Otto Calculus and Gradient Flows on the Infinite-Dimensional Manifold of Probability Measures

Mara-Ioana Postolache

October 14, 2024

Thank you to

- My supervisors, Prof Clément Mouhot and his PhD student, Amélie Loher, for their guidance and support throughout this entire project.
- My college, Peterhouse, for the free accomodation over the summer.
- Sally Baume, for being a legend and keeping this whole programme running.
- The donors of the SRIM funding, without whom this project would not have been possible.
- Everyone in this room today, for having the patience to listen to me for the next 20-ish minutes!

Overview

PDEs we want to consider

We care about evolutionary PDEs of the form:

$$\frac{\partial \rho}{\partial t} = \nabla \cdot \left(\rho \nabla \left(U'(\rho) + V + W * \rho\right)\right)$$

Here, the quantities are as follows:

- x is a spatial variable in $\mathbb{R}^d$
- t is a time variable in $[0, \infty)$
- $\rho(t, x)$ is a time-dependent (non-negative) probability density on $\mathbb{R}^d$
- $U : \mathbb{R}^+ \to \mathbb{R}$ is a density of internal energy
- $V : \mathbb{R}^d \to \mathbb{R}$ is a confinement potential
- $W : \mathbb{R}^d \to \mathbb{R}$ is an interaction potential, which we assume to be symmetric: $W(z) = W(-z)$

The Porous Medium Equation

The Porous Medium Equation (PME) is the following:

$$\frac{\partial \rho}{\partial t} = \Delta \rho^m \qquad \text{(PME)}$$

For simplicity, we'll only consider the case $m > 1$ in this talk.
This equation arises as the general PDE we described before by considering

$$U(s) = \frac{1}{m-1} s^m, \qquad V(x) = 0, \qquad W = 0$$

However, since we will need V to be strictly convex, we will introduce a change of variables: set $\lambda = \frac{1}{d(m-1)+2}$, $\widetilde{x} = x t^{-\lambda}$, $\widetilde{t} = \ln t$, and define

$$\hat{\rho}\left(\widetilde{t}, \widetilde{x}\right) = e^{d\lambda \widetilde{t}} \rho(t, x) = t^{d\lambda} \rho(t, x)$$

The rescaled PME

Dropping the $\sim$'s, the PME becomes:

$$\frac{\partial \hat{\rho}}{\partial t} = \Delta \hat{\rho}^m + \lambda \nabla \cdot (\hat{\rho} x) = 0 \tag{PME-R}$$

This now arises by considering

$$U(s) = \frac{1}{m-1} s^m, \qquad V(x) = \frac{1}{2}\lambda |x|^2, \qquad W = 0$$

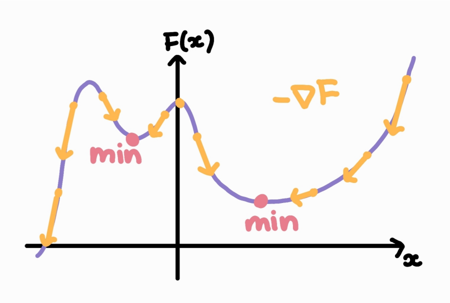
Clearly now V is strictly convex, with $\mathrm{D}^2 V = \lambda \, \mathrm{Id}$, $\lambda > 0$. We also claim that all the other assumptions we need for the result hold for these specific choices of U, V, W.
Thus, our convergence results (which we'll introduce later) will hold for this equation.

Gradient flows in $\mathbb{R}^n$

Definition (Gradient flow in $\mathbb{R}^n$)

Let $F : \mathbb{R}^n \to \mathbb{R}$ be a smooth function. A **gradient flow** of F is a curve given by:

$$\dot{x}(t) = -\nabla F(x(t))$$



This is also sometimes called the **steepest descent** curve of F.

The manifold of probability densities

Consider the set:

$$\mathcal{M} = \mathcal{P}_2^a(\mathbb{R}^d) = \left\{ \rho : \mathbb{R}^d \to [0, \infty) \,\middle|\, \int \rho = 1, \int |x|^2 \, \mathrm{d}\rho(x) < \infty \right\}$$

By handwaving, we can consider this as an *infinite-dimensional* version of a manifold, with the tangent space at any point given by

$$T_\rho \mathcal{M} = \left\{ s : \mathbb{R}^d \to \mathbb{R} \,\middle|\, \int s = 0 \right\}$$

This makes sense since for any curve $\rho(t)$ on $\mathcal{M}$, the derivative $\frac{\mathrm{d}\rho}{\mathrm{d}t}$ must satisfy

$$\int \frac{\mathrm{d}\rho}{\mathrm{d}t} = \frac{\mathrm{d}}{\mathrm{d}t} \int \rho = 0$$

The tangent space revisited and the metric

With some further handwaving, we use the equation $s = -\nabla \cdot (\rho \nabla p)$ identify the tangent space at ρ with the following space:

$$T_\rho \mathcal{M} = \left\{ p : \mathbb{R}^d \to \mathbb{R} \right\} / \sim$$

where $p_1 \sim p_2$ if and only if they differ by a constant.
We can now define a *Riemannian-like* metric on the "manifold" $\mathcal{M}$ by:

$$g_\rho(s_1, s_2) = \int_{\mathbb{R}^d} \nabla p_1 \cdot \nabla p_2 \, \mathrm{d}\rho$$

This metric is defined in such a way that we have the relation:

$$g_\rho(s_1, s_2) = \int_{\mathbb{R}^d} s_1 p_2 \, \mathrm{d}x = \int_{\mathbb{R}^d} s_2 p_1 \, \mathrm{d}x$$

Gradient flows $(\mathcal{M}, g)$

Definition (Gradient flow on a manifold)

Let $F : \mathcal{M} \to \mathbb{R}$ be a smooth functional. A **gradient flow** of F is a curve given by the evolution equation:

$$\dot{\gamma}(t) = -\operatorname{grad} F \,|_{\gamma(t)}$$

The gradient of F is defined as the unique vector field along γ such that

$$g_{\gamma}\left(\operatorname{grad} F\,|_{\gamma}, s\right) = \operatorname{diff} F\,|_{\gamma} .s \qquad \forall \text{vector fields } s \text{ along } \gamma$$

and hence the definition of gradient flow can be re-written as:

$$g_{\gamma}\left(\dot{\gamma}(t), s\right) = -\operatorname{diff} F\,|_{\gamma} .s \qquad \forall \text{vector fields } s \text{ along } \gamma$$

The free energy functional

It turns out that solutions to the PDE

$$\frac{\partial \rho}{\partial t} = \nabla \cdot \left(\rho \nabla \left(U'(\rho) + V + W * \rho\right)\right) \tag{1}$$

are the same as gradient flows on $\mathcal{M}$ of the following functional:

Definition (Free energy functional)

The **free energy functional** of the density ρ is given by

$$F(\rho) = \int_{\mathbb{R}^d} U(\rho)\mathrm{d}x + \int_{\mathbb{R}^d} V(x)\,\mathrm{d}\rho(x) + \frac{1}{2}\int_{\mathbb{R}^d \times \mathbb{R}^d} W(x-y)\,\mathrm{d}\rho(x)\,\mathrm{d}\rho(y)$$

Proof of being a gradient flow

By IBP and differentiation under the integral sign, we have

$$\operatorname{diff} F\,|_{\rho}\,.s = -\int_{\mathbb{R}^d} \nabla \cdot \left(\rho \nabla \left(U'(\rho) + V + W * \rho\right)\right) p \,\mathrm{d}x \quad \forall \text{v.f. } s \text{ along } \rho$$

Thus, using the formula $g_\rho(s_1, s_2) = \int_{\mathbb{R}^d} s_1 p_2 \,\mathrm{d}x$, the gradient flow equation

$$g_\gamma\left(\frac{\mathrm{d}\rho}{\mathrm{d}t}, s\right) = -\operatorname{diff} F\,|_{\rho}\,.s \qquad \forall \text{vector fields } s \text{ along } \rho$$

becomes

$$\int_{\mathbb{R}^d} \frac{\mathrm{d}\rho}{\mathrm{d}t} p \,\mathrm{d}x = \int_{\mathbb{R}^d} \nabla \cdot \left(\rho \nabla \left(U'(\rho) + V + W * \rho\right)\right) p \,\mathrm{d}x \quad \forall p$$

But since p is arbitrary, this is equivalent to the PDE we started with.

Geodesics and the Wasserstein distance

It is known that between any two points ρ_0 and ρ_1 on the manifold $(\mathcal{M}, g)$, there is a unique geodesic connecting them, which we denote by ρ_s (for $s \in [0, 1]$).

Furthermore, the most important property of the geometric structure we introduced is the fact that the geodesic distance (i.e. the length of ρ_s) on $(\mathcal{M}, g)$ recovers a known quantity, namely the Wassterstein distance W_2.

Statement of the convergence result

Assume that a number of small assumptions on the convexity of U and W and the strict convexity of V (i.e. there exists some $\lambda > 0$ such that $\mathrm{D}^2 V \geq \lambda \, \mathrm{Id}$) are satisfied.
Let ρ_∞ be the (unique) steady-state (i.e. time-independent) solution of the PDE, and denote $F(\rho|\rho_\infty) = F(\rho) - F(\rho_\infty)$.
Then any solution $\rho(t)$ of the PDE satisfies the following convergence results:

$$F(\rho|\rho_\infty) \leq e^{-2\lambda t} F(\rho_0|\rho_\infty)$$

and

$$W_2(\rho(t), \rho_\infty) \leq e^{-\lambda t} \sqrt{\frac{2F(\rho_0|\rho_\infty)}{\lambda}}$$

So, under the well-known Wasserstein distance, any solution of the PDE converges exponentially (in time) to the steady state.

Idea of the proof

Key fact

Under certain convexity assumptions on the potentials U, V, W, the free energy functional F becomes *uniformly convex* on $\mathcal{M}$ i.e. there exists some $\lambda > 0$ such that

$$\frac{\mathrm{d}^2 F(\rho_s)}{\mathrm{d}s^2} \geq \lambda W_2^2(\rho_0, \rho_1)$$

for any geodesic ρ_s between ρ_0 and ρ_1.

More precisely, we can show that, assuming U obeys a convexity under rescaling property and that W is a convex function, the terms of the functional F which correspond to U and W are *convex*, while assuming that V is uniformly convex then implies uniform convexity of the corresponding term in F, and thus of F itself.

Idea of the proof (cont)

We will not go into details of why a minimiser of F exists, but we note that F being uniformly convex means that any minimiser must be unique. We denote it by ρ_∞.

Proof (Uniqueness).

If ρ_0 and ρ_1 are two distinct minimisers, then $s \mapsto F(\rho_s)$ is uniformly convex on $[0,1]$ and achieves it's minimum at the two endpoints, which is nonsense - contradiction! □

The remaining of the proof relies on Taylor expanding $F(\rho)$ to second order around ρ_∞ and using the uniform convexity of F to bound the second derivative of F.

This leads us to a bound on the first derivative:

$$-\frac{\mathrm{d}}{\mathrm{d}t}F(\rho(t)|\rho_\infty) \geq 2\lambda F(\rho|\rho_\infty)$$

which, used with Gronwall's inequality, gives us the desired exponential convergence in F.

Closing remarks

- The geometric picture inspired a gradient flow approach to the PDE, leading us to conclude convergence results from bounds of the Hessians of a functional.
- However, the definition of $(\mathcal{M}, g)$ significantly restricts the type of PDEs we can consider while keeping the functional F still *nice*.
- When looking at other types of PDEs, such as the Landau equation, we will need a different structure on the space. Some progress has been made abstractly, by replacing the Wasserstein distance with a more complicated distance, but (as far as I am aware) this doesn't yet have a nice manifold interpretation.

Thank you for your attention!

Any questions?

[ABS21] Luigi Ambrosio, Elia Brué, and Daniele Semola. *Lectures on Optimal Transport*. Vol. 130. UNITEXT. Cham: Springer International Publishing, 2021. ISBN: 978-3-030-72161-9 978-3-030-72162-6. DOI: 10.1007/978-3-030-72162-6.

[CMV03] José A. Carrillo, Robert J. McCann, and Cédric Villani. "Kinetic Equilibration Rates for Granular Media and Related Equations: Entropy Dissipation and Mass Transportation Estimates". In: *Revista Matemática Iberoamericana* 19.3 (Dec. 2003), pp. 971–1018. ISSN: 0213-2230, 2235-0616. DOI: 10.4171/rmi/376.

[Ott01] Felix Otto. "THE GEOMETRY OF DISSIPATIVE EVOLUTION EQUATIONS: THE POROUS MEDIUM EQUATION". In: *Communications in Partial Differential Equations* 26.1-2 (Jan. 2001), pp. 101–174. ISSN: 0360-5302, 1532-4133. DOI: 10.1081/PDE-100002243.