NUMERICAL ANALYSIS OF A STEEPEST-DESCENT PDE MODEL FOR SURFACE RELAXATION BELOW THE ROUGHENING TEMPERATURE

R. V. KOHN* AND H. M. VERSIEUX†

Abstract. We study the numerical solution of a PDE describing the relaxation of a crystal surface to a flat facet. The PDE is a singular, nonlinear, fourth order evolution equation, which can be viewed as the gradient flow of a convex but non-smooth energy with respect to the H^{-1}_{per} inner product. Our numerical scheme uses implicit discretization in time and a mixed finite-element approximation in space. The singular character of the energy is handled using regularization, combined with a primal-dual method. We study the convergence of this scheme, both theoretically and numerically.

Key words. H^{-1} steepest descent, crystal growth, surface relaxation, Galerkin approximation, mixed finite element methods

AMS subject classifications. 65M55, 65M12, 35G25, 35R70, 74H15

1. Introduction.

The relaxation of crystalline surfaces has been an area of active research in recent years, motivated by the many applications of nanodevices. At such small scales the properties of a device depend on its nanoscale features. However, nanoscale features are easily changed by surface diffusion. An understanding of this relaxation process is therefore important for the modeling and fabrication of nanoscale devices.

This paper addresses a widely-used PDE model for the relaxation of a crystalline surface below the roughening temperature; see e.g. [18, 21, 23, 25, 26]. We introduce a numerical solution scheme (using finite elements in space and implicit discretization in time) and study its convergence.

The PDE we want to study is formally

$$u_t = -\Delta\nabla\cdot\left(\beta\frac{\nabla u}{|\nabla u|} + |\nabla u|^{p-2}\nabla u\right) \quad \text{with } u = u^0 \text{ at } t = 0 \tag{1.1}$$

(the proper interpretation of $\frac{\nabla u}{|\nabla u|}$ will be discussed soon). Unless otherwise stated we assume $p > 1$, and $\beta \in \mathbb{R}$. We work with periodic boundary conditions, writing $\Omega = \prod_{i=1}^{d}(a_i, b_i)$ for the period cell. Our initial data $u(0,x) = u^0(x)$ has mean value zero, $\int_\Omega u^0 = 0$, and this property is preserved by the dynamics. The analysis presented here could presumably be extended to the solution of (1.1) on a polygonal domain Ω with a suitable boundary conditions.

We expect a solution with facets, where $\nabla u = 0$, so the PDE (1.1) is purely formal. What we really mean is that u evolves by "H^{-1}_{per} steepest descent" for the functional

$$E(u) = \int_\Omega \beta|\nabla u| + \frac{1}{p}|\nabla u|^p dx.$$

*Courant Institute of Mathematical Sciences, *kohn@courant.nyu.edu*. The support of NSF through grant DMS-0313744 is gratefully acknowledged.

†IM, Universidade Federal do Rio de Janeiro, *henrique@im.ufrj.br*. This research was mainly done while this author was a Visiting Member at the Courant Institute of Mathematical Sciences, supported primarily by a grant from CNPq (Brazil). Additional support from NSF through grant DMS-0313744 and CAPES-PRODOC fellowship is gratefully acknowledged.

We shall review what this means in Section 2. From the results in [19, 20] one can see that the steepest-descent solution is the same as the one defined e.g. in [26, 18, 21] via continuity of the chemical potential at the edge of the facet.

We are naturally not the first to consider the numerical solution of this PDE. Numerous authors have relied on regularization, but other alternatives have also been considered; see [18, 25, 22]. None of these methods have, to our knowledge, been studied from a numerical analysis point of view; in other words there are no rigorous results on their convergence rates.

In this paper we use implicit time-stepping, combined with a "mixed" finite element scheme (see e.g. [11]) for spatial approximation. Like many other authors (see e.g. [12, 13, 14]) we use regularization to handle the singular character of the surface energy. Since the PDE is H^{-1}_{per} steepest-descent, the time-step problem minimizes a regularized and discretized version of $E(v) + \frac{1}{2\Delta t}\|v - u^{n-1}\|^2_{-1}$. When the regularization parameter δ is small it is important to choose a good scheme for this minimization problem. We use a primal-dual method introduced in [1, 7], which has the advantage of being very efficient even when the regularization parameter is quite small (see Section 5).

Our convergence analysis relies mostly on standard arguments for the numerical analysis of parabolic problems. The overall strategy is to estimate separately the errors associated with regularization, time-stepping, and spatial discretization. We do this by first comparing u to u_δ (the solution of the regularized problem), then comparing u_δ to its discrete-in-time approximation u_δ^τ (obtained by solving a variational problem at each time step), then finally comparing u_δ^τ to its discrete-in-space approximation $u_\delta^{\tau,h}$. Our main convergence results is Theorem 4.10, which demonstrates convergence in the natural (but rather weak) space $L^\infty(0,T;H^{-1}_{per})$. The methods needed to compare u to u_δ and u_δ to u_δ^τ are well-established; we follow [17] for the former and [24] for the latter. The analysis needed to compare u_δ^τ to $u_\delta^{\tau,h}$ draws some of its ideas from the work of Barrett and Liu concerning the parabolic p-Laplacian [3].

Besides proving results about convergence, we also solve the PDE numerically. As often happens, the numerically-observed convergence is somewhat better than what we can prove.

Notation. Throughout the paper we use the notation $\|\cdot\|_{s,q}$ and $|\cdot|_{s,q}$ for $s \in [0,\infty)$ and $q \in (0,\infty]$ to denote, respectively, standard norms and seminorms associated to Sobolev spaces $W^{s,q}(\Omega)$. We also use the notation $\|\cdot\|_s$ and $|\cdot|_s$, for $s \in [0,\infty)$ to denote respectively, standard norms and seminorms associated to Hilbert spaces $H^s(\Omega)$. We use c to denote an arbitrary constant independent of mesh parameters and δ. Finally, given a sequence of number a^n, we introduce the notation $d_t a^n = (a^n - a^{n-1})/\tau$, τ representing the size of the time step.

Acknowledgement. We thank the anonymous referees for their helpful comments, which led to substantial improvement of the paper.

2. Steepest Descent Framework. As already noted in the introduction, the PDE (1.1) is not to be taken literally, since $\nabla u/|\nabla u|$ is apparently undefined on the facets, where $\nabla u = 0$. Our continuous-time, continuous-space solution $u(x,t)$ is really defined as the evolution of its initial data $u^0(x)$ under H^{-1}_{per} steepest-descent for the functional E. We explain briefly what this means in Section 2.1. Then we discuss implicit time-stepping in Section 2.2 and the use of regularization in Section 2.3.

2.1. The steepest-descent interpretation of (1.1). The goals of this paper are very concrete: numerical algorithms and convergence theorems for the solution of (1.1). The interpretation of (1.1) is by contrast a bit abstract: it requires defining the Hilbert space H^{-1}_{per} and discussing the subgradient of E. The reader who finds this discussion uncomfortably abstract should skip to Sections 2.2 and 2.3, since as a practical matter the only equations we ever study numerically are discrete-time, regularized versions of (1.1).

The function space H^{-1}_{per} is the dual of $H^1_{per}/\mathbb{R}$. This space is equipped with the norm associated with the inner product

$$\langle f, g\rangle_{-1} = \int_\Omega \langle \nabla(-\Delta^{-1} f), \nabla(-\Delta^{-1} g)\rangle \, dx.$$

Here Δ^{-1} denotes the inverse of the Laplacian, and we use the fact that the Laplacian is an isomorphism from $H^1_{per}/\mathbb{R}$ to H^{-1}_{per}.

We are interested in the H^{-1}_{per} steepest descent of the functional

$$E(u) = \int_\Omega \beta|\nabla u| + \frac{1}{p}|\nabla u|^p dx = \int_\Omega \Phi(\nabla u)\, dx, \tag{2.1}$$

defined as a special case of a much more general theory (see e.g. [6]). It is conventional to define the *domain* of E by

$$D(E) = \{v \in H^{-1}_{per} \,:\, E(v) < \infty\}. \tag{2.2}$$

Since E is not differentiable, the steepest-descent evolution cannot be expressed as $u_t = -\nabla_{H^{-1}_{per}} E$. Rather, it must be expressed in terms of the H^{-1}_{per} *subdifferential*, defined by

$$\partial_{H^{-1}_{per}} E(u) = \{v \in H^{-1}_{per} \,:\, \langle v, z-u\rangle_{-1} \le E(z) - E(u), \text{ for all } z \in H^{-1}_{per}\}.$$

Kashima showed in [19] that this subgradient can be made quite explicit:

$$\partial_{H^{-1}_{per}} E(u) = \{\Delta\nabla\cdot\xi \,:\, \xi(x) \in \partial\Phi(\nabla u(x))\}$$

where $\partial\Phi$ is the subgradient of the function $\Phi(\eta) = \beta|\eta| + \frac{1}{p}|\eta|^p$, namely

$$\partial\Phi(\eta) = \begin{cases} \beta\eta/|\eta| + |\eta|^{p-2}\eta & \text{if } \eta \neq 0 \\ \{|\xi| \le \beta\} & \text{if } \eta = 0. \end{cases}$$

Thus: the steepest-descent framework interprets (1.1) by permitting $\nabla u/|\nabla u|$ to be replaced by any vector of length ≤ 1 when $\nabla u = 0$.

The general theory [6, 19] shows that for any $u^0 \in D(E)$ there is a unique steepest-descent evolution starting from u^0. The energy E decreases with time, and

$$-u_t \in \partial_{H^{-1}_{per}} E(u). \tag{2.3}$$

2.2. Implicit-in-time approximation. A basic fact about the steepest-descent evolution is that it can be approximated by implicit time-stepping.

Fixing a timestep $\tau > 0$ and a time interval $[0, T]$, let N be the smallest integer such that $N\tau \ge T$. For $n \in \{0, 2.., N\}$ we define the functions u^n recursively by letting u^0 be the initial data and letting u^n solve the minimization problem

$$\min_{v\in D(E)} E(v) + \frac{\|v - u^{n-1}\|^2_{-1}}{2\tau}. \tag{2.4}$$

Now define $u^\tau(x,t)$ by piecewise-linear interpolation in time:

$$u^\tau = u^{n-1} + \frac{t-(n-1)\tau}{\tau}\left(u^n - u^{n-1}\right) \text{ for } t \in [(n-1)\tau, n\tau). \tag{2.5}$$

The general theory assures that $u^\tau \to u$ as $\tau \to 0$; the error is linear in τ, as one naturally expects [24].

2.3. Regularization. Our numerical scheme relies on regularization. We now examine in detail the associated error. Let $\varphi_\delta(x)$ be a regularization of $|x|$, for example

$$\varphi_\delta(\nabla v) = \sqrt{|\nabla v|^2 + \delta} \tag{2.6}$$

(the regularization we used for our numerics) and consider the regularized energy

$$E_\delta(v) = \int_\Omega \beta\varphi_\delta(\nabla v) + \frac{1}{p}|\nabla v|^p dx = \int_\Omega \Phi_\delta(\nabla u_\delta)dx. \tag{2.7}$$

The associated regularized evolution u_δ solves

$$u_{\delta t} = -\Delta\nabla\cdot(\Phi'_\delta(\nabla u_\delta)) \text{ in } \Omega. \tag{2.8}$$

with the δ-independent initial data u^0 at $t=0$. (Our notation is a bit informal: $\Phi'_\delta(\nabla u_\delta) = \beta\varphi'_\delta(\nabla u) + |\nabla u|^{p-2}\nabla u$ represents the vector-valued function $\partial\Phi_\delta/\partial\nabla u$.) The PDE (2.8) equation amounts to H^{-1}_{per} steepest-descent for E_δ.

We want to estimate the difference between u_δ and u. Rather than focus on the example (2.6), let us say more generally what constitutes a "reasonable" regularization φ_δ. We shall assume that

$$\begin{gathered} 0<\delta<1;\ \varphi_\delta \text{ is convex; } \varphi_\delta(x) \le C(|x| + |x|^p + 1) \text{ for all } x \text{ and} \\ \text{with } C \text{ independent of } \delta\text{; there exists } \alpha>0 \text{ such that } |\varphi_\delta(x) - |x|| \le C\delta^\alpha \\ \text{for all } x, \text{ with } C \text{ independent of } \delta \text{ and } \alpha. \end{gathered} \tag{2.9}$$

The second condition guarantees that the functionals E_δ and E have the same domain, since if $E(u)$ is finite then *a fortiori* $\nabla u \in L^p$. The third condition specifies a rate for the convergence $\varphi_\delta(x) \to |x|$.

Theorem 2.1. *Let u and u_δ be the solutions of Equations (1.1) and (2.8), respectively. Assume that the regularization satisfies (2.9). Then*

$$\operatorname*{ess\,sup}_{t\in[0,T]} \|u(t) - u_\delta(t)\|_{-1} \le c\sqrt{T}\delta^{\alpha/2}. \tag{2.10}$$

Proof. The argument is almost the same as used in [17] to prove that paper's Theorem 2; the only difference is that we are considering an H^{-1}_{per} steepest descent rather than an L^2 steepest descent. □

Remark 2.2. *The regularization we use in our numerical work, (2.6), satisfies $|\varphi_\delta(x) - |x|| \le \delta^{1/2}$; thus (2.9) holds with $\alpha = 1/2$.*

3. Discretization using Finite Elements. This section introduces a convenient spatial discretization using piecewise-linear finite elements and a mixed formulation. Section 3.1 lays the foundation introducing a mixed formulation of the time-step variational problem. Section 3.2 discusses the associated finite element scheme. Finally, Section 3.3 introduces modifications associated with the primal-dual scheme.

3.1. A mixed variational formulation. Section 2 discussed implicit time stepping and regularization separately, but we want to use them together. So our goal is to discretize the timestep variational problems, which define u_δ^n recursively by solving

$$\min_{v\in D(E)} \frac{\|v-u_\delta^{n-1}\|_{-1}^2}{2\tau} + E_\delta(v) \tag{3.1}$$

with $u_\delta^0 = u^0$. The functions u_δ^n determine a spatially-continuous approximate solution u_δ^τ of our PDE by linear interpolation:

$$u_\delta^\tau = u_\delta^{n-1} + \frac{t-(n-1)\tau}{\tau}\left(u_\delta^n - u_\delta^{n-1}\right) \text{ for } t\in[(n-1)\tau, n\tau). \tag{3.2}$$

The optimality condition for (3.1) is

$$\langle \frac{u_\delta^n - u_\delta^{n-1}}{\tau}, v\rangle_{-1} = -\int_\Omega \Phi_\delta'(\nabla u_\delta^n)\cdot\nabla v\,dx \quad \forall v\in D(E). \tag{3.3}$$

One might be tempted to ask that the finite-element version of u_δ^n satisfy (3.3) for all v in the finite-element space. But this is not convenient, because the H^{-1} inner product of two finite-element functions is nonlocal and laborious to compute.

This difficulty is familiar: the same issue arises when discretizing the biharmonic equation. The solution is also familiar: one can avoid the use of negative norms by introducing a "mixed formulation," see e.g. [11, 16]. In the present setting the mixed formulation of (3.3) is this: given $u_\delta^{n-1}\in D(E)$, find $\tilde u_\delta^n \in D(E)$ and $w_\delta^n \in H^1_{per}(\Omega)/\mathbb{R}$ such that

$$\begin{aligned} \int_\Omega \frac{\tilde u_\delta^n - u_\delta^{n-1}}{\tau} v\,dx &= \int_\Omega \nabla w_\delta^n\cdot\nabla v\,dx \quad \forall v\in H^1_{per}(\Omega)/\mathbb{R} \\ \int_\Omega w_\delta^n \phi\,dx &= -\int_\Omega \Phi_\delta'(\nabla\tilde u_\delta^n)\cdot\nabla\phi\,dx \quad \forall \phi\in D(E). \end{aligned} \tag{3.4}$$

We state the equivalence as a lemma:

Lemma 3.1. *If u_δ^n solves (3.3), then the unique solutions $\tilde u_\delta^n$ and w_δ^n of (3.4) are $\tilde u_\delta^n = u_\delta^n$ and $w_\delta^n = \nabla\cdot\Phi_\delta'(\nabla u_\delta^n)$.*

Proof. The result follows easily from the definitions. □

3.2. FEM approximation. Let $\mathcal{T}^h(\Omega)$ be a regular partition of the domain Ω by triangular elements K_i. Our finite element space is

$$V^h(\Omega) = \left\{\phi\in H^1_{per}(\Omega) : \phi|_{K_i}\in\mathcal{P}_1(K_i) \text{ for all } i \text{ and } \textstyle\int_\Omega \phi\,dx = 0\right\},$$

where $\mathcal{P}_1(K_i)$ is the space of polynomials of degree less than or equal to 1. Note that $V^h(\Omega)\subset D(E)$.

The Galerkin approximation $u_\delta^{n,h}$ of u_δ^n is defined recursively as follows. When $n=0$, $u_\delta^{0,h} = u^{0,h} = \mathcal{I}^h u^0 - \int_\Omega \mathcal{I}^h u^0\,dx$, where $\mathcal{I}^h : D(E)\to V^h(\Omega)$ is a suitable finite element interpolation operator. Here depending on the regularity of v, the interpolation $\mathcal{I}^h v$ in $V^h(\Omega)$ is obtained either by a standard pointwise interpolation (if v is continuous) or by a local averaging procedure (if v is not continuous; see [9]). Given $u_\delta^{n-1,h}$, we determine $u_\delta^{n,h}$ by asking that $u_\delta^{n,h}, w_\delta^{n,h}\in V^h(\Omega)$ solve

$$\begin{aligned} \int_\Omega \frac{u_\delta^{n,h} - u_\delta^{n-1,h}}{\tau} v^h\,dx &= \int_\Omega \nabla w_\delta^{n,h}\cdot\nabla v^h\,dx \quad \forall v^h\in V^h(\Omega) \\ \int_\Omega w_\delta^{n,h}\phi^h\,dx &= -\int_\Omega \Phi_\delta'(\nabla u_\delta^{n,h})\cdot\nabla\phi^h\,dx \quad \forall\phi^h\in V^h(\Omega). \end{aligned} \tag{3.5}$$

To show that this problem has a unique solution, we argue as Barrett, Blowey, and Garcke did in [4] for a different nonlinear fourth-order problem. Define the discrete inverse Laplacian $\Delta^{-1,h} : H^{-1}_{per} \to V^h(\Omega)$ by

$$\int_\Omega \nabla(-\Delta^{-1,h} v) \cdot \nabla \phi^h \, dx = \int v\phi^h dx, \quad \forall \phi^h \in V^h(\Omega). \tag{3.6}$$

Note that $\Delta^{-1,h} v \in V^h(\Omega)$ exists and is unique for any $v \in H^{-1}_{per}$ (we use here the fact that the functions in H^{-1}_{per} have mean value zero). We also define the inner product $\langle \cdot, \cdot \rangle_{-1,h}$ on H^{-1}_{per}, and the norm associated with it by

$$\langle \phi, v \rangle_{-1,h} = \int_\Omega \nabla(\Delta^{-1,h}\phi) \cdot \nabla(\Delta^{-1,h} v) \, dx, \quad \|v\|_{-1,h} = \langle v, v \rangle^{1/2}_{-1,h} \tag{3.7}$$

Lemma 3.2. *Given $u^{n-1,h}_\delta \in V^h(\Omega)$ the problem (3.5) has a unique solution $u^{n,h}_\delta$ and $w^{n,h}_\delta \in V^h(\Omega)$; moreover $u^{n,h}_\delta$ solves the variational problem*

$$\min_{v \in V^h(\Omega)} \frac{\|v - u^{n-1,h}_\delta\|^2_{-1,h}}{2\tau} + E_\delta(v). \tag{3.8}$$

and satisfies, for every $v^h \in V^h(\Omega)$

$$\langle d_t u^{n,h}_\delta, v^h \rangle_{-1,h} = -\int_\Omega (\beta \varphi'_\delta(\nabla u^{n,h}_\delta) + |\nabla u^{n,h}_\delta|^{p-2} \nabla u^{n,h}_\delta) \cdot \nabla v^h dx. \tag{3.9}$$

Proof. One verifies using the definitions and injectivity of the map $\Delta^{-1,h}$: $V^h(\Omega) \to V^h(\Omega)$ that (3.5) and (3.9) are equivalent, and that they are the first-order optimality conditions for the variational problem (3.8). Existence and uniqueness follow, using the strict convexity of (3.8). ∎

As usual, the functions $u^{n,h}_\delta$ determine an approximate solution $u^{\tau,h}_\delta$ of our PDE by linear interpolation:

$$u^{\tau,h}_\delta = u^{n-1,h}_\delta + \frac{t - (n-1)\tau}{\tau} \left(u^{n,h}_\delta - u^{n-1,h}_\delta \right) \text{ for } t \in [(n-1)\tau, n\tau). \tag{3.10}$$

The rest of this subsection develops some properties of the inner product $\langle \cdot, \cdot \rangle_{-1,h}$ and the norm $\| \cdot \|_{-1,h}$ which will be needed for our convergence analysis.

Lemma 3.3. *Let $v \in L^2(\Omega)$ and $\phi \in H^{-1}_{per}$. Then*

$$|\langle \phi, v \rangle_{-1} - \langle \phi, v \rangle_{-1,h}| \le ch \|\phi\|_{-1} \|v\|_0, \tag{3.11}$$

and

$$|\|v\|_{-1} - \|v\|_{-1,h}| = \|v\|_{-1} - \|v\|_{-1,h} \le ch \|v\|_0. \tag{3.12}$$

Proof. From the definition of the inner product we have

$$\langle \phi, v \rangle_{-1} - \langle \phi, v \rangle_{-1,h} = \int_\Omega \phi(-\Delta^{-1} v + \Delta^{-1,h} v) \, dx \le \|\phi\|_{-1} \|\Delta^{-1} v - \Delta^{-1,h} v\|_1.$$

Inequality (3.11) follows from easily, using the standard finite element error estimate

$$\|\Delta^{-1} v - \Delta^{-1,h} v\|_1 \le ch \|\Delta^{-1} v\|_2 \le ch \|v\|_0.$$

Next we show that $\|v\|_{-1,h} \leq \|v\|_{-1}$. Indeed, $-\Delta^{-1}v$ and $-\Delta^{-1,h}v$ minimize the functional $I(z) = |z|_1^2/2 - \langle v, z\rangle$ on $H^1_{per}(\Omega)/\mathbb{R}$ and $V^h(\Omega)$ respectively. Therefore $I(-\Delta^{-1}v) \leq I(-\Delta^{-1,h}v)$. Since $I(-\Delta^{-1}v) = -\|v\|_{-1}^2/2$ and $I(-\Delta^{-1,h}v) = -\|v\|_{-1,h}^2/2$ it follows that $\|v\|_{-1,h} \leq \|v\|_{-1}$.

Finally we show (3.12). Arguing as for the proof of (3.11) we find that

$$\left|\|v\|_{-1}^2 - \|v\|_{-1,h}^2\right| \leq ch^2\|v\|_0^2.$$

This implies (3.12) since $(\|v\|_{-1} - \|v\|_{-1,h})^2 \leq \|v\|_{-1}^2 - \|v\|_{-1,h}^2$ (using the fact that $0 \leq \|v\|_{-1,h} \leq \|v\|_{-1}$). □

3.3. Numerical Implementation. We now discuss how to solve the discretized problem (3.5). When δ is relatively large this can be done by minimizing (3.8) using an iterative optimization scheme such as Newton's method. When δ gets small however that works poorly, due to the nearly-singular character of the energy. We obtained better results using a version of the primal-dual method introduced in [1, 7]. The basic idea in a continuous-space setting is to introduce the new unknown

$$z_\delta^n = \frac{\nabla u_\delta^n}{\sqrt{|\nabla u_\delta^n|^2 + \delta}}.$$

The system (3.3) can then be written as

$$\begin{aligned} \langle \frac{u_\delta^n - u_\delta^{n-1}}{\tau}, v\rangle_{-1} = -\int_\Omega (\beta z_\delta^n + |\nabla u_\delta^n|^{p-2}\nabla u_\delta^n)\cdot\nabla v\,dx \quad \forall v \in D(E) \qquad (3.13)\\ z_\delta^n\sqrt{|\nabla u_\delta^n|^2 + \delta} = \nabla u_\delta^n. \end{aligned}$$

and we can use Newton's method to solve for z_δ^n and u_δ^n simultaneously. The advantage of this scheme is that it remains robust when δ is small. In particular, the number of Newton iterations required to solve (3.13) is almost independent of δ; see Subsection 5.1.

To implement this idea in our discrete finite-element setting, we take advantage of the fact that our finite elements are piecewise linear. Therefore

$$z_\delta^{n,h} = \frac{\nabla u_\delta^{n,h}}{\sqrt{|\nabla u_\delta^{n,h}|^2 + \delta}} \qquad (3.14)$$

is constant on each element.

The discrete version of (3.13) is obtained as follows. We focus for simplicity on the case $p = 3$ (the exponent of primary physical interest) and $\Omega = (0, 1)$ (one-dimensional dynamics, representing the evolution of a two-dimensional staircase; the case $\Omega = (0, 1)\times(0, 1)$ is similar). Our finite-element space $V^h(\Omega)$ consists of piecewise linear, mean-zero, periodic functions on a uniform mesh of size $h = 1/M$; each function in $V^h(\Omega)$ has the form $\sum_{i=1}^M \alpha_i\phi_i$ where $\sum_{i=1}^M \alpha_i = 0$ and $\{\phi_i\}_{i=1}^M$ is the periodic piecewise linear function that equals 1 at the ith node and 0 at the other nodes for $i \neq 1$ and ϕ_1 has value 1 at $x = 0$ and $x = 1$ and vanishes at the other nodes. The functions $z^{n,h}$ belong to the space of functions that are constant on each interval; the general form of such a function is $\sum_{i=1}^M \eta_i\sigma_i$ where σ_i is equal to 1 on the ith interval and 0 on the others:

Suppose $u_\delta^{n,h} = \sum_{i=1}^M \alpha_i^n \phi_i$ with $\sum_{i=1}^M \alpha_i^n = 0$ and $z_\delta^{n,h} = \sum \eta_i^n \sigma_i$. Then, for $p = 3$ we obtain the following discrete version of (3.13)

$$\sum_{i=1}^M \left(d_t \alpha_i^n \langle \phi_i, \phi_j \rangle_{-1,h} + \int_\Omega \beta \eta_i^n \sigma_i \cdot \nabla \phi_j + \alpha_i^n |\sum_k \alpha_k^n \nabla \phi_k | \nabla \phi_i \cdot \nabla \phi_j \, dx \right) = 0$$

$$\left(\sum_i \eta_i \sigma_i \right) \left(\sqrt{|\sum_k \alpha_k^n \nabla \phi_k|^2 + \delta} \right) = \sum_k \alpha_k^n \nabla \phi_k$$

for all j. We solve this nonlinear system (subject to the constraint $\sum \alpha_i = 0$) by Newton's method. The implementation is straightforward, since its Jacobian is easily accessible.

4. Convergence Analysis for the FEM. This section studies the convergence of our finite element scheme. We shall assume that the discretized problem (3.9) is solved exactly. Our main result is Theorem 4.10, which estimates the error between the the exact solution u (with $\delta = 0$) and its numerical approximation $u_\delta^{\tau,h}$, in the norm $\| \cdot \|_{L^\infty(0,T;H^{-1}_{per})}$.

There are three sources of error: regularization, time-stepping, and spatial discretization. We shall estimate them separately, using the triangle inequality

$$\|u - u_\delta^{\tau,h}\| \le \|u - u_\delta\| + \|u_\delta - u_\delta^\tau\| + \|u_\delta^\tau - u_\delta^{\tau,h}\|.$$

We already estimated the first term on the right, in Theorem 2.1. An estimate for the second term is available from the existing literature (see Theorem 4.1). The main task of this section is thus to handle the spatial discretization error.

To analyze the effect of discretization in time, we observe that for $\delta > 0$ the H^{-1} subgradient $\partial_{H^{-1}_{per}} E_\delta$ has just one element, namely $\Delta \nabla \cdot \Phi'_\delta(\nabla u)$. Abusing notation slightly, we shall write $\partial E_\delta(u) = \Delta \nabla \cdot \Phi'_\delta(\nabla u)$ in what follows. The standard tools for controlling the time discretization error of a steepest descent are the resolvent and the Yosida approximation of the associated operator, which in the present setting are respectively

$$J_\lambda^\delta = (I + \lambda \partial E_\delta)^{-1} \quad \text{and} \quad A_\lambda^\delta = \frac{1}{\lambda}(I - J_\lambda^\delta). \tag{4.1}$$

Here I is the identity operator and $\lambda > 0$; for more properties of these operators see [6, 15]. The following theorem estimates the time discretization error, i.e. the difference between u_δ and u_δ^τ:

Theorem 4.1. *Let u_δ and u_δ^τ be defined by Equations (2.8) and (3.2), and set $\underline{u}_\delta^\tau(s) = u_\delta^{n-1}$ for $(n-1)\tau \le s < n\tau$. Then*

$$\|u_\delta(t) - u_\delta^\tau(t)\|_{-1}^2 + 4 \int_0^t (A_\tau^\delta(\underline{u}_\delta^\tau(t)) - \partial E_\delta(u_\delta(t)), J_{\Delta t}^\delta(\underline{u}_\delta^\tau(t)) - u_\delta(t)) \, dt$$
$$+ \tau \int_0^t \|\partial E_\delta(u_\delta(t)) - A_\tau^\delta(\underline{u}_\delta^\tau(t))\|_{-1}^2 \, dt \le C\tau^2 \|\partial E_\delta(u^0)\|_{-1}^2 \tag{4.2}$$

for all $t \in [0, T]$. The constant C depends on T but not on δ or Δt.

Proof. This is essentially Theorem 5 of [24]. The only difference is that in the result just cited, the term $\|u_\delta(t) - u_\delta^\tau(t)\|_{-1}^2$ on the left side of (4.2) is replaced by

$\|u_\delta(t) - \underline{\mathrm{u}}_\delta^\tau(t)\|_{-1}^2$. However, it easy to see that the argument in [24] also proves our assertion. ◻

While the constant in (4.2) is independent of δ, the right hand side nevertheless depends on δ through the term $\|\partial E_\delta(u^0)\|_{-1}$. The following Lemma shows that this dependence is at worst proportional to $1/\delta$ when the initial data are smooth enough. (This would not be the case for faceted initial data, but it would be the case for example if $u^0(x) = c\sin(x)$.)

Lemma 4.2. *Assume φ_δ is given by (2.6), $0 < \delta \leq 1$, and $p \geq 3$. If the spatial dimension is $d \leq 4$ then for any $v \in H^3(\Omega)$ we have*

$$\|\partial E_\delta(v)\|_{-1} \leq \frac{b(\|v\|_3)}{\delta} \tag{4.3}$$

where $b(\cdot)$ is a polynomial of degree less or equal than p. Moreover in any space dimension we have a similar statement for any $v \in W^{3,4}(\Omega)$, with the RHS of the last inequality replaced by $\frac{b(\|v\|_{3,4})}{\delta}$. Where $b(\cdot)$ is a polynomial of degree less than or equal to p.

Proof. To simplify the exposition we focus on the proof of (4.3) in space dimension one (the arguments for $d > 1$ are similar). We have

$$\begin{aligned}
\|\partial E_\delta(v)\|_{-1} &= \left[\int_\Omega (\Phi'_\delta(v_x))^2_{xx}\,dx\right]^{1/2} \\
&= \left[\int_\Omega \left(\frac{\beta v_{xxx}}{(v_x^2+\delta)^{1/2}} - \frac{\beta v_x^2 v_{xxx}}{(v_x^2+\delta)^{3/2}} - 3\frac{\beta v_x v_{xx}^2}{(v_x^2+\delta)^{3/2}}\right.\right. \\
&\qquad \left.\left. + 3\frac{\beta v_x^3 v_{xx}^2}{(v_x^2+\delta)^{5/2}} + (|v_x|^{p-2}v_x)_{xx}\right)^2 dx\right]^{1/2} \\
&\leq 2\frac{\beta|v|_3}{\delta^{1/2}} + 6\frac{\beta|v|_{2,4}^2}{\delta} + \|(|v_x|^{p-2}v_x)_{xx}\|_0.
\end{aligned} \tag{4.4}$$

Here we have used the triangle inequality and an L^∞ bound for the terms of the form $v_x^{a_0}/(|v_x|^2+\delta)^{a_1}$ to obtain the last line. For instance, the term $|\beta v_{xxx}/(v_x^2+\delta)^{1/2}|$ is bounded by $|\beta v_{xxx}\delta^{-1/2}|$. To estimate the last term of the above inequality, we observe that the function $f(s) = |s|^{p-2}s$ has derivative

$$f'(s) = (p-1)|s|^{p-2} \text{ if } s \neq 0.$$

Therefore f' is continuous if $p \geq 2$ and f'' is bounded at $s = 0$ if $p \geq 3$, and

$$\|(|v_x|^{p-2}v_x)_{xx}\|_0 = (p-1)(p-2)\|v_x^{p-3}\|_0.$$

Combining these observations with (4.4) we easily obtain (4.3) from the hypothesis $v \in H^3(\Omega)$ and an application of the Sobolev inequality. ◻

Remark 4.3. *Theorem 4.1 and Lemma 4.2 give $\|u_\delta(t) - u_\delta^\tau(t)\|_{-1} \leq c\frac{\tau}{\delta}$.*

We turn now to the main task of this section: estimation of the spatial discretization error.

The following result is Lemma 2.2 from [3]. It will be used to prove Lemma 4.5, and also for handling the term $|\nabla u|^{p-2}\nabla u$ in the proof of Lemma 4.7.

Lemma 4.4. *For any $p \in (1,\infty)$, $\epsilon \geq 0$ and $d \geq 1$ there exist positive constants c_1 and c_2 such that: for all $\xi, \eta \in \mathbb{R}^d$,*

$$||\xi|^{p-2}\xi - |\eta|^{p-2}\eta| \leq c_1|\xi-\eta|^{1-\epsilon}(|\xi|+|\eta|)^{p-2+\epsilon} \tag{4.5}$$

and

$$(|\xi|^{p-2}\xi - |\eta|^{p-2}\eta) \cdot (\xi - \eta) \geq c_2|\xi - \eta|^{2+\epsilon}(|\xi| + |\eta|)^{p-2-\epsilon}.$$

Our next result bounds $d_t u_\delta^n = (u_\delta^n - u_\delta^{n-1})/\tau$ and $d_t u_\delta^{n,h} = (u_\delta^{n,h} - u_\delta^{n-1,h})/\tau$ in the norms $\|\cdot\|_{-1}$ and $\|\cdot\|_{-1,h}$. We will use it in the proof of Lemma 4.7.

Lemma 4.5. *Let u_δ^n and $u_\delta^{n,h}$ be defined by (3.3) and (3.9). Assume φ_δ satisfies (2.9). Then there exists a constant $c > 0$ independent of τ, h and δ such that*

$$\|d_t u_\delta^n\|_{-1} \leq \|\partial E_\delta(u^0)\|_{-1}, \quad \|d_t u_\delta^{n,h}\|_{-1,h} \leq \|d_t u_\delta^{1,h}\|_{-1,h} \leq \frac{c}{\tau^{1/2}} \tag{4.6}$$

Furthermore, if u^0 satisfies Lemma 4.2's hypothesis, and φ_δ is given by (2.6), then

$$\|d_t u_\delta^{1,h}\|_{-1,h} \leq c\kappa_\delta^{\tau,h} \tag{4.7}$$

where

$$\kappa_\delta^{\tau,h} = \min\left[\frac{|u^{0,h} - u^0|_{1,\infty}}{h^{2-(d/2)}}(\frac{1}{\delta^{1/2}} + |u^0|_{1,\infty}^{p-2}) + \|\nabla \cdot \Phi'_\delta(u_0)\|_1, \frac{c}{\tau^{1/2}}\right] \tag{4.8}$$

Proof. To prove the first inequality of (4.6), recall the nonlinear resolvent operator defined by (4.1). Clearly

$$u_\delta^n = (J_\tau^\delta)^n u^0,$$

and applying the contraction property of J_τ^δ (see for example Theor. 2 pg 526 of [15]) we obtain

$$\|u_\delta^n - u_\delta^{n-1}\|_{-1} \leq \|(J_\tau^\delta)^{n-1}u_\delta^0 - (J_\tau^\delta)^{n-2}u_\delta^0\|_{-1} \leq \|J_\tau^\delta u_\delta^0 - u_\delta^0\|_{-1}.$$

We now observe that $J_\tau^\delta u_\delta^0 - u_\delta^0 = \tau A_\tau^\delta u^0$. Finally, from the properties of the Yosida approximation we have $\|A_\tau^\delta u^0\|_{-1} \leq \|\partial E_\delta(u^0)\|_{-1}$.

Turning now to the second inequality of (4.6), we begin by showing that $\|u_\delta^{n,h} - u_\delta^{n-1,h}\|_{-1,h} \leq \|u_\delta^{n-1,h} - u_\delta^{n-2,h}\|_{-1,h}$. Recall from (3.9) that

$$\langle d_t u_\delta^{n,h}, v^h\rangle_{-1,h} = -\int_\Omega \Phi'_\delta(\nabla u_\delta^{n,h}) \cdot \nabla v^h \, dx$$

and

$$\begin{aligned}
&\|u_\delta^{n-1,h} - u_\delta^{n-2,h}\|_{-1,h}^2 = \|u_\delta^{n,h} - u_\delta^{n-1,h} - (u_\delta^{n,h} - u_\delta^{n-1,h}) + u_\delta^{n-1,h} - u_\delta^{n-2,h}\|_{-1,h}^2 \\
&\quad = \|u_\delta^{n,h} - u_\delta^{n-1,h}\|_{-1,h}^2 + \|u_\delta^{n,h} - 2u_\delta^{n-1,h} + u_\delta^{n-2,h}\|_{-1,h}^2 \\
&\qquad + 2\tau \int_\Omega (\Phi'_\delta(\nabla u_\delta^{n,h}) - \Phi'_\delta(\nabla u_\delta^{n-1,h})) \cdot \nabla(u_\delta^{n,h} - u_\delta^{n-1,h}) \, dx \\
&\quad \geq \|u_\delta^{n,h} - u_\delta^{n-1,h}\|_{-1,h}^2,
\end{aligned}$$

where we have used (3.9) to obtain the second equation, and the last inequality follows from the convexity of Φ_δ. Finally, we estimate $\|d_t u_\delta^{1,h}\|_{-1,h}$ using the steepest descent feature of the problem ($\|d_t u_\delta^{1,h}\|_{-1,h}^2 \le E_\delta(u^{0,h})/\tau$; see (3.8)), and the stability of the FEM interpolation operator.

Next we prove inequality (4.7). Let $w_\delta^{0,h} \in V^h(\Omega)$ be defined by

$$\int_\Omega w_\delta^{0,h} \phi^h dx = -\int_\Omega \Phi_\delta'(\nabla u^{0,h}) \cdot \nabla \phi^h \, dx \quad \forall \phi^h \in V^h(\Omega), \tag{4.9}$$

and let $w_\delta^0 = \nabla \cdot \Phi_\delta'(\nabla u^0)$. Using Equation (3.9), and adding and subtracting the term $\int_\Omega \Phi_\delta'(\nabla u^{0,h}) \cdot \nabla d_t u_\delta^{1,h} \, dx$ we obtain

$$\begin{aligned}
\|d_t u_\delta^{1,h}\|_{-1,h}^2 &= \int_\Omega -(\Phi_\delta'(\nabla u_\delta^{1,h}) - \Phi_\delta'(\nabla u^{0,h})) \cdot \nabla d_t u_\delta^{1,h} - \Phi_\delta'(\nabla u^{0,h}) \cdot \nabla d_t u_\delta^{1,h} \, dx \\
&\le -\int_\Omega \Phi_\delta'(\nabla u^{0,h}) \cdot \nabla d_t u_\delta^{1,h} \, dx, \quad \text{by the convexity of } \Phi_\delta \\
&\le \int_\Omega (w_\delta^{0,h} - P_0^h w_\delta^0) d_t u_\delta^{1,h} + P_0^h w_\delta^0 \cdot d_t u_\delta^{1,h} \, dx, \quad \text{by (4.9)} \\
&\le (\|w_\delta^{0,h} - P_0^h w_\delta^0\|_1 + \|P_0^h w_\delta^0\|_1) \|d_t u_\delta^{1,h}\|_{-1,h}
\end{aligned} \tag{4.10}$$

where P_0^h denotes the L^2 projection on $V^h(\Omega)$. The last inequality comes from the fact that if $f, g \in V^h(\Omega)$, then $\int_\Omega fg \, dx = \int_\Omega \nabla f \cdot \nabla(-\Delta^{-1,h} g) \, dx \le |f|_1 \|g\|_{-1,h}$.

The first term on the right hand side of the last inequality is estimated as follows. We first apply an inverse estimate (see Lemma 4.5.3 of [5]):

$$\|v\|_{m,r} \le c h^{l-m+(d/r)-(d/s)} \|v\|_{l,s} \quad \forall v \in V^h(\Omega) \tag{4.11}$$

to obtain

$$\|w_\delta^{0,h} - P_0^h w_\delta^0\|_1 \le c h^{-1} \|w_\delta^{0,h} - P_0^h w_\delta^0\|_0.$$

Then we use a basic property of the L^2 projection, and integration by parts to obtain

$$\begin{aligned}
\|w_\delta^{0,h} - P_0^h w_\delta^0\|_0^2 &= \int_\Omega (w_\delta^{0,h} - w_\delta^0)(w_\delta^{0,h} - P_0^h w_\delta^0) \, dx \\
&= -\int_\Omega (\Phi_\delta'(\nabla u^{0,h}) - \Phi_\delta'(\nabla u^0)) \cdot \nabla (w_\delta^{0,h} - P_0^h w_\delta^0) \, dx \\
&\le c|u^{0,h} - u^0|_{1,\infty} \left(\frac{1}{\delta^{1/2}} + (|u^{0,h}|_{1,\infty} + |u^0|_{1,\infty})^{p-2}\right) \|w_\delta^{0,h} - P_0^h w_\delta^0\|_{1,1} \quad \text{by (4.5)} \\
&\le c|u^{0,h} - u^0|_{1,\infty} \left(\frac{1}{\delta^{1/2}} + |u^0|_{1,\infty}^{p-2}\right) h^{-1+d/2} \|w_\delta^{0,h} - P_0^h w_\delta^0\|_0.
\end{aligned}$$

To obtain the last inequality we used the stability of the finite element interpolation operator, i.e. the estimate $|u^{0,h}|_{1,\infty} \le c|u^0|_{1,\infty}$, combined with (4.11).

The second term on right hand side of (4.10) can be estimated as follows. From the stability property of the L^2 projector we have $\|P_0^h w_\delta^0\|_1 \le c\|w_\delta^0\|_1$. Substituting the preceding results into (4.10) we confirm (4.8). ☐

The following standard result will be needed to handle the term $|\nabla u|^{p-2} \nabla u$ in the proof of Lemma 4.7.

Lemma 4.6. *For any $p \in (1, \infty)$ there exists $\epsilon_0 > 0$ with the following property: for any $\epsilon \in (0, \epsilon_0)$ and any $a, b, c \ge 0$ we have*

$$(a+b)^{p-2} bc \le \epsilon (a+b)^{p-2} b^2 + C(\epsilon, p)(a+c)^{p-2} c^2$$

for some constant $C(\epsilon, p)$ (independent of a, b, c).

Proof. This is Lemma 2.3 from [3]. □

The heart of any convergence theory for a Galerkin method is an estimate of the error in terms of the best approximation of the solution in the Galerkin space. Our next result provides such a estimate. The first term on the RHS of (4.13) comes from the fact that we use the norm $\|\cdot\|_{-1,h}$ to approximate the norm $\|\cdot\|_{-1}$ in (3.1).

Lemma 4.7. *Let u_δ^n and $u_\delta^{n,h}$ be defined as usual by (3.3) and (3.9). The regularization need not be $\varphi_\delta(z) = \sqrt{|z|^2+\delta}$, but we assume it satisfies*

$$|\varphi_\delta'(z)| \le c_0, \quad \forall z \in \mathbb{R}^d, \tag{4.12}$$

for some constant c_0 and (2.9). Let $e_\delta^n = u_\delta^n - u_\delta^{n,h}$. Then there exists constants $c, \gamma > 0$ and independent of τ, h and δ, such that for any $v^h \in V^h(\Omega)$ we have

$$\begin{aligned} &\frac{\|e_\delta^n\|_{-1,h}^2}{\tau} + \gamma \int_\Omega (|\nabla u_\delta^n| + |\nabla e_\delta^n|)^{p-2} |\nabla e_\delta^n|^2 \, dx \\ &\le ch\|\partial E_\delta(u^0)\|_{-1} + c\nu_\delta^{\tau,h}\|u_\delta^n - v^h\|_{-1} + c\int_\Omega |\nabla(u_\delta^n - v^h)|\, dx \\ &\quad + c\int_\Omega (|\nabla u_\delta^n| + |\nabla(u_\delta^n - v^h)|)^{p-2}|\nabla(u_\delta^n - v^h)|^2\, dx + \frac{\|e_\delta^{n-1}\|_{-1,h}}{\tau}\|e_\delta^n\|_{-1,h} \end{aligned} \tag{4.13}$$

where $\nu_\delta^{\tau,h} = c\tau^{-1/2}$. Furthermore, if u^0 satisfies Lemma 4.2's hypothesis, and φ_δ is given by (2.6), then $\nu_\delta^{\tau,h} = \kappa_\delta^{\tau,h}$; see (4.8).

Proof. From (3.3) and (3.9) we obtain

$$\begin{aligned} \langle d_t u_\delta^n - d_t u_\delta^{n,h}, e_\delta^n\rangle_{-1,h} &= \langle d_t u_\delta^n, e_\delta^n\rangle_{-1,h} - \langle d_t u_\delta^n, e_\delta^n\rangle_{-1} + \langle d_t u_\delta^n, e_\delta^n\rangle_{-1} - \langle d_t u_\delta^{n,h}, e_\delta^n\rangle_{-1,h} \\ &= \langle d_t u_\delta^n, e_\delta^n\rangle_{-1,h} - \langle d_t u_\delta^n, e_\delta^n\rangle_{-1} - \int_\Omega (\Phi_\delta'(\nabla u_\delta^n) - \Phi_\delta'(\nabla u_\delta^{n,h}))\cdot\nabla e_\delta^n\, dx \\ &\qquad - \langle d_t u_\delta^{n,h}, u_\delta^n\rangle_{-1,h} - \int_\Omega \Phi_\delta'(\nabla u_\delta^{n,h})\cdot \nabla u_\delta^n\, dx. \end{aligned}$$

Let v^h be an arbitrary element of $V^h(\Omega)$. Adding and subtracting $\int_\Omega (\Phi_\delta'(\nabla u_\delta^n) - \Phi_\delta'(\nabla u_\delta^{n,h}))\cdot\nabla v^h\, dx$ on the right hand side of the preceding equation, moving the third term on the right hand side to the left side, and using the convexity of φ_δ we obtain

$$\begin{aligned} &\langle d_t u_\delta^n - d_t u_\delta^{n,h}, e_\delta^n\rangle_{-1,h} + \int_\Omega (|\nabla u_\delta^n|^{p-2}\nabla u_\delta^n - |\nabla u_\delta^{n,h}|^{p-2}\nabla u_\delta^{n,h})\cdot\nabla e_\delta^n\, dx \\ &\qquad \le \langle d_t u_\delta^n, e_\delta^n\rangle_{-1,h} - \langle d_t u_\delta^n, e_\delta^n\rangle_{-1} + \langle d_t u_\delta^n, u_\delta^n - v^h\rangle_{-1} \\ &\qquad - \langle d_t u_\delta^{n,h}, u_\delta^n - v^h\rangle_{-1,h} - \int_\Omega (\Phi_\delta'(\nabla u_\delta^n) - \Phi_\delta'(\nabla u_\delta^{n,h}))\cdot\nabla(u_\delta^n - v^h)\, dx. \end{aligned}$$

We now use Lemma 4.4, the fact that the terms $|\nabla u_\delta^n| + |\nabla e_\delta^n|$ and $|\nabla u_\delta^n| + |\nabla u_\delta^{n,h}|$ are equivalent, and (4.12) to estimate the second term on LHS and the fifth term on the RHS of the above inequality, obtaining

$$\langle d_t u_\delta^n - d_t u_\delta^{n,h}, e_\delta^n\rangle_{-1,h} + \gamma \int_\Omega (|\nabla u_\delta^n| + |\nabla e_\delta^n|)^{p-2} |\nabla e_\delta^n|^2 \, dx$$
$$\leq \langle d_t u_\delta^n, e_\delta^n\rangle_{-1,h} - \langle d_t u_\delta^n, e_\delta^n\rangle_{-1} + \langle d_t u_\delta^n, u_\delta^n - v^h\rangle_{-1}$$
$$- \langle d_t u_\delta^{n,h}, u_\delta^n - v^h\rangle_{-1,h} + c\int_\Omega (1 + (|\nabla u_\delta^n| + |\nabla e_\delta^n|)^{p-2} |\nabla e_\delta^n|) |\nabla (u_\delta^n - v^h)| \, dx.$$

Next we use Lemmas 3.3 and 4.5 to estimate the sum of the first four terms on the right hand side of the last inequality, obtaining

$$\frac{\|e_\delta^n\|_{-1,h}^2}{\tau} + \gamma \int_\Omega (|\nabla u_\delta^n| + |\nabla e_\delta^n|)^{p-2} |\nabla e_\delta^n|^2 \, dx$$
$$\leq ch\|\partial E_\delta(u^0)\|_{-1} \|e_\delta^n\|_1 + c\nu_\delta^{\tau,h} \|u_\delta^n - v^h\|_{-1}$$
$$+ c\int_\Omega (1 + (|\nabla u_\delta^n| + |\nabla e_\delta^n|)^{p-2} |\nabla e_\delta^n|) |\nabla (u_\delta^n - v^h)| \, dx + \frac{\|e_\delta^{n-1}\|_{-1,h}}{\tau} \|e_\delta^n\|_{-1,h}.$$

Now we estimate the term $\|e_\delta^n\|_1$ that appears in the first term on the RHS of the last inequality. Taking $v = u_\delta^{n-1}$ in (3.1) we obtain

$$\|d_t u_\delta^n\|_{-1}^2 + 2E_\delta(u_\delta^n) \leq 2E_\delta(u_\delta^{n-1}), \tag{4.14}$$

hence $E_\delta(u_\delta^n) \leq E_\delta(u^0) \leq c(E(u^0) + 1)$. From (3.8) we obtain similar bound for $E_\delta(u_\delta^{n,h})$. Therefore

$$\|e_\delta^n\|_1 \leq \|u_\delta^n\|_1 + \|u_\delta^{n,h}\|_1 \leq c(E(u^0) + E(u^{0,h}) + 1) \leq c.$$

where we have used a Poincare and Sobolev inequality, and (for the last inequality) the stability of the finite element interpolation operator.

We finally obtain

$$\frac{\|e_\delta^n\|_{-1,h}^2}{\tau} + \gamma \int_\Omega (|\nabla u_\delta^n| + |\nabla e_\delta^n|)^{p-2} |\nabla e_\delta^n|^2 \, dx$$
$$\leq ch\|\partial E_\delta(u^0)\|_{-1} + c\nu_\delta^{\tau,h} \|u_\delta^n - v^h\|_{-1}$$
$$+ c\int_\Omega (1 + (|\nabla u_\delta^n| + |\nabla e_\delta^n|)^{p-2} |\nabla e_\delta^n|) |\nabla (u_\delta^n - v^h)| \, dx + \frac{\|e_\delta^{n-1}\|_{-1,h}}{\tau} \|e_\delta^n\|_{-1,h}. \tag{4.15}$$

Applying Lemma 4.6 to the third term on the right hand side of (4.15) gives the desired result (4.13). ◻

The following auxiliary Lemma will be used in the proof of Proposition 4.9.

Lemma 4.8. *Let $\{a_i\}_{i=0}^M$ be a sequence of positive real numbers satisfying, for some $\gamma > 0$, $a_i^2 \leq \tau\gamma + a_i a_{i-1}$, for $i = 1, 2, .., M$. Assume furthermore that $a_0 \leq \gamma^{1/2}$. Then $a_i \leq (i\tau + 1)\gamma^{1/2}$, for $i = 0, 1, .., M$.*

Proof. We argue by induction. The result holds for $i = 0$ by hypothesis. Assuming the result is true for i, we now prove it for $i+1$: If $a_{i+1} \leq \gamma^{1/2}$ we are done. If $a_{i+1} > \gamma^{1/2}$ then from hypothesis we have $a_{i+1}^2 \leq \tau a_{i+1}\gamma^{1/2} + a_{i+1}a_i$; dividing both sides by a_{i+1} and using the induction hypothesis to estimate a_i we conclude that $a_{i+1} \leq (\tau(i+1) + 1)\gamma^{1/2}$. ◻

The following proposition is our main estimate for the spatial discretization error. It controls $e_\delta^n = u_\delta^n - u_\delta^{n,h}$ in the norm $\|\cdot\|_{-1,h}$ and the seminorm $|\cdot|_{1,p}$. It also allows us to estimate e_δ^n in the H_{per}^{-1} norm through Lemma 3.3.

Proposition 4.9. *Let u_δ^n and $u_\delta^{n,h}$ be defined by as usual by (3.3) and (3.9). Assume φ_δ satisfies (2.9) and (4.12). Then there exists a constant $c > 0$ independent of n, τ, h and δ, such that*

$$\|e_\delta^n\|_{-1,h}^2 \leq c\rho_\delta^{\tau,h} \tag{4.16}$$

where

$$\rho_\delta^{\tau,h} = h\|\partial E_\delta(u^0)\|_{-1} + \max_{n\leq N} \inf_{v^h\in V^h(\Omega)} \Bigg(\nu_\delta^{\tau,h}\|u_\delta^n - v^h\|_{-1} \tag{4.17}$$
$$+ \int_\Omega |\nabla(u_\delta^n - v^h)| + (|\nabla u_\delta^n| + |\nabla(u_\delta^n - v^h)|)^{p-2}|\nabla(u_\delta^n - v^h)|^2\,dx\Bigg)$$

where $\nu_\delta^{\tau,h} = c\tau^{-1/2}$. Furthermore, if u^0 satisfies Lemma 4.2's hypothesis, and φ_δ is given by (2.6), then $\nu_\delta^{\tau,h} = \kappa_\delta^{\tau,h}$; see (4.8). Also,

$$\|\nabla e_\delta^n\|_{L^p}^p \leq c\nu_\delta^{\tau,h}(\rho_\delta^{\tau,h})^{1/2} + c\rho_\delta^{\tau,h}. \tag{4.18}$$

Proof. The first inequality (4.16) is an immediate consequence of (4.13) and Lemma 4.8. As for the second inequality (4.18): subtracting $\|e_\delta^n\|_{-1,h}^2/\tau$ from both sides of (4.13) gives

$$\int_\Omega |\nabla e_\delta^n|^p\,dx \leq \|d_t e_\delta^n\|_{-1,h}\|e_\delta^n\|_{-1,h} + ch\|\partial E_\delta(u^0)\|_{-1} + c\nu_\delta^{\tau,h}\|u_\delta^n - v^h\|_{-1}$$
$$+ c\int_\Omega |\nabla(u_\delta^n - v^h)| + (|\nabla u_\delta^n| + |\nabla(u_\delta^n - v^h)|)^{p-2}|\nabla(u_\delta^n - v^h)|^2\,dx. \tag{4.19}$$

We estimate the first term on the RHS of the last inequality using (4.16) and the following inequality:

$$\|d_t e_\delta^n\|_{-1,h} \leq \|d_t u_\delta^n\|_{-1,h} + \|d_t u_\delta^{n,h}\|_{-1,h} \leq c\nu_\delta^{\tau,h}.$$

Here, to obtain the last inequality we have used Lemma (4.5), the steepest descent estimate $\|d_t u_\delta^n\|_{-1} \leq c\tau^{-1/2}$, and the fact that $\|\partial E_\delta(u^0)\|_{-1} \leq \|\nabla\cdot\Phi_\delta'(u^0)\|_1$. We estimate the second and third terms on the RHS of (4.19) using (4.16) and (4.17). This leads directly to (4.18). ◻

We now combine our estimates for the errors due to regularization, implicit time stepping, and spatial discretization. The following theorem bounds the error between the unregularized continuum solution u and its numerical approximation $u_\delta^{\tau,h}$, in the $\|\cdot\|_{L^\infty(0,T;H_{per}^{-1})}$ norm.

Theorem 4.10. *Let u and $u_\delta^{\tau,h}$ be defined by (1.1) and (3.10), respectively. Assume φ_δ satisfies (2.9) and (4.12). Then there exists a constant c independent of τ, h, δ, such that*

$$\operatorname*{ess\,sup}_{t\in[0,T]} \|u(t) - u_\delta^{\tau,h}(t)\|_{-1} \leq c\left(T^{1/2}\delta^{1/4} + \tau\|\partial E_\delta(u^0)\|_{-1} + (\rho_\delta^{\tau,h})^{1/2}\right)$$

where $\rho_\delta^{\tau,h}$ is defined by Equation (4.17). Furthermore, if u^0 satisfies Lemma 4.2's hypothesis, and φ_δ is given by (2.6), then we may take $\nu_\delta^{\tau,h} = \kappa_\delta^{\tau,h}$ in the definition of $\rho_\delta^{\tau,h}$; see (4.8) and (4.17).

Proof. We use the triangle inequality to obtain

$$\begin{aligned}\|u(t) - u_\delta^{\tau,h}(t)\|_{-1} &\le \|u(t) - u_\delta(t)\|_{-1} + \|u_\delta(t) - u_\delta^\tau(t)\|_{-1} \\ &\quad + \|u_\delta^\tau(t) - u_\delta^{\tau,h}(t)\|_{-1} - \|u_\delta^\tau(t) - u_\delta^{\tau,h}(t)\|_{-1,h} + \|u_\delta^\tau(t) - u_\delta^{\tau,h}(t)\|_{-1,h}.\end{aligned}$$

The first, second, and the fourth terms on the RHS of the last inequality can be estimated using Theorem 2.1 and Remark 2.2, Theorem 4.1, and Proposition 4.9 respectively. The third term on the RHS of the last inequality is estimated as follows: first use inequality (3.12) to obtain

$$|\|u_\delta^\tau(t) - u_\delta^{\tau,h}(t)\|_{-1} - \|u_\delta^\tau(t) - u_\delta^{\tau,h}(t)\|_{-1,h}| \le ch\|u_\delta^\tau(t) - u_\delta^{\tau,h}(t)\|_1. \quad (4.20)$$

Next use a Sobolev inequality, the steepest descent feature of the Equation (2.8), (4.14), and a Poincare inequality to obtain

$$\|u_\delta^\tau(t) - u_\delta^{\tau,h}(t)\|_1 \le c\|u_\delta^\tau(0) - u_\delta^{\tau,h}(0)\|_{1,p} \le c\|u^0\|_{1,p}$$

(for the last inequality we used the stability property of the finite element interpolation operator, i.e. the fact that $\|u^{0,h}\|_{1,p} \le c\|u^0\|_{1,p}$). This gives the desired result. ☐

As usual in the finite element method, the convergence rate depends on the regularity of the exact solution. Alas, we do not know very much about the undiscretized time-step problem (3.3). In the next Corollary we assume some reasonable-sounding hypotheses about the regularity of u_δ^n. Using those hypotheses, we provide an estimate for the term $\rho_\delta^{\tau,h}$.

Corollary 4.11. *Let u and $u_\delta^{\tau,h}$ be the solutions of Equations (1.1) and (3.10), respectively, and d be the dimension of Ω. Assume $p \ge 3$, $u^0 \in H^3(\Omega) \cap W^{2,\infty}(\Omega)$, $d \le 4$, φ_δ is defined by (2.6), and $u_\delta^n \in W^{r,l}(\Omega) \cap W^{1,\infty}(\Omega)$, $1 < r \le 2$, with*

$$\|u_\delta^n\|_{1,\infty}, \|u_\delta^n\|_{r,l}, \|u^0\|_3, \|u^0\|_{2,\infty} \le \gamma_{max}$$

for some constant γ_{max} independent of δ, τ and n. Then there exists a constant c independent of τ, h, δ, such that

$$\operatorname*{ess\,sup}_{t\in[0,T]} \|u(t)-u_\delta^{\tau,h}(t)\|_{-1} \le T^{1/2}\delta^{1/4}+\frac{\tau}{\delta}b(\gamma_{max})+\left[\zeta_\delta^{\tau,h}h^{\theta_1} + \frac{h}{\delta} + h^{\theta_2}\right]^{\frac{1}{2}} b(\gamma_{max}). \quad (4.21)$$

Here

$$\zeta_\delta^{\tau,h} = \min\left[\left(\frac{h^{d/2-1}}{\delta^{1/2}} + \frac{1}{\delta}\right)c, \frac{c}{\tau^{1/2}}\right], \quad (4.22)$$

$$\theta_1 = \frac{d}{s_d} - \frac{d}{l} + r \quad \textit{and} \quad \theta_2 = \min\left[d - \frac{d}{l} + r - 1, d - \frac{2d}{l} + 2r - 2\right], \quad (4.23)$$

where $s_d = 1$ if $d = 1$, $s_d > 1$ if $d = 2$, or $s_d = 2d/(2+d)$ if $d \ge 3$; $b(\cdot)$ denotes a suitable polynomial of degree less or equal than p, whose coefficients are independent *of τ, h and δ. (The value of s_d changes with dimension due to a Sobolev inequality used in the proof. We obtain a better estimate in lower dimensions. The value of $b(\gamma_{max})$ depends on the choice of s_d, since it incorporates the constant of the Sobolev inequality. When $d = 2$ we are free to choose any $s_d > 1$. However, in this case $b(\gamma_{max}) \to \infty$ as $s_d \to 1$.)*

Proof. We first estimate the term $\rho_\delta^{\tau,h}$, defined by (4.17). The first term on the right side of (4.17) is estimated by Lemma 4.2: $h\|\partial E_\delta(u^0)\|_{-1} \le h\delta^{-1} b(\gamma_{max})$. To estimate the second term on the RHS of (4.17), we use the following property of the standard finite element interpolation operator $\mathcal{I}^h$:

$$\|v - \mathcal{I}^h v\|_{s,q} \le c h^{(d/q)-(d/l)+r-s} \|v\|_{r,l} \ \text{ for } s \le 1, \text{ if } W^{s,q}(\Omega) \hookrightarrow W^{r,l}(\Omega) \qquad (4.24)$$

(see e.g. Theorem 3.1.5 of [10]. For the term $\kappa_\delta^{\tau,h}$ we use (4.24) and Lemma 4.2 to obtain

$$\kappa_\delta^{\tau,h} = \min\left[\frac{|u^{0,h} - u^0|_{1,\infty}}{h^{2-d/2}}\left(\frac{1}{\delta^{1/2}} + |u^0|_{1,\infty}^{p-2}\right) + \|\nabla\cdot\Phi_\delta'(u^0)\|_1, \frac{c}{\tau^{1/2}}\right] \qquad (4.25)$$
$$\le \min\left[\left(\frac{h^{d/2-1}}{\delta^{1/2}} + \frac{1}{\delta}\right) b(\gamma_{max}), \frac{c}{\tau^{1/2}}\right].$$

Here the estimate $\|\nabla\cdot\Phi_\delta'(u^0)\|_1 \le c/\delta$ is obtained proceeding as in the proof of Lemma 4.2. Let $M_v = \int_\Omega v dx$; preparing to choose $v^h = \mathcal{I}^h u_\delta^n - M_{\mathcal{I}^h u_\delta^n}$ in (4.17) we observe that:

$$\|u_\delta^n - \mathcal{I}^h u_\delta^n + M_{\mathcal{I}^h u_\delta^n}\|_{-1} \le c\|u_\delta^n - \mathcal{I}^h u_\delta^n\|_{0,s_d} + c\|M_{\mathcal{I}^h u_\delta^n}\|_{0,s_d}.$$

Here we used a Sobolev and a triangle inequality to obtain the last inequality. To estimate the second term on the RHS of the last inequality first we use an inverse inequality and the fact that u_δ^n has mean value zero to obtain

$$\|M_{\mathcal{I}^h u_\delta^n}\|_{0,s_d} \le c h^{d/s_d - d}\|M_{\mathcal{I}^h u_\delta^n}\|_{0,1} = c h^{d/s_d-d}|M_{\mathcal{I}^h u_\delta^n}| \le c h^{d/s_d-d}\|\mathcal{I}^h u_\delta^n - u_\delta^n\|_{0,1}.$$

Next, we use (4.24) to estimate $\|\mathcal{I}^h u_\delta^n - u_\delta^n\|_{0,s_d}$ and $\|\mathcal{I}^h u_\delta^n - u_\delta^n\|_{0,1}$. Hence

$$\|u_\delta^n - \mathcal{I}^h u_\delta^n + M_{\mathcal{I}^h u_\delta^n}\|_{-1} \le c h^{d/s_d - d/l + r}\gamma_{max} \text{ by } (4.24)$$

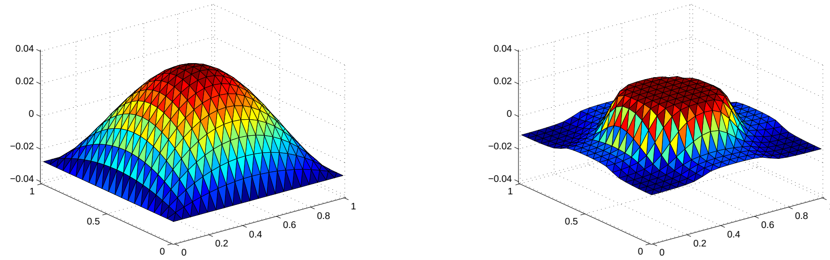
We now observe that $\|\nabla(u_\delta^n - \mathcal{I}^h u_\delta^n)\|_{0,1} \le c\|\nabla(u_\delta^n - \mathcal{I}^h u_\delta^n)\|_0$, and from the stability of the interpolation operator we have $\|\nabla(u_\delta^n - \mathcal{I}^h u_\delta^n)\|_{0,\infty} \le c\gamma_{max}$. Hence the choice $v^h = \mathcal{I}^h u_\delta^n - M_{\mathcal{I}^h u_\delta^n}$ gives

$$\int_\Omega |\nabla(u_\delta^n - v^h)| + (|\nabla u_\delta^n| + |\nabla(u_\delta^n - v^h)|)^{p-2}|\nabla(u_\delta^n - v^h)|^2 dx$$
$$\le b(\gamma_{max})(|u_\delta^n - v^h|_{1,1} + |u_\delta^n - v^h|_1^2) \le (h^{d-\frac{d}{l}+r-1} + h^{d-\frac{2d}{l}+2r-2}) b(\gamma_{max}) \text{ by } (4.24)$$

We thus conclude that $\rho_\delta^{\tau,h} \le \left[\zeta_\delta^{\tau,h} h^{\theta_1} + h^{\theta_2}\right]^{\frac{1}{2}} b(\gamma_{max})$. Finally, we obtain (4.21) by combining the last inequality with Theorem 4.10 and Lemma 4.2. ▯

The preceding estimates all controlled the error in the H^{-1} norm. That was natural, because continuum model is a steepest descent in the H^{-1} inner product. However estimates in stronger norms are also possible, by interpolation.

Remark 4.12. *We see from the last corollary that the order of convergence depends on the regularity of the solution u. We know that u develops facets with time. Near the edge of a facet we expect u_x to behave like the square root of the distance to the facet's edge; see [18]. This suggests that $u \in W^{2,l}$ for every $l < 2$.*



(a) $t = 0$ (b) $t = 3 \times 10^{-5}$

Fig. 5.1: Plot of $u_\delta^{h,\tau}(t,\cdot)$ for $t = 0$ and $t = 10^{-3}$, and $u^0(x) = x(x-1)y(y-1) - 1/36$. Here $h = 1/160$, $\tau = 10^{-6}$ and $\delta = 10^{-6}$.

5. Numerical Results. We implemented the finite element scheme discussed in Section 3 in one and two dimensions. This section reports on the results, emphasizing the observed convergence rates as the regularization parameter δ tends to 0, the spatial discretization gets finer ($h \to 0$), and the time step tends to 0. For all the one dimension simulations reported here, we solved the PDE with period cell $\Omega = (0, 2)$ and initial condition $u^0(x) = 0.1\cos(\pi x)$. For the two dimensions experiments we considered $\Omega = [0, 1] \times [0, 1]$, and $u^0(x) = x(x-1)y(y-1) - 1/36$. The exponent p was always taken to be 3 (the case of primary interest for surface relaxation).

It is well-known that the solution develops facets near the maxima and minima of the initial data, and that it reaches $u = 0$ in finite time. Figure 5.1 shows the initial data $u^0(x) = x(x-1)y(y-1) - 1/36$ and the profile at time 3×10^{-5}. All our 1D tests used a final time T on the order of 10^{-3} (long enough to show significant evolution, but short enough that the surface has not yet flattened).

5.1. Superiority of the Primal-Dual Method. In order to show the superiority of the Primal-Dual Newton Method, we implemented another version of the FEM method using regular Newton Method. For fixed values of $\tau = 10^{-6}$, error tolerance of 2×10^{-9}, and maximum number of 400 iterations for the Primal-Dual and regular Newton method, we observed the following results:

Primal-Dual Method. For $h = 1/10$: Converged for $\delta = 9 \times 10^{-6}$ and did not converge for $\delta = 8 \times 10^{-6}$. For $h = 1/20$: Converged for $\delta = 2 \times 10^{-5}$ and did not converge for $\delta = 1.5 \times 10^{-5}$.

Regular Newton Method. For $h = 1/10$: Converged for $\delta = 1.5 \times 10^{-3}$ and did not converge for $\delta = 10^{-3}$. For $h = 1/20$: Converged for $\delta = 4.5 \times 10^{-3}$ and did not converge for $\delta = 4 \times 10^{-3}$

Table 5.1 shows the maximum number of iterations to solve the nonlinear system for different values of δ for the regular (RNM) and Primal Dual Newton Method (PDNM). We used $\tau = 10^{-6}$ or $\tau = 10^{-3}$, $h = 1/20$, error tolerance of 2×10^{-9} and allowed a maximum number of 400 iterations for the both methods.

Finally, we note that for fixed values $\delta = 2.5 \times 10^{-3}$, $h = 1/20$ and error tolerance of 2×10^{-9} and maximum number of 400 iterations for the Primal-Dual and regular Newton method, we observed the following results:

Primal-Dual Method. Converged for any choice of $\tau \leq 2 \times 10^{-2}$.

Regular Newton Method. Converged for $\tau = 10^{-7}$. Did not converge for $\tau \geq 2 \times 10^{-7}$.

Table 5.1: Maximum number of iterations to solve the nonlinear system, for both regular (RNM) and Primal Dual Newton Method (PDNM). Here $h = 1/20$.

$\tau = 10^{-6}$

δ	2×10^{-2}	10^{-2}	5×10^{-3}	2.5×10^{-3}	1.25×10^{-3}	6.25×10^{-4}
PDNM	6	6	7	7	7	7
RNM	5	11	28	DNC	DNC	DNC

$\tau = 10^{-3}$

δ	10^{-4}	10^{-5}	10^{-6}	10^{-7}	10^{-8}	10^{-9}
PDNM	9	9	10	10	10	10

Similar results were also observed for different values of δ and h. All results in this subsection were obtained in dimension 2.

5.2. Regularization error. The functional analysis of steepest descent assures us that there is a well-defined solution in the limit $\delta \to 0$. The simulations bear this out; for example, Table 5.2 demonstrates that for fixed values $h = 1/160$, and $\tau = 10^{-}6$, $\max_x u_\delta^{\tau,h}$ is virtually independent of δ once this parameter is less than 10^{-6}.

Table 5.2: The value of $\max_x u_\delta^{\tau,h}$ at time $t = 2.8 \times 10^{-3}$, for various values of δ. Here $h = 1/160$ and $\tau = 10^{-6}$.

δ	10^{-3}	10^{-4}	10^{-5}	10^{-6}	10^{-7}	10^{-9}
	0.0080	0.0040	0.0029	0.0026	0.0025	0.0025

Theorem 2.1 and Remark 2.2 show that regularization error, measured in the H^{-1} norm, is bounded by $C\delta^{1/4}$. The regularization error we observed numerically was actually much smaller, more like $C\delta^{6/10}$. Indeed, Table 5.3 gives the value of $\sup_{t\in(0,T)} \|u_\delta^{\tau,h} - u_{\delta/2}^{\tau,h}\|_{-1}$ for selected choices of δ, when $h = 1/640$ and $\tau = 10^{-6}$. A bit of arithmetic using the data in the Table reveals that

Table 5.3: The value of $\sup_{t\in(0,T)} \|u_\delta^{\tau,h} - u_{\delta/2}^{\tau,h}\|_{-1}$ for various choices of δ. Here $h = 1/640$ and $\tau = 10^{-6}$.

δ	8×10^{-3}	4×10^{-3}	2×10^{-3}	10^{-3}	5×10^{-4}
	13.01e-4	8.92e-04	5.89e-04	3.69e-04	2.30e-04

$$\sup_{t\in(0,T)} \|u_{\delta/2}^{\tau,h} - u_{\delta/4}^{\tau,h}\|_{-1} \approx 0.65 \sup_{t\in(0,T)} \|u_\delta^{\tau,h} - u_{\delta/2}^{\tau,h}\|_{-1}.$$

If $u_\delta^{\tau,h} = u_0^{\tau,h} + C\delta^\alpha + o(\delta^\alpha)$, then $2^{-\alpha} \approx .65$, whence $\alpha \approx .62$. Thus, the observed H^{-1} error due to regularization is about $\delta^{6/10}$.

We did the same test as in Table 5.3 but using the L^2 norm instead of the H^{-1} norm. We found that

$$\sup_{t\in(0,T)} \|u_{\delta/2}^{\tau,h} - u_{\delta/4}^{\tau,h}\|_0 \approx 0.73 \sup_{t\in(0,T)} \|u_\delta^{\tau,h} - u_{\delta/2}^{\tau,h}\|_0.$$

Since $2^{-\alpha} = .73 \Rightarrow \alpha \approx .47$, the observed L^2 error due to regularization is about $\delta^{1/2}$.

Our analysis of the convergence as $\tau \to 0$ gave a bound of the form $C\tau/\delta$ (see Theorem 4.1, Lemma 4.2, and Remark 4.3). The bound is proportional to δ^{-1} due to the presence of $\|\partial E_\delta(u^0)\|_{-1}$ on the right side of (4.2). Our convergence estimates as $h \to 0$ also have terms proportional to δ^{-1}, whose origin is essentially the same (for example, in Proposition 4.9 the error estimate $B_\delta^{N,h}$ includes a term $h^2\|\partial E_\delta(u^0)\|_{-1}$). Therefore it is interesting to assess the sharpness of Lemma 4.2, which showed that $\|\partial E_\delta(u^0)\|_{-1} \leq C/\delta$. In fact the estimate appears not to be sharp when $u^0(x) = .1\cos(\pi x)$. For this specific choice of u^0, our numerics shows that $\|\partial E_\delta(u^0)\|_{-1} \sim \delta^{-3/4}$.

5.3. Time discretization error. Fixing $\delta = 10^{-3}$ and $h = 1/320$. We observed that

$$\frac{\int_0^T \|u_\delta^{\tau/2,h} - u_\delta^{\tau/4,h}\|_{-1}\, ds}{\int_0^T \|u_\delta^{\tau,h} - u_\delta^{\tau/2,h}\|_{-1}\, ds} = .60 \text{ or } .53$$

depending on the choice of $\tau = 2.5 \times 10^{-4}$ or $\tau = 1.25 \times 10^{-4}$. The anticipated linear behavior in τ corresponds to a ratio of 1/2.

5.4. Finite element discretization error. The numerically-observed convergence rate as $h \to 0$ was $O(h^2)$, far better than the $h^{1/2}$ behavior suggested by Corollary 4.11 and Remark 4.12. Indeed, we computed $\int_0^T \|u_\delta^{\tau,h} - u_\delta^{\tau,h/2}\|_{-1}\, ds$ for different values of $h \in \{1/10, 1/20, 1/40, 1/80\}$, and a few different choices of $\delta \in \{10^{-4}, 10^{-5}, 10^{-6}\}$ and $\tau \in \{10^{-6}, 10^{-8}, 10^{-9}\}$. Our results suggest that

$$\int_0^T \|u_\delta^{\tau,h/2} - u_\delta^{\tau,h/4}\|_{-1}\, ds \approx 0.25 \int_0^T \|u_\delta^{\tau,h} - u_\delta^{\tau,h/2}\|_{-1}\, ds.$$

Since $2^{-\alpha} = 1/4 \Rightarrow \alpha = 2$, the observed discretization error in the $L^1(0,T;H^{-1}_{per})$ norm is about h^2. This same convergence rate was also observed for the error in the norms $L^\infty(0,T;H^{-1}_{per})$ and $L^1(0,T;L^2)$.

6. Conclusions. We have discussed the numerical solution of a widely-used PDE model for surface relaxation below the roughening temperature. We use implicit time-stepping and a mixed finite-element spatial discretization. The singular surface energy is regularized, and the time-step problem is solved using a primal-dual scheme. Our convergence analysis is the first rigorous analysis of any numerical scheme for solving (1.1). Our estimates may not be optimal. Indeed, the numerically-observed convergence as $\delta \to 0$ and $h \to 0$ for the 1D problem with $u_0(x) = .1\cos(\pi x)$ is considerably better than our estimates would suggest.

REFERENCES

[1] K. Andersen, E. Christiansen, A. Conn, M. Overton; *An efficient primal-dual interior-point method for minimizing a sum of Euclidean norms*, SIAM J. Sci. Comput. 22 (2000) no. 1, 243–262

[2] V. Barbu; *Nonlinear semigroups and differential equations in Banach spaces*, Noordhoff International Publishing, Leiden (1976)

[3] J.W. Barrett, W. B. Liu; *Finite element approximation of the parabolic p-Laplacian*, SIAM J. Numer. Anal. 31 (1994) no. 2, 413–428

[4] J.W. Barrett, J.F. Blowey, H. Garcke; *Finite element approximation of a fourth order nonlinear degenerate parabolic equation*, Numer. Math. 80 (1998) no. 4, 525–556

[5] S. Brenner, R. Scott; *The Mathematical Theory of Finite Element Methods*, Springer-Verlag (1994)

[6] H. Brezis; *Opérateurs maximaux monotones et semi-groupes de contractions dans les espaces de Hilbert*, North-Holland Mathematics Studies, No. 5 (1973)

[7] T. Chan, G. Golub, P Mulet; *A nonlinear primal-dual method for total variation-based image restoration*, SIAM J. Sci. Comput. 20 (1999) no. 6, 1964–1977

[8] W.L. Chan, A. Ramasubramaniam, V.B. Shenoy, E. Chason; *Relaxation kinetics of nano-ripples on Cu(001) surfaces*, Phys. Rev. B 70 (2004) 245403

[9] P. Clement; *Approximation by finite element functions using local regularization*, RAIRO Anal. Numér. 9 (1975), 33–75

[10] P. Ciarlet; *The Finite Element Method for Elliptic Problems*, SIAM (2002)

[11] P. Ciarlet, P. Raviart; *A mixed finite element method for the bi-harmonic equation*, in Mathematical Aspects of Finite Elements in Partial Differential Equations, C. De Boor, ed., Academic Press, New York (1974) 125–143

[12] C. Elliot, A. Mikelic; *Existence for the Cahn-Hilliard model of phase separation with a non-differentiable energy*. Annali Matematica Pura ed Applicata CLVIII (1991) 181–203.

[13] C. Elliott and S. Smitheson *Analysis of the TV regularization and H^{-1} fidelity model for decomposing an image into cartoon plus texture* Comm. Pure Appl. Analysis 6 (2007), no 4, 917–936

[14] C. Elliott and S. Smitheson *Numerical analysis of the TV regularization and H^{-1} fidelity model for decomposing an image into cartoon plus texture* IMA J. Num. Anal (2009), 29, 651–689

[15] L.C. Evans; *Partial Differential Equations*, American Mathematical Society (1998)

[16] R. Falk, J. Osborn; *Error estimates for mixed methods*, RAIRO Anal. Numér. 14 (1980) no. 3, 249–277

[17] X. Feng, M. von Oehsen, A. Prohl; *Rate of convergence of regularization procedures and finite element approximations for the total variation flow*, Numer. Math. 100 (2005) no. 3, 441–456

[18] J. Hager, H. Spohn; *Self-similar morphology and dynamics of periodic surface profiles below the roughening transition*, Surf. Sci. 324 (1995) 365–372

[19] Y. Kashima; *A subdifferential formulation of fourth order singular diffusion equations*, Adv. Math. Sci. Appl. 14 (2004) no. 1, 49-74

[20] Y. Kashima; *A variational approach to very singular gradient flow equations*, in Proc. Czech-Japanese Seminar in Applied Mathematics 2004, Czech Technical University in Prague, pages 79-84

[21] D. Margetis, M. J. Aziz, H. A. Stone ; *Continuum approach to self-similarity and scaling in morphological relaxation of a crystal with a facet*, Phys. Rev. B 71 (2005) 1–22.

[22] I. Odisharia; *Simulation and Analysis of the Relaxation of a Crystalline Surface*, PhD thesis, Dep. of Phys. NYU (2006)

[23] M.V. Ramana Murty; *Morphological stability of nanostructures*, Phys. Rev. B 62 (2000) 17004-17011

[24] J. Rulla; *Error analysis for implicit approximations to solutions to Cauchy problems*, SIAM J. Numer. Anal. 33 (1996) no. 1, 68-87

[25] V. B. Shenoy, A. Ramasubramaniam, L. B. Freund; *A variational approach to nonlinear dynamics of nanoscale surface modulations*, Surf. Sci. 529 (2003) 365–383

[26] H. Spohn; *Surface dynamics below the roughening transition*, J. Phys. I. France 3 (1993) 69–81