A Lagrangian Scheme for the solution of nonlinear diffusion equations

Bertram Düring

Mathematics Institute
University of Warwick

Joint work with J.A. Carrillo,
D. Matthes & D. McCormick

Nonlinear diffusion equations

Nonlinear diffusion partial differential equations with a Wasserstein gradient flow structure have received rapidly growing attention.

Well-known examples are

- porous medium equation
- fast diffusion equation
- lubrication equations describing thin viscous films
- fluid-type quantum models for semiconductors

Apart from their obvious relevance in theoretical physics and engineering applications, they are of great interest in mathematical analysis:

- behaviour of their solutions is very rich
- open questions on qualitative properties of the solutions
- accurate and efficient numerical solution is challenging

Motivation: structure-preserving discretisations

When it comes to solving nonlinear diffusion equations numerically, it is natural to ask for schemes which preserve certain properties at a discrete level:

- positivity-preserving
- mass-preserving
- energy-dissipating
- gradient-flow structure
- ...

Large numerics literature (non-exhaustive list)

Many(!) approaches tackling nonlinear diffusions numerically,

- FEM, in part. Cahn-Hilliard, Allen-Cahn, ... [Elliott '86–], [Barrett], [Garcke], [Styles] & co-workers
- Particle methods based on suitable regularizations of the flux of the continuity equation [Degond-Mustieles '90], [Russo '90], [Lions-Mas-Gallic '01], [Mas-Gallic '02]
- discrete self-similar solutions for PME [Budd et al. '98, ...]
- high-resolution schemes for nonlinear convection-diffusion problems [Kurganov-Tadmor '00].
- high-order relaxation schemes [Cavalli et al. '07]
- FV methods preserving decay of energy at semi-discrete level (non-negativity, mass conservation) [Bessemoulin-Chatard-Filbet '12], [Cances-Guichard '16], [Carrillo et al. '15].
- blob methods [Carrillo et al. '17a,'17b]

Class of nonlinear diffusion equations

In this talk consider the following class of equations

$$\partial_t \rho = \Delta P(\rho) + \nabla \cdot (\rho \nabla V) \qquad \text{on } \mathbb{R}_{>0} \times \mathbb{R}^d,$$
$$\rho(\cdot, 0) = \rho^0 \qquad \text{on } \mathbb{R}^d.$$

where $P(r) = rh'(r) - h(r)$ for all $r \geq 0$ with some non-negative and convex $h \in C^1(\mathbb{R}_{\geq 0}) \cap C^\infty(\mathbb{R}_{>0})$, and a non-negative potential $V \in C^2(\mathbb{R}^d)$.

This encompasses large class of diffusion equations, e.g.

- $P(r) = r$: heat equation
- $P(r) = r^m, m > 1$: porous medium equation
- $P(r) = r^m, m < 1$: fast diffusion equation

Lagrangian formulation

Equation can be written as a transport equation,

$$\partial_t \rho + \nabla \cdot \big(\rho \, \mathbf{v}[\rho]\big) = 0,$$

with a velocity field $\mathbf{v}$ that depends on the solution ρ itself,

$$\mathbf{v}[\rho] = -\nabla\big(h'(\rho) + V\big).$$

Note: further evolution equations can be written in this form, e.g. non-local aggregation equations [Ambrosio et al. '08], Keller-Segel type models [Blanchet et al. '08], fourth-order thin film equations [Otto '98] or quantum diffusion equations [Gianazza et al. '09].

Variational structure: Wasserstein gradient flow

A celebrated results is (see [Otto '98] or [Ambrosio et al.'01]) that this problem is a gradient flow for the relative Renyi entropy functional

$$\mathcal{E}(\rho) = \int_{\mathbb{R}^d} \left[h(\rho(x)) + V(x)\rho(x) \right] \mathrm{d}x,$$

with respect to the L^2-Wasserstein metric on the space $\mathcal{P}_2^{\mathrm{ac}}(\mathbb{R}^d)$ of probability densities on $\mathbb{R}^d$ with finite second moment.

An important consequence (see [JKO 98], [Ambrosio et al. '08]) is that the unique flow can be obtained as the limit for $\tau \searrow 0$ of the time-discrete minimizing movement scheme

$$\rho_\tau^n := \operatorname*{argmin}_{\rho \in \mathcal{P}_2^{\mathrm{ac}}(\mathbb{R}^d)} \mathcal{E}_\tau(\rho; \rho_\tau^{n-1}), \quad \mathcal{E}_\tau(\rho, \hat{\rho}) := \frac{1}{2\tau} \mathrm{W}_2(\rho, \hat{\rho})^2 + \mathcal{E}(\rho).$$

Numerical scheme based on minimizing movements

The minimizing movement scheme

$$\rho_\tau^n := \operatorname*{argmin}_{\rho \in \mathcal{P}_2^{\mathrm{ac}}(\mathbb{R}^d)} \mathcal{E}_\tau(\rho; \rho_\tau^{n-1}), \quad \mathcal{E}_\tau(\rho, \hat{\rho}) := \frac{1}{2\tau} \mathrm{W}_2(\rho, \hat{\rho})^2 + \mathcal{E}(\rho).$$

has originally been used as a tool for the analysis of the equations.

Q: can it be the basis for a practical, structure-preserving discretisation to approximate solutions of the nonlinear diffusion equations?

Related results in the literature

The numerical approximation of the minimizing movement scheme has been tackled by different methods:

- using pseudo-inverse distributions in one dimension, e.g. in [Carrillo-Toscani '05], [Blanchet-Calvez-Carrillo '08], [Carrillo-Moll '09], [Westdickenberg-Wilkening '10]
- solving for the optimal map in a minimizing movement step [Benamou et al. '15], [Junge et al. '15]
- methods in one dimension for higher-order, drift diffusion and Fokker–Planck equations in [Düring et al. '10], [Matthes-Osberger '14,'15a,'15b]
- and many more...

$\rightsquigarrow$ remains challenging in higher space dimensions

Aim: Developing a structure-preserving algorithm based on minimizing movement scheme in multiple space dimensions

Lagrangian formulation

Let ρ be a smooth positive solution of the transport equation, and $\overline{\rho}$ a *reference density*, i.e. a probability density supported on some compact set $K \subset \mathbb{R}^d$. Let $G_{\#}\overline{\rho}$ denote the *push-forward* of $\overline{\rho}$ under a map $G\colon K \to \mathbb{R}^d$.
Now, let $G^0\colon K \to \mathbb{R}^d$ be a given map such that $G^0_{\#}\overline{\rho} = \rho^0$.
Further, let $G\colon [0,T] \times \mathbb{R}^d \to \mathbb{R}^d$ be the flow map associated to the transport

$$\partial_t G_t = \mathbf{v}[\rho_t] \circ G_t, \quad G(0,\cdot) = G^0,$$

where $\rho_t := \rho(t,\cdot)$ and $G_t := G(t,\cdot)\colon \mathbb{R}^d \to \mathbb{R}^d$.
Then, one can show that at any $t \in [0,T]$,

$$\rho_t = (G_t)_{\#}\overline{\rho}$$

$\rightsquigarrow$ solution G is a Lagrangian map for the solution ρ

Evolution equation for G and L^2 gradient flow

We can now insert $\rho_t = (G_t)_{\#}\overline{\rho}$ for ρ in the expression for the velocity, $\mathbf{v}[\rho] = -\nabla\big(h'(\rho) + V\big)$, and obtain an evolution equation for G:

$$\partial_t G_t = -\nabla \left[h' \left(\frac{\overline{\rho}}{\det \mathrm{D}G_t} \right) \right] \circ G_t - \nabla V \circ G_t.$$

Moreover (see [Evans et al. '05], [Carrillo/Moll '09], [Carrillo/Lisini '10]), this is also a gradient flow, namely for the functional

$$\mathbf{E}(G|\overline{\rho}) := \mathcal{E}(G_{\#}\overline{\rho}) = \int_K \left[\widetilde{h} \left(\frac{\det \mathrm{D}G}{\overline{\rho}} \right) + V \circ G \right] \overline{\rho} \, \mathrm{d}\omega,$$

with $\widetilde{h}(s) := s\, h(s^{-1})$ on the Hilbert space $L^2(K \to \mathbb{R}^d; \overline{\rho})$ of square integrable maps from K to $\mathbb{R}^d$.

$\rightsquigarrow$ related approach in [Carrillo-Moll '09], [Carrillo et al. '16] who discretise the above equation by FD/FEM

Minimizing movement scheme for L^2 gradient flow

[Ambrosio, Lisini and Savaré '06] proved that the gradient flow for

$$\mathbf{E}(G|\overline{\rho}) := \mathcal{E}(G_{\#}\overline{\rho}) = \int_K \left[\widetilde{h}\left(\frac{\det \mathrm{D}G}{\overline{\rho}}\right) + V \circ G\right] \overline{\rho}\, \mathrm{d}\omega,$$

is globally well-defined, and can again be approximated by the minimizing movement scheme:

$$G_\tau^n := \underset{G \in L^2(K \to \mathbb{R}^d; \overline{\rho})}{\operatorname{argmin}} \mathbf{E}_\tau\big(G; G_\tau^{n-1}\big),$$

$$\mathbf{E}_\tau(G; \hat{G}) = \frac{1}{2\tau} \int_K \|G - \hat{G}\|^2 \, \mathrm{d}\overline{\rho} + \mathbf{E}(G|\overline{\rho}).$$

$\rightsquigarrow$ in the following we present a discretize-then-optimize where we adapt this minimizing movement scheme for a numerical algorithm

Idea of the discretize-then-optimize algorithm in 2d

For simplicity; restrict ourselves to 2d in the following

- Spatial discretisation: triangulation in $\mathbb{R}^2$
- ansatz space $\mathcal{A}_{\mathscr{T}}$ for G: on each triangle $\Delta_m \subset K$, let $G(\omega) = A_m\omega + b_m$ for some matrix $A_m \in \mathbb{R}^{2\times 2}$ and some vector $b_m \in \mathbb{R}^2$
- this affine ansatz for G corresponds to piecewise constant ansatz for its derivatives $g := \mathrm{D}G$
  $\rightarrow$ density function ρ is piecewise constant
- define inductively discrete maps $G_{\boxplus}^n \in \mathcal{A}_{\mathscr{T}}$ by solution of the minimisation problems

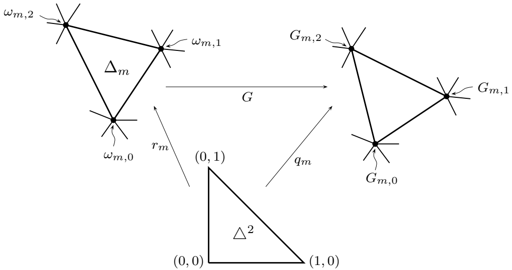
$$G_{\boxplus}^n := \operatorname*{argmin}_{G\in\mathcal{A}_{\mathscr{T}}} \mathbf{E}_{\boxplus}\big(G; G_{\boxplus}^{n-1}\big),$$

$$\text{with } \mathbf{E}_{\boxplus}(G; G^*) = \frac{1}{2\tau}\|G - G^*\|^2_{L^2(K;\overline{\rho}_{\mathscr{T}})} + \mathbf{E}(G|\overline{\rho}_{\mathscr{T}}).$$

Discrete maps in 2d

reference triangulation (fixed)

triangulation related to $G^n_{\boxplus}$ (changes in time)



standard triangle

We introduce the linear interpolation maps

$r_m\colon \triangle^2 \to K$, $r_m(\xi) = \omega_{m,0} + \sum_{j=1}^{2} (\omega_{m,j} - \omega_{m,0})\xi_j$, and

$q_m\colon \triangle^2 \to \mathbb{R}^2$, $q_m(\xi) = G_{m,0} + \sum_{j=1}^{2} (G_{m,j} - G_{m,0})\xi_j$

$\rightsquigarrow$ the affine map equals to $G_m(\omega) = q_m \circ r_m^{-1}$

Derivation of the discrete minimisation problem

We introduce the linear interpolation maps
$r_m \colon \triangle^2 \to K$, $r_m(\xi) = \omega_{m,0} + \sum_{j=1}^{2}(\omega_{m,j} - \omega_{m,0})\xi_j$, and
$q_m \colon \triangle^2 \to \mathbb{R}^2$, $q_m(\xi) = G_{m,0} + \sum_{j=1}^{2}(G_{m,j} - G_{m,0})\xi_j$
$\rightsquigarrow$ the affine map equals to $G_m(\omega) = q_m \circ r_m^{-1}$

We have

$$\det A_m = \frac{\det \mathrm{D}q_m}{\det \mathrm{D}r_m} = \frac{\det Q_{\mathscr{T}}^m[G]}{2|\Delta_m|}$$

where $Q_{\mathscr{T}}^m[G] := \big(G_{m,1} - G_{m,0} \big| G_{m,2} - G_{m,0}\big)$

$\rightsquigarrow$ we can express the maps via the coordinates of the triangulation

Derivation of the discrete minimisation problem II

Substitution of the special form $G(\omega) = A_m\omega + b_m$ produces

$\mathbf{E}(G|\overline{\rho}_{\mathscr{T}}) = \sum_{\Delta_m \in \mathscr{T}} \mu^m_{\mathscr{T}} \left[\mathbb{H}^m_{\mathscr{T}}(G) + \mathbb{V}^m_{\mathscr{T}}(G)\right]$

with internal energy $\mathbb{H}^m_{\mathscr{T}}(G) := \widetilde{h}\left(\frac{\det A_m}{\overline{\rho}^m_{\mathscr{T}}}\right) = \widetilde{h}\left(\frac{\det Q^m_{\mathscr{T}}[G]}{2\mu^m_{\mathscr{T}}}\right)$

and potential energy

$$\mathbb{V}^m_{\mathscr{T}}(G) = ⨍_{\Delta_m} V(A_m\omega + b_m)\,\mathrm{d}\omega = ⨍_{\triangle} V\big(r_m(\omega)\big)\,\mathrm{d}\omega.$$

Further, we can show

$$\|G - G^*\|^2_{L^2(K;\overline{\rho}_{\mathscr{T}})} = \int_K \|G - G^*\|^2 \overline{\rho}_{\mathscr{T}}\,\mathrm{d}\omega = \sum_{\Delta_m \in \mathscr{T}} \mu^m_{\mathscr{T}} \mathbb{L}^m_{\mathscr{T}}(G, G^*)$$

where

$$\mathbb{L}^m_{\mathscr{T}}(G, G^*) := ⨍_{\Delta_m} \|G(\omega) - G^*(\omega)\|^2\,\mathrm{d}\omega = ⨍_{\triangle} \|r_m(\omega) - r^*_m(\omega)\|^2\,\mathrm{d}\omega$$
$$= \frac{1}{6} \sum_{0 \le i \le j \le 2} (G_{m,i} - G^*_{m,i}) \cdot (G_{m,j} - G^*_{m,j})$$

Practical algorithm for finding minimizers

- Now we can compute the discrete Euler-Lagrange equations (for each node in the triangulation)
- outer (time stepping) and inner (Newton) iteration
- initialising $G^{(0)} := G^{n-1}$ with G^{n-1}, the solution at previous time step, define inductively

$$G^{(s+1)} := G^{(s)} + \delta G^{(s+1)},$$

where the update $\delta G^{(s+1)}$ is the solution to

$$\mathbf{H}[G^{(s)}]\delta G^{(s+1)} = -\mathbf{Z}[G^{(s)}; G^{n-1}].$$

- if norm of $\delta G^{(s+1)}$ drops below given stopping criterion, define $G^n := G^{(s+1)}$ as approximate solution in the nth time step.
- effort of each inner iteration step is essentially determined by the effort to invert the sparse Hessian matrix $\mathbf{H}$

Numerical example: porous medium equation

Example: porous medium equation

$$\partial_t u = \operatorname{div}\Big(u\nabla\Big(\frac{m}{m-1}u^{m-1}\Big)\Big) = \Delta u^m$$

Behaviour in the long-time limit $t \to \infty$:

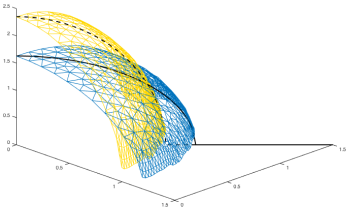
$$u^*(t,x) = t^{-d\alpha}\mathcal{B}\big(t^{-\alpha}x\big) \quad \text{with} \quad \alpha = \frac{1}{d+(m-1)},$$

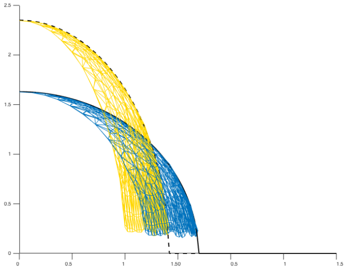
where $\mathcal{B}$ is the Barenblatt profile

$$\mathcal{B}(z) = \big(C - \frac{(m-1)\alpha}{4m}\|z\|^2\big)_+^{\frac{1}{m-1}}$$

Numerical example: porous medium equation

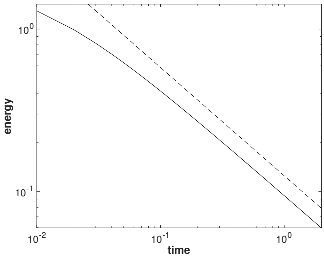
Evolution of the support of the Barenblatt profile ($m = 3$):

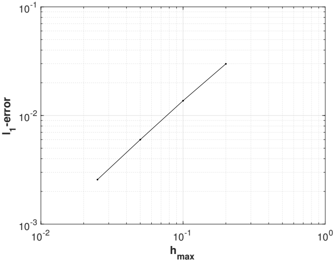




→ good agreement between analytical and numerical solution, at least visually

Numerical example: porous medium equation





Decay of the energy of the discrete solution in comparison with the analytical decay $t^{-2/3}$ of the Barenblatt solution (left). Numerical convergence for fixed ratio $\tau/h^2_{\max} = 0.4$ (right).

Numerical analysis: overview of results

- sequence of fully discrete minimization problems is well-posed, we obtain sequence $(G^n_\boxplus)_{n=0,1,\ldots}$ for each sufficiently fine discretization $\boxplus$.
- induced densities $\widetilde{\rho}_\boxplus$ converge weakly to an absolutely continuous limit trajectory ρ
- fluxes $\widetilde{\rho}_\boxplus \widetilde{\mathbf{v}}_\boxplus$ converge weakly to a limit of the form $\rho\mathbf{v}$
- identification of the limit velocity $\mathbf{v}$, however, is only possible under strong additional hypotheses
- for $d = 2$, we prove numerical consistency, i.e. if G is a smooth solution, then its restriction to the mesh $\boxplus$ satisfies fully discrete Euler–Lagrange equations associated to the minimizing movement scheme, with a quantifiable error that vanishes in suitable continuous limit (requires certain assumptions on triangulation)

Numerical example: porous medium equation

A few more numerical results...

Summary & Outlook

Summary:

- variational numerical scheme for non-linear diffusion equations that respects their gradient flow structure
- applicable to a wide range of nonlinear diffusion problems
- dissipates energy 'as fast as possible' — just like the original gradient flow
- built-in conservation of mass and non-negativity
- efficient solution also for two-dimensional problems

Outlook:

- modified scheme with improved convergence analysis

MERCI BEAUCOUP!