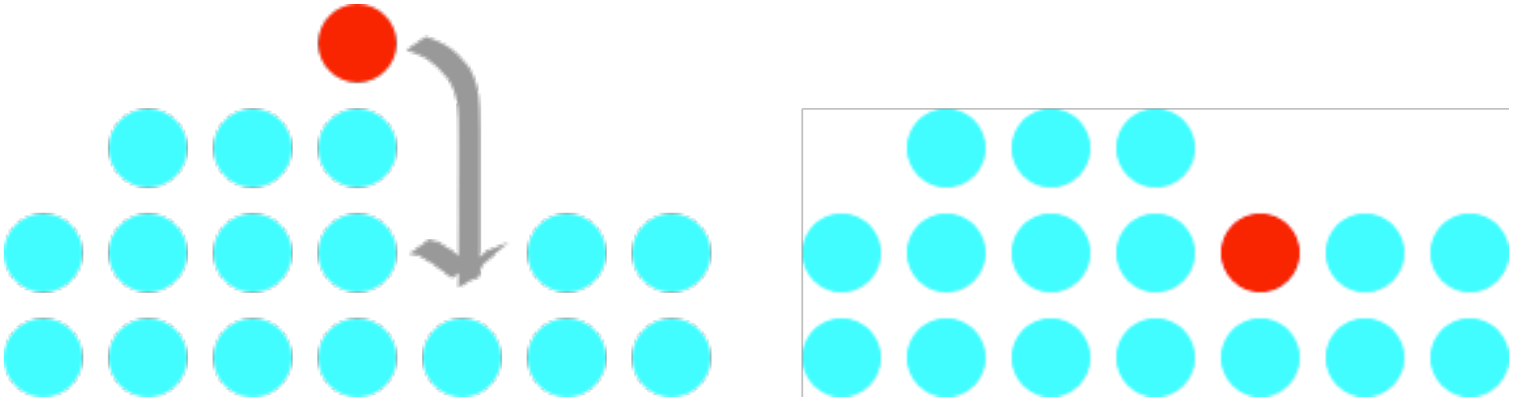
On 4th order nonlinear thin-film like PDEs describing crystal surface evolution

Joint work with Jon Weare (NYU), Anya Katsevich (NYU), Dio Margetis (Maryland), Katy Craig (UCSB), Li Wang (Minnesota), Jianfeng Lu (Duke), Jian-Guo Liu (Duke)

Jeremy L. Marzuola

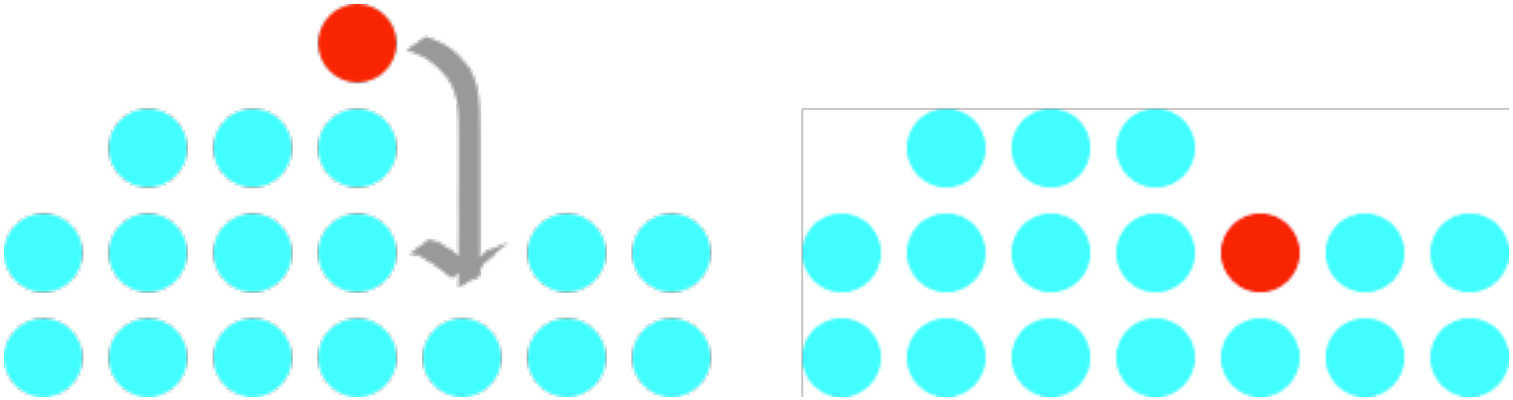
Mathematics Department University of North Carolina, Chapel Hill

Giga Lab Interfaces & Free Boundaries Conference - June 8, 2022



Each surface atom breaks the bonds with its nearest (horizontal) neighbors at a random (exponentially distributed) time and then jumps (uniformly) to one of its neighboring sites.

The rate at which those bonds are broken depends on the energy barrier that must be overcome in breaking them.... but atoms with lots of horizontal neighbors will move infrequently and neighbors with no (adatoms) or few horizontal neighbors will move more frequently.



Each surface atom breaks the bonds with its nearest (horizontal) neighbors at a random (exponentially distributed) time and then jumps (uniformly) to one of its neighboring sites.

The rate at which those bonds are broken depends on the energy barrier that must be overcome in breaking them.... but atoms with lots of horizontal neighbors will move infrequently and neighbors with no (adatoms) or few horizontal neighbors will move more frequently.

Macroscopically you might expect convex regions to move more slowly than concave regions.

We'll come back to this later.

We model the crystal surface as an indexed set of integer valued random variables $h_N(t, \alpha) \in \mathbb{Z}$ where $\alpha \in \mathbb{T}_N^d$ (the d-dimensional torus).

The jumps:

$h_N \to J_\alpha J^\beta h_N$ add an atom at β and subtract one from α

where

$$J_\alpha h_N(\gamma) = \begin{cases} h_N(\alpha) - 1, & \gamma = \alpha \\ h_N(\gamma), & \gamma \neq \alpha \end{cases} \quad \text{and} \quad J^\alpha h_N(\gamma) = \begin{cases} h_N(\alpha) + 1, & \gamma = \alpha \\ h_N(\gamma), & \gamma \neq \alpha \end{cases}$$

Now we need the **rates**...

We need an estimate of the energy barrier that must be overcome in breaking an atom's neighbor bonds.

If we define the total energy of the surface to be:

$$H(h_N) = \sum_{\alpha \in \mathbb{T}_N} \sum_{i \leq d} V(\nabla_i^+ h_N(\alpha))$$

where

$$\nabla_i^+ g(\alpha) = g(\alpha + \mathbf{e}_i) - g(\alpha) \quad \text{and} \quad \nabla_i^- g(\alpha) = g(\alpha) - g(\alpha - \mathbf{e}_i)$$

then the (symmetrized) change in energy that results from removing the atom at site α :

$$n_\alpha(h_N) = \frac{1}{2} \sum_{i \leq d} \big(V(\nabla_i^+ J_\alpha h_N(\alpha)) - V(\nabla_i^+ h_N(\alpha)) + V(\nabla_i^- J_\alpha h_N(\alpha)) - V(\nabla_i^- h_N(\alpha)) \big).$$

The rate at which the atom at site α breaks its bonds and becomes free to move is

$$e^{-2Kn_\alpha(h_N)}$$

The rate at which an atom breaks its bonds and moves to any particular neighbor is

$$r_N(\alpha) = \frac{1}{2d} e^{-2Kn_\alpha(h_N)}$$

The generator is

$$\mathcal{A}_N f(h_N) = \sum_{\substack{\alpha,\beta \in \mathbb{T}_N^d \\ |\alpha-\beta|=1}} r_N(\alpha) \left(f(J_\alpha^\beta h_N) - f(h_N) \right).$$

$h(t,\cdot)$ is reversible with respect to the distribution

$$\rho_N^m \propto \begin{cases} e^{-K\sum_{\alpha\in\mathbb{T}_N}\sum_{i\le d} V(\nabla_i^+ h_N(\alpha))}, & \text{if } \sum_{\alpha\in\mathbb{T}_N} h_N(\alpha) = m \\ 0, & \text{otherwise} \end{cases}$$

with $m = \sum_{\alpha\in\mathbb{T}_N} h_N(0,\alpha)$

Our objective is to characterize the non-equilibrium continuum (large N) behavior of the surface.

But first...

This is oversimplified but physically reasonable:

The atoms mostly jiggle around on their sites but eventually climb out of the potential well (from bonds with neighbors) confining them. The escape event is rare and it's reasonable to assume that

1. the rate is dominated by the change in energy and
2. the projection of the system onto the site occupation number remains Markovian.

A couple of examples

1. Suppose $V(z) = |z|$. Then

$$n_\alpha(h_N) + 2^{d-1} = \sum_{\substack{\beta \in \mathbb{T}_N^d \\ |\alpha - \beta| = 1}} \mathbf{1}_{(h_N(\alpha) \leq h_N(\beta))},$$

   i.e. up to an additive constant (a time rescaling), the generalized coordination number is the number of bonds that need to be broken to free the atom at lattice site α.

2. Suppose $V(z) = z^2$. Then

$$n_\alpha(h_N) - 2d = \sum_{i \leq d} \nabla_i^+ h_N(\alpha) - \nabla_i^- h_N(\alpha),$$

   i.e. up to an additive constant, the generalized coordination number is the discrete Laplacian of the surface at lattice site α.

Why this model?

This model is far from perfect:

1. no substrate interactions (e.g. elasticity),
2. no defects (e.g. vacancies),
3. many systems require modeling the electronic structure (probably via DFT) to get correct barrier heights for estimating jump rates,
4. . . .

But it's remarkably close to models used in very large scale simulations.

These models have not been considered by mathematicians but close relatives have been studied:

Funaki and Spohn studied the hydrodynamic limit of

$$dh_N(t,\alpha) = -\frac{\partial H(h_N)}{\partial h_N(\alpha)}dt + \sqrt{2K^{-1}}dW(t)$$

where here $h_N(t,\alpha) \in \mathbb{R}$

They show that if

$$\bar{h}_N(t,x) = N^{-1} h_N(N^4 t, \alpha) \text{ where } Nx \in \cap_{i=1}^d \left[\alpha_i - \frac{1}{2}, \alpha_i + \frac{1}{2}\right)$$

then $\bar{h}_N$ converges to the solution of the PDE

$$\partial_t h = K \text{div}\left[\sigma_C(\nabla h)\right]$$

The surface tension is defined by $\sigma_C(u) = \nabla \mathcal{F}_C(u)$ where

$$\mathcal{F}_C(u) = \frac{1}{K} \sup_{\sigma \in \mathbb{R}^d} \left\{ \sigma^{\mathsf{T}} u - \log \int_{w \in \mathbb{R}^d} e^{-K \sum_{i \le d} V(w_i) + K \sigma^{\mathsf{T}} w} dw \right\}$$

$$u = \frac{\int w\, e^{-K \sum_{i \le d} V(w_i) + K \sigma(u)^{\mathsf{T}} w} dw}{\int e^{-K \sum_{i \le d} V(w_i) + K \sigma(u)^{\mathsf{T}} w} dw}$$

They show that if

$$\bar{h}_N(t,x) = N^{-1} h_N(N^4 t, \alpha) \text{ where } Nx \in \cap_{i=1}^d \left[\alpha_i - \frac{1}{2}, \alpha_i + \frac{1}{2}\right)$$

then $\bar{h}_N$ converges to the solution of the PDE

$$\partial_t h = K \text{div}\left[\sigma_C(\nabla h)\right]$$

The surface tension is defined by $\sigma_C(u) = \nabla \mathcal{F}_C(u)$ where

$$\mathcal{F}_C(u) = \frac{1}{K} \sup_{\sigma \in \mathbb{R}^d} \left\{ \sigma^{\mathsf{T}} u - \log \int_{w \in \mathbb{R}^d} e^{-K \sum_{i \le d} V(w_i) + K \sigma^{\mathsf{T}} w} dw \right\}$$

$$u = \frac{\int w\, e^{-K \sum_{i \le d} V(w_i) + K \sigma(u)^{\mathsf{T}} w} dw}{\int e^{-K \sum_{i \le d} V(w_i) + K \sigma(u)^{\mathsf{T}} w} dw}$$

Nishikawa extended Funaki and Spohn's approach to

$$dh(t,\alpha) = \Delta \frac{\partial H(h)}{\partial h(\alpha)} dt + \sqrt{-2K^{-1}\Delta} dW(t)$$

where

$$\Delta g(\alpha) = \sum_{i \leq d} \nabla_i^+ g(\alpha) - \nabla_i^- g(\alpha).$$

He proved that $\bar{h}_N$ converges to the solution of the PDE

$$\partial_t h = -K\Delta \left[\text{div}\left[\sigma_C(\nabla h)\right]\right]$$

Recent work of Armstrong-Wu (2019) shows that $\sigma_C \in C^{2,\beta}$ for some $\beta > 0$. What can we say about σ_D??

This last overdamped Langevin equation for continuous height variables can be derived from the KMC model by introducing a vanishing lattice constant.

Replace $V(z)$ by $V(az)$, rescaling time appropriately, and taking $a \to 0$.

One can then apply Nishikawa's result to make an approximate statement about the scaling limit of the KMC equations.

This was pointed out by Haselwandter and Vvedensky.

The small lattice constant limit sacrifices interesting features.

For example, if V is symmetric and if h_N is a solution of the SDE then $-h_N$ is also a solution.

Remember that in the KMC model convex regions are much stickier than concave regions.

Krug, Dobbs, and Majaniemi studied the evolution of a model very similar to our KMC models in 1D when $V(z) = |z|$

They first argue for

$$\partial_t h = -K\Delta\left[\operatorname{div}\left[\sigma_D(\nabla h)\right]\right]$$

with $\sigma_D(u) = \nabla \mathcal{F}_D(u)$ where

$$\mathcal{F}_D(u) = \frac{1}{K} \sup_{\sigma \in \mathbb{R}^d} \left\{ \sigma^{\mathsf{T}} u - \log \sum_{z \in \mathbb{Z}^d} e^{-K \sum_{i \leq d} V(z_i) + K\sigma^{\mathsf{T}} z} \right\}$$

Then at the end of the paper they give arguments supporting

$$\partial_t h = \Delta \left[e^{-K \mathrm{div}[\sigma_D(\nabla h)]} \right]$$

As they point out... this PDE gives the first PDE when curvature of the surface is small.

- As for special solutions, the stationary states will be part of the following class:
  1. Rough and Smooth: h constant, ∇h constant;
  2. Rough: p-harmonic functions $-\Delta_p h = 0$;
  3. Rough: p-elliptic functions $-\Delta_p h = \log \phi$, $\Delta \phi = 0$.

- For the rough pde with $p = 1$ and the continuous surface tension, we are working jointly with Jian-Guo Liu, Jianfeng Lu and Dio Margetis on a version of facet dynamics related to the work of Giga-Kohn, Giga-Giga on Total Variation flows.
- The PDE takes the form

$$\partial_t h = \partial_{xx} e^{-\partial_x \left(\frac{\partial_x h}{|\partial_x h|} \right)}, \quad h(x,0) = h_0(x) \, ,$$

- For the rough pde with $p = 1$ and the continuous surface tension, we are working jointly with Jian-Guo Liu, Jianfeng Lu and Dio Margetis on a version of facet dynamics related to the work of Giga-Kohn, Giga-Giga on Total Variation flows.
- The PDE takes the form

$$\partial_t h = \partial_{xx} e^{-\partial_x \left(\frac{\partial_x h}{|\partial_x h|} \right)}, \quad h(x,0) = h_0(x) \ ,$$

- For a non-exponential version of this model,

$$\partial_t h = -\partial_{xx}\partial_x \left(\frac{\partial_x h}{|\partial_x h|} \right), \quad h(x,0) = h_0(x) \,,$$

Giga proved the existence of shock-like solutions satisfying

$$\begin{cases} \dot{h}_f = -\frac{3}{x_f^3(t)} \,, \\ \dot{x}_f (h_0(x_f) - h_f) = -3x_f^{-2} \,. \end{cases}$$

- Note, the case $p = 1$ technically is forbidden in our scaling limit in the adatom case, however, by either doing a temperature rescaling or coming from the BCF model, the rough PDE with $p = 1$ can be derived
- There are implicit facet dynamics that can be derived and relate to a jump forming at the point of minima and maxima.
- Numerical methods also need to be carefully developed to capture shocks.

- Note, the case $p = 1$ technically is forbidden in our scaling limit in the adatom case, however, by either doing a temperature rescaling or coming from the BCF model, the rough PDE with $p = 1$ can be derived
- There are implicit facet dynamics that can be derived and relate to a jump forming at the point of minima and maxima.
- Numerical methods also need to be carefully developed to capture shocks.

- Note, the case $p = 1$ technically is forbidden in our scaling limit in the adatom case, however, by either doing a temperature rescaling or coming from the BCF model, the rough PDE with $p = 1$ can be derived
- There are implicit facet dynamics that can be derived and relate to a jump forming at the point of minima and maxima.
- Numerical methods also need to be carefully developed to capture shocks.

- In this case, the energy supports the formation of shock discontinuities. The front of a shock (x_f, h_f) can be shown to obey the Differential-Algebraic system of equations

$$\begin{cases} \dot{X}_f = \frac{\dot{x}_f F(X_f)}{1 - x_f F'(X_f)} \ , \\ \dot{h}_f = -2F(X_f)^2 \ , \\ \dot{x}_f (h_0(x_f) - h_f) = -2x_f F(X_f)^2 \ , \end{cases}$$

where the variable X_f is an extra variable included to enforce an algebraic condition.

- These PDEs can be related to the total variation flow in H^{-1} for the given energy.

$$E[h] = \gamma \int_\Omega \left(|\nabla h| + \frac{g}{3} |\nabla h|^3 \right) \, \mathrm{d}x \qquad (\Omega \subset \mathbb{R}^2) \, ,$$

- For $g > 0$, due to presence of the (less singular) term $|\partial_x h|^3$ in the surface energy, the solution to this PDE no longer develops jumps in the height profile. This is expected from other studies in the non-weighted H^{-1} total variation flow.

- These PDEs can be related to the total variation flow in H^{-1} for the given energy.

$$E[h] = \gamma \int_\Omega \left(|\nabla h| + \frac{g}{3} |\nabla h|^3 \right) \, \mathrm{d}x \qquad (\Omega \subset \mathbb{R}^2) \, ,$$

- For $g > 0$, due to presence of the (less singular) term $|\partial_x h|^3$ in the surface energy, the solution to this PDE no longer develops jumps in the height profile. This is expected from other studies in the non-weighted H^{-1} total variation flow.

- Recent progress on Exponential PDE with $p = 2$ by J.G. Liu, Y. Gao, J. Lu, X. Xu, R. Granero-Belinchon, M. Magliocca, R. Strain, D. Ambrose, JLM, ...
- In this case, one can even include 2nd order effects of evaporation and deposition to arrive at at PDE of the form:

$$\partial_t h = \Delta e^{-\Delta h} + (1 - e^{-\Delta h}).$$

See Gao-Liu-Lu-M (2020).

- Define

$$u := e^{-\Delta h}.$$

The equation can be formally recast as

$$\partial_t u = -u\Delta[\Delta u + (1 - u)].$$

- Recent progress on Exponential PDE with $p = 2$ by J.G. Liu, Y. Gao, J. Lu, X. Xu, R. Granero-Belinchon, M. Magliocca, R. Strain, D. Ambrose, JLM, ...
- In this case, one can even include 2nd order effects of evaporation and deposition to arrive at at PDE of the form:

$$\partial_t h = \Delta e^{-\Delta h} + (1 - e^{-\Delta h}).$$

See Gao-Liu-Lu-M (2020).

- Define

$$u := e^{-\Delta h}.$$

The equation can be formally recast as

$$\partial_t u = -u\Delta[\Delta u + (1 - u)].$$

- Recent progress on Exponential PDE with $p = 2$ by J.G. Liu, Y. Gao, J. Lu, X. Xu, R. Granero-Belinchon, M. Magliocca, R. Strain, D. Ambrose, JLM, ...
- In this case, one can even include 2nd order effects of evaporation and deposition to arrive at at PDE of the form:

$$\partial_t h = \Delta e^{-\Delta h} + (1 - e^{-\Delta h}).$$

  See Gao-Liu-Lu-M (2020).
- Define

$$u := e^{-\Delta h}.$$

  The equation can be formally recast as

$$\partial_t u = -u\Delta[\Delta u + (1 - u)].$$

- Consider the second order equation from this formulation

$$\partial_t u = u \Delta u,$$

which resembles a porous medium equation, but without the gradient structure.

- Self-similar solutions of the form

$$u(x,t) = \left(\frac{x^2 + \alpha_0}{2(t_0 - t)} \right),$$

can be found, but they do not behave well with respect to boundary conditions.

- Consider the second order equation from this formulation

$$\partial_t u = u \Delta u,$$

which resembles a porous medium equation, but without the gradient structure.

- Self-similar solutions of the form

$$u(x,t) = \left(\frac{x^2 + \alpha_0}{2(t_0 - t)} \right),$$

can be found, but they do not behave well with respect to boundary conditions.

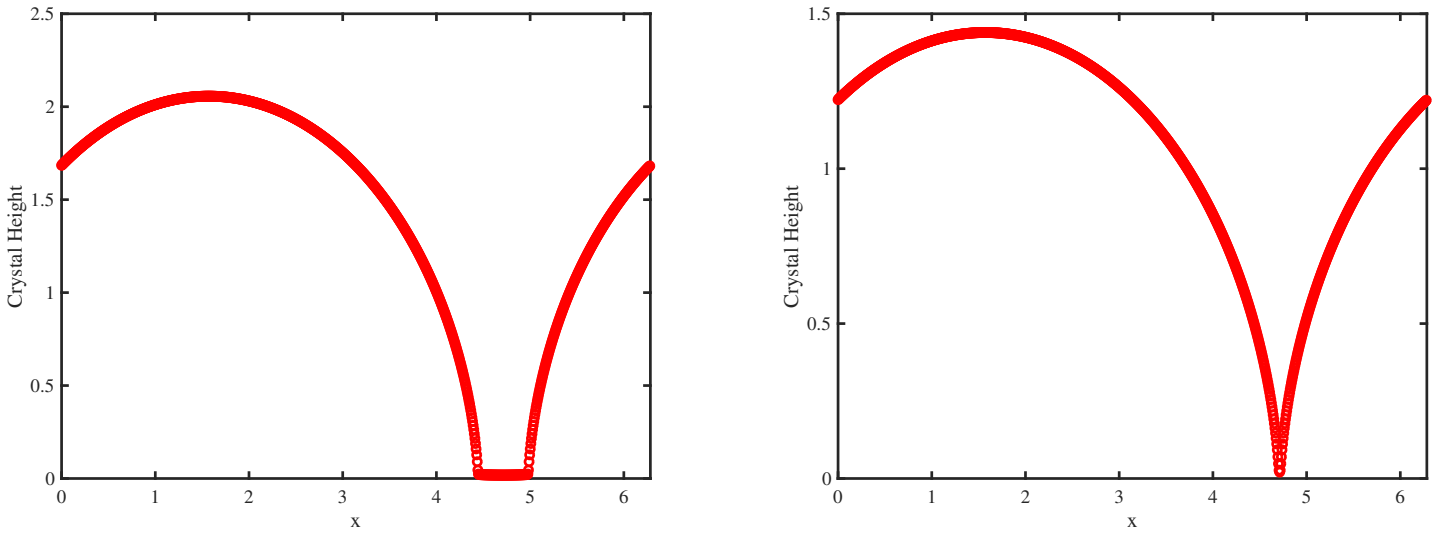


Figure: Snapshots of surface height evolution of the second order equation.

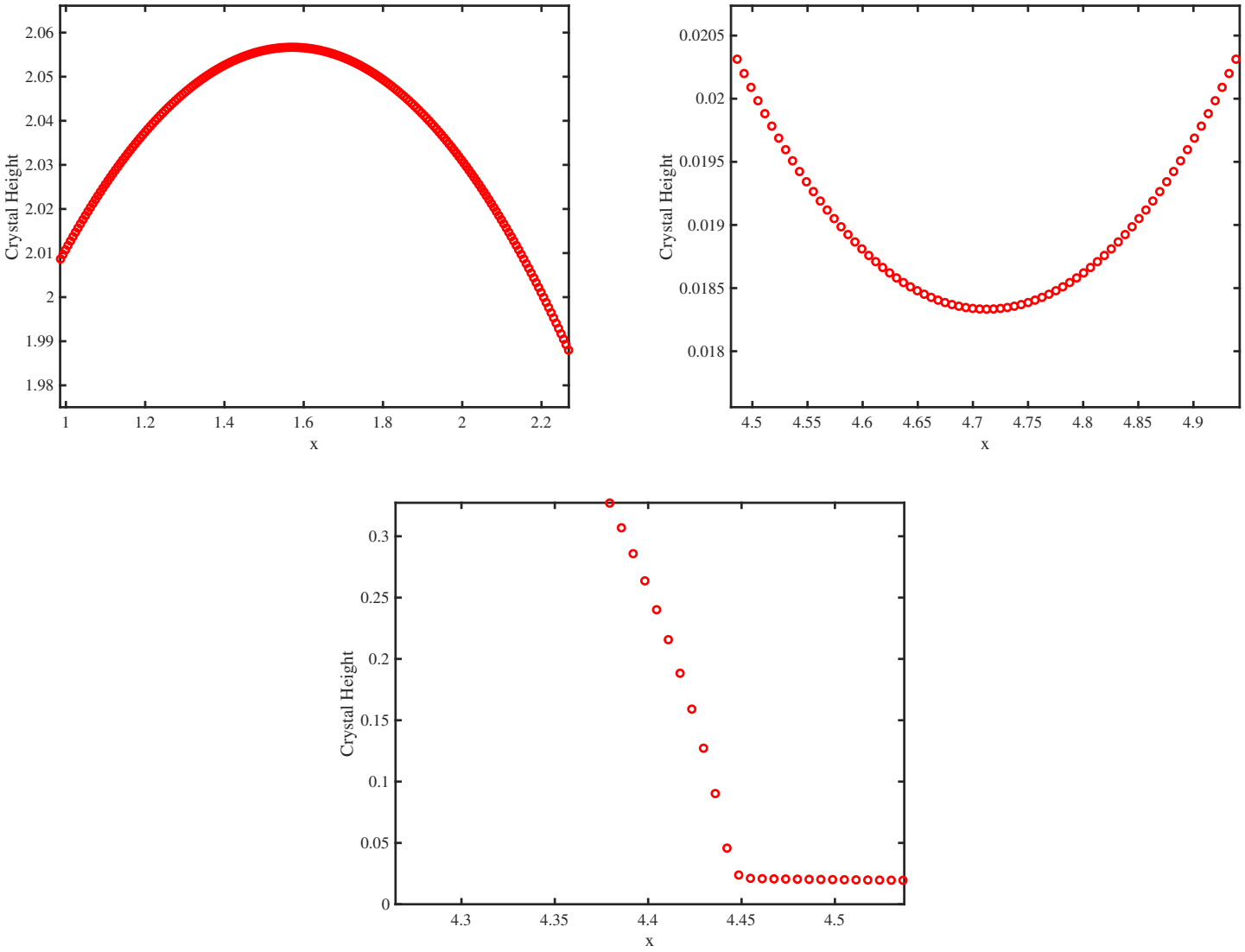


Figure: Snapshots of surface height evolution of the second order equation.

- Note, setting $v = u_x$, we get

$$\partial_t v = v v_x + u v_{xx},$$

which resembles an Euler equation with variable coefficient viscosity. With David Ambrose and Doug Wright, we are now pursuing this as a means of capturing both the transition layer between self-similar profiles and building solutions that do indeed form singularities.

- With Yuan Gao, Jian-Guo Liu, Jianfeng Lu and Anya Katsevich, we considered Metropolis-style rates and analyzing dynamics in comparison to the Adatom KMC models. These rates have some distinct differences and a temperature dependence in the form of the limiting PDEs can be seen. We focus here on the case $p = 2$ as many calculations become more explicit. The PDE limit proposed here is of the form

$$\partial_t h(t,x) = \partial_x \left(e^{\left[-\partial_x^3 h(t,x)\right]} - e^{\left[\partial_x^3 h(t,x)\right]} \right).$$

- We are looking at the relative entropy methods with Amarjit Budhiraja and Anya Katsevich for these conservative KMC systems with $p = 2$. To start, we are trying to prove a rigorous convergence between our KMC processes and SDE limits on which people like Funaki, Spohn et al have developed a great deal of tools for relative entropy calculations.

- Recent progress on Exponential PDE with $p = 2$ by J.G. Liu, Y. Gao, J. Lu, X. Xu, R. Granero-Belinchon, M. Magliocca, R. Strain, D. Ambrose, JLM, ...
- There are three main approaches:
  1. A small data theory in the Weiner Algebra space (Granero-Belinchon, Magiliocca; Liu-Strain; Ambrose, GLLM).
  2. A large data weak solution theory similar to the entropy solutions of Bernis-Friedman (Liu-Xu, GLLM).
  3. A small data, quasilinear approach for the u equation, see certain recent results of Quoc Hung Nguyen on quasilinear parabolic models.

- Recent progress on Exponential PDE with $p = 2$ by J.G. Liu, Y. Gao, J. Lu, X. Xu, R. Granero-Belinchon, M. Magliocca, R. Strain, D. Ambrose, JLM, ...
- There are three main approaches:
  1. A small data theory in the Weiner Algebra space (Granero-Belinchon, Magiliocca; Liu-Strain; Ambrose, GLLM).
  2. A large data weak solution theory similar to the entropy solutions of Bernis-Friedman (Liu-Xu, GLLM).
  3. A small data, quasilinear approach for the u equation, see certain recent results of Quoc Hung Nguyen on quasilinear parabolic models.

- The Weiner Algebra norms are of the form

$$\|f\|^p_{\dot{\mathcal{F}}^{s,p}}(t) := \int_{\mathbb{R}^d} |\xi|^{sp} |\hat{f}(\xi, t)|^p d\xi, \quad s > -d/p, \quad 1 \leq p \leq 2.$$

  We note that the Wiener algebra $A(\mathbb{R}^d)$ is $\dot{\mathcal{F}}^{0,1}$, and the condition $\Delta h_0 \in A(\mathbb{R}^d)$ is given by $h_0 \in \dot{\mathcal{F}}^{2,1}$.

- Using variations on this space, small data solutions can be constructed in the setting of analytic functions with increasing radius of analyticity.

- The Weiner Algebra norms are of the form

$$\|f\|^p_{\dot{\mathcal{F}}^{s,p}}(t) := \int_{\mathbb{R}^d} |\xi|^{sp} |\hat{f}(\xi,t)|^p d\xi, \quad s > -d/p, \quad 1 \leq p \leq 2.$$

  We note that the Wiener algebra $A(\mathbb{R}^d)$ is $\dot{\mathcal{F}}^{0,1}$, and the condition $\Delta h_0 \in A(\mathbb{R}^d)$ is given by $h_0 \in \dot{\mathcal{F}}^{2,1}$.
- Using variations on this space, small data solutions can be constructed in the setting of analytic functions with increasing radius of analyticity.

- For weak solutions, use a regularized method to first prove the existence and strict positivity for regularized solution u_ε to a properly modified equation below, then take limit $\varepsilon \to 0$. For $0 < \alpha < 1$, the regularization of the u equation we consider is

$$\begin{cases} \partial_t u_\varepsilon = -\dfrac{u_\varepsilon^{1+\alpha}}{u_\varepsilon^\alpha + \varepsilon^\alpha}(\partial_x^4 u_\varepsilon - \partial_x^2 u_\varepsilon), & \text{for } t \in [0, T],\ x \in \mathbb{T}; \\ u_\varepsilon(0, x) = u_0(x) + \varepsilon, & \text{for } x \in \mathbb{T}. \end{cases} \tag{1}$$

Basic questions remain unanswered:

1. Can we prove that the exponential PDE has a solution (in some space) for other p values, especially $p = 1$?
2. Can we prove convergence of the microscopic model to the PDE in our scaling regimes?
3. Can we establish finite time singularity formation in the 2nd derivative in the case $p = 2$?
4. Full global strong solutions with $g > 0$?

There are also lots of great questions about the qualitative behavior that can be asked w.r.t. wetting and self-similarity.

Basic questions remain unanswered:

1. Can we prove that the exponential PDE has a solution (in some space) for other p values, especially $p = 1$?
2. Can we prove convergence of the microscopic model to the PDE in our scaling regimes?
3. Can we establish finite time singularity formation in the 2nd derivative in the case $p = 2$?
4. Full global strong solutions with $g > 0$?

There are also lots of great questions about the qualitative behavior that can be asked w.r.t. wetting and self-similarity.

PDE models

- Recall, we are generally interested in the models we consider of the form:

$$h_t = \Delta e^{-\Delta_p h}$$

for Δ_p the p-Lapalcian for $p \geq 1$ given by the Euler-Lagrange equation for the surface energy

$$E(h) = \frac{1}{p} \int |\nabla h|^p dx.$$

Gradient Flow Framework

- Recently, with Craig, Liu, Lu and Wang, we have introduced a gradient flow structure numerical flow that allows one to numerically solve

$$h_t = \Delta e^{-\Delta_1 h}$$

using the FORMAL gradient flow structure

$$h_t + \nabla \cdot \left(M(h) \nabla \frac{\partial \mathcal{E}}{\partial h} \right) = 0,$$
$$M(h) = e^{-\Delta_1 h}, \;\; \mathcal{E}(h) = \|\nabla h\|_{L^1}.$$

Gradient Flow Framework

- Examples:
  1. $M(h) = 1 \to H^{-1}$ gradient flow,
  2. $M(h) = h \to W_2$ gradient flow,
  3. $M(h)$ non-negative, concave $\to$ weighted W_2 gradient flow (See Carrillo-Lisini-Savaré-Slepcev (2009), Dolbeault-Nazaret-Savaré (2009), Lisini-Matthew-Savaré (2019), ...)
  4. $M(h) \in \text{Lin}(\mathbb{R}^{\ell \times d}, \mathbb{R}^{\ell \times d}) \to$ gradient system (Liero-Mielke (2013))
- Our mobility is a nonlinear function of the 1-Laplacian and hence falls well outside existing theories, but....

Gradient Flow Framework

- Examples:
  1. $M(h) = 1 \to H^{-1}$ gradient flow,
  2. $M(h) = h \to W_2$ gradient flow,
  3. $M(h)$ non-negative, concave $\to$ weighted W_2 gradient flow (See Carrillo-Lisini-Savaré-Slepcev (2009), Dolbeault-Nazaret-Savaré (2009), Lisini-Matthew-Savaré (2019), ...)
  4. $M(h) \in \text{Lin}(\mathbb{R}^{\ell \times d}, \mathbb{R}^{\ell \times d}) \to$ gradient system (Liero-Mielke (2013))
- Our mobility is a nonlinear function of the 1-Laplacian and hence falls well outside existing theories, but....

Recent Advances with formal H^{-1} gradient flow for $p = 1$ using ideas of Otto (2001), Giga-Giga (2010), ...

- Suppose $0 \leq M(h) \in L^1(\mathbb{T}^d)$. Define

$$\Delta_h v = \nabla \cdot (M(h) \nabla v).$$

- We define the weighted Hilbert space

$$\|v\|_{H_h^1}^2 = \int_{\mathbb{T}^d} M(h) |\nabla v|^2 dx = -\int_{\mathbb{T}^d} v \Delta_h v dx.$$

- We have the dual space:

$$\|\psi\|_{H_h^{-1}}^2 = -\int_{\mathbb{T}^d} \psi \Delta_h^{-1} \psi dx.$$

- And the sub-differential

$$\partial_{H_h^{-1}} \mathcal{E}$$

and gradient

$$\nabla_{H_h^{-1}} \mathcal{E}(\psi).$$

Recent Advances with formal H^{-1} gradient flow for $p = 1$ using ideas of Otto (2001), Giga-Giga (2010), ...

- Suppose $0 \leq M(h) \in L^1(\mathbb{T}^d)$. Define

$$\Delta_h v = \nabla \cdot (M(h) \nabla v).$$

- We define the weighted Hilbert space

$$\|v\|_{H_h^1}^2 = \int_{\mathbb{T}^d} M(h) |\nabla v|^2 dx = -\int_{\mathbb{T}^d} v \Delta_h v dx.$$

- We have the dual space:

$$\|\psi\|_{H_h^{-1}}^2 = -\int_{\mathbb{T}^d} \psi \Delta_h^{-1} \psi dx.$$

- And the sub-differential

$$\partial_{H_h^{-1}} \mathcal{E}$$

and gradient

$$\nabla_{H_h^{-1}} \mathcal{E}(\psi).$$

Recent Advances with formal H^{-1} gradient flow for $p = 1$ using ideas of Otto (2001), Giga-Giga (2010), ...

- Suppose $0 \leq M(h) \in L^1(\mathbb{T}^d)$. Define

$$\Delta_h v = \nabla \cdot (M(h) \nabla v).$$

- We define the weighted Hilbert space

$$\|v\|_{H_h^1}^2 = \int_{\mathbb{T}^d} M(h) |\nabla v|^2 dx = -\int_{\mathbb{T}^d} v \Delta_h v dx.$$

- We have the dual space:

$$\|\psi\|_{H_h^{-1}}^2 = -\int_{\mathbb{T}^d} \psi \Delta_h^{-1} \psi dx.$$

- And the sub-differential

$$\partial_{H_h^{-1}} \mathcal{E}$$

and gradient

$$\nabla_{H_h^{-1}} \mathcal{E}(\psi).$$

Recent Advances with formal H^{-1} gradient flow for $p = 1$ using ideas of Otto (2001), Giga-Giga (2010), ...

- Suppose $0 \leq M(h) \in L^1(\mathbb{T}^d)$. Define

$$\Delta_h v = \nabla \cdot (M(h) \nabla v).$$

- We define the weighted Hilbert space

$$\|v\|_{H_h^1}^2 = \int_{\mathbb{T}^d} M(h) |\nabla v|^2 dx = -\int_{\mathbb{T}^d} v \Delta_h v dx.$$

- We have the dual space:

$$\|\psi\|_{H_h^{-1}}^2 = -\int_{\mathbb{T}^d} \psi \Delta_h^{-1} \psi dx.$$

- And the sub-differential

$$\partial_{H_h^{-1}} \mathcal{E}$$

and gradient

$$\nabla_{H_h^{-1}} \mathcal{E}(\psi).$$

Recent Advances with formal H^{-1} gradient flow for $p = 1$ using ideas of Otto (2001), Giga-Giga (2010), ...

- To get to a well-defined gradient flow, we require that the time derivative of $h(t)$ makes sense wrt $\|\cdot\|_{H^{-1}_{h(t)}}$ and that $M(h(t))$ remains in L^1 and ≥ 0 for all t.

$$\partial_t h = -\nabla_{H_h^{-1}} \mathcal{E}(h) \Leftrightarrow \partial_t h + \Delta_h \frac{\partial E}{\partial h} = 0 \Leftrightarrow \partial_t h + \nabla \cdot M(h) \nabla \frac{\partial E}{\partial h} = 0$$

Recent Advances with formal H^{-1} gradient flow for $p = 1$ using ideas of Otto (2001), Giga-Giga (2010), ...

- To get to a well-defined gradient flow, we require that the time derivative of $h(t)$ makes sense wrt $\|\cdot\|_{H^{-1}_{h(t)}}$ and that $M(h(t))$ remains in L^1 and ≥ 0 for all t.
- $$\partial_t h = -\nabla_{H_h^{-1}} \mathcal{E}(h) \Leftrightarrow \partial_t h + \Delta_h \frac{\partial E}{\partial h} = 0 \Leftrightarrow \partial_t h + \nabla \cdot M(h) \nabla \frac{\partial E}{\partial h} = 0$$

Our proposed numerical method - Mobility

- Even with a smooth profile, $e^{-\Delta_1 h}$ is not really well-defined since $\Delta_1(\sin(x)) \sim -2\delta_{\pi/2}(x) + 2\delta_{3\pi/2}(x)$.
- Prior works have considered the mobility formulation of the Total Variation flow, see (Giga-Giga (2010)) in which they consider a Mobility that is of the form given by the Taylor expansion $e^x \approx 1 + x$.
- Given that the exponential breaks convex/concave symmetry, we consider of mollified mobility. Given

$$\phi \in C_c^\infty(\mathbb{T}^d), \quad \phi \geq 0, \quad \int \phi dx = 1, \quad \phi_\epsilon(x) = \epsilon^{-d}\phi(x/\epsilon),$$

we define

$$M_\epsilon(h) := e^{-\phi_\epsilon * \Delta_1 h}.$$

Our proposed numerical method - Mobility

- Even with a smooth profile, $e^{-\Delta_1 h}$ is not really well-defined since $\Delta_1(\sin(x)) \sim -2\delta_{\pi/2}(x) + 2\delta_{3\pi/2}(x)$.
- Prior works have considered the mobility formulation of the Total Variation flow, see (Giga-Giga (2010)) in which they consider a Mobility that is of the form given by the Taylor expansion $e^x \approx 1 + x$.
- Given that the exponential breaks convex/concave symmetry, we consider of mollified mobility. Given

$$\phi \in C_c^\infty(\mathbb{T}^d), \quad \phi \geq 0, \quad \int \phi dx = 1, \quad \phi_\epsilon(x) = \epsilon^{-d}\phi(x/\epsilon),$$

we define

$$M_\epsilon(h) := e^{-\phi_\epsilon * \Delta_1 h}.$$

Our proposed numerical method - Mobility

- Even with a smooth profile, $e^{-\Delta_1 h}$ is not really well-defined since $\Delta_1(\sin(x)) \sim -2\delta_{\pi/2}(x) + 2\delta_{3\pi/2}(x)$.
- Prior works have considered the mobility formulation of the Total Variation flow, see (Giga-Giga (2010)) in which they consider a Mobility that is of the form given by the Taylor expansion $e^x \approx 1 + x$.
- Given that the exponential breaks convex/concave symmetry, we consider of mollified mobility. Given

$$\phi \in C_c^\infty(\mathbb{T}^d), \quad \phi \geq 0, \quad \int \phi dx = 1, \quad \phi_\epsilon(x) = \epsilon^{-d}\phi(x/\epsilon),$$

we define

$$M_\epsilon(h) := e^{-\phi_\epsilon * \Delta_1 h}.$$

Our proposed numerical method - time discretization

- We consider a semi-implicit time stepping scheme similar to for instance the JKO framework, see also Murphy-Walkington (2019) in the Porous Medium Equation.
- $$h^{n+1} \in \arg\min_h \mathcal{E}(h) + \frac{1}{2\tau}\|h - h^n\|^2_{H^{-1}_{h^n}}$$
- Or...
$$\frac{h^{n+1} - h^n}{\tau} = -\nabla \cdot \left(M(h^n) \nabla \frac{\partial \mathcal{E}}{\partial h^{n+1}} \right).$$
- For the Total Variation Energy (L^1), if $h^n \in D(E)$ and $0 \leq M(h^n) \in L^1$, there exists a unique solution to h^{n+1}!!

Our proposed numerical method - time discretization

- We consider a semi-implicit time stepping scheme similar to for instance the JKO framework, see also Murphy-Walkington (2019) in the Porous Medium Equation.
- $$h^{n+1} \in \arg\min_h \mathcal{E}(h) + \frac{1}{2\tau}\|h - h^n\|^2_{H^{-1}_{h^n}}$$
- Or...
$$\frac{h^{n+1} - h^n}{\tau} = -\nabla \cdot \left(M(h^n) \nabla \frac{\partial \mathcal{E}}{\partial h^{n+1}} \right).$$
- For the Total Variation Energy (L^1), if $h^n \in D(E)$ and $0 \leq M(h^n) \in L^1$, there exists a unique solution to h^{n+1}!!

Our proposed numerical method - time discretization

- We consider a semi-implicit time stepping scheme similar to for instance the JKO framework, see also Murphy-Walkington (2019) in the Porous Medium Equation.
- $$h^{n+1} \in \arg\min_h \mathcal{E}(h) + \frac{1}{2\tau}\|h - h^n\|^2_{H^{-1}_{h^n}}$$
- Or...
$$\frac{h^{n+1} - h^n}{\tau} = -\nabla \cdot \left(M(h^n) \nabla \frac{\partial \mathcal{E}}{\partial h^{n+1}} \right).$$
- For the Total Variation Energy (L^1), if $h^n \in D(E)$ and $0 \leq M(h^n) \in L^1$, there exists a unique solution to h^{n+1} !!

Our proposed numerical method - time discretization

- We consider a semi-implicit time stepping scheme similar to for instance the JKO framework, see also Murphy-Walkington (2019) in the Porous Medium Equation.
- $$h^{n+1} \in \arg\min_h \mathcal{E}(h) + \frac{1}{2\tau}\|h - h^n\|^2_{H^{-1}_{h^n}}$$
- Or...
$$\frac{h^{n+1} - h^n}{\tau} = -\nabla \cdot \left(M(h^n) \nabla \frac{\partial \mathcal{E}}{\partial h^{n+1}} \right).$$
- For the Total Variation Energy (L^1), if $h^n \in D(E)$ and $0 \leq M(h^n) \in L^1$, there exists a unique solution to h^{n+1}!!

Our proposed numerical method - space discretization

- We have the optimization

$$h^{n+1} \in \arg\min_h \mathcal{E}(h) + \frac{1}{2\tau}\|h - h^n\|^2_{H^{-1}_{h^n}}$$
$$= \arg\min_h (f(Kh) + g(h))$$

for Z, H Hilbert spaces to be explained, $f : Z \to \mathbb{R}$, $g : H \to \mathbb{R}$ convex, $K : H \to Z$ a bounded linear operator.

- Such schemes can be regularized and solved using a primal-dual algorithm to solve inner and outer optimization problems!! See (Laborde-Benamou-Carlier (2016), Carrillo-Craig-Wang-Wei (2019)).

Our proposed numerical method - space discretization

- We have the optimization

$$h^{n+1} \in \arg\min_h \mathcal{E}(h) + \frac{1}{2\tau}\|h - h^n\|^2_{H^{-1}_{h^n}}$$
$$= \arg\min_h (f(Kh) + g(h))$$

  for Z, H Hilbert spaces to be explained, $f : Z \to \mathbb{R}$, $g : H \to \mathbb{R}$ convex, $K : H \to Z$ a bounded linear operator.
- Such schemes can be regularized and solved using a primal-dual algorithm to solve inner and outer optimization problems!! See (Laborde-Benamou-Carlier (2016), Carrillo-Craig-Wang-Wei (2019)).

Our proposed numerical method - space discretization

- For our case, we are motivated by the work of Jacobs-Léger-Li-Osher (2019).
- Gradient descent of smooth, convex function $F(u) = f(Ku) + g(u)$ with a unique min at u^*.
- Convergence rate:

$$F(u_n) \leq F(u^*) + 2L_{\mathcal{H}} \frac{\|u^* - u_0\|_{\mathcal{H}}^2}{n+4}.$$

Our proposed numerical method - space discretization

- For our case, we are motivated by the work of Jacobs-Léger-Li-Osher (2019).
- Gradient descent of smooth, convex function $F(u) = f(Ku) + g(u)$ with a unique min at u^*.
- Convergence rate:

$$F(u_n) \leq F(u^*) + 2L_{\mathcal{H}} \frac{\|u^* - u_0\|_{\mathcal{H}}^2}{n+4}.$$

Our proposed numerical method - space discretization

- For our case, we are motivated by the work of Jacobs-Léger-Li-Osher (2019).
- Gradient descent of smooth, convex function $F(u) = f(Ku) + g(u)$ with a unique min at u^*.
- Convergence rate:

$$F(u_n) \leq F(u^*) + 2L_{\mathcal{H}} \frac{\|u^* - u_0\|_{\mathcal{H}}^2}{n+4}.$$

Our proposed numerical method - space discretization

- Nesterov:

$$F(u_n) \leq F(u^*) + 4L_{\mathcal{H}} \frac{R^2}{(n+2)^2} + \min_{\|u-u_0\|_{\mathcal{H}} \leq R} (F(u) - F(u^*)).$$

Punchline: $\|u^* - u_0\|_{\mathcal{H}} = +\infty$ OK as long as

$$\min_{\|u-u_0\|_{\mathcal{H}} \leq R} (F(u) - F(u^*)) \to 0$$

as $R \to \infty$.

Our proposed numerical method - space discretization

- The primal-dual algorithm is thus:

$$u_{n+1} = \arg\min_{u \in \mathcal{H}} g(u) + (u, K^T \overline{p}_n)_{\mathcal{H}} + \frac{1}{2\tau} \|u - u_n\|^2_{\mathcal{H}},$$

$$p_{n+1} = \arg\max_{p \in \mathcal{Z}} -f^*(p) + (Ku_{n+1}, p)_{\mathcal{Z}} - \frac{1}{2\sigma} \|p - p_n\|^2_{\mathcal{Z}},$$

$$\overline{p}_n = 2p_{n+1} - p_n.$$

- Cf. JLLO, Chambolle-Pock PDHG method, ...

Our proposed numerical method - space discretization

- The primal-dual algorithm is thus:

$$u_{n+1} = \arg\min_{u \in \mathcal{H}} g(u) + (u, K^T \overline{p}_n)_{\mathcal{H}} + \frac{1}{2\tau} \|u - u_n\|^2_{\mathcal{H}},$$

$$p_{n+1} = \arg\max_{p \in \mathcal{Z}} -f^*(p) + (Ku_{n+1}, p)_{\mathcal{Z}} - \frac{1}{2\sigma} \|p - p_n\|^2_{\mathcal{Z}},$$

$$\overline{p}_n = 2p_{n+1} - p_n.$$

- Cf. JLLO, Chambolle-Pock PDHG method, ...

Our proposed numerical method - space discretization

- For us, we get the following:
- Outer time iteration:

$$h^{n+1} = \arg\min_h \mathcal{E}(h) + \frac{1}{2\tau}\|h - h^n\|^2_{H^{-1}_{h^n}}.$$

- Inner time iteration ($(\sigma\lambda) < 1$):

$$h^{m+1} = \left(\frac{\tau}{\lambda}\Delta_{h^n}\Delta + I\right)^{-1}\left(\frac{\tau}{\lambda}\Delta_{h^n}\Delta h^{(m)} - \tau\Delta_{h^n}\nabla\cdot\phi^{(m)} + h^n\right),$$

$$\overline{h}^{(m+1)} = 2h^{(m+1)} - h^{(m)},$$

$$\phi^{(m+1)} = (I + \sigma\partial F^*)^{-1}(\phi^{(m)} + \sigma\nabla\overline{h}^{(m+1)})$$

where

$$(I + \sigma\partial F^*)^{-1}(u(x)) = \min\{u(x), 1\}\,\mathrm{sign}(u(x)).$$

Our proposed numerical method - space discretization

- For us, we get the following:
- Outer time iteration:

$$h^{n+1} = \arg\min_h \mathcal{E}(h) + \frac{1}{2\tau}\|h - h^n\|^2_{H^{-1}_{h^n}}.$$

- Inner time iteration ($(\sigma\lambda) < 1$):

$$h^{m+1} = \left(\frac{\tau}{\lambda}\Delta_{h^n}\Delta + I\right)^{-1}\left(\frac{\tau}{\lambda}\Delta_{h^n}\Delta h^{(m)} - \tau\Delta_{h^n}\nabla\cdot\phi^{(m)} + h^n\right),$$

$$\overline{h}^{(m+1)} = 2h^{(m+1)} - h^{(m)},$$

$$\phi^{(m+1)} = (I + \sigma\partial F^*)^{-1}(\phi^{(m)} + \sigma\nabla\overline{h}^{(m+1)})$$

where

$$(I + \sigma\partial F^*)^{-1}(u(x)) = \min\{u(x), 1\}\,\mathrm{sign}(u(x)).$$

Our proposed numerical method - space discretization

- For us, we get the following:
- Outer time iteration:

$$h^{n+1} = \arg\min_{h} \mathcal{E}(h) + \frac{1}{2\tau}\|h - h^n\|^2_{H^{-1}_{h^n}}.$$

- Inner time iteration ($(\sigma\lambda) < 1$):

$$h^{m+1} = \left(\frac{\tau}{\lambda}\Delta_{h^n}\Delta + I\right)^{-1}\left(\frac{\tau}{\lambda}\Delta_{h^n}\Delta h^{(m)} - \tau\Delta_{h^n}\nabla\cdot\phi^{(m)} + h^n\right),$$

$$\overline{h}^{(m+1)} = 2h^{(m+1)} - h^{(m)},$$

$$\phi^{(m+1)} = (I + \sigma\partial F^*)^{-1}(\phi^{(m)} + \sigma\nabla\overline{h}^{(m+1)})$$

where

$$(I + \sigma\partial F^*)^{-1}(u(x)) = \min\{u(x), 1\}\,\mathrm{sign}(u(x)).$$

Our proposed numerical method - space discretization

- What do we gain???
- We do not invert the 1-Laplacian!
- Can choose λ quite large!

Our proposed numerical method - space discretization

- What do we gain???
- We do not invert the 1-Laplacian!
- Can choose λ quite large!

Our proposed numerical method - space discretization

- What do we gain???
- We do not invert the 1-Laplacian!
- Can choose λ quite large!

Theorem [CMLLW 2020]

- Let $d = 1$. Suppose the PDHG algorithm is initialized with

$$h^{(0)} = h^n, \ \phi^{(0)} = 0, \ \text{ where } \ \mathcal{E}(h^n) < \infty.$$

Then for all $\delta > 0$, there exist $\tilde{M}, \lambda, \sigma$ such that

$$F(h^{(M)} - F(h^{n+1}) \leq \delta, \ \forall M \geq \tilde{M}, \ F(h) = \mathcal{E}(h) + \frac{1}{2\tau}\|h - h^n\|^2_{H^{-1}_{h^n}}.$$

Remarks

- Extends to general $M(h)$ provided $M(h^n), 1/M(h^n) \in L^1$ (Cancés-Gallouët-Todeschi (2019)).
- The restriction $d = 1$ is to apply a Sobolev embedding. Different integrability on $1/M$ required to go to higher dimensions.
- If $\nabla h^n \in BV$, the estimates are quantitative: $\tilde{M} \sim \delta^{-2}, \lambda \sim \delta^{-1}, \sigma \sim \delta$.
- $$\min_{\|h^n - h\|_{\mathcal{H}} \leq R} (F(h) - F(h^{n+1})) \to 0$$

  as $R \to \infty$.

Remarks

- Extends to general $M(h)$ provided $M(h^n), 1/M(h^n) \in L^1$ (Cancés-Gallouët-Todeschi (2019)).
- The restriction $d = 1$ is to apply a Sobolev embedding. Different integrability on $1/M$ required to go to higher dimensions.
- If $\nabla h^n \in BV$, the estimates are quantitative: $\tilde{M} \sim \delta^{-2}, \lambda \sim \delta^{-1}, \sigma \sim \delta$.
- $$\min_{\|h^n - h\|_{\mathcal{H}} \leq R} (F(h) - F(h^{n+1})) \to 0$$

  as $R \to \infty$.

Remarks

- Extends to general $M(h)$ provided $M(h^n), 1/M(h^n) \in L^1$ (Cancés-Gallouët-Todeschi (2019)).
- The restriction $d = 1$ is to apply a Sobolev embedding. Different integrability on $1/M$ required to go to higher dimensions.
- If $\nabla h^n \in BV$, the estimates are quantitative: $\tilde{M} \sim \delta^{-2}, \lambda \sim \delta^{-1}, \sigma \sim \delta$.
- $$\min_{\|h^n - h\|_{\mathcal{H}} \leq R} (F(h) - F(h^{n+1})) \to 0$$

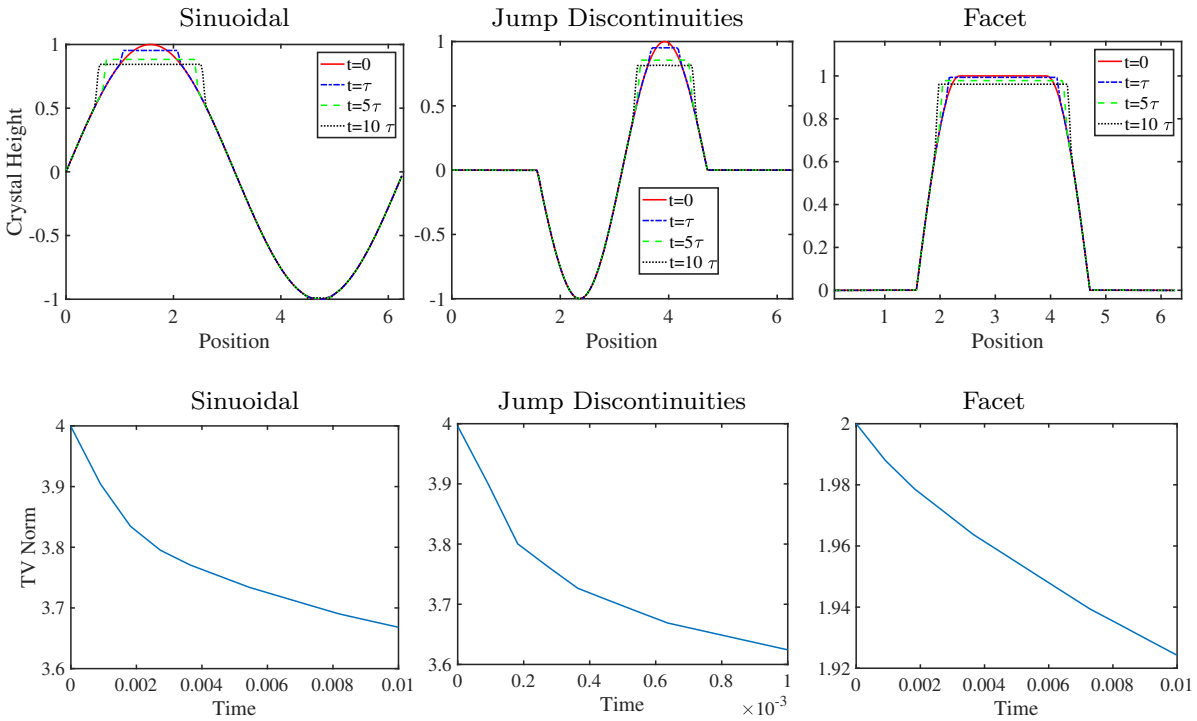
  as $R \to \infty$.

Remarks

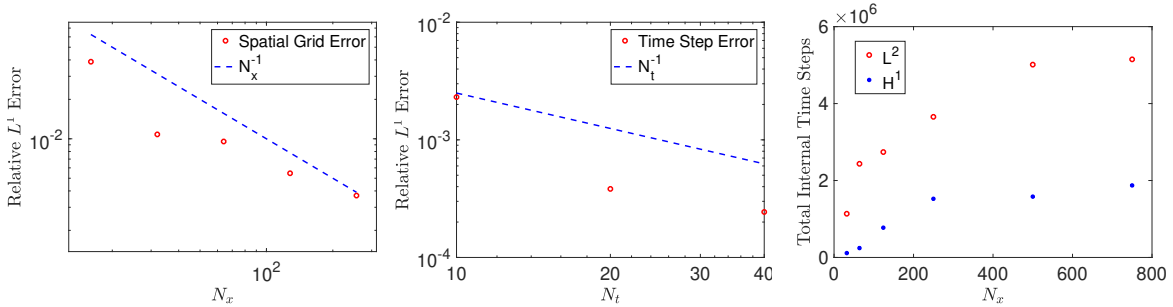
- Extends to general $M(h)$ provided $M(h^n), 1/M(h^n) \in L^1$ (Cancés-Gallouët-Todeschi (2019)).
- The restriction $d = 1$ is to apply a Sobolev embedding. Different integrability on $1/M$ required to go to higher dimensions.
- If $\nabla h^n \in BV$, the estimates are quantitative: $\tilde{M} \sim \delta^{-2}, \lambda \sim \delta^{-1}, \sigma \sim \delta$.
- $$\min_{\|h^n - h\|_{\mathcal{H}} \leq R} (F(h) - F(h^{n+1})) \to 0$$

  as $R \to \infty$.

Numerical Results: energy decrease



Numerical Results: convergence



Observations:

- Error vs Nx: slightly sublinear convergence (low spatial regularity)
- Error vs Nt: first order (semi-implicit Euler)
- Internal time steps vs Nt: importance of selecting correct Hilbert space

sinusoidal, (Nx = 200), (Nt = 10), $\sigma = 0.0005$, $\lambda = 500$, $\varepsilon = 0.05$, $T = 10^{-4}$

Open Questions

- Appropriate notion of weak solution for $p = 1$??
- Better time discretization/Gradient Flow formulation to prove existence of wider class of weak solutions in other cases?
- Convergence as $\tau \to 0$? Convergence as $\tau, \epsilon \to 0$??

Open Questions

- Appropriate notion of weak solution for $p = 1$??
- Better time discretization/Gradient Flow formulation to prove existence of wider class of weak solutions in other cases?
- Convergence as $\tau \to 0$? Convergence as $\tau, \epsilon \to 0$??

Open Questions

- Appropriate notion of weak solution for $p = 1$??
- Better time discretization/Gradient Flow formulation to prove existence of wider class of weak solutions in other cases?
- Convergence as $\tau \to 0$? Convergence as $\tau, \epsilon \to 0$??