM J. Sci. Comput.*, 41(3), 2019.

[6] Feydy, J. (2019). *GEOMLOSS https://www.kernel-operations.io/geomloss/.*

Curvature and relaxation for discrete mean-field dynamics

Matthias Erbar

(joint work with Max Fathi, Vaios Laschos, André Schlichting)

We consider non-linear evolution equations arising from mean-field limits of particle systems on discrete spaces. We investigate a notion of curvature bounds for these dynamics based on the convexity of the free energy along interpolations in a discrete transportation distance related to the gradient flow structure of the dynamics. We show that positive curvature bounds entail several functional inequalities controlling the convergence to equilibrium of the dynamics. We establish explicit curvature bounds for several examples of mean-field limits of various classical models from statistical mechanics.

This work is about longtime behavior for mean-field systems on discrete spaces. Mean-field equations describe the large-scale limit of interacting particle systems where the total force exerted on any given particle is the average of the forces exerted by all other particles on the tagged particle. They are used to describe collective behavior in many areas of sciences. One of the important questions in the mathematical analysis of these equations is their longtime behavior. In [2], Carrillo, McCann and Villani obtained quantitative bounds on the rate of convergence to equilibrium for McKean-Vlasov equations in a continuous setting of the form

$$\partial_t \rho = \nabla \cdot [\rho \nabla (S'(\rho) + V + W * \rho)]$$

under strong convexity assumptions on the potentials S, V and W. The core idea underlying their method is the fact that the PDE has a gradient flow structure, i.e. it can be recast as a gradient descent equation $\partial_t \rho = -\nabla F(\rho)$ of the free energy functional $F(\rho) = \int S(\rho) + \int V dx \rho + \int W * \rho dx \rho$ in the space of probability measure with respect to the Kantorovitch-Wasserstein distance W_2. The use of such structures in the study of longtime behavior stems from the fact that uniform strict convexity of the driving functional (with respect to the particular metric structure), implies exponentially fast convergence of the gradient flow. Moreover, convexity can be used to derive strong functional inequalities relating the distance, free energy and dissipation [9].

Our main motivation here is to adapt the approach of [2] to mean-field equations in a discrete setting. We consider discrete mean-field dynamics of the form

$$\dot{\mu}(t) = \mu(t) Q\big(\mu(t)\big) \,, \tag{1}$$

where μ is a flow of probability measures on a finite set $\mathcal{X}$ and $(Q(\mu)_{xy})_{x,y\in\mathcal{X}}$ is a parametrized collection of Markov kernels. These dynamics naturally arise as scaling limits of interacting particles systems on graphs where the interaction only depends on the normalized empirical measure of the system (which indeed corresponds to mean-field interactions). They generalize linear Markov chains on discrete spaces, which correspond to the case where Q is a constant Markov kernel, independent of μ.

More precisely, consider a mean-field particle system $\boldsymbol{X}_t = (X_t^1, \dots, X_t^N)$ on $\mathcal{X}$ modeled as a reversible continuous time Markov chain dynamics on $\mathcal{X}^N$ with invariant measure and transition kernel given by

$$\pi^N(\boldsymbol{x}) \sim \exp\big(-U(\mu^N)\big), \qquad Q^N(\boldsymbol{x}, \boldsymbol{x}^{i,y}) = Q_{x^i,y}(\mu^n) \,,$$

where $\mu^N = 1/N \sum_i \delta_{x^i}$ denotes the empirical measure of the configuration $\boldsymbol{x} = (x^1, \dots, x^n)$, $U(\mu) = \sum_x \mu_x K_x(\mu)$ for a kernel $K : \mathcal{X} \times \mathcal{P}(\mathcal{X}) \to \mathbb{R}$ and $\boldsymbol{x}^{i,y}$ is the new configuration where particle i has jumped to site y. Then, under suitable assumptions on Q and K, the empirical distribution of $\boldsymbol{X}_t$ converges to the deterministic evolution μ_t solving (1). Typical choices of U are $U(\boldsymbol{x}) = \frac{1}{N}\sum_i V(x^i) + \frac{1}{N^2}\sum_{ij} W(x^i, x^j)$ for an external potential V and an interaction potential W. An archetypical example is the classical Curie-Weiss model, which corresponds to a mean-field dynamic on a two-point space. Already this easy model

exhibits interesting behavior, such as a phase transition at an explicit critical value of a temperature parameter.

In the work [4], we derived a gradient flow structure for (1) by replacing the role of the Wasserstein distance with a distance $\mathcal{W}$ constructed via a suitable modification of the Benamou-Brenier formula for optimal transport, extending similar earlier results for linear reversible Markov chains obtained in [7, 8, 3].

Note that the rates $Q(\mu)$ are reversible with respect to a local Gibbs measure of the form $\pi_x(\mu) \sim \exp\big(-H_x(\mu)\big)$, with $H_x(\mu) = \partial_{\mu_x} U(\mu)$. The dynamics can be recast as a gradient flow of the free energy

$$\mathcal{F}(\mu) = \sum_{x\in\mathcal{X}} \mu_x \log \mu_x + U(\mu), \qquad \text{with} \qquad U(\mu) = \sum_{z\in\mathcal{X}} \mu_z\, K_z(\mu). \tag{2}$$

Namely, we can rewrite (1) as

$$\partial_t \mu = -K_\mu^{MF} D\mathcal{F}(\mu)\ ,$$

with the Onsager operator

$$K_\mu^{MF}\psi(x) = -\sum_y \big(\psi_y - \psi_x\big)\Lambda\big(\mu_x Q_{xy}(\mu), \mu_y Q_{yx}(\mu)\big)\ ,$$

where $\Lambda(s,t) = (s-t)/(\log s - \log t)$ is the logarithmic mean. This built up on previous works (e.g. [1]) that showed that $\mathcal{F}$ is indeed a Lyapunov functional for the flow. Moreover, we showed that the gradient flow structure of the mean-field equation (1) naturally arises in the limit $N \to \infty$ from the gradient structure of the N-particle Markov chain given in [7, 8, 3] in the sense of evolutionary Γ-convergence.

In the work [5], we exploit this gradient flow structure to analyze the long-term behavior of (1) inspired by the approach in [2] by investigating convexity properties of the free energy along discrete optimal transport paths for a non-linear Markov triple $(\mathcal{X}, Q, \pi)$ as above. Following the works of Lott, Villani, and Sturm for metric measure spaces and [6, 8] for linear Markov chains, we make the following definition:

We say that the non-linear Markov chain $\big(\mathcal{X}, Q(\cdot), \pi(\cdot)\big)$ has *entropic Ricci curvature bounded below by* $\kappa \in \mathbb{R}$ if for any $\mathcal{W}$-geodesic $(\mu_t)_{t\in[0,1]}$:

$$\mathcal{F}(\mu_t) \le (1-t)\mathcal{F}(\mu_0) + t\mathcal{F}(\mu_1) - \frac{\kappa}{2} t(1-t)\mathcal{W}(\mu_0,\mu_1)^2\ .$$

We show that Ricci curvature lower bounds can be characterized in terms of a discrete Bochner-type inequality by deriving the Hessian of $\mathcal{F}$ in the Riemannian structure $\mathcal{W}$, as well as in terms of the Evolution Variational Inequality for the solutions to (1). Further, we show that a positive lower bound on the Ricci curvature entails a number of functional inequalities that control the convergence to equilibrium of the mean-field systems. These involve a discrete Fisher information functional $\mathcal{I} : \mathcal{P}(\mathcal{X}) \to [0,\infty]$ given by

$$\mathcal{I}(\mu) = \frac{1}{2}\sum_{x,y} \Theta\big(\mu_x Q_{xy}(\mu), \mu_y Q_{yx}(\mu)\big)\ , \qquad \Theta(a,b) = (a-b)(\log a - \log b)\ ,$$

which arises from the dissipation of $\mathcal{F}$ along solutions to (1) as $\frac{d}{dt}\mathcal{F}(\mu_t) = -\mathcal{I}(\mu_t)$. One of our main results is the following theorem which can be seen as a discrete analog of [2, Thm. 2.1].

Theorem 1. *Assume that* $\mathsf{Ric}(\mathcal{X}, Q, \pi) \geq \lambda$ *for some* $\lambda > 0$. *Then the following hold:*

(i) *there exists a unique stationary point* π^* *for the evolution* (1), *it is the unique minimizer of* $\mathcal{F}$. *Let* $\mathcal{F}^*(\cdot) := \mathcal{F}(\cdot) - \mathcal{F}(\pi^*)$;

(ii) *the* modified logarithmic Sobolev inequality *with constant* $\lambda > 0$ *holds, i.e. for all* $\mu \in \mathcal{P}(\mathcal{X})$,

$$\mathcal{F}^*(\mu) \leq \frac{1}{2\lambda}\mathcal{I}(\mu) \; ;$$

(iii) *for any solution* $(\mu_t)_{t\geq 0}$ *to* (1) *we have exponential decay of the free energy:*

$$\mathcal{F}^*(\mu_t) \leq e^{-2\lambda t}\mathcal{F}^*(\mu_0) \; ;$$

(iv) *the* entropy-transport inequality *with constant* $\lambda > 0$ *holds, i.e. for all* $\mu \in \mathcal{P}(\mathcal{X})$,

$$\mathcal{W}(\mu, \pi^*) \leq \sqrt{\frac{2}{\lambda}\mathcal{F}^*(\mu)} \; .$$

We establish explicit curvature bounds for several examples of (relatively simple) mean-field dynamics, such as the Curie-Weiss model, zero-range mean-field dynamics, and misanthrope processes.

References

[1] Amarjit Budhiraja, Paul Dupuis, Markus Fischer and Kavita Ramanan. Limits of relative entropies associated with weakly interacting particle systems. *Electron. J. Probab.* **20**, no. 80, 22 (2015) .

[2] José A. Carrillo, Robert J. McCann and Cédric Villani. Kinetic equilibration rates for granular media and related equations: entropy dissipation and mass transportation estimates. *Rev. Mat. Iberoamericana* **19** (3), 971–1018 (2003).

[3] Shui-Nee Chow, Wen Huang, Yao Li and Haomin Zhou. Fokker-Planck equations for a free energy functional or Markov process on a graph. *Arch. Ration. Mech. Anal.* **203** (3), 969–1008 (2012).

[4] Matthias Erbar, Max Fathi, Vaios Laschos and André Schlichting. Gradient flow structure for McKean-Vlasov equations on discrete spaces. *Discrete Contin. Dyn. Syst.* **36** (12), 6799–6833 (2016).

[5] Matthias Erbar, Max Fathi and André Schlichting. Entropic curvature and convergence to equilibrium for mean-field dynamics on discrete spaces. *ALEA Lat. Am. J. Probab. Math. Stat.* **17** (1), 445–471, (2020).

[6] Matthias Erbar and Jan Maas. Ricci Curvature of Finite Markov Chains via Convexity of the Entropy. *Archive for Rational Mechanics and Analysis* **206** (3), 997–1038 (2012).

[7] Jan Maas. Gradient flows of the entropy for finite Markov chains. *Journal of Functional Analysis* **261** (8), 2250–2292 (2011).

[8] Alexander Mielke. Geodesic convexity of the relative entropy in reversible Markov chains. *Calc. Var. Partial Differential Equations* **48** (1-2), 1–31 (2013).

[9] Felix Otto and Cédric Villani. Generalization of an Inequality by Talagrand and Links with the Logarithmic Sobolev Inequality. *Journal of Functional Analysis* **173** (2), 361–400 (2000).

Initial value problems by convex minimization and matrix-valued generalizations of some OT, MFG and Schrödinger problems

Yann Brenier

Few years ago, at a previous edition of the MFO workshop on "Variational Methods for Evolution Equations", I explained how to solve the Initial Value Problem (IVP) by convex minimization for the class of first-order systems of conservation laws enjoying a convex entropy [3]. I was asked by Ulisse Stefanelli whether our approach could be extended to the Navier-Stokes (NS) equations, for instance as an alternative to the method introduced by Alexander Mielke and himself [8]. In the present talk, we try to answer this question and start with simpler parabolic models, such as the porous medium equation and the viscous Hamilton-Jacobi equations. Then, it can be established that all solutions of the IVP on a given time interval $[0,T]$, with $T < +\infty$, can be recovered from our convex minimization method. In the more challenging case of the NS equations, the corresponding convex minimization problem can be easily computed and presents interesting features. (However, it is not true a priori that *all* solutions of the IVP can be recovered this way. There might be some restriction on the size of the time interval, depending on the initial conditions.) Indeed, in the case of a periodic box $D = \mathbb{T}^d$, with given initial velocity field v_0, our convex minimization problem for the NS equations can be written as a kind of "generalized" Schrödinger problem:

$$\inf_{M,j} \int_Q \frac{(j - \epsilon \nabla \cdot M) \cdot M^{-1} \cdot (j - \epsilon \nabla \cdot M)}{2} + j \cdot v_0$$

where $Q = [0,T] \times D$ is the space-time domain, the symmetric matrix-valued "density" field $M = M(t,x) g_\epsilon 0$ and the "current" $j = j(t,x) \in \mathbb{R}^d$ are subject to a generalized version of the "continuity equation", namely:

$$\partial_t M + \mathcal{L} j = 0, \quad M(T,\cdot) = I_d,$$

where $\mathcal{L}$ is the constant coefficient first-oder pseudo-differential operator

$$\mathcal{L} j = \nabla j + \nabla j^T - 2D^2 \Delta^{-1} \nabla \cdot j.$$

This minimization problem, which is closely related to the so-called "Bredinger" (or "Brödinger") problem adressed in [1] and more recently [2], differs from the standard Schrödinger problem (for which we refer to [6]) in two essential ways:
1) the "density" field M is not scalar but matrix-valued, a little bit like in Quantum Mechanics or in non-commutative Geometry and also in the recent theory of matrix-valued optimal transportation (for which we refer to [4] as a recent reference);
2) the time boundary conditions are of backward-forward type, just as in the theory of Mean-Field Games (MFG) à la Lasry-Lions [7]: M is only prescribed at the final time $t = T$, while v_0 substitutes for an initial condition for j.
Finally, let us observe that our minimization problem involves a matrix-valued version of the Fisher information

$$(\nabla \cdot M) \cdot M^{-1} \cdot (\nabla \cdot M), \quad M = M^T g_\epsilon 0,$$

very roughly similar to the Einstein-Hilbert Lagrangian, which reads, in 4 space-time dimension, up to a null Lagrangian [5],

$$(\Gamma^m_{ij}\ g^{ij}\ \Gamma^k_{km} - \Gamma^m_{ik}\ g^{ij}\ \Gamma^k_{jm})\sqrt{-\det\ g}$$

for which g_{ij} is a Lorentzian metric (of inverse g^{ij}) and Γ is its Levi-Cività connection:

$$\Gamma^i_{jk} = g^{im}(g_{km,j} + g_{jm,k} - g_{kj,m})/2.$$

Yann Brenier, CNRS, DMA-Ecole Normale Supérieure, UPSL, Paris.

References

[1] M. Arnaudon, A. B. Cruzeiro, Ch. Léonard, J.-C. Zambrini *An entropic interpolation problem for incompressible viscid fluids* preprint arXiv:1704.02126.

[2] A. Baradat, L. Monsaingeon, *Small noise limit and convexity for generalized incompressible flows, Schrödinger problems, and optimal transport,* Arch. Ration. Mech. Anal. 235 (2020) 1357–1403.

[3] Y. Brenier, *The initial value problem for the Euler equations of incompressible fluids viewed as a concave maximization problem,* Comm. Math. Phys. 364 (2018) 579-605.

[4] Y. Brenier, D. Vorotnikov, *On optimal transport of matrix-valued measures,* SIAM J. Math. Anal. 52 (2020) 2849-2873.

[5] B. Dobrouvine, S. Novikov, A. Fomenko, *Géométrie contemporaine, 1re partie, Chapitre 6, paragraphe 4,* Mir, Moscou 1982.

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

[7] J.-M. Lasry, P.-L. Lions, *Mean field games,* Japan J. Math. 2 (2007) 229-260.

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

A Proximal-Gradient Algorithm for Crystal Surface Evolution

Katy Craig

(joint work with Jian-Guo Liu, Jianfeng Lu, Jeremy Marzuola, Li Wang)

As a counterpoint to recent numerical methods for crystal surface evolution, which agree well with microscopic dynamics but suffer from significant stiffness, we develop a new numerical method based on the macroscopic partial differential equation, leveraging its formal structure as the gradient flow of the total variation energy with respect to a weighted H^{-1} norm. This gradient flow structure relates to several metric space gradient flows of recent interest, including 2-Wasserstein gradient flows and its generalizations to nonlinear, concave mobilities. We develop a novel semi-implicit time discretization of the gradient flow, inspired by the classical minimizing movement scheme (known as the JKO scheme in the 2-Wasserstein case). We then use a primal dual hybrid gradient (PDHG) operator splitting method to compute each element of the semi-implicit scheme. In one dimension, we prove convergence of the PDHG method to the semi-implicit scheme, under general integrability assumptions on the mobility and its reciprocal. Finally, by taking finite difference approximations of our PDHG method, we arrive at a

fully discrete numerical algorithm, with iterations that converge at a rate independent of the spatial discretization: in particular, the convergence properties do not deteriorate as we refine our spatial grid. We close with several numerical examples illustrating the properties of our method, including facet formation at local maxima, pinning at local minima, and convergence as the spatial and temporal discretizations are refined.

References

[1] J.A. Carrillo, K. Craig, L. Wang, C. Wei, *Primal dual methods for wasserstein gradient flows*, arXiv:1901.08081.

[2] M.Jacobs, F. Léger, W. Li, S. Osher, *Solving large-scale optimization problems with a convergence rate independent of grid size*, SIAM Journal on Numerical Analysis **57** (3), 1100–1123 (2019).

[3] J.-G. Liu, J. Lu, D. Margetis, J.L. Marzuola, *Asymmetry in crystal facet dynamics of homoepitaxy by a continuum model*, Physica D: Nonlinear Phenomena **393**, 54–67 (2019).

[4] J.L. Marzuola, J. Weare, *Relaxation of a family of broken-bond crystal-surface models*, Physical Review E **88** (3), 032403 (2013).

Asymptotic limits of random walks via generalized gradient flows

Anastasiia Hraivoronska

(joint work with Oliver Tse)

We consider the law of continuous-time random walks on graphs as an approximation for solutions of diffusion equations. The motivation comes from developing structure-preserving numerical schemes. Our approach is based on exploiting the gradient structure of both evolution types and passing to the discrete-to-continuous limit in the corresponding variational formulation. We postulate that the limit may depend wildly on the graph geometry and can be surprisingly nontrivial.

Let $G = (\mathcal{T}, \Sigma)$ be a graph with the set of vertices $\mathcal{T}$ and edges Σ. We assume that the law of the random walk on G satisfies the forward Kolmogorov equation:

$$\partial_t \rho_t = Q^* \rho_t, \tag{1}$$

where $t \mapsto \rho_t$ is a curve in the space of probability measures $\mathcal{P}(\mathcal{T})$. The infinitesimal generator is given for any bounded function $\varphi \in B(\mathcal{T})$ by

$$(Q\varphi)(K) = \sum_{L \in \mathcal{T}_K} (\overline{\nabla}\varphi)(K, L)\kappa(K, L), \qquad K \in \mathcal{T},$$

where $\overline{\nabla} : B(\mathcal{T}) \to B(\Sigma)$ is the discrete gradient $(\overline{\nabla}\varphi)(K, L) = \varphi(L) - \varphi(K)$, and the sum is taken over the collection of adjacent vertices $\mathcal{T}_K = \{L \in \mathcal{T} : (K, L) \in \Sigma\}$ of K. The jump kernel $\kappa : \Sigma \to [0, \infty)$ defines the jump rates for each edge in Σ.

To relate the discrete equation to the continuous space, we follow a finite-volume approach. Let the discretization of $\mathbb{R}^d$ be given by a family $\mathcal{T}^h$ of control volumes and a family Σ^h of faces. We introduce the parameter h as a characteristic size of the discretization $h = \max_{K \in \mathcal{T}^h} \operatorname{diam}(K)$.

We aim to determine the law of the limit process as $h \to 0$ with the help of a gradient structure. The works [1, 2] introduced a distance on $\mathcal{P}(\mathcal{T})$, and showed that that the solution of (1) evolves as the gradient flow of the relative entropy:

$$\mathcal{F}^h(\rho^h) = \sum_{K \in \mathcal{T}^h} \phi\left(u^h(K)\right) \pi^h(K), \qquad \phi(s) = s \log s - s + 1,$$

where u^h is the density of the measure ρ^h with respect to an invariant measure π^h. Since then, gradient structures have been introduced for jump process on more general spaces (see for instance [3] and the recently created functional framework for generalized gradient systems [4]).

We follow [4] to characterize the solution of (1) via the energy-dissipation principle (EDP). We say that a pair (ρ^h, j^h) is an $(\mathcal{F}_h, \mathcal{R}_h, \mathcal{R}_h^*)$-generalized gradient flow solution of (1) if it satisfies the continuity equation

$$\partial_t \rho_t^h + \overline{\nabla} \cdot j_t^h = 0 \tag{CE$_h$}$$

and it is a minimizer of the EDP functional $\mathcal{I}_h$, i.e.

$$\mathcal{I}_h(\rho^h, j^h) = 0, \tag{EDP$_h$}$$

where

$$\mathcal{I}_h(\rho^h, j^h) := \int_0^T \mathcal{R}_h(\rho_t^h, j_t^h) + \mathcal{R}_h^*(\rho_t^h, -\overline{\nabla}\phi'(u_t^h))\, dt + \mathcal{F}_h(\rho_T^h) - \mathcal{F}_h(\rho_0^h).$$

Here $\mathcal{R}_h$ and $\mathcal{R}_h^*$ are the dissipation potential and its dual. We focus on the cosh-structure that has been shown to arise from large-deviation principles for Markov processes [5].

Our goal is to recover the corresponding limits of (CE$_h$) and (EDP$_h$) as $h \to 0$. The first step is to embed the discrete objects into the continuous space beginning with the pairs of measures and fluxes:

$$\hat{\rho}^h(dx) = \sum_{K \in \mathcal{T}^h} \frac{\rho^h(K)}{|K|} \mathbb{I}_K(x) dx, \qquad \hat{j}^h(dx) = \sum_{K|L \in \Sigma^h} j^h(K,L) \sigma_{KL}(dx). \tag{2}$$

Remarkably, σ_{KL} can be chosen such that the pair $(\hat{\rho}^h, \hat{j}^h)$ satisfies the continuity equation on $\mathbb{R}^d$:

$$\partial_t \hat{\rho}^h + \nabla \cdot \hat{j}^h = 0 \qquad \text{in distributions.}$$

For the EDP functional we prove the lower-semicontinuity

$$\liminf_{h \to 0} \mathcal{I}_h(\rho^h, j^h) \geq \mathcal{I}_0(\rho, j)$$

separately for all components: the dissipation potential $\mathcal{R}_h$, its dual $\mathcal{R}_h^*$, and the driving energy $\mathcal{F}_h$. While the results for the energy being standard and the limit for $\mathcal{R}_h$ being rather straightforward, the convergence for $\mathcal{R}_h^*(\rho^h, -\overline{\nabla}\phi'(u^h))$ is more delicate and requires the use of localization methods for Γ-convergence.

We conclude that the geometry of the discretization is of crucial importance for the limit evolution. In general, we obtain the Γ-limit EDP functional with a

weighted dual dissipation potential

$$\mathcal{R}_0^*(\rho, \nabla\varphi) = \int \langle \nabla\varphi, \mathbb{T}\,\nabla\varphi\rangle\, d\rho,$$

with some tensor $\mathbb{T} \in \mathbb{R}^{d\times d}$ that depends on the limiting geometry of the family of discretizations. For $\mathbb{T}$ to be equal to the identity matrix the sequence of finite-volume tessellation must satisfy strong asymptotic assumptions. Note that similar assumptions on graphs are required in [6] for proving the convergence of discrete transport metrics to the Wasserstein distance. However, we show as well that the limit exists with the weighted $\mathcal{R}_0^*$ under relaxed assumptions.

References

[1] J. Maas, *Gradient flows of the entropy for finite Markov chains*, Journal of Functional Analysis **261** (2011), 8:2250–2292.

[2] A. Mielke, *Geodesic convexity of the relative entropy in reversible Markov chains*, Calc. Var. Partial Differential Equations **48** (2013), 11–31.

[3] M. Erbar, *Gradient flows of the entropy for jump processes*, Annales de l'IHP Probabilités et statistiques **50** (2014), 3:920–945.

[4] M.A. Peletier, R. Rossi, G. Savaré, O. Tse, *Jump processes as Generalized Gradient Flows*, arXiv:2006.10624.

[5] A. Mielke, M.A. Peletier, M.D.R. Renger, *On the relation between gradient flows and the large-deviation principle, with applications to Markov chains and diffusion*, Potential Analysis **41** (2014), 4:1293–1327.

[6] P. Gladbach, E. Kopfer, J. Maas, *Scaling limits of discrete optimal transport*, SIAM Journal on Mathematical Analysis **52** (2020), 3:2759–2802.

EDP convergence for nonlinear fast-slow reaction systems

Artur Stephan

(joint work with Alexander Mielke, Mark A. Peletier)

The study of nonlinear reaction systems with different time scales has attracted much attention over the last decades, see e.g. [Bot03, DLZ18, MiS19] and the references therein. In this work we consider the simplest case of fast-slow reaction systems with mass-action kinetics that have only two time scales, namely 1 and ε,

$$\dot{c} = \boldsymbol{R}_{\mathrm{sl}}(c) + \frac{1}{\varepsilon}\boldsymbol{R}_{\mathrm{fa}}(c), \tag{1}$$

where $c \in \mathbf{C} := [0,\infty[^{i_*}$ denotes the vector of the concentrations c_i of the ith species X_i. The typical aim of the above-mentioned work is to derive the limiting equation for the evolution of c on the slow time scale, while the fast reactions are in equilibrium. Under suitable assumptions the limiting equation can be formulated in three equivalent ways, namely

$$\begin{aligned}
\text{constraint dynamics:}&\quad \dot{c}(t) = \boldsymbol{R}_{\mathrm{sl}}(c(t)) + \lambda(t), \quad \lambda(t) \in \Gamma_{\mathrm{fa}} \subset \mathbb{R}^{i_*}, \quad \boldsymbol{R}_{\mathrm{fa}}(c(t)) = 0,\\
\text{projected dynamics:}&\quad \dot{c}(t) = (I - \mathbb{P}(c(t)))\boldsymbol{R}_{\mathrm{sl}}(c(t)), \quad \boldsymbol{R}_{\mathrm{fa}}(c(0)) = 0,\\
\text{reduced dynamics:}&\quad \dot{\mathsf{q}}(t) = Q_{\mathrm{fa}}\boldsymbol{R}_{\mathrm{sl}}(\Psi(\mathsf{q}(t))), \quad c(t) = \Psi(\mathsf{q}(t)).
\end{aligned}$$

The goal of this work is to revisit the same limit process, but now from the point of view of variational evolution. Our starting point is that reaction-rate equations such as (1) can be written as a gradient-flow equation if the reactions occur in pairs of forward and backward reactions and that these pairs satisfy the *detailed-balance condition* ([Yon08, Mie11]). A different gradient structure has its origin in the thermodynamic considerations in [Mar15] from 1915. The latter gradient structure, which we will call the *cosh-type gradient structure* following [MiS19], was mathematically derived in [MPR14, MP*17] from microscopic chemical master equations via a large-deviation principle.

To be specific, we assume that the species X_i, $i \in I := \{1, \dots, i_*\}$ undergo r_* forward-backward reactions of mass-action type

$$\alpha_1^r X_1 + \cdots + \alpha_{i_*}^r X_{i_*} \quad \rightleftharpoons \quad \beta_1^r X_1 + \cdots + \beta_{i_*}^r X_{i_*},$$

where $\alpha^r = (\alpha_i^r)_{i\in I}$ and $\beta^r = (\beta_i^r)_{i\in I}$ are the stoichiometric vectors in $\mathbb{N}_0^{i_*}$. The reaction-rate equation (1) takes the form

$$\dot c = -\sum_{r=1}^{r_*} \big(k_r^{\mathrm{fw}} c^{\alpha^r} - k_r^{\mathrm{bw}} c^{\beta^r}\big)\big(\alpha^r - \beta^r\big), \quad \text{where } c^\alpha = c_1^{\alpha_1} \cdots c_{i_*}^{\alpha_{i_*}}. \tag{2}$$

The *detailed-balance condition* asks for the existence of a positive concentration vector $c_* = (c_i^*)_{i\in I} \in \mathbf{C}_+ :=]0,\infty[^{i_*}$ such that all r_* reactions are in

$$\exists\, c_* = (c_i^*)_{i\in I} \in \mathbf{C}_+ \ \forall\, r \in R := \{1, \dots, r_*\} : \quad k_r^{\mathrm{fw}} c_*^{\alpha^r} = k_r^{\mathrm{bw}} c_*^{\beta^r}. \tag{3}$$

Throughout this work we will assume that c_* will not depend on the small parameter ε measuring the ratio between the slow and the fast time scale. The set of reaction pairs will be decomposed into slow and fast reactions, and, introducing the scalars $\widehat\kappa_r = k_r^{\mathrm{fw}}\big(c_*^{\alpha^r}/c_*^{\beta^r}\big)^{1/2} = k_r^{\mathrm{bw}}\big(c_*^{\beta^r}/c_*^{\alpha^r}\big)^{1/2}$, we assume $\widehat\kappa_r = \kappa_r$ for the slow reactions and $\widehat\kappa_r = \kappa_r/\varepsilon$ for the fast reactions. With this we obtain

$$\dot c = \boldsymbol{R}_{\mathrm{sl}}(c) + \frac{1}{\varepsilon}\boldsymbol{R}_{\mathrm{fa}}(c) \quad \text{with } \boldsymbol{R}_{\mathrm{xy}}(c) := -\sum_{r\in R_{\mathrm{xy}}} \kappa_r\, \delta_r^* \left(\frac{c^{\alpha^r}}{c_*^{\alpha_r}} - \frac{c^{\beta^r}}{c_*^{\beta_r}}\right)(\alpha^r - \beta^r). \tag{4}$$

The cosh-type gradient structure is now defined in terms of a gradient system $(\mathbf{C}, \mathcal{E}, \mathcal{R}_\varepsilon^*)$, where the energy functional is given in terms of the relative Boltzmann entropy

$$\mathcal{E}(c) = \sum_{i\in I} c_i^* \lambda(c_i/c_i^*), \quad \text{where } \lambda(\rho) := \rho\log\rho - \rho + 1,$$

and the dual dissipation potential $\mathcal{R}_\varepsilon^*$ in the form

$$\mathcal{R}_\varepsilon^*(c,\xi) = \mathcal{R}_{\mathrm{sl}}^*(c,\xi) + \frac{1}{\varepsilon}\mathcal{R}_{\mathrm{fa}}^*(c,\xi) \text{ with } \mathcal{R}_{\mathrm{xy}}^*(c,\xi) = \sum_{r\in R_{\mathrm{xy}}} \kappa_r \big(c^{\alpha^r} c^{\beta^r}\big)^{\frac12} \mathsf{C}^*\big((\alpha^r - \beta^r)\cdot\xi\big),$$

where $\mathsf{C}^*(\zeta) = 4\cosh(\zeta/2) - 4$. The fast-slow reaction-rate equation (4) can now be written as the gradient flow equation $\dot c(t) = \partial_\xi \mathcal{R}_\varepsilon^*\big(c(t), -\mathrm{D}\mathcal{E}(c(t))\big)$.

In the talk, we construct the effective gradient system $(\mathbf{C}, \mathcal{E}_{\mathrm{eff}}, \mathcal{R}_{\mathrm{eff}}^*)$ for the given family $(\mathbf{C}, \mathcal{E}, \mathcal{R}_\varepsilon^*)$ in the limit $\varepsilon \to 0^+$. Here we use the notion of *convergence of gradient system in the sense of the energy-dissipation principle*, shortly

called *EDP-convergence*. This convergence notion was introduced in [DFM19] and further developed in [MMP19, MiS19] and is based on the dissipation functionals

$$\mathfrak{D}_\varepsilon(c) := \int_0^T \big\{\mathcal{R}_\varepsilon(c,\dot c) + \mathcal{R}_\varepsilon^*(c,-\mathrm{D}\mathcal{E}(c))\big\}\,\mathrm{d}t,$$

which are defined for all curves $c \in \mathrm{L}^1([0,T];\mathbf{C})$. The notion of *EDP-convergence* now asks that the two Γ-convergences $\mathcal{E}_\varepsilon \xrightarrow{\Gamma} \mathcal{E}_{\mathrm{eff}}$ and $\mathfrak{D}_\varepsilon \xrightarrow{\Gamma} \mathfrak{D}_0$ (in suitable topologies) and that the limit $\mathfrak{D}_0$ has the form $\mathfrak{D}_0(c) = \int_0^T \mathcal{R}_{\mathrm{eff}}(c,\dot c) + \mathcal{R}_{\mathrm{eff}}^*(c,-\mathrm{D}\mathcal{E}(c))\,\mathrm{d}t$.

Our main theorem reads as follows.

Main Theorem. [MPS20] *Let the Unique Fast Equilibrium Condition (UFEC) be satisfied. Then the gradient systems* $(\mathbf{C},\mathcal{E},\mathcal{R}_\varepsilon)$ *EDP-converge to the gradient system* $(\mathbf{C},\mathcal{E},\mathcal{R}_{\mathrm{eff}})$*, where*

$$\mathcal{E}_{\mathrm{eff}} = \mathcal{E} \qquad \textit{and} \qquad \mathcal{R}_{\mathrm{eff}}^*(c,\xi) = \mathcal{R}_{\mathrm{sl}}^*(c,\xi) + \chi_{\Gamma_{\mathrm{fa}}^\perp}(\xi),$$

with $\Gamma_{\mathrm{fa}} = \mathrm{span}\big\{\alpha^r - \beta^r \mid r \in R_{\mathrm{fa}}\big\}$ *and* $\Gamma_{\mathrm{fa}}^\perp := \big\{\xi \in \mathbb{R}^{i_*} \mid \forall\,\gamma \in \Gamma_{\mathrm{fa}} :\ \gamma\cdot\xi = 0\big\}$.

The proof of the theorem relies on the following observations:

(1) The primal dissipation potentials $\mathcal{R}_\varepsilon$ decrease monotonically to their limit $\mathcal{R}_{\mathrm{eff}}$, which is degenerate. Defining $Q_{\mathrm{fa}} : \mathbb{R}^{i_*} \to \mathbb{R}^{m_{\mathrm{fa}}}$ such that $\ker Q_{\mathrm{fa}} = \Gamma_{\mathrm{fa}}$ and $\mathrm{im}\, Q_{\mathrm{fa}}^\top = \Gamma_{\mathrm{fa}}^\perp$, the bound $\mathfrak{D}_\varepsilon(c^\varepsilon) \le M_{\mathrm{diss}} < \infty$ does not provide a uniform bound on $\dot c^\varepsilon$, but we are able to show weak compactness of $Q_{\mathrm{fa}}c^\varepsilon$ in $\mathrm{W}^{1,1}([0,T];\mathbb{R}^{m_{\mathrm{fa}}})$.

(2) The bound $\mathfrak{D}_\varepsilon(c^\varepsilon) \le M_{\mathrm{diss}} < \infty$ implies $\int_0^T \mathcal{R}_{\mathrm{fa}}^*(c^\varepsilon, -\mathrm{D}\mathcal{E}(c^\varepsilon))\,\mathrm{d}t \le \varepsilon\, M_{\mathrm{diss}}$, which forces c^ε into the set of equilibria of the fast equation, namely $\mathscr{E}_{\mathrm{fa}} := \big\{c \in \mathbf{C} \mid \boldsymbol{R}_{\mathrm{fa}}(c) = 0\big\}$.

An important assumption is the UFEC, which states that the fast reaction system $c'(\tau) = \boldsymbol{R}_{\mathrm{fa}}(c(\tau))$ has a unique equilibrium (denoted by $\Psi(\mathsf{q})$) in each invariant subset $\mathbf{C}_{\mathsf{q}}^{\mathrm{fa}} := \{c \in \mathbf{C} \mid Q_{\mathrm{fa}}c = \mathsf{q}\}$.

The main difficulty is to show that the two complementary information in points (1) and (2) is enough to obtain the compactness necessary for deriving liminf estimate for the Γ-convergence $\mathfrak{D}_\varepsilon \xrightarrow{\Gamma} \mathfrak{D}_0$ for the non-convex functionals $\mathfrak{D}_\varepsilon$. On the local level, one sees that (1) provides partial control of the temporal oscillations of $\dot c^\varepsilon$ via the bound on $Q_{\mathrm{fa}}\dot c^\varepsilon$ in $\mathrm{L}^1([0,T];\mathbb{R}^{m_{\mathrm{fa}}})$, whereas (2) provides strong convergence towards $\mathscr{E}_{\mathrm{fa}}$, which is locally defined via $\mathrm{D}\mathcal{E}(c) \in \Gamma_{\mathrm{fa}}^\perp$. In fact, both information yield strong convergence $c^{\varepsilon_n} \to \widetilde c$, where $\widetilde c(t) = \Psi(\mathsf{q}(t))$ with $\mathsf{q} \in \mathrm{W}^{1,1}([0,T];\mathbb{R}^{m_{\mathrm{fa}}})$.

As a corollary we obtain that the limiting evolution lies in $\mathscr{E}_{\mathrm{fa}}$ and is governed by the reduced (or coarse grained) equation $\dot{\mathsf{q}} = Q_{\mathrm{fa}}\boldsymbol{R}_{\mathrm{sl}}(\Psi(\mathsf{q}))$ described by the slow variables $\mathsf{q} \in Q_{\mathrm{fa}}\mathbf{C}$ and a natural gradient structure $(Q_{\mathrm{fa}}\mathbf{C},\mathsf{E},\mathsf{R})$.

References

[Bot03] D. Bothe. Instantaneous limits of reversible chemical reactions in presence of macroscopic convection. *J. Diff. Eqns.*, 193(1), 27–48, 2003.

[DFM19] P. Dondl, T. Frenzel, and A. Mielke. A gradient system with a wiggly energy and relaxed EDP-convergence. *ESAIM Control Optim. Calc. Var.*, 25(68), 45 pp, 2019.

[DLZ18] K. Disser, M. Liero, and J. Zinsl. Evolutionary Γ-convergence of gradient systems modeling slow and fast chemical reactions. *Nonlinearity*, 31(8), 3689–3706, 2018.

[Mar15] R. Marcelin. Contribution a l'étude de la cinétique physico-chimique. *Annales de Physique*, III, 120–231, 1915.

[Mie11] A. Mielke. A gradient structure for reaction-diffusion systems and for energy-drift-diffusion systems. *Nonlinearity*, 24, 1329–1346, 2011.

[Mie17] A. Mielke. Uniform exponential decay for reaction-diffusion systems with complex-balanced mass-action kinetics. In P. Gurevich, J. Hell, B. Sandstede, and A. Scheel, editors, *Pattern of Dynamics*, Springer Proc. in Math. & Stat. Vol. 205, pages 149–171. Springer, 2017.

[MiS19] A. Mielke and A. Stephan. Coarse graining via EDP-convergence for linear fast-slow reaction systems. *Math. Models Meth. Appl. Sci.*, 30(9), 1765–1807, 2020.

[MMP19] A. Mielke, A. Montefusco, and M. A. Peletier. Exploring families of energy-dissipation landscapes via tilting — three types of EDP convergence. *Contin. Mech. Thermodyn.*, 2020. To appear. WIAS preprint 2668 (arXiv:2001.01455).

[MP*17] A. Mielke, R. I. A. Patterson, M. A. Peletier, and D. R. M. Renger. Non-equilibrium thermodynamical principles for chemical reactions with mass-action kinetics. *SIAM J. Appl. Math.*, 77(4), 1562–1585, 2017.

[MPR14] 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(4), 1293–1327, 2014.

[MPS20] A. Mielke, M. A. Peletier, and A. Stephan. EDP-convergence for nonlinear fast-slow reaction systems with detailed balance, *submitted*, arXiv:2010.08458.

[Yon08] W.-A. Yong. An interesting class of partial differential equations. *J. Math. Phys.*, 49, 033503, 21, 2008.

Fast reaction limits via Γ-convergence of the Flux Rate Functional

Michiel Renger
(joint work with Mark Peletier)

This work is about fast reaction limits a linear systems, studied from a variational perspective. We are given a weighted directed network $(\mathcal{V}, \mathcal{R}, \kappa^\epsilon)$, under the assumptions, see Figure 1:

(1) $\mathcal{R} = \mathcal{R}_{\rm slow} \cup \mathcal{R}_{\rm fast}$ where $\kappa^\epsilon_r = \begin{cases} \kappa_r, & r \in \mathcal{R}_{\rm slow}, \\ \frac{1}{\epsilon}\kappa_r, & r \in \mathcal{R}_{\rm fast}, \end{cases}$

(2) the network is irreducible, and so there exists a unique positive invariant measure $\pi^\epsilon \in \mathcal{P}(\mathcal{V})$,

(3) $\mathcal{V} = \mathcal{V}_0 \cup \mathcal{V}_1$ where $\begin{cases} \pi^\epsilon_x \to \pi_x > 0, & x \in \mathcal{V}_0, \\ \epsilon^{-1}\pi^\epsilon_x \to \pi_x > 0, & x \in \mathcal{V}_1, \end{cases}$.

The flow of mass is described by the discrete continuity equation $\dot\rho^\epsilon(t) = -{\rm div} j^\epsilon(t)$ together with the constitutive law for the fluxes $j^\epsilon_r(t) = \kappa^\epsilon_r \rho^\epsilon_{r-}(t)$, where $r-$ denotes the source node of edge r. Instead of studying the upscaled limit equation, our goal is to prove the Γ-limit of the variational formulation $\mathcal{I}^\epsilon_0(\rho(0)) + \mathcal{J}^\epsilon(\rho, j)$ that corresponds to the flux large deviations of independent

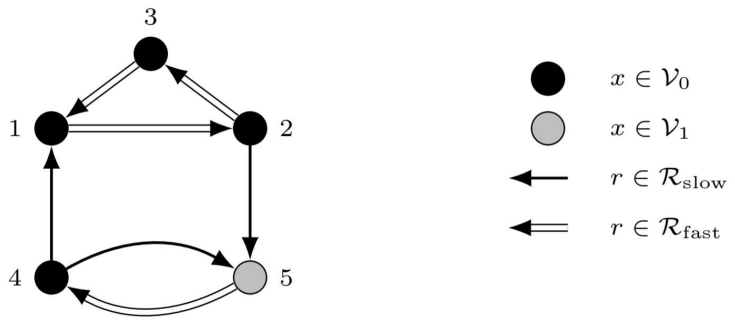

FIGURE 1. An example of a network with slow and fast edges.

copies of a Markov chain on the graph [1, 2, 5]:

$$\mathcal{I}_0^\epsilon(\rho) := \sum_{x\in\mathcal{V}} s(\rho_x \mid \pi_x^\epsilon), \qquad \mathcal{J}^\epsilon(\rho, j) := \sum_{r\in\mathcal{R}} \int_0^T s(j_r(t) \mid \kappa_r^\epsilon \rho^\epsilon(t)),$$

with the usual relative entropy function $s(a|b) = a \log a/b - a + b$. We implicitly set $\mathcal{I}_0^\epsilon + \mathcal{J}^\epsilon = \infty$ if the continuity equation is violated.

We stress that we do not assume that the network is reversible, so that our technique is more general than the gradient flow techniques used for example in [3] and [4].

In order to capture the intrinsic scaling behaviour of the variables we subdivide $\mathcal{R}_{\text{fast}}$ into $\mathcal{R}_{\text{fcyc}}$ for all edges going out of a $\mathcal{V}_0$-node and $\mathcal{R}_{\text{damp}}$ for all edges going out of a $\mathcal{V}_1$-node. All edges in $\mathcal{R}_{\text{fcyc}}$ are part of a connected component/cycle of fast edges. As such we can also subdivide $\mathcal{V}_0$ into all nodes $\mathcal{V}_{0\text{fcyc}}$ that are part of such cycle, and all nodes $\mathcal{V}_{0\text{slow}}$ that are not. The concentrations can easily be rescaled by taking densities, and the fast cycle fluxes will need to be scaled down to prevent blowup:

$$u_x^\epsilon(t) := \frac{\rho_x^\epsilon(t)}{\pi_x^\epsilon}, x \in \mathcal{V}, \qquad j_r^\epsilon(t) =: \tfrac{1}{\epsilon}\kappa_r \rho^\epsilon(t) + \tfrac{1}{\sqrt{\epsilon}}\tilde{\jmath}_r^\epsilon, r \in \mathcal{R}_{\text{fcyc}}.$$

In addition we introduce the density in each connected component $\mathfrak{c} \subset \mathfrak{C}$ of fast (cycle) edges $u_{\mathfrak{c}}^\epsilon(t) := \frac{1}{\pi_{\mathfrak{c}}^\epsilon} \sum_{x\in\mathfrak{c}} \pi_x^\epsilon u_x^\epsilon(t)$, with $\pi_{\mathfrak{c}}^\epsilon = \sum_{x\in\mathfrak{c}} \pi_x^\epsilon$. The reason is that this total density in the component has much better topological properties than the densities u_x^ϵ in each of the nodes $x \in \mathfrak{c} \subset \mathcal{V}_{0\text{slow}}$.

We prove the following equicoercivity result for the rescaled flux-density functional, using a tilde to denote that the variables are rescaled:

Theorem 1. *Let $(u^\epsilon, j^\epsilon)_{\epsilon>0}$ such that $\tilde{\mathcal{I}}_0^\epsilon\big(u^\epsilon(0)\big) + \tilde{\mathcal{J}}^\epsilon(u^\epsilon, j^\epsilon) \le C$ for some $C > 0$. Then there exists a subsequence so that:*

1. *$u^\epsilon_{\mathcal{V}_{0\text{slow}}}$ converges strongly in $C([0,T]; \mathbb{R}^{\mathcal{V}_{0\text{slow}}})$,*
2. *$u^\epsilon_{\mathcal{V}_{0\text{fcyc}}}$ converges weakly-* in $L^\infty([0,T]; \mathbb{R}^{\mathcal{V}_{0\text{fcycle}}})$,*
3. *$u^\epsilon_{\mathfrak{C}}$ converges strongly in $C([0,T]; \mathbb{R}^{\mathfrak{C}})$,*
4. *$u^\epsilon_{\mathcal{V}_1}$ converes narrowly in $\mathcal{M}([0,T]; \mathbb{R}^{\mathcal{V}_1})$*

(5) $j^\epsilon_{\mathcal{R}_{\rm slow}}$ *converges weakly-* in the Orlicz space* $L^{\mathcal{C}}([0,T];\mathbb{R}^{\mathcal{R}_{\rm slow}})$,
(6) $j^\epsilon_{\mathcal{R}_{\rm damp}}$ *converges narrowly in* $\mathcal{M}([0,T];\mathbb{R}^{\rm damp})$,
(7) $j^\epsilon_{\mathcal{R}_{\rm fcyc}}$ *converges weakly-* in the Orlicz space* $L^{\mathcal{C}}([0,T];\mathbb{R}^{\mathcal{R}_{\rm fcyc}})$.

In the topology defined by this result, we prove the convergence of the rescaled functional:

Theorem 2. $\tilde{\mathcal{I}}^\epsilon_0 + \tilde{\mathcal{J}}^\epsilon \xrightarrow[\epsilon\to 0]{\Gamma} \tilde{\mathcal{I}}^0_0 + \tilde{\mathcal{J}}^0$, *where*

$$\tilde{\mathcal{I}}^0_0\big(u(0)\big) := \sum_{x\in\mathcal{V}_{0\rm slow}} s\big(\pi^\epsilon_x u_x(0) \mid \pi^\epsilon_x\big) + \sum_{\mathfrak{c}\in\mathfrak{C}} s\big(\pi^\epsilon_{\mathfrak{c}} u_{\mathfrak{c}}(0) \mid \pi^\epsilon_{\mathfrak{c}}\big),$$

$$\tilde{\mathcal{J}}^0(u,j) := \sum_{r\in\mathcal{R}_{\rm slow}} \int_0^T s\big(j_r(t) \mid \kappa_r \pi_{r\_} u_{r\_}(t)\big)\, dt + \sum_{r\in\mathcal{R}_{\rm damp}} \int_0^T s\big(j_r \mid \kappa_r \tilde{\pi}_{r\_} u_{r\_}\big)\,(dt) + \frac{1}{2} \sum_{r\in\mathcal{R}_{\rm fcycle}} \int_0^T \frac{\tilde{j}_r(t)^2}{\kappa_r \pi_{r\_} u_{r\_}(t)}\, dt,$$

where we implicitly set $\tilde{\mathcal{I}}^0_0 + \tilde{\mathcal{J}}^0 = \infty$ *if the limiting continuity equations (not written here) are violated.*

As a corollary we obtain convergence of the minimisers describing the solutions of the ODE.

Observe that in the above result $s(\cdot \mid \cdot)(dt)$ is a straight-forward generalisation of $s(\cdot \mid \cdot)$ needed to allow for densities $u_{\mathcal{V}_1}$ and fluxes $j_{\mathcal{R}_{\rm damp}}$ that may be measure-valued in time, see Theorem 1. A natural question is then whether these singularities in time can really occur or not. The answer to this question is related to the possible occurrence of damped fluxes that appear in a cycle, see Figure 2:

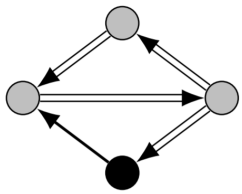

FIGURE 2. An example of a network with slow and fast edges.

The next result shows that the possible occurrence of these singularities or 'spikes' are related to damped cycles.

Theorem 3. *For any convergent sequence* $(u^\epsilon, j^\epsilon) \to (u, j)$ *such that* $\tilde{\mathcal{I}}^\epsilon_0\big(u^\epsilon(0)\big) + \tilde{\mathcal{J}}^\epsilon(u^\epsilon, j^\epsilon) \le C$ *for some* $C > 0$, *we have* $j_r \in L^{\mathcal{C}}([0,T]))$ *if* $r \in \mathcal{R}_{\rm damp}$ *is not part of a cycle of damped fluxes.*

On the other hand, if the network contains a cycle of damped fluxes, then one can construct a convergent sequence $(u^\epsilon, j^\epsilon) \to (u, j)$ *with* $\tilde{\mathcal{I}}^\epsilon_0\big(u^\epsilon(0)\big) + \tilde{\mathcal{J}}^\epsilon(u^\epsilon, j^\epsilon) \le$

C and $\tilde{\mathcal{I}}_0\big(u^\epsilon(0)\big)+\tilde{\mathcal{J}}(u,j)\le C$ for some $C>0$ such that j_r are singular measures in time for all $r\in\mathcal{R}_{\mathrm{damp}}$ in the damped cycle.

References

[1] S. Feng, *Large Deviations for Empirical Process of Mean-Field Interacting Particle System with Unbounded Jumps*, The Annals of Probability **22** (1994), 1679–2274.

[2] C. Léonard, *Large deviations for long range interacting particle systems with jumps*, Annales de l'Institut Henri Poincaré, section B **31** (1995).

[3] K. Disser, M. Liero and J. Zinsl, *Evolutionary Γ-convergence of gradient systems modeling slow and fast chemical reactions*, Nonlinearity **31** (2018), 3689.

[4] A. Mielke and A. Stephan, *Coarse-graining via EDP-convergence for linear fast-slow reaction systems*, ArXiv **1911.06234** (2019).

[5] D.R.M. Renger, *Flux large deviations of independent and reacting particle systems, with implications for Macroscopic Fluctuation Theory*, Journal of Statistical Physics **172** (2018).

The singular set in the Stefan problem.

Alessio Figalli

(joint work with Xavier Ros-Oton, Joaquim Serra)

The Stefan problem, dating back to the XIXth century, is among the most classical and well known free boundary problems and describes phase transitions, such as ice melting to water. More precisely, in its simplest case, the problem consists in finding the evolution of the temperature $\theta(x,t)$ of the water when a block of ice is submerged inside. Then, the function $\theta\ge 0$ satisfies $\partial_t\theta=\Delta\theta$ in the region $\{\theta>0\}$, while the evolution of the free boundary $\partial\{\theta>0\}$ is dictated by the Stefan condition $\partial_t\theta=|\nabla_x\theta|^2$ on $\partial\{\theta>0\}$ —where the gradient is computed from inside $\{\theta>0\}$.

Let χ_A denotes the characteristic function of a set A. After the transformation $u(x,t):=\int_0^t\theta(x,\tau)d\tau$, one can note that $\{u>0\}=\{\theta>0\}$ and the Stefan problem becomes equivalent to the so-called "parabolic obstacle problem":

$$\begin{cases}\partial_t u-\Delta u=-\chi_{\{u>0\}}\\ u\ge 0\\ \partial_t u\ge 0\\ \partial_t u>0 \qquad \text{inside}\quad \{u>0\}\end{cases}\qquad\text{in}\quad \Omega\times(0,T)\subset\mathbb{R}^n\times\mathbb{R}.$$

The regularity of free boundaries for the Stefan problem was developed in 1977 in Caffarelli's groundbreaking paper. The main result therein establishes that the free boundary (i.e., the interface $\partial\{u>0\}$) is C^∞ in space and time, outside some closed set $\Sigma\subset\Omega\times(0,T)$ of *singular points* at which the contact set $\{u=0\}$ has zero density.

A first natural question in understanding the singular set Σ is to estimate its parabolic Hausdorff dimension. In our work in progress, we prove the following sharp bound:

Theorem: *Let $\Omega \subset \mathbb{R}^n$ be an open set, let $u \in L^\infty(\Omega \times (0,T))$ solve the Stefan problem, and let $\Sigma \subset \Omega \times (0,T)$ be the set of singular points. Then* $\dim_{\mathrm{par}}(\Sigma) \leq n-1$.

One could then wonder "how often" singular points arise. In the physical space $\mathbb{R}^3$, we can prove the following:

Theorem: *Let $\Omega \subset \mathbb{R}^3$, let $u \in L^\infty(\Omega \times (0,T))$ solve the Stefan problem, and let*

$$\mathcal{S} := \big\{ t \in (0,T) \ : \ \exists\, (x,t) \in \Sigma \big\}$$

denote the set of "singular times". Then $\dim_{\mathcal{H}}(\mathcal{S}) \leq \frac{1}{2}$. *In particular, for almost every time $t \in (0,T)$, the free boundary is an $(n-1)$-dimensional submanifold of $\mathbb{R}^n$ of class C^∞.*

Reporter: Anastasiia Hraivoronska

Participants

Dr. Jean-David Benamou
INRIA Paris
2, Rue Simone Iff
75012 Paris Cedex
FRANCE

Prof. Dr. Yann Brenier
Département de Mathématiques et Applications,
Ecole Normale Supérieure
Ecole Normale Supérieure
45 rue d'Ulm
FR75005 Palaiseau
FRANCE

Prof. Dr. Martin Burger
Department Mathematik
Universität Erlangen-Nürnberg
Cauerstrasse 11
91058 Erlangen
GERMANY

Prof. Dr. Jose Antonio Carrillo de la Plata
Mathematical Institute
University of Oxford
Oxford OX2 6GG
UNITED KINGDOM

Prof. Dr. Katy Craig
Department of Mathematics
University of California at Santa Barbara
South Hall
Santa Barbara, CA 93106
UNITED STATES

Dr. Sara Daneri
Gran Sasso Science Institute (GSSI)
Viale Francesco Crispi, 7
67100 L'Aquila
ITALY

Dr. Matthias Erbar
Institut für Angewandte Mathematik
Universität Bonn
Endenicher Allee 60
53115 Bonn
GERMANY

Dr. Antonio Esposito
Department Mathematik
Universität Erlangen-Nürnberg
Cauerstrasse 11
91058 Erlangen
GERMANY

Dr. Max Fathi
Institut de Mathématiques de Toulouse
Université Paul Sabatier
118, route de Narbonne
31062 Toulouse Cedex 9
FRANCE

Prof. Dr. Jin Feng
Department of Mathematics
University of Kansas
405 Snow Hall
Lawrence KS 66045-7594
UNITED STATES

Prof. Dr. Alessio Figalli
Department Mathematik
ETH Zürich
ETH-Zentrum, HG G 45.2
Rämistrasse 101
8092 Zürich
SWITZERLAND

Prof. Dr. Adriana Garroni
Dipartimento di Matematica
"Guido Castelnuovo"
Università di Roma "La Sapienza"
Piazzale Aldo Moro, 2
00185 Roma
ITALY

Dr. Martin Heida
Weierstraß-Institut für
Angewandte Analysis und Stochastik
Mohrenstrasse 39
10117 Berlin
GERMANY

Jasper Hoeksema
Department of Mathematics and
Computer Science
Eindhoven University of Technology
MetaForum 7.123
P.O. Box 513
5600 MB Eindhoven
NETHERLANDS

Dr. Franca Hoffmann
Department of Computing and
Mathematical Sc.
California Institute of Technology
1200 E. California Blvd., MC 305-16
Pasadena CA, 91125-2100
UNITED STATES

Katharina Hopf
Weierstraß-Institut für
Angewandte Analysis und Stochastik
Mohrenstrasse 39
10117 Berlin
GERMANY

Anastasiia Hraivoronska
Department of Mathematics and
Computer Science
Eindhoven University of Technology
P.O. Box 513
5600 MB Eindhoven
NETHERLANDS

Prof. Dr. Mikaela Iacobelli
HG G 48.2
Departement Mathematik
ETH-Zentrum
Rämistr. 101
8092 Zürich
SWITZERLAND

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

Prof. Dr. Dorothee Knees
FB 10 - Mathematik und
Naturwissenschaften
Institut für Mathematik
Universität Kassel
Heinrich-Plett-Strasse 40
34132 Kassel
GERMANY

Prof. Dr. Christian Leonard
Université Paris-Nanterre
Bâtiment Maurice Allais (bât. G)
200, Avenue de la République
92000 Nanterre
FRANCE

Dr. Matthias Liero
Weierstraß-Institut für
Angewandte Analysis und Stochastik
Mohrenstrasse 39
10117 Berlin
GERMANY

Prof. Dr. Jian-Guo Liu
Department of Mathematics
Duke University
P.O.Box 90320
3806 Northampton Road
Durham NC 27708-0320
UNITED STATES

Prof. Jan Maas
IST Austria
Am Campus 1
3400 Klosterneuburg
AUSTRIA

Prof. Dr. Daniel Matthes
Zentrum Mathematik
Technische Universität München
Boltzmannstrasse 3
85747 Garching bei München
GERMANY

Prof. Dr. Alexander Mielke
Weierstraß-Institut für
Angewandte Analysis und Stochastik
Mohrenstrasse 39
10117 Berlin
GERMANY

Prof. Dr. Maria Giovanna Mora
Dipartimento di Matematica
Università di Pavia
Via Ferrata, 1
27100 Pavia
ITALY

Dr. Robert I.A. Patterson
Weierstraß-Institut für
Angewandte Analysis und Stochastik
Mohrenstrasse 39
10117 Berlin
GERMANY

Prof. Dr. Mark A. Peletier
Department of Mathematics and
Computer Science
Eindhoven University of Technology
MetaForum 7.070
P.O. Box 513
5600 MB Eindhoven
NETHERLANDS

Dr. Dirk Peschka
Weierstraß-Institut für Angewandte
Analysis und Stochastik
Mohrenstrasse 39
10117 Berlin
GERMANY

Dr. Michiel Renger
Weierstraß-Institut für
Angewandte Analysis und Stochastik
Mohrenstrasse 39
10117 Berlin
GERMANY

Prof. Riccarda Rossi
Department of Mechanical and
Industrial Engineering
University of Brescia
Via Branze 38
25133 Brescia
ITALY

Prof. Tomas Roubicek
Mathematical Institute
Charles University
Sokolovska 83
186 75 Praha 8
CZECH REPUBLIC

Dr. Matthias Ruf
École Polytechnique Fédérale de
Lausanne
MA B1 457 (Bâtiment MA)
Station 8
1015 Lausanne
SWITZERLAND

Prof. Giuseppe Savare
Dipartimento di Matematica
"F. Casorati"
Universita di Pavia
Via Ferrata, 1
27100 Pavia
ITALY

Dr. Lucia Scardia
Department of Mathematics
Heriot-Watt University
Riccarton
Edinburgh EH14 4AS
UNITED KINGDOM

Prof. Dr. André Schlichting
Institut für Angewandte Mathematik
Universität Bonn
Endenicher Allee 60
53115 Bonn
GERMANY

Mikola Schlottke
Department of Mathematics
Eindhoven University of Technology
P.O.Box 513
5600 MB Eindhoven
NETHERLANDS

Dr. Upanshu Sharma
Department of Mathematics and Computer Science, Institute of Mathematics, Freie Universität Berlin
Arnimallee 9
14195 Berlin
GERMANY

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
Weierstraß-Institut für Angewandte Analysis und Stochastik
Mohrenstraße 39
10117 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 and Computer Science
Eindhoven University of Technology
P.O. Box 513
5600 MB Eindhoven
NETHERLANDS

Prof. Dr. Maria G. Westdickenberg
Lehrstuhl für Angewandte Analysis
RWTH Aachen
Gebäude 1950, Rm. 214
Pontdriesch 14-16
52062 Aachen
GERMANY

Dr. Marie-Therese Wolfram
Mathematics Institute
University of Warwick
Zeeman Building
Coventry CV4 7AL
UNITED KINGDOM

Prof. Dr. Johannes Zimmer
Fakultät für Mathematik (M6)
TU München
Boltzmannstr. 3
85748 Garching bei München
GERMANY