# An Analysis of the Perona-Malik Scheme

SELIM ESEDOḠLU
*Courant Institute*

**Abstract**

We investigate how the Perona-Malik scheme evolves piecewise smooth initial data in one dimension. By scaling a natural parameter that appears in the scheme in an appropriate way with respect to the grid size, we obtain a meaningful continuum limit. The resulting evolution can be seen as the gradient flow for an energy, just as the discrete evolutions are gradient flows for discrete energies. It involves, except at special isolated times, solving a system of heat equations coupled to each other through nonlinear boundary conditions. At the special times, the solutions experience gradient blowup; nevertheless, there is a natural continuation for the solutions beyond these singular times. © 2001 John Wiley & Sons, Inc.

## 1 Introduction

In [16] Perona and Malik proposed a numerical method for selectively smoothing digital images, designed to keep "edges" in pictures sharp. The essence of their method is contained in the following discretization:

$$\frac{u_{i,j}^{t+\Delta t} - u_{i,j}^{t}}{\Delta t} = \frac{1}{\Delta x}\left(c_{E_{i,j}}^{t} \nabla_E u_{i,j}^{t} - c_{W_{i,j}}^{t} \nabla_W u_{i,j}^{t}\right) + \frac{1}{\Delta y}\left(c_{N_{i,j}}^{t} \nabla_N u_{i,j}^{t} - c_{S_{i,j}}^{t} \nabla_S u_{i,j}^{t}\right) \tag{1.1}$$

where $N$, $S$, $E$, and $W$ denote north, south, east, and west, the symbol $\nabla$ denotes the nearest-neighbor difference quotient in the direction of its subscript, and the remaining coefficients are given by

$$c_{N_{i,j}}^{t} = g_k\left(|\nabla_N u_{i,j}^{t}|^2\right), \qquad c_{S_{i,j}}^{t} = g_k\left(|\nabla_S u_{i,j}^{t}|^2\right),$$
$$c_{E_{i,j}}^{t} = g_k\left(|\nabla_E u_{i,j}^{t}|^2\right), \qquad c_{W_{i,j}}^{t} = g_k\left(|\nabla_W u_{i,j}^{t}|^2\right),$$

where $g_k$ is a function with certain important properties, as we shall presently explain. In applications, the computational domain is ordinarily just a rectangle, and one imposes either periodic or homogeneous Neumann boundary conditions.

In this paper we focus on the one-dimensional version of scheme (1.1). Our purpose is to recognize a continuum (PDE) problem that it solves in the limit as the grid size $\Delta x$ goes to 0. As indicated, the function $g_k$ comes equipped with a

Communications on Pure and Applied Mathematics, Vol. LIV, 0001–0046 (2001)

parameter $k$; we obtain our continuum limit by choosing a specific relation between $k$ and $\Delta x$. The resulting evolution is unusual: It involves solving a system of heat equations coupled to each other through nonlinear boundary conditions that become singular at special times, leading to gradient blowup for the solutions. However, the scheme suggests a natural continuation beyond each one of these singular times that involves a change in the PDE system. Our convergence proof applies on any bounded interval of time, which might include singular times (under some technical restrictions). And our continuum limit has some of the features observed in applications of the numerical scheme.

It is natural to think of discretization (1.1) as a candidate for the numerical solution of the continuum problem:

$$u_t = \operatorname{div}(g_k(|\nabla u|^2)\nabla u)\,. \tag{1.2}$$

To be more precise, and as Perona and Malik note in their paper, the discretization (1.1) is suggestive of the similar but more anisotropic equation

$$u_t = \left(g_k\left(u_x^2\right)u_x\right)_x + \left(g_k\left(u_y^2\right)u_y\right)_y\,. \tag{1.3}$$

In fact, the authors propose their numerical scheme with this intention.

An essential feature of the method is the choice of the function $g_k(\xi)$. For Perona and Malik's choices, equation (1.2) (or (1.3)) is not parabolic: In regions of high enough gradient (depending on the parameter $k$), the diffusion coefficient becomes negative. Our approach avoids trying to make sense of equations (1.2) or (1.3). It instead concentrates on the scheme (1.1) itself.

## 1.1 Background

Image segmentation and edge detection are two fundamental procedures of computer vision that rely on image smoothing as an important first step. Their goal is to decompose a given image into regions that are essentially homogeneous (with little variation in color or brightness). These regions should be separated by sharp boundaries (edges). Such an operation forms an early stage of interpreting and extracting useful information from digital pictures, since it helps recognize parts of the scene that belong to different objects [14].

An image is described mathematically by a real-valued, bounded function defined on a subset of the plane; the value of the function at a point represents the gray-scale intensity, or brightness, at that point in the image. We think of edges in the image as places where the intensity function has high gradient or discontinuity due to an abrupt change. Abrupt changes in an image occur, however, not only because of a transition from one distinct region in the scene to another, but also because of the presence of noise or fine detail within regions. Those can appear as redundant edges. The natural approach of thresholding the gradient, therefore, is not a satisfactory method of locating edges. As a cure, a preprocessing step is often introduced. It involves smoothing the image—for instance, by some averaging technique—in order to remove noise and fine detail.

A common way of "de-noising" is to convolve the original image with the Gauss kernel, or equivalently, to solve the heat equation with the original image as initial data [12]. In that case, the variance of the kernel (or the time variable $t$ of the heat equation) plays the role of a coarseness parameter. This method has an obvious disadvantage: Edges in the image, which are the ultimate goal, get blurred. Ideally, as $t$ gets large, we would like edges to remain sharp, and hence well-defined and localized, until they disappear. One therefore wishes for a more selective smoothing procedure: one that smoothes the interior of individual regions but not their boundaries.

Various methods have been suggested to avoid the disadvantages of Gaussian smoothing; a recurring theme is to replace the heat equation by a nonlinear diffusion equation. One such approach is directional diffusion, a typical example of which is the equation $u_t = |\nabla u| \operatorname{div}(\nabla u / |\nabla u|)$ that models "motion by curvature" and also appears in other contexts [11]; it is degenerate along the gradient direction, and so has the effect of smoothing the image along but not across the edges. Perona and Malik proposed another procedure in [16]. Their idea is to coarsen the image using a nonlinear heat equation whose constitutive function decreases rapidly for large values of the gradient and thus suppresses diffusion near edges. There are also methods based on modifications of Perona and Malik's idea [2] and methods that combine their idea with the degeneracy in the motion by curvature equation [1].

## 1.2 The Perona-Malik Method

In [16] Perona and Malik report numerical experiments with their scheme using

$$g_k(\xi) = \frac{1}{1 + \frac{\xi}{k}} \quad \text{and} \quad g_k(\xi) = \exp\left(-\frac{\xi}{2k}\right).$$

Other choices used in practice include

$$g_k(\xi) = \left(1 + \frac{\xi}{k}\right)^{(\beta-1)} \quad \text{where } \beta \in \left(0, \tfrac{1}{2}\right).$$

These choices have the following common characteristics, as noted in [10]:

(1) $g_k(\xi) > 0$ for all $\xi \geq 0$.
(2) The parameter $k$ defines a positive critical value $z(k)$ such that $\partial_\xi(\xi g_k(\xi^2)) > 0$ for $|\xi| < z(k)$ and $\partial_\xi(\xi g_k(\xi^2)) < 0$ for $|\xi| > z(k)$.
(3) Both $g_k(\xi)$ and $\partial_\xi(\xi g_k(\xi^2))$ tend to 0 as $\xi$ goes to infinity.

Figure 1.1 illustrates $\xi g_k(\xi^2)$ for such a choice of $g_k(\xi)$.

In light of these properties, the parameter $k$ constitutes a threshold for the intensity gradient: In regions where the gradient is small compared to $k$, the equation is parabolic. On the other hand, if the gradient is large compared to $k$, not only does the diffusion coefficient vanish, but it actually becomes negative. This is an alarming situation since backwards heat equations are notoriously ill-posed. Nevertheless, experiments with the scheme yield visually impressive segmentations [15, 16].

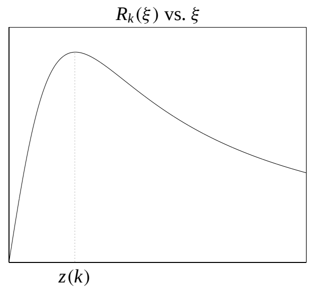


FIGURE 1.1. Graph of $R_k(\xi) := \xi g_k(\xi^2)$ for a typical choice of the function $g_k(\xi)$ in the Perona-Malik scheme. In this case, $g_k(\xi) = 1/(1 + \xi/k)$.

Many previous authors have reported numerical experiments with the Perona-Malik scheme. Some of the important features observed are as follows: Large-scale oscillations that one expects to see (as an indication of ill-posedness) are prominently absent. Instead, unstable behavior seems to be confined to regions that are thick with high gradient. Such regions are uncommon in real pictures, but do arise in very blurred ones. In one dimension, the intensity function in these regions goes through a transition period during which it develops terraces separated by sudden jumps; this is the effect referred to as “staircasing” in [8, 10, 15]. We understand by a terrace any maximal subinterval of the domain in which the discrete derivatives (difference quotients) are small enough, compared to the parameter $k$, so that the scheme is parabolic. The sharp transition from one terrace to the next occurs over a single grid cell, and the gradient across this transition exceeds the parabolicity threshold. The configuration of steps that emerge from regions of high gradient is very sensitive to perturbations and has global influence on the evolution of the image [19]. Our own experiments agree with these observations.

The formation and subsequent interaction of steps (or terraces) mentioned in the previous paragraph is a major characteristic of the scheme and seems to be related to coarsening. Indeed, in one dimension we observed that the transitions between terraces do not move. Furthermore, the scheme does not introduce new transitions: Neighboring terraces can merge, but a terrace never breaks into smaller ones. Remarkably, this property holds even at the level of a few grid cells (one can think of a single grid cell as a very small terrace). As a result, terraces quickly merge to form larger ones, and the image evolves into one that looks piecewise smooth (see Figure 1.2). From the point of view of image segmentation, these properties are very desirable.

The success of the Perona-Malik scheme at its intended purpose and its better-than-expected stability have led to some recent work on obtaining a continuum theory that might explain the major characteristics observed in numerical experiments.

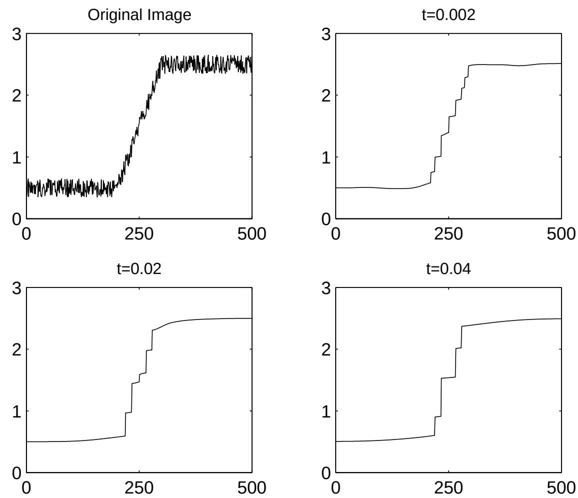


FIGURE 1.2. Evolution of a ramp with noise superposed under the Perona-Malik scheme with $g_k(\xi) = 1/(1 + \xi/k)$. The grid size is $h = 1/500$ and the threshold value of slope is $z(h) = \sqrt{80}$. The ramp has slope 10. The noisy original image turns very quickly (by $t = 0.002$) into one that looks piecewise smooth.

## 1.3 Previous Work

Some previous mathematical work deals with understanding whether equation (1.2) can be given an existence and uniqueness theory. In [10], Kichenassamy provides an argument for why this equation does not possess a weak solution in the usual sense (if the initial data is not analytic in a neighborhood of high-gradient regions). His argument draws on the regularization property of parabolic equations (in divergence form, with possibly discontinuous coefficients). Specializing to the one-dimensional version of (1.2), he then introduces a new notion of weak solution that allows for discontinuities. Naturally, this leads to a jump condition that relates the speed of a "shock" to the jump in the value of the function and its space derivative across the discontinuity. He also proposes an entropy condition with the intention of obtaining uniqueness. These considerations lead him to a continuum problem for piecewise smooth initial data with small derivatives. It consists of a system of parabolic PDEs (one equation for each smooth piece) coupled through their boundary conditions. Our goal—a well-posed PDE capturing the essential behavior of the Perona-Malik scheme—is similar to Kichenassamy's. However, our treatment is different in two important ways: (1) We specify a relation between the parameter $k$ and the grid size $\Delta x$, and (2) we prove a rigorous convergence theorem linking the discrete and continuous schemes.

The paper by Kawohl and Kutev [8] is also about the one-dimensional Perona-Malik PDE rather than the numerical scheme. It concerns weak $C^1$ solutions. (The set of such solutions is not empty, since if we start with initial data with small slope, the equation remains parabolic for all time; such solutions therefore exist and are well-behaved). Among the results presented is nonexistence of global-in-time weak $C^1$ solutions whose initial data have regions with slope larger than the parabolicity threshold. So even for analytic initial data such solutions break down in finite time (if they exist at all; local-in-time existence is not known). Other results the authors obtain include maximum and comparison principles and a uniqueness result for certain $C^1$ solutions. Our continuum limit is rather different from that considered by Kawohl and Kutev. Still, we do make use of ideas from their paper in deriving a suitable maximum principle (Proposition 2.4).

In [3], Chambolle proves Γ-convergence of a class of discrete approximations to the Mumford-Shah functional in two dimensions. Let us recall the form of this functional:

$$\mathrm{MS}(u) := \int_{\Omega - S_u} |\nabla u|^2 \, dx + \alpha H^1(S_u) + \lambda \int_{\Omega} |u - u^0|^2 \, dx \,. \tag{1.4}$$

It is defined for functions $u$ in GSBV, the space of generalized special functions of bounded variation. $H^1$ denotes the one-dimensional Hausdorff measure, $S_u$ is the jump set of $u$ and $\nabla u$ its approximate gradient, and $u^0$ is the original image. Chambolle's approximations to MS, with $H^1$ replaced by an anisotropic version (cab driver length), are defined on uniform rectangular grids; they look like

$$\begin{aligned} E_h(u) := \sum_{i,j} h^2 \left\{ W_k\left( \frac{u_{i+1,j} - u_{i,j}}{h} \right) + W_k\left( \frac{u_{i,j+1} - u_{i,j}}{h} \right) \right\} \\ + \lambda \sum_{i,j} h^2 |u_{i,j} - u^0_{i,j}|^2 \end{aligned}$$

where the function $W_k(x) := \min\{x^2, k\}$ and $h > 0$ is the grid size. The function $W_k$ is convex for $|x| \leq \sqrt{k}$. In that sense, the parameter $k$ plays the same kind of thresholding role as it does in the Perona-Malik scheme. Chambolle shows that if $k$ is scaled as $k = \alpha / h$ with respect to the grid size, this family of discrete functionals Γ-converge to MS. The Perona-Malik method is dynamic, while the Mumford-Shah variational problem is static. There is, however, a very strong link between our work and that of Chambolle: We follow his lead in assuming that the parameter $k$ must scale with $h$.

Paper [6] by Gobbino concerns the same kind of problem with a similar approach as in Chambolle's work. It establishes Γ-convergence to MS (with $\lambda = 0$, an inessential difference) of a class of approximations that in one dimension have the form

$$F_\varepsilon(u) := \frac{1}{\varepsilon} \int_{\mathbb{R}} \arctan\left( \frac{(u(x+\varepsilon) - u(x))^2}{\varepsilon} \right) dx \,. \tag{1.5}$$

Gobbino's result in fact holds for the $n$-dimensional analogue of the problem.

The more recent paper by Gobbino [7] is dynamic rather than static and very closely related to the present work. This paper looks at the one-dimensional approximations (1.5) as defined on spaces of piecewise constant functions

$$\mathrm{PC}^2_\varepsilon := \left\{ u \in L^2(\mathbb{R}) : u(x) = u\left(\varepsilon\left[\frac{x}{\varepsilon}\right]\right), \forall x \in \mathbb{R} \right\}$$

where $\varepsilon > 0$ is the grid size and $[\cdot]$ denotes the integer part of its argument. The paper is devoted to defining a gradient flow for MS as the limit of gradient flows to (1.5), which are defined by the relation

$$u'_\varepsilon(x) = -(\nabla F_\varepsilon)(u_\varepsilon(t)) \quad \text{with } u_\varepsilon(0) = u^0_\varepsilon\,. \tag{1.6}$$

The initial condition $u^0$ is required to be $L^\infty_{\mathrm{loc}}$ and have finite Mumford-Shah energy. In one dimension, this stipulation implies that $u^0$ is piecewise $W^{1,2}$. Also, the approximate initial data $u^0_\varepsilon$ (which are piecewise constant) must converge to $u^0$ in $L^2$ and in energy. Gobbino shows that the flows generated by (1.6) converge, and for a large class of initial data the limit is independent of the approximating sequence.

Gobbino's paper is related to our work because the gradient flows (1.5) are given precisely by the semidiscrete (continuous-in-time) version of the Perona-Malik scheme (1.1) with $g_k(\xi) = 1/(1 + \xi^2/k^2)$ and subject to the scaling $k = 1/\Delta x$. The limiting evolution he obtains is similar to ours, consisting of solving a system of linear heat equations in a variable domain. However, it differs from ours in the boundary conditions that couple these equations to each other: His equations have a homogeneous Neumann boundary condition at each "interface," whereas our limit involves nonlinear boundary conditions that strongly couple equations on neighboring intervals to each other. Our analysis is therefore similar to that of Gobbino at many points but also requires new ideas.

## 1.4 Our Approach

We now turn to the central task of this paper: understanding the Perona-Malik scheme (1.1) as the grid size $h$ goes to 0 with $k = k(h)$ scaled appropriately. We shall address the semidiscrete version of the scheme (discrete in space, continuous in time), and we restrict our attention to one space dimension. Thus, the scheme to be analyzed is

$$\frac{d}{dt}u_j(t) = \frac{1}{h}\big(R_k(\nabla_E u_j(t)) - R_k(\nabla_W u_j(t))\big) \tag{1.7}$$

where $R_k(\xi) := \xi g_k(\xi^2)$. We will work with the specific family of nonlinearities

$$g_{\beta,k}(\xi) = \left(1 + \frac{\xi}{k}\right)^{(\beta-1)} \quad \text{with } \beta \in \left[0, \tfrac{1}{2}\right).$$

Let $z(h) = z_\beta(h)$ denote the threshold value of the slope for this family; more explicitly,

$$z(h) = \frac{1}{\sqrt{1-2\beta}}\sqrt{k(h)}$$

so that

$$\partial_\xi R_{\beta,h}(\xi) \begin{cases} > 0 & \text{for } |\xi| < z(h) \\ \leq 0 & \text{otherwise.} \end{cases}$$

For later reference, we record two important properties of the function $R$. First, $R(\xi) := R_k(\xi)$ is a one-to-one, increasing function on $[-z(h), z(h)]$; it therefore has an increasing inverse with domain $[-R(z(h)), R(z(h))]$. We shall denote this function $R_*^{-1}$, i.e.,

$$R_*^{-1}(\xi) := \left(R(\xi)\Big|_{[-z(h),z(h)]}\right)^{-1}. \tag{1.8}$$

Second, since $R_h(x)/x$ is a function of only $x/z$, we have the following bound from below, which is independent of $h$:

$$\theta(\beta) := \inf_{\substack{|x|<z(h)/2 \\ |y|<z(h)}} \left| \frac{R_h(x) - R_h(y)}{x-y} \right| > 0\,. \tag{1.9}$$

Let us now try to understand how the one-dimensional scheme (1.7) operates on an initial image that is smooth except at a point $p$, at which it has a jump of height $J$. Let $p$ be located between the two grid points $x_j$ and $x_{j+1}$. The scheme then reads

$$\dot{u}_j = \frac{1}{h}\left(R_k\left(\frac{J}{h}\right) - R_k\left(\frac{u_j - u_{j-1}}{h}\right)\right),$$
$$\dot{u}_{j+1} = \frac{1}{h}\left(R_k\left(\frac{u_{j+2} - u_{j+1}}{h}\right) - R_k\left(\frac{J}{h}\right)\right).$$

Roughly speaking, we interpret this to mean that the scheme imposes the condition

$$R_k\left(\frac{u_j - u_{j-1}}{h}\right) = R_k\left(\frac{J}{h}\right) = R_k\left(\frac{u_{j+2} - u_{j+1}}{h}\right). \tag{1.10}$$

In words, the slopes on either side of a jump are equal and are related to the jump height by the above formula. One way to understand why this is so is to note that unless these three quantities are within $O(h)$ of each other, a process that operates at a faster time scale will adjust them until this is the case. Kichenassamy also observed this property as a "note added in proof" of his paper [10].

We therefore expect the scheme to impose (possibly inhomogeneous) Neumann boundary conditions at jump locations of a piecewise smooth image.

Second, we note that the difference quotients $(u_{j+1} - u_j)/h$ scale as $O(1)$ at differentiable regions in the image and as $O(1/h)$ across jumps. We are thus led to look for a way to adjust the thresholding parameter $k$ that appears in the scheme with respect to grid size $h$ so that relation (1.10) translates into a nontrivial boundary condition in the limit as $h \to 0^+$.

When we try a scaling of the form $k(h) = h^\alpha$ and look for $\alpha$, we see that for each $\beta \in [0, \frac{1}{2})$ there is only one value for $\alpha$ that leads to a nontrivial limit: We find that

$$\alpha = \frac{2\beta - 1}{1 - \beta} \quad \text{leads to} \quad \lim_{h \to 0^+} R_{\beta,k}\left(\frac{J}{h}\right) = J|J|^{2\beta-2} . \tag{1.11}$$

Such scalings that depend on the discretization appear in a different context in the works of A. Chambolle [3] and Chambolle and Dal Maso [4] on the $\Gamma$-convergence of discrete approximations to the Mumford-Shah functional.

The threshold value $z(h)$ of slope for a given grid size $h$ thus becomes

$$z(h) = \frac{1}{\sqrt{1 - 2\beta}} h^{(2\beta-1)/(2-2\beta)}$$

and has the important property that $z(h) \to \infty$ and $hz(h) \to 0$ as $h \to 0^+$. As a consequence, for small enough $h$ the scheme becomes diffusive at all regions in which the image is differentiable, no matter how high the slope there is. The only features in a piecewise differentiable image that “feel” the backwards nature of the scheme are jumps, at which the backwardness manifests itself as boundary conditions.

In fact, scaling in the manner indicated by (1.11) leads to

$$\lim_{h \to 0^+} R_{\beta,k}(x) = x ,$$

which means in the limit we should expect the scheme to solve the standard heat equation wherever the image is differentiable. We have thus obtained enough clues as to what kind of continuum limit, defined for piecewise smooth images, we should put forth.

## 1.5 Proposed Limit

$$\begin{aligned}
&\frac{\partial u_i}{\partial t} = \Delta u_i && \text{for } p_{i-1} < x < p_i \, , \\
&\frac{\partial u_i}{\partial x}(p_i, t) = \frac{\partial u_{i+1}}{\partial x}(p_i, t) = J_i |J_i|^{2\beta-2} && \text{for } i = 1, 2, \ldots, N-1 \, , \\
&\frac{\partial u_1}{\partial x}(p_0, t) = \frac{\partial u_N}{\partial x}(p_N, t) = 0 \, ,
\end{aligned} \tag{1.12}$$

where $\beta \in [0, \frac{1}{2})$, $p_0 < p_1 < \cdots < p_N$, and $J_i = u_{i+1}(p_i, t) - u_i(p_i, t)$. For the function $\{u_i(x, t)\}_{i=1}^N$ we prescribe piecewise continuous initial conditions with jumps at $\{p_i\}_{i=1}^{N-1}$. This is our proposed limit for the Perona-Malik scheme provided that we scale $R_{\beta,k}(\xi) = \xi(1 + \xi^2/k)^{\beta-1}$ according to the prescription

$$k = h^{(2\beta-1)/(1-\beta)} . \tag{1.13}$$

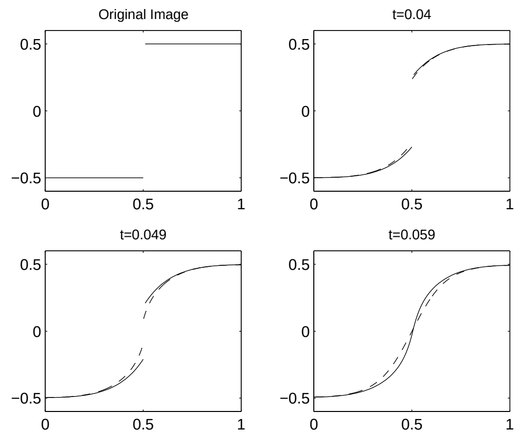


FIGURE 1.3. Evolution up to and beyond the quenching time for a symmetric step. The solid line is the solution generated by the Perona-Malik scheme with $g_k(\xi) = 1/(1 + \xi/k)$, $k = 100$, and mesh size $h = 1/100$. It quenches just before $t = 0.059$. The dashed line is the solution to the proposed continuum limit, computed on a very fine mesh. It quenched a little after $t = 0.049$. There is a discrepancy between the quenching times, but they coincide in the limit as $h \to 0^+$.

We single out the case $\beta = 0$ that gives the most singular boundary condition

$$\frac{\partial u_i}{\partial x}(p_i, t) = \frac{\partial u_{i+1}}{\partial x}(p_i, t) = \frac{1}{J_i} \quad \text{for } i = 1, 2, \ldots, N-1 \tag{1.14}$$

because it requires special treatment in some of our claims.

The system (1.12) is meaningful until one (or more) of the jump heights $J_i$ vanish, since according to the boundary conditions the slope at a jump location goes to infinity as the jump height goes to 0. We will refer to such a breakdown as *quenching*. It is easy to see that quenching has to happen in finite time (see Proposition 2.8). Therefore, this PDE system is only part of our proposed limit; we will explain how to continue the solution beyond quenching times.

## 1.6 Numerical Experiments and Experience

As we mentioned earlier, at least under some circumstances the Perona-Malik scheme turns a general image quickly into one that looks piecewise smooth: The number of jumps in the picture becomes small compared to the number of pixels, and the terraces become wide. This numerical observation is supported by Proposition 3.2 and Lemma 3.3, and is illustrated by Figure 1.2. Since our approach to understanding the Perona-Malik scheme, as expressed in Theorem 3.13, is limited to piecewise smooth initial data, we cannot expect it to describe what happens during this initial transition.

Figure 1.3 illustrates how the Perona-Malik scheme behaves after an interface

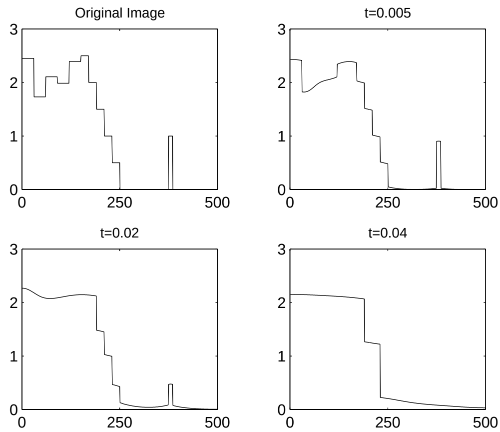


FIGURE 1.4. Interaction of steps under the Perona-Malik scheme. The small terrace at the right (at about $x = 380$) washes out quickly despite the large jumps on either side of it.

heals. The solid line is generated by the original scheme. The dashed line represents our proposed limit, which in this symmetric situation can be expressed in terms of the solution to the *single* nonlinear boundary value problem

$$u_t(x,t) = u_{xx}(x,t) \quad \text{on } x \in (0, 0.5)$$
$$\text{with } u_x(0,t) = 0 \text{ and } u_x(0.5,t) = \frac{-1}{2u(0.5,t)}\,.$$

In order to compute an accurate solution to the PDE above, we followed a suggestion in [9] and discretized in a standard way the equation satisfied by $u^2$ (which involves a constant Neumann boundary condition) instead, and used a very fine grid. As explained in Section 2.4, after the quenching time of about $t = 0.049$ for the proposed continuum limit, the continuation of the solution beyond the blowup in this case calls for the solution of the standard heat equation on the entire interval $(0, 1)$. This was accomplished by a straightforward finite difference discretization of the heat equation, again on the very fine grid.

Figure 1.4 shows how the Perona-Malik scheme evolves piecewise smooth data. Neighboring terraces interact and merge to form fewer and bigger terraces separated by larger jumps.

## 2 Analysis of the Limit Problem

This section is devoted to the study of the PDE system (1.12). Section 2.1 is standard: It recalls some simple facts about the heat equation to establish well-posedness for the system while jump heights remain bounded away from 0. In Section 2.2 we obtain some fundamental estimates for (1.12). Among them, the result regarding Hölder continuity in time allows us to extend the solutions up to the singular (quenching) times. That paves the way to Section 2.4, where we explain how to modify the PDE system and continue the solution after an interface heals. In Section 2.3 we look more carefully into what happens in a neighborhood of quenching times. Our results show that the solutions do not spend too much time with large gradients: Under suitable conditions, quenching time for a jump can be estimated from above in terms of the jump height.

The results we obtain here have discrete analogues for the Perona-Malik scheme and will be derived also in that context in Section 3. Together, they will eventually be used in our convergence argument.

## 2.1 Existence, Uniqueness, and Regularity

We first consider the following linear Neumann problem on an interval:

Given continuous and bounded $f(t)$ and $g(t)$ and continuous

$$\phi(x) : [p_{i-1}, p_i] \to \mathbb{R},$$

find $u \in C^{2,1}([p_{i-1}, p_i] \times (0,\infty)) \cap C([p_{i-1}, p_i] \times [0,\infty))$ such that

$$\begin{aligned} u_t(x,t) &= u_{xx}(x,t) \ \text{ on } x \in (p_{i-1}, p_i) && \text{for } t > 0\,, \\ u_x(p_{i-1},t) &= f(t) \ \text{ and } \ u_x(p_i,t) = g(t) && \text{for } t > 0\,, \\ u(x,0) &= \phi(x)\,. \end{aligned} \tag{2.1}$$

The solution to this problem can be represented via the method of images, as follows: Let

$$P(x,y,t) = \sum_{n=-\infty}^{\infty} \frac{1}{\sqrt{2\pi t}} \left\{ e^{-(x-y-2n)^2/(2t)} + e^{-(x+y-2n)^2/(2t)} \right\},$$

$$P_i(x,y,t) = \frac{1}{p_i - p_{i-1}} P\left( \frac{x - p_{i-1}}{p_i - p_{i-1}}, \frac{y - p_{i-1}}{p_i - p_{i-1}}, \frac{t}{(p_i - p_{i-1})^2} \right).$$

Then the solution to problem (2.1) is given by

$$\begin{aligned} u(x,t) = &\int_0^t P_i(x, p_{i-1}, t-s) f(s)ds + \int_0^t P_i(x, p_i, t-s) g(s)ds \\ &+ \int_{p_{i-1}}^{p_i} P_i(x,y,t)\phi(y)dy\,. \end{aligned} \tag{2.2}$$

**Some Basic Estimates**

From the explicit formula (2.2) we can compute various derivatives of the solution. That yields estimates such as the following:

$$\sup_{\substack{x \in [p_{i-1}, p_i] \\ T \geq t \geq \varepsilon}} \sum_{\substack{2\alpha + \beta \leq 2n \\ \alpha,\beta \in N}} |D_t^\alpha D_x^\beta u(x,t)| \leq C(\varepsilon, T) \left\{ |f|_{C^n_{\varepsilon,T}} + |g|_{C^n_{\varepsilon,T}} + |\phi|_{L^1} \right\} \tag{2.3}$$

where $|f|_{C^n_{\varepsilon,T}} := |f|_{C^n([\varepsilon,T])}$. If we also restrict $x$ to remain bounded away from the spatial boundary, we can bound interior derivatives by low-order norms of $f$, $g$, and $\phi$:

$$\sup_{\substack{x\in[p_{i-1}+\delta,\, p_i-\delta] \\ T\geq t\geq\varepsilon}} |D_t^\alpha D_x^\beta u(x,t)| \leq C(\varepsilon,\delta,T)\left\{|f|_{L^\infty} + |g|_{L^\infty} + |\phi|_{L^1}\right\}. \tag{2.4}$$

The case of the first $x$-derivative $u_x(x,t)$ is slightly better in that for positive time it is controlled by the lower norms up to the spatial boundary

$$\sup_{\substack{x\in[p_{i-1},\, p_i] \\ T\geq t\geq\varepsilon}} |u_x(x,t)| \leq C(\varepsilon,T)\left\{|f|_{L^\infty} + |g|_{L^\infty} + |\phi|_{L^1}\right\}. \tag{2.5}$$

## Solution of the Nonlinear System

We establish local-in-time existence and uniqueness for continuous initial data.

PROPOSITION 2.1 *The system of equations*

$$\begin{aligned}
\frac{\partial u_i}{\partial t} &= \frac{\partial^2 u_i}{\partial x^2} && \textit{for } p_{i-1} < x < p_i\,,\\
\frac{\partial u_i}{\partial x}(p_i,t) &= \frac{\partial u_{i+1}}{\partial x}(p_i,t) \\
&= f_i(u_{i+1}(p_i,t) - u_i(p_i,t)) && \textit{for } i \neq 0, N\,,\\
\frac{\partial u_1}{\partial x}(p_0,t) &= \frac{\partial u_N}{\partial x}(p_N,t) = 0\,,\\
u_i(x,0) &= \phi_i(x)\,,
\end{aligned} \tag{2.6}$$

*where the $f_i(x)$ are Lipschitz-continuous and the functions $\phi_i(x)$ in the initial condition are continuous, has a unique (local-in-time) solution.*

PROOF: Let $X_i = \{u \in C([0,T]\times I_i) : u(x,0) = \phi_i(x)\}$ with $I_i := (p_{i-1}, p_i)$, and set $X = X_1 \times X_2 \times \cdots \times X_N$. On this set we take the metric

$$d(u,v) := \max_{i=1,2,\ldots,N} \sup_{\substack{x\in I_i \\ t\in[0,T]}} |u_i(x,t) - v_i(x,t)| \quad \text{for } u, v \in X.$$

Now consider the mapping $S : X \to X$ defined in the following manner: For $v \in X$, $S(v) = u$ where $u_i$ is the solution of the problem:

$$\begin{aligned}
\frac{\partial u_i}{\partial t} &= \frac{\partial^2 u_i}{\partial x^2} \quad \text{in } I_i\,, \\
\frac{\partial u_i}{\partial x}(p_i, t) &= \frac{\partial u_{i+1}}{\partial x}(p_i, t) = f_i(v_{i+1}(p_i, t) - v_i(p_i, t)) \quad \text{for } i \neq 0, N\,, \\
\frac{\partial u_1}{\partial x}(p_0, t) &= \frac{\partial u_N}{\partial x}(p_N, t) = 0\,, \\
u_i(x, 0) &= \phi_i(x)\,,
\end{aligned}$$

which is a linear decoupled system whose solutions can be expressed by formula (2.2). Indeed, if we let $J_i[u](t) := u_{i+1}(p_i, t) - u_i(p_i, t)$, we obtain the formula

$$S(u)_i = \int_{p_{i-1}}^{p_i} P_i(x, y, t)\phi(y)dy + \sum_{j=i-1,i} \int_0^t P_i(x, p_j, s)\, f_j(J_j[u](t-s))ds\,.$$

We will show that the mapping $S$ can be made contractive by choosing $T > 0$ small enough. To this end, let $u, v \in X$ and set $L := \max_i \mathrm{Lip}(f_i)$. Then the representation obtained above implies

$$\begin{aligned}
|S(u)_i - S(v)_i| \le L &\int_0^t P_i(x, p_{i-1}, t-s)|u_{i-1}(p_{i-1}, s) - v_{i-1}(p_{i-1}, s)|ds \\
&+ L \int_0^t P_i(x, p_{i-1}, t-s)|u_i(p_{i-1}, s) - v_i(p_{i-1}, s)|ds \\
&+ L \int_0^t P_i(x, p_i, t-s)|u_i(p_i, s) - v_i(p_i, s)|ds \\
&+ L \int_0^t P_i(x, p_i, t-s)|u_{i+1}(p_i, s) - v_{i+1}(p_i, s)|ds\,,
\end{aligned}$$

which by the elementary bound

$$\left| \int_0^t P_i(x, y, t-s)ds \right| \le C\sqrt{t}$$

implies the inequality

$$\sup_{\substack{p_{i-1} \le x \le p_i \\ 0 \le t \le T}} |S(u)_i - S(v)_i| \le C\sqrt{T} d(u, v)$$

where $C$ depends on $L$. But then taking the maximum over $i$ we get

$$d(S(u), S(v)) \le C\sqrt{T} d(u, v)\,,$$

which of course means we have a contraction for a sufficiently small choice of $T > 0$. It follows that the mapping $S$ has a fixed point $u(x, t)$ in the (complete) metric space $X$. This is our candidate for the solution to the system.

Next, we note that $u(x,t)$ can be recognized as the limit of a sequence $\{u^{(n)}\}_{n=0}^{\infty}$ where $u_i^{(0)}(x,t) := \phi_i(x)$ and $u^{(n+1)}(x,t) := S(u^{(n)})$ for $n = 0, 1, \ldots,$ by definition. By virtue of our argument, this sequence converges uniformly as soon as we ensure that $S$ is contractive by taking $T > 0$ suitably small. Applying estimate (2.4) to $u^{(n)}(x,t) - u^{(m)}(x,t)$, we see that the sequence of derivatives $\{D_t^\alpha D_x^\beta u^{(n)}(x,t)\}_{n=1}^{\infty}$ converges uniformly on every compactly included subset of $(p_0, p_1) \times (p_1, p_2) \times \cdots \times (p_{N-1}, p_N) \times (0, T)$, and therefore the limiting function $u(x,t)$ is smooth on this domain and satisfies the heat equation there just like every term in the sequence.

We also need to check that the boundary condition makes sense (i.e., the limit possesses one derivative in $x$ up to the boundary for positive time) and is satisfied. This is a consequence of (2.5) applied once again to $u^{(n)}(x,t) - u^{(m)}(x,t)$; this time we see that $\{u_x^{(n)}(x,t)\}_{n=1}^{\infty}$ converges uniformly on every set of the form $[p_0, p_1] \times [p_1, p_2] \times \cdots \times [p_{N-1}, p_N] \times [\varepsilon, T - \varepsilon]$. So the limit $u(x,t)$ possesses an $x$-derivative up to the spatial boundary, and since the sequence of boundary values $\{f_j(u^{(n)}(p_{j+1}, t)) - f_j(u^{(n)}(p_j, t))\}_{n=1}^{\infty}$ converge to $f_j(u(p_{j+1}, t) - u(p_j, t))$, it satisfies the correct boundary condition.

Finally, the candidate $u(x,t)$ assumes the correct initial value as $t \to 0^+$ as a consequence of its continuity and the manner in which the sequence has been constructed. We hence see that $u(x,t)$ is the unique solution of the nonlinear system. □

*Remark.* The choice of $T > 0$ in the existence argument is constrained only by the size of the Lipschitz constants of the functions $f_i$. Therefore, in case the functions are globally Lipschitz, by iteration of the argument we can obtain global-in-time existence.

**Higher Regularity**

We need bounds on higher derivatives (e.g., $u_{xxx}$) on the domain $[p_0, p_1] \times [p_1, p_2] \times \cdots \times [p_{N-1}, p_N] \times (0, T)$; in other words, we need higher regularity up to the spatial boundary for positive time. This will be needed for the convergence argument later on, where we shall need to estimate how well difference quotients approximate first and second derivatives of the solution to the system.

PROPOSITION 2.2 *Let $f_1, f_2, \ldots, f_{N-1} \in C^\infty$, and let $u(x,t) = \{u_i(x,t)\}_{i=1}^N$ be the solution to the system with nonlinear boundary conditions* (2.6). *Then*

$$u_i(x,t) \in C^\infty([p_{i-1}, p_i] \times (0, \infty)) \quad \text{for } i = 1, 2, \ldots, N\,.$$

PROOF: We recall some fundamental properties of heat potentials; for details, see [5] and [13]. First, if $f_i(t)$ are continuous functions, then the single layer potential

$$\sum_{j=i-1,i} \int_0^t P_i(x, p_j, s) f_j(s) ds \tag{2.7}$$

is $C^{\gamma,\gamma/2}([p_{i-1}, p_i] \times [\delta, T - \delta])$ for any $\gamma \in (0, \frac{1}{2})$. To fix ideas, take $\gamma = \frac{1}{4}$. It is easy to see by a uniqueness argument that this implies $u(x,t)$ has the same Hölder continuity.

Second, if $f_i(t) \in C^{\nu/2}$ where $\nu$ is a noninteger positive number, then the single-layer potential given in (2.7) is in fact $C^{\nu+1,(\nu+1)/2}$. In other words, convolution with the heat potential $P_i$ as in (2.7) allows us to gain (at least) one full derivative in the $x$-direction and half a derivative (in the Hölder sense) in the $t$-direction.

The proof of regularity can now proceed by induction. Assume $u \in C^{\gamma,\gamma/2}$ so that $u(p_i,t)$ are $C^{\gamma/2}$-functions of time. Then the jump heights $J_i[u](t) \in C^{\gamma/2}$. Since the functions $f_i$ are $C^\infty$, we get

$$\partial_x u_i(p_i,t) = \partial_x u_{i+1}(p_i,t) = f_i(J_i[u](t)) \in C^{\gamma/2}\,.$$

Our remarks in the previous paragraph imply $u_i(x,t) \in C^{\gamma+1,(\gamma+1)/2}$. By induction, $u \in C^\infty$. □

COROLLARY 2.3 *System* (1.12)*, proposed as a continuum limit for the Perona-Malik scheme, has a unique solution with good regularity properties while the jump heights $J_i(t)$ remain bounded away from* 0*.*

PROOF: In terms of the notation employed in the existence proof, the proposed continuum limit (1.12) is nothing other than system (2.6) with $f_i(\xi) := \xi|\xi|^{2\beta-2}$ for $i = 1, 2, \dots, N-1$.

Let $m := \min_{i=1,2,\dots,N-1} J_i(0) > 0$, and fix $\varepsilon \in (0,m)$. Let $f^{(\varepsilon)}(\xi)$ be a $C^\infty$-function such that $f^{(\varepsilon)}(\xi) = \xi|\xi|^{2\beta-2}$ for $|\xi| > \varepsilon$. Apply the existence theorem (2.1) with the choice of functions $f_i(\xi) = f^{(\varepsilon)}(\xi)$ for $i = 1, 2, \dots, N-1$. That yields a (global-in-time) solution; call it $u^{(\varepsilon)}(x,t)$. But then $u^{(\varepsilon)}(x,t)$ is a solution to the system (1.12) as long as $\min_{i=1,2,\dots,N-1} u^{(\varepsilon)}_{i+1}(p_i,t) - u^{(\varepsilon)}_i(p_i,t) \ge \varepsilon$. Note that if $0 < \varepsilon' < \varepsilon$, then $u^{(\varepsilon)} = u^{(\varepsilon')}$ while $\min_i u^{(\varepsilon')}_{i+1}(p_i,t) - u^{(\varepsilon')}_i(p_i,t) \ge \varepsilon$. That proves our claim, since $\varepsilon > 0$ can be chosen arbitrarily small. □

## 2.2 Properties of Solutions

Here we discuss some important properties of solutions to the proposed continuum limit: maximum principle, bounds on gradients, and Hölder continuity in time. We will denote $u(x,t)$ the piecewise continuous function on $[p_0, p_N]$ where $u(x,t) := u_i(x,t)$ for $x \in (p_{i-1}, p_i)$.

PROPOSITION 2.4 (Maximum Principle) *Let $u(x,t) := \{u_i(x,t)\}_{i=1}^N$ be a solution to the proposed continuum limit* (1.12) *for $0 \le t \le T$ with $\beta \in [0, \frac{1}{2})$. Then for all $t \ge 0$,*

$$\sup_x |u(x,t)| \le \sup_x |u(x,0)|\,.$$

PROOF: A convenient way of showing this statement is to follow an argument given in [8]. Set $f(\xi) = \xi|\xi|^{2\beta-2}$, the boundary conditions in (1.12). Then

$$\begin{aligned}\frac{1}{p}\frac{d}{dt}\int_{p_{i-1}}^{p_i} |u_i|^p\,dx &= \int_{p_{i-1}}^{p_i} |u_i|^{p-2}u_i\partial_x^2 u_i\,dx \\ &\le |u_i|^{p-2}u_i\partial_x u_i\Big|_{p_{i-1}}^{p_i} \\ &= |u_i|^{p-2}u_i f(u_{i+1}-u_i)\Big|_{p_i} - |u_i|^{p-2}u_i f(u_i - u_{i-1})\Big|_{p_{i-1}}.\end{aligned}$$

Summing over $i = 1, 2, \ldots, N$ we see that

$$\frac{1}{p}\frac{d}{dt}\sum_{i=1}^{N}\int_{p_{i-1}}^{p_i} |u_i|^p\,dx \le \sum_{i=1}^{N-1}\left[|u_i|^{p-2}u_i - |u_{i+1}|^{p-2}u_{i+1}\right] f(u_{i+1}-u_i)\Big|_{p_i},$$

which is negative because $xf(x) \ge 0$ for our specific choice of boundary conditions, and

$$(|x|^{p-2}x - |y|^{p-2}y)(x-y) \ge 0$$

for all $x, y$ provided that $p > 1$. Letting $p \to \infty$ gives the desired result. □

For $\varepsilon > 0$ and $p > 1$ let

$$F_{p,\varepsilon}(x) = (x^2 + \varepsilon^2)^{p/2}$$

so that

$$F''_{p,\varepsilon}(x) \ge \begin{cases} 0 & \text{for all } p > 1 \\ p(p-1)(x^2+\varepsilon^2)^{(p-2)/2} & \text{for } p \in (1,2]. \end{cases} \tag{2.8}$$

Now we compute

$$\begin{aligned}&\frac{d}{dt}\sum_{i=1}^{N}\int_{p_{i-1}}^{p_i} F_{p,\varepsilon}(\partial_x u_i)dx \\ &\quad= \sum_{i=1}^{N}\int_{p_{i-1}}^{p_i} F'_{p,\varepsilon}(\partial_x u_i)\partial_{xt}u_i\,dx \\ &\quad= -\sum_{i=1}^{N}\int_{p_{i-1}}^{p_i} F''_{p,\varepsilon}(\partial_x u_i)\partial_x^2 u_i\partial_t u_i\,dx + \sum_{i=1}^{N} F'_{p,\varepsilon}(\partial_x u_i)\partial_t u_i\Big|_{p_{i-1}}^{p_i}\end{aligned}$$

where we integrated by parts in the last step. Invoking the boundary conditions (1.12), we find

$$\begin{aligned}\frac{d}{dt}\sum_{i=1}^{N}\int_{p_{i-1}}^{p_i} F_{p,\varepsilon}(\partial_x u_i)dx = &-\sum_{i=1}^{N}\int_{p_{i-1}}^{p_i} F''_{p,\varepsilon}(\partial_x u_i)\partial_x^2 u_i\partial_t u_i\,dx \\ &-\sum_{i=1}^{N} F'_{p,\varepsilon}(J_i(t)|J_i(t)|^{2\beta-2})\partial_t J_i(t)\end{aligned} \tag{2.9}$$

where we rearranged the terms in the last sum. Integrating in $t$ over $(t_1, t_2)$ and making the change of variable $y = J_i(t)$ in the boundary terms that make up the last sum, we obtain the equation

$$\begin{aligned}\sum_{i=1}^{N}\int_{p_{i-1}}^{p_i} F_{p,\varepsilon}(\partial_x u_i)dx\Big|_{t_1}^{t_2} = &-\sum_{i=1}^{N}\int_{t_1}^{t_2}\int_{p_{i-1}}^{p_i} F''_{p,\varepsilon}(\partial_x u_i)\partial_x^2 u_i \partial_t u_i \, dx\, dt \\ &-\sum_{i=1}^{N-1}\int_{J_i(t_1)}^{J_i(t_2)} F'_{p,\varepsilon}(y|y|^{2\beta-2})dy\,. \end{aligned} \tag{2.10}$$

From this formula we obtain the following estimates:

PROPOSITION 2.5 ($L^p$-Bound for Derivatives) *Let* $u(x,t) = \{u_i(x,t)\}_{i=1}^{\infty}$ *be a solution to system* (1.12) *with* $\beta \in [0, \frac{1}{2})$ *for* $0 \le t < T$. *We then have*

$$\sup_{0<t\le T}\sum_{i=1}^{N}\int_{p_{i-1}}^{p_i}\left|\frac{\partial u_i}{\partial x}(x,t)\right|^p dx = C(p) < \infty$$

*where* $1 < p < 2(1-\beta)/(1-2\beta)$. *Furthermore, the constant* $C(p)$ *depends only on the piecewise* $W^{1,p}$*-norm of the initial condition and the initial jump heights, in addition to* $p$ *and* $\beta$.

PROOF: Since $\partial_t u_i = \partial_x^2 u_i$ and $F''_{p,\varepsilon}(\xi) \ge 0$ for all $\xi$ as seen in (2.8), we get

$$F''_{p,\varepsilon}(\partial_x u_i)\partial_x^2 u_i \partial_t u_i = F''_{p,\varepsilon}(\partial_x u_i)(\partial_t u_i)^2 \ge 0\,.$$

So the first term in the right-hand side of formula (2.10) is negative; once we let $\varepsilon$ go to 0 in this formula, we therefore get the inequality

$$\begin{aligned}&\sum_{i=1}^{N}\|\partial_x u_i(\cdot,t_2)\|^p_{L^p(p_{i-1},p_i)} \le \\ &\qquad \sum_{i=1}^{N}\|\partial_x u_i(\cdot,t_1)\|^p_{L^p(p_{i-1},p_i)} + p\sum_{i=1}^{N-1}\int_{J_i(t_1)}^{J_i(t_2)}|y|^{(2\beta-1)(p-1)}\,dy\end{aligned}$$

The integrands on the right-hand side are locally integrable for $p < 2(1-\beta)/(1-2\beta)$. Furthermore, the intervals of integration can be bounded in terms of the initial condition by using, for example, Proposition 2.4 (the maximum principle). That proves the claim, since the right-hand side is shown to be controlled completely in terms of the initial condition. □

A most important property of solutions to the proposed limit (1.12) with the less singular boundary conditions that correspond to $\beta \in (0, \frac{1}{2})$ is that they are the steepest descent for an energy. This is merely a special case of equation (2.9):

PROPOSITION 2.6 (Steepest Descent) *Let* $u(x,t) = \{u_i(x,t)\}_{i=1}^N$ *be a solution to system* (1.12) *with* $\beta \in (0, \frac{1}{2})$ *for* $0 \le t < T$. *Define the energy*

$$\mathbf{E}_u(t) := \frac{1}{2} \sum_{i=1}^{N} \int_{p_{i-1}}^{p_i} \left( \frac{\partial u_i}{\partial x}(x,t) \right)^2 dx + \frac{1}{2\beta} \sum_{i=1}^{N-1} |J_i(t)|^{2\beta} . \tag{2.11}$$

*Then the following relation holds*:

$$\frac{d}{dt}\mathbf{E}_u(t) = -\sum_{i=1}^{N} \int_{p_{i-1}}^{p_i} \left( \frac{\partial u_i}{\partial t}(x,t) \right)^2 dx . \tag{2.12}$$

PROOF: In equation (2.9) take $p = 2$ and $\varepsilon = 0$. Noting that

$$\frac{d}{dt} \frac{1}{2\beta} |J_i(t)|^{2\beta} = F_{2,0}\big(J_i(t)|J_i(t)|^{2\beta-2}\big) ,$$

we obtain the promised formula. □

*Remark.* The case $\beta = 0$ decreases the energy

$$\mathbf{E}_u(t) := \frac{1}{2} \sum_{i=1}^{N} \int_{p_{i-1}}^{p_i} \left( \frac{\partial u_i}{\partial x}(x,t) \right)^2 dx + \sum_{i=1}^{N-1} \log(|J_i(t)|) ,$$

which, however, is not bounded from below as $J_i \to 0$.

As a consequence of the estimates above, we obtain the following Hölder-continuity-in-time result, which shows that solutions to the continuum limit evolve slowly all the way up to the singular times.

COROLLARY 2.7 (Hölder Continuity) *Let* $u(x,t) = \{u_i(x,t)\}_{i=1}^\infty$ *be a solution to system* (1.12) *with* $\beta \in (0, \frac{1}{2})$ *for* $0 \le t < T$. *Then for* $i = 1, 2, \ldots, N$,

$$u_i(t, \cdot) \in C^{1/2}\big([0,T); L^2((p_{i-1}, p_i))\big)$$

*and*

$$u_i(t, \cdot) \in C^{1/4}\big([0,T); L^\infty((p_{i-1}, p_i))\big) .$$

*For the more singular case* $\beta = 0$, *we instead have*

$$u_i(t, \cdot) \in C^{\mu}\big([0,T); L^2((p_{i-1}, p_i))\big)$$

*and*

$$u_i(t, \cdot) \in C^{\nu}\big([0,T); L^\infty((p_{i-1}, p_i))\big)$$

*for any* $\mu \in (0, \frac{1}{2})$ *and* $\nu \in (0, \frac{1}{4})$. *Furthermore, in all cases the Hölder constants involved depend only on the appropriate piecewise* $W^{1,p}$*-norm and jump heights of the initial condition.*

PROOF: By an application of the Hölder inequality followed by switching the order of integration, we have

$$\begin{aligned}\sum_{i=1}^{N}\int_{p_{i-1}}^{p_i}|u_i(x,t_2)-u_i(x,t_1)|^2\,dx &= \sum_{i=1}^{N}\int_{p_{i-1}}^{p_i}\left|\int_{t_1}^{t_2}\frac{\partial u_i}{\partial t}(x,s)ds\right|^2 dx\\ &\le |t_2-t_1|\sum_{i=1}^{N}\int_{t_1}^{t_2}\int_{p_{i-1}}^{p_i}\left|\frac{\partial u_i}{\partial t}(x,s)\right|^2 dx\,ds\,.\end{aligned}$$

The right-hand side of the above inequality can now be bounded by the integral in $t$ over $[t_1, t_2]$ of the energy identity (2.12) to get

$$\begin{aligned}\sum_{i=1}^{N}\int_{p_{i-1}}^{p_i}\left|u_i(x,t_2)-u_i(x,t_1)\right|^2 dx &\le |t_2-t_1|\big(\mathbf{E}_u(t_1)-\mathbf{E}_u(t_2)\big)\\ &\le |t_2-t_1|\mathbf{E}_u(t_1)\,,\end{aligned}$$

which is exactly the definition of $C^{1/2}$ Hölder continuity in time with values in $L^2$ of space, and the Hölder constant $(\mathbf{E}_u(t_1))^{1/2}$ depends on conditions at the beginning of the time interval, as promised.

To get Hölder continuity with values in $L^\infty$ of space, first note that by Proposition 2.5 the $L^2$-norm of derivatives $\partial_x u_i$ are bounded:

$$\sup_{t\ge 0}\sum_{i=1}^{N}\|\partial_x u_i(\cdot,t)\|_{L^p((p_{i-1},p_i))}<\infty\,.$$

We can therefore apply the interpolation lemma (Lemma 4.1) to $f = u_i(x,t_2) - u_i(x,t_1)$ with $p = q = r = 2$ and $\theta = \frac{1}{2}$ to get the desired result.

For the case $\beta = 0$, the boundary terms are not integrable for $p = 2$, so we are forced to work with $p < 2$. To that end, we take $t_1 \le t_2$ and write equation (2.9) as

$$\begin{aligned}&\sum_{i=1}^{N}\int_{t_1}^{t_2}\int_{p_{i-1}}^{p_i}F''_{p,\varepsilon}(\partial_x u_i)(\partial_t u_i)^2\,dx\,dt\le\\ &\qquad\sum_{i=1}^{N}\int_{p_{i-1}}^{p_i}F_{p,\varepsilon}(\partial_x u_i)dx\bigg|_{t_1}-\sum_{i=1}^{N-1}\int_{J_i(t_1)}^{J_i(t_2)}F'_{p,\varepsilon}(y|y|^{2\beta-2})dy\,,\end{aligned}$$

where we note that as before the right-hand side can be bounded in terms of the initial condition. So we have that, for each $i$,

$$\int_{t_1}^{t_2}\int_{p_{i-1}}^{p_i}F''_{p,\varepsilon}(\partial_x u_i)(\partial_t u_i)^2\,dx\,dt$$

is bounded. Apply now the Hölder inequality with exponents $2/p$ and $2/(2-p)$ to get

$$\begin{aligned}&\int_{t_1}^{t_2}\int_{p_{i-1}}^{p_i} |\partial_t u_i|^p\,dx\,dt\\ &\quad= \int_{t_1}^{t_2}\int_{p_{i-1}}^{p_i} |\partial_t u_i|^p F''_{p,\varepsilon}(\partial_x u_i)^{\frac{p}{2}} F''_{p,\varepsilon}(\partial_x u_i)^{-\frac{p}{2}}\,dx\,dt\\ &\quad\le \left(\int_{t_1}^{t_2}\int_{p_{i-1}}^{p_i} |\partial_t u_i|^2 F''_{p,\varepsilon}(\partial_x u_i)dx\,dt\right)^{\frac{p}{2}}\left(\int_{t_1}^{t_2}\int_{p_{i-1}}^{p_i} F''_{p,\varepsilon}(\partial_x u_i)^{\frac{p}{p-2}}\,dx\,dt\right)^{\frac{2-p}{2}}.\end{aligned}$$

The first term in the right-hand side is bounded by our comments above; as for the second term, by (2.8) we have

$$|F''_{p,\varepsilon}(x)|^{\frac{p}{p-2}} \le C(p)(x^2+\varepsilon^2)^{\frac{p}{2}} \le C(p)(|x|^p+|\varepsilon|^p)$$

and therefore

$$\begin{aligned}\left(\int_{t_1}^{t_2}\int F''_{p,\varepsilon}(\partial_x u_i)^{\frac{p}{p-2}}\,dx\,dt\right)^{\frac{2-p}{2}} &\le \left(\int_{t_1}^{t_2}\int C(p)(|\partial_x u_i|^p+|\varepsilon|^p)dx\,dt\right)^{\frac{2-p}{p}}\\ &\le C(p)|t_2-t_1|^{\frac{2-p}{2}}\left(\sup_{t\ge 0}\|\partial_x u_i\|_{L^p}^p+|\varepsilon|^p\right)^{\frac{2-p}{p}}.\end{aligned}$$

But by Proposition 2.5 the term $\sup_{t\ge 0}\|\partial_x u_i\|_{L^p}$ is bounded in terms of the initial condition. Hence we finally get

$$\int_{t_1}^{t_2}\int_{p_{i-1}}^{p_i} |\partial_t u_i|^p\,dx\,dt \le C|t_2-t_1|^{\frac{2-p}{2}}$$

where the constant $C$ depends only on the initial condition and the exponent $p$. Proceeding now as in the case for $p=2$, another application of the Hölder inequality gives

$$\begin{aligned}\int_{p_{i-1}}^{p_i} |u_i(x,t_2)-u_i(x,t_1)|^p\,d &= \int_{p_{i-1}}^{p_i}\left|\int_{t_1}^{t_2}\partial_t u_i(x,s)ds\right|^p dx\\ &\le |t_2-t_1|^{p-1}\int_{t_1}^{t_2}\int_{p_{i-1}}^{p_i}|\partial_t u_i|^p\,dx\,dt\\ &\le C|t_2-t_1|^{\frac{p}{2}}\end{aligned}$$

which means $u_i(\cdot,t)\in C^{1/2}([0,T];L^p((p_{i-1},p_i)))$ with the Hölder constant depending only on initial data as before. But now judicious use of the interpolation lemma, as in the $p=2$ case, shows that the solutions also lie in the spaces quoted in the proposition. □

## 2.3 Healing of Interfaces

We start by showing that for solutions of the proposed continuum limit, jump heights $J_i(t)$ at the discontinuity points $p_i$ converge to 0 in finite time, thereby leading to gradient blowup at the jump locations.

PROPOSITION 2.8 *Let $u(x,t)$ be the solution to the proposed continuum limit* (1.12) *with $\beta \in [0, \frac{1}{2})$. Then there exists $T > 0$ such that*

(i) $|J_i(t)| > 0$ *for all $i = 1, 2, \ldots, N-1$ and $t \in [0, T)$.*
(ii) *There is $j \in \{1, 2, \ldots, N-1\}$ such that* $\liminf_{t\to T^-} |J_j(t)| = 0$.
(iii) *For all $i \in 1, 2, \ldots, N-1$ with* $\liminf_{t\to T^-} |J_i(t)| = 0$, *we have in fact* $\lim_{t\to T^-} |J_i(t)| = 0$.

In order to prove this proposition, we first show the following lemma, which establishes the preliminary result that jump heights $J_i(t)$ cannot remain bounded away from 0, so that no solution to the proposed limit with discontinuous initial data can exist for all time.

LEMMA 2.9 *There are no global-in-time solutions to the system given in* (1.12) *if the initial condition has jumps. In particular, jump heights cannot remain bounded away from* 0.

PROOF: For the most singular boundary condition (case $\beta = 0$) given in (1.14), the statement is particularly easy to show: Suppose that $u(x,t) = \{u_i(x,t)\}_{i=1}^N$ is a global-in-time solution to (1.12) with $N > 1$. We will obtain a contradiction. Compute

$$\begin{aligned}\frac{1}{2}\frac{d}{dt}\sum_{i=1}^{N}\int_{p_{i-1}}^{p_i} u_i^2(x,t)dx &= -\sum_{i=1}^{N}\int_{p_{i-1}}^{p_i}(\partial_x u_i(x,t))^2\,dx - \sum_{i=1}^{N-1}\left(\partial_x u_{i+1}u_{i+1} - \partial_x u_i u_i\right)\Big|_{(p_i,t)}\\ &\le -\sum_{i=1}^{N-1} 1 = -(\text{no. of jumps}),\end{aligned}$$

where we integrated by parts in $x$ and then employed the boundary conditions. We thus see that the $L^2$-norm of the solution decays at a definite rate in the presence of jumps and therefore would become negative in finite time if the jumps persisted; this is a contradiction. The evolution will necessarily be interrupted, and that can happen only if jump heights vanish.

For the less singular boundary conditions (case $\beta \in (0, \frac{1}{2})$), one can proceed as follows: We consider the jump at $p_1$. Without loss of generality, assume that $u_2(p_1, 0) > u_1(p_1, 0)$ so that the jump $J_1(t)$ is positive. By Proposition 2.4 (maximum principle), we are assured that there is a constant $M > 0$ such that

$|u_1(x,t)|, |u_2(x,t)| \le M$; so in particular, $J_1(t) = u_2(p_1,t) - u_1(p_1,t) \le 2M$. But then we compute

$$\begin{aligned} \frac{d}{dt}\int_{p_0}^{p_1} |M - u_1(x,t)|dx &= \frac{d}{dt}\int_{p_0}^{p_1} (M - u_1(x,t))dx \\ &= -\int_{p_0}^{p_1} \partial_{xx} u_1(x,t)dx \\ &= -\partial_x u_1(x,t)\Big|_{p_0}^{p_1} = -\partial_x u_1(p_1,t) \\ &= -J_1^{2\beta-1} \le -(2M)^{2\beta-1} < 0\,. \end{aligned}$$

So the $L^1$-norm of $M - u_1$ decays at a definite rate, and it would become negative if the jump at $p_1$ survived. This is a contradiction; jumps cannot remain bounded away from 0. That concludes the proof. □

PROOF OF PROPOSITION 2.8: Let $T > 0$ be the maximal time of existence for the solution. By Lemma 2.9 we know that $T < \infty$. Furthermore, the local-in-time existence result verifies the second assertion of the proposition, since if all jump heights remained bounded away from 0 up to $T$, the solution could be continued a little further.

For the third assertion, we use the Hölder continuity in time with values in $L^\infty$ of space property given by Corollary 2.7. As a consequence, the limit

$$\hat{u}_i := \lim_{t\to T^-} u_i(\cdot,t)$$

exists in the uniform sense on $[p_{i-1}, p_i]$ for every $i$, and so $\liminf_{t\to T^-} |J_i(t)| = 0$ for any $i$ implies $\limsup_{t\to T^-} |J_i(t)| = 0$ as well. Moreover, uniform convergence also means that the functions $\hat{u}_i$ are continuous up to the boundary on their respective domains. □

If only one jump height vanishes at the maximal time of existence, we can say more about the behavior of the solution. The next lemma shows that under this circumstance, the jump height in question strictly decreases once it becomes small enough. We will employ the following notation:

$$\begin{aligned} \Omega_i^T &= (p_{i-1}, p_i) \times (0,T]\,, \\ \Gamma_i^T &= [p_{i-1}, p_i] \times \{0\} \cup \{p_{i-1}\} \times [0,T] \cup \{p_i\} \times [0,T]\,, \end{aligned}$$

for $i = 1, 2, \dots, N$. So $\Gamma_i^T$ is the parabolic boundary of the cylindrical domain $\Omega_i^T$.

LEMMA 2.10 *Let* $\{u_i(x,t)\}_{i=1}^N$ *be a solution to the PDE system in* (1.12)*, and let* $T_q > 0$ *be the first quenching time. Assume*

$$m := \min_{i\ne k} \inf_{0\le t\le T_q} |J_i(t)| > 0$$

*so that $p_k$ is the only quenching point at $t = T_q$. Let*

$$M := \max\left\{ \sup_{p_{k-1}<x<p_k} |\partial_x u_k(x,0)|, \sup_{p_k<x<p_{k+1}} |\partial_x u_{k+1}(x,0)| \right\}.$$

*Then there exists a $\delta_J > 0$, depending only on $m$ and $M$, such that if $|J_k(t_0)| \le \delta_J$ for some $t_0 \in [0, T_q]$, then $|J_k(t)|$ is decreasing for $t \in [t_0, T_q]$.*

PROOF: Without loss of generality, we will take $J_k(0) > 0$. Choose $\delta_J > 0$ so small that $\delta_J < m$ and $\delta_J^{2\beta-1} > M$. Let $t_0 := \inf\{t \ge 0 : J_k(t) = \delta_J\}$. We need to show that $\partial_t J_k(t) < 0$ for $t \in [t_0, T_q)$. Assume not; then we can let $T_* := \inf\{t \in [t_0, T_q) : \partial_t J_k(t) \ge 0\}$. The choice of $\delta_J$ and the definitions of $t_0$ and $T_*$ give

$$\partial_x u_k(p_k, T_*) = \partial_x u_{k+1}(p_k, T_*) > \max\{M, m^{2\beta-1}\}$$

so that

$$\partial_x u_k(p_k, T_*) > \partial_x u_k(x,t) \qquad \text{for all } (x,t) \in \Gamma_k^{T_*} - (p_k, T_*)$$

and

$$\partial_x u_{k+1}(p_k, T_*) > \partial_x u_{k+1}(x,t) \quad \text{for all } (x,t) \in \Gamma_{k+1}^{T_*} - (p_k, T_*)\,.$$

By the strict maximum principle applied to $\partial_x u_k$ and $\partial_x u_{k+1}$ we must have

$$\partial_x u_k(p_k, T_*) > \partial_x u_k(x,t) \qquad \text{for all } (x,t) \in \Omega_k^{T_*}$$

and

$$\partial_x u_{k+1}(p_k, T_*) > \partial_x u_{k+1}(x,t) \quad \text{for all } (x,t) \in \Omega_{k+1}^{T_*}\,,$$

which leads to

$$\partial_t u_k(p_k, T_*) = \partial_{xx} u_k(p_k, T_*) \ge 0$$

and

$$\partial_t u_{k+1}(p_k, T_*) = \partial_{xx} u_{k+1}(p_k, T_*) \le 0\,.$$

In fact, by the parabolic analogue of the Hopf lemma (Lemma 4.2), we must have

$$\partial_t u_k(p_k, T_*) > 0 \quad \text{and} \quad \partial_t u_{k+1}(p_k, T_*) < 0\,,$$

which implies $\partial_t J_k(T_*) < 0$, a contradiction. □

Under the circumstances of the last lemma, we can establish an upper bound on the quenching time $T_q$ in terms of the jump height $J_k(t)$; that is the content of the next lemma.

LEMMA 2.11 *Let* $\{u_i(x,t)\}_{i=1}^N$ *and* $m$, $M$, *and* $\delta_J$ *be as in Lemma* 2.10. *Let* $T_q > 0$ *be the quenching time of the* $k^{th}$ *jump, and assume* $t_0 < T_q$ *is such that* $|J_k(t_0)| < \delta_J$. *Then*

$$T_q < t_0 + \frac{2K(p_k - p_{k-1})}{|J_k(t_0)|^{2\beta-1} - m^{2\beta-1}}$$

*where* $K := \sup_{i,x} |u_i(x,0)|$.

PROOF: By Proposition 2.4 we know that $\sup_{i,x,t} |u_i(x,t)| = K$. Without loss of generality, we will take $J_k(0) > 0$. By Lemma 2.10, the hypothesis $J_k(t_0) < \delta_J$ means that $J_k'(t) < 0$ for all $t \in (t_0, T_q)$. So in particular, $\partial_x u_k(p_k, t) > J_k(t_0)^{2\beta-1}$ for all $t \in (t_0, T_q)$. Consequently,

$$\begin{aligned}\frac{d}{dt}\int_{p_{k-1}}^{p_k} |K - u_k(x,t)|dx &= \frac{d}{dt}\int_{p_{k-1}}^{p_k} K - u_k(x,t)dx \\ &= -\partial_x u_k(p_k,t) + \partial_x u_k(p_{k-1},t) \le m^{2\beta-1} - J_k(t_0)^{2\beta-1}\end{aligned}$$

which, after integration in $t$ over $[t_0, T_q]$, implies that we have

$$\begin{aligned}0 &\le \int |K - u_k(x,T_q)|dx \\ &\le \int |K - u_k(x,t_0)|dx + \big(m^{2\beta-1} - J_k(t_0)^{2\beta-1}\big)(T_q - t_0) \\ &\le 2K(p_k - p_{k-1}) + \big(m^{2\beta-1} - J_k(t_0)^{2\beta-1}\big)(T_q - t_0)\,,\end{aligned}$$

and that is exactly the inequality required. □

## 2.4 Continuation Beyond Blowup

We have seen how the proposed PDE limit breaks down in finite time; to give a global-in-time candidate for the continuum limit, we must supplement the description afforded by (1.12).

Proposition 2.8 characterizes the manner in which the breakdown occurs: One or more of the jump heights converge to 0. Let $T_q$ be the first of these quenching times. In view of the results of the last section, it is easy to show that solutions $\{u_i(x,t)\}_{i=1}^N$ to the PDE system (1.12) have well-behaved limits as $t \to T_q^-$. Let $\phi_i(x) := \lim_{t\to T_q^-} u_i(x,t)$; what we meant is that this limit exists, and $\phi_i(x)$ are continuous up to the boundary on their respective domains $(p_{i-1}, p_i)$. What is more, according to Proposition 2.5, if the initial data are in $W^{1,p}$ (where the exponent $p$ is related to the parameter $\beta$ of the scheme as described in that proposition), then so are $\phi_i(x)$. We continue the evolution beyond the merging time $T_q$ in the following manner:

> If at the quenching time $t = T_q$ the jump located at $p_k$ vanishes so that $\lim_{t\to T_q^-} J_k(t) = 0$, then there is no longer a discontinuity across $p_k$ (in other words, $\phi_k(p_k) = \phi_{k+1}(p_k)$). We therefore

remove $p_k$ from our list of jump locations and merge the two intervals $(p_{k-1}, p_k)$ and $(p_k, p_{k+1})$ on either side of $p_k$ into a single, longer interval $(p_{k-1}, p_{k+1})$. That brings us back to a setting where the PDE-based evolution given in (1.12) makes sense, although we now have a different system with fewer jump points. We continue the evolution as the solution to the new PDE system.

## 3 Perona-Malik as a Numerical Scheme

In this section, we first obtain for the Perona-Malik scheme discrete analogues of the results from the last section. Section 3.3 then pulls together all that we know to prove the promised convergence result. It is worth emphasizing that our purpose has not been to propose an efficient numerical method for the proposed continuum problem; rather, it has been to show that the Perona-Malik scheme, although not intended for this purpose, in effect solves our proposed limit. And since the original purpose of Perona and Malik was quite different, we cannot expect the scheme to be particularly efficient in solving our limit.

### 3.1 Definitions and Hypothesis

DEFINITIONS (i) For $h = 1/m$ with $m \in \mathbb{N}$, define the grid $G_h$ to be the collection of points $\{x_0, x_1, \dots, x_m\}$ where $x_0 = 0$, $x_m = 1$, and $x_j = x_{j-1} + h$.

(ii) For $k \geq j+2$ let

$$\Omega^T_{j,k} = \{x_{j+1}, x_{j+2}, \dots, x_{k-1}\} \times (0, T]\,,$$
$$\Gamma^T_{j,k} = \{x_j, x_{j+1}, \dots, x_k\} \times \{0\} \cup \{x_j\} \times [0, T] \cup \{x_k\} \times [0, T]\,.$$

(iii) For a function $\phi$ defined on the interval $[0, 1]$, let $S(\phi)$ denote the set of its discontinuity points.

(iv) For a function $\phi^h$ defined on the grid $G_h$, let

$$\phi^h_j := \phi^h(x_j)\,, \quad D^+\phi^h_j := \frac{\phi^h_{j+1} - \phi^h_j}{h}\,, \quad \text{and} \quad D^-\phi^h_j := \frac{\phi^h_j - \phi^h_{j-1}}{h}\,.$$

(v) For a function $\phi^h$ defined on the grid $G_h$, define $S(\phi^h)$ to be the collection of indices $j \in \{0, 1, \dots, m-1\}$ such that $|D^+\phi^h_j| \geq z(h)$.

**The Numerical Scheme**

Let us recall the one-dimensional semidiscrete version of the Perona-Malik scheme; it can be written as

$$\begin{aligned}
&\frac{d}{dt}v_j(t) = D^- R_h(D^+ v_j(t)) \quad \text{for } 1 \leq j \leq m-1\,,\\
&\frac{d}{dt}v_0(t) = \frac{1}{h} R_h(D^+ v_0(t))\,,\\
&\frac{d}{dt}v_m(t) = -\frac{1}{h} R_h(D^+ v_{m-1}(t))\,.
\end{aligned} \tag{3.1}$$

**Assumptions on Initial Data**

Assume that we are given piecewise $W^{1,p}$ initial data with discontinuities at $p_1, p_2, \dots, p_{n-1} \in (0,1)$. More precisely, we require

(1) $\phi \in W^{1,p}((p_{i-1}, p_i))$ where $p = 2$ for $\beta \in (0, \frac{1}{2})$ and $p \in (1,2)$ for $\beta = 0$.
(2) $\lim_{x \to p_i^-} \phi(x) \neq \lim_{x \to p_i^+} \phi(x)$.
(3) $p_i \notin G_h$ for any $i$ and $h > 0$.

*Note.* Condition 1 implies continuity up to the boundary in each interval $(p_{i-1}, p_i)$. Condition 3 is purely for convenience; see also remark 2 below.

**Assumptions on Approximate Initial Data**

It is required that the numerical approximations $\phi^h$ to the initial condition $\phi$ have jump sets compatible with that of $\phi$. More precisely, assume the following:

(1) For all $p_i \in S(\phi)$ there exists $j \in S(\phi^h)$ such that $p_i \in (x_j, x_{j+1})$. For all $j \in S(\phi^h)$ there exists a unique $p_i \in S(\phi)$ such that $p_i \in (x_j, x_{j+1})$.
(2) $\max_j |\phi(x_j) - \phi^h(x_j)| \to 0$ as $h \to 0^+$.
(3) $\sup_{h>0} \sum_{j \notin S(\phi^h)} h|D^+\phi_j^h|^p < \infty$ for the same $p$ as in the assumptions on initial data above.

*Remarks.* (1) By the assumptions on continuum initial data, such an approximating sequence is easy to generate. For instance, one can take a piecewise $C^1$ sequence $\phi_n(x)$ that converges to $\phi(x)$ in the $W^{1,p}$-norm on each one of the intervals $(p_{i-1}, p_i)$, with $|\phi_n'(x)| < z(1/n)$, and then let $\phi_j^{1/n} := \phi_n(x_j)$.
(2) It is possible to be less restrictive about how well the jump *locations* of discrete data should match those of the continuum data. In fact, it should not be hard to show that both the continuum and the discrete evolutions are stable (in the $L^2$-norm) under changes in jump positions.
(3) Our assumptions impose a one-to-one correspondence between jump sets of the continuum initial condition and its discrete approximations. In Gobbino's work [7], this condition is automatically satisfied by requiring convergence of energies. Our energies do not impose compatibility of jump sets; we therefore made the necessary assumptions explicitly.

## 3.2 Qualitative Properties and Estimates

PROPOSITION 3.1 (Maximum Principle) *Let* $\{v_j(t)\}_{j=0}^m$ *be the solution generated by scheme* (3.1) *on* $G_h \times [0,\infty)$ *from initial data* $\phi^h$. *Then*

$$\sup_{t \geq 0} \max_{0 \leq j \leq m} |v_j(t)| \leq \max_{0 \leq j \leq m} |\phi_j^h| .$$

PROOF: This property was noticed by a number of previous authors; a proof appears in [17], for instance. Here, we mimic the proof of Proposition 2.4. Note that $xR(x) \geq 0$ for all $x$. We now compute, using summation by parts,

$$\begin{aligned}\frac{d}{dt}\sum_{j=0}^{m} h|v_j|^p &= \sum_{j=0}^{m} hp|v_j|^{p-2}v_j\dot{v}_j = \sum_{j=0}^{m} hp|v_j|^{p-2}v_j D^-(R(D^+v_j)) \\ &= -\sum_{j=0}^{m-1} hpD^+\big(|v_j|^{p-2}v_j\big)R(D^+v_j) \\ &\quad - pR(D^+v_{-1})|v_0|^{p-2} + pR(D^+v_m)|v_m|^{p-2} \\ &= -\sum_{j=0}^{m-1} hp(p-1)(D^+v_j)R(D^+v_j)|\xi_j|^{p-2}\end{aligned}$$

where $\xi_j$ is between $v_j$ and $v_{j+1}$. Therefore, whenever $p > 1$,

$$\sum_{j=0}^{m} h|v_j(t)|^p \leq \sum_{j=0}^{m} h|\phi_j|^p\,.$$

Sending $p \to \infty$ proves the claim. □

We next show that the difference quotients generated by the discretization satisfy a strict maximum principle; in particular, scheme (3.1) does not generate new jump locations. A proof of this fact first appeared in Gobbino's paper [7].

PROPOSITION 3.2 *Let* $\{v_j(t)\}$ *be the solution generated by scheme* (3.1) *on the grid* $G_h$. *We have*

(i) $S(v_j(t_2)) \subseteq S(v_j(t_1))$ *whenever* $t_2 \geq t_1$.

(ii) *Let* $\{x_\alpha, \ldots, x_{\alpha'+1}\}$ *be a subset of* $G_h$ *with* $\alpha' \geq \alpha + 2$. *Assume that*

$$\sup_{\alpha+1\leq j\leq\alpha'-1} |D^+v_j(0)| < z(h)$$

*and set* $M := \sup_{(j,t)\in\Gamma^T_{\alpha,\alpha'}} |R(D^+v_j(t))|$. *Then*

$$\sup_{(j,t)\in\Omega^T_{\alpha,\alpha'}} |D^+v_j(t)| \leq R_*^{-1}(M)\,.$$

*Moreover, if there exists* $(j_0, t_0) \in \Omega^T_{\alpha,\alpha'}$ *such that* $D^+v_{j_0}(t_0) = \pm R_*^{-1}(M)$, *then* $D^+v_j(0) = \pm R_*^{-1}(M)$ *for all* $j$.

PROOF: To show the first claim, we follow the argument in Gobbino's paper [7]: For all $j$, $D^+v_j(t)$ satisfies an ODE of the form

$$\frac{d}{dt}D^+v_j(t) = \frac{2}{h^2}\big(A(t) - R(D^+v_j(t))\big) \tag{3.2}$$

where $|A(t)| \leq R(z(h))$ for all $t \geq 0$. If $|D^+v_j(0)| < z(h)$, a comparison argument immediately shows that $|D^+v_j(t)| < z(h)$ for all $t \geq 0$.

To show the second claim, let

$$\overline{M} := \sup_{\Gamma^T_{\alpha,\alpha'} \cup \Omega^T_{\alpha,\alpha'}} |R(D^+ v_j(t))| .$$

Assume there is $(j_0, t_0) \in \Omega^T_{\alpha,\alpha'}$ such that $R(D^+ v_{j_0}(t_0)) = \overline{M}$; then $\partial_t D^+ v_{j_0}(t_0) \geq 0$. Discretization (3.1) implies

$$R(D^+ v_{j_0})(t_0) \leq \frac{1}{2}\Big(R(D^+ v_{j_0-1})(t_0) + R(D^+ v_{j_0+1})(t_0)\Big) ,$$

which of course means $R(D^+ v_{j_0-1}) = R(D^+ v_{j_0+1}) = \overline{M}$ at $t = t_0$. Repetition of this reasoning leads to

$$R(D^+ v_j(t_0)) = \overline{M} \quad \text{for all } j \in \{\alpha, \dots, \alpha'\} .$$

Therefore, $\overline{M} = M$. Revisiting formula (3.2), we see that $|A(t)| \leq M$ and $|D^+ v_j(0)| \leq R_*^{-1}(M)$. The same comparison argument shows that, under our assumption, $D^+ v_j(t) = R_*^{-1}(M)$ for all $t \in [0, t_0]$. In case $R(D^+ v_{j_0}(t_0)) = -\overline{M}$, the argument is the same, leading to $\overline{M} = -M$ and $D^+ v_j(t) = -R_*^{-1}(M)$ for all $t \in [0, t_0]$. □

Furthering the similarities between the evolutions of jump sets in the discrete and continuous settings, we next show that all jumps of a discrete solution vanish in finite time.

LEMMA 3.3 *Let* $\{\phi^h\}$ *be a sequence of discrete approximations* (*each defined on* $G_h$) *to a given piecewise continuous initial data* $\phi$ *with* $\phi^h$ *and* $\phi$ *subject to the usual assumptions. Let* $\{v^h(t)\}$ *be the corresponding discrete solutions generated by the scheme in* (3.1) *in which the constitutive functions are scaled with respect to* $h$ *as prescribed in* (1.13). *Let* $T_h := \inf\{t \geq 0 : S(v^h(t))$ *is empty*$\}$. *Then*

$$\limsup_{h \to 0^+} T_h < \infty .$$

PROOF: Let $S(\phi^h) := \{l_1(h), \dots, l_n(h)\}$ be ordered. By induction, it is sufficient to show that the leftmost jump (located at $x_{l_1}$) will "vanish" in finite time; in precise terms this means we will show that

$$\limsup_{h \to 0^+} \inf\Big\{t \geq 0 : |D^+ v^h_{l_1}(t)| = z(h)\Big\} < \infty .$$

Since the approximating sequence $\{\phi_h\}$ is required to converge to $\phi$ in a uniform sense on the grid as $h \to 0^+$, for $h$ small enough we have that

$$\max_j |\phi^h_j| \leq M := 2 \max_j |\phi(x_j)| < \infty .$$

By Proposition 3.1, we then have $\sup_{j,t\geq 0}|v_j^h(t)| \leq M$. Without loss of generality, we assume that $D^+ v_{l_1}^h > 0$. Let us write $v := v^h$ and compute

$$\begin{aligned}\frac{d}{dt}\sum_{j=0}^{l_1} h|M - v_j| &= \frac{d}{dt}\sum_{j=0}^{l_1} h(M - v_j)\\ &= -\sum_{j=0}^{l_1} h\dot{v}_j = -\sum_{j=0}^{l_1} hD^- R_h(D^+ v_j) = -R_h(D^+ v_{l_1})\,.\end{aligned}$$

But if $l_1 \in S(v^h(t))$, then $D^+ v_{l_1} \geq z(h)$ by definition, and by the bound on the maximum norm $D^+ v_{l_1} \leq 2M/h$. That gives

$$R_h(D^+ v_{l_1}) \geq C(\beta) M^{2\beta - 1}\,,$$

which means

$$\frac{d}{dt}\sum_{j=0}^{l_1} h|M - v_j| \leq -C(\beta) M^{2\beta-1}\,,$$

which is a definite decay rate for the $L^1$-norm that is independent of $h$. As in the proof of Lemma 2.9, that implies the leftmost jump located at $x_{l_1}$ vanishes in finite time. The argument can be iterated to show that all the other jumps also collapse in finite time. That concludes the proof. □

We thus have seen that quenching is also inevitable in the discrete setting. We now specialize to the case where only one jump is eliminated at a given quenching time. Under this assumption, and in the continuum setting of the proposed limiting evolution, Lemmas 2.10 and 2.11 gave us an upper bound on the quenching time in terms of the jump height. They have very simple discrete analogues.

LEMMA 3.4 *Let* $\{v_j(t)\}$ *be the solution generated by scheme* (3.1) *on the grid* $G_h$, *and let* $\{J_i(t)\}_{i=1}^n$ *be the jump heights. Assume* $T > 0$ *is such that quenching does not occur on* $t \in [0, T]$, *and let*

$$m := \min_{i\neq k}\inf_{0\leq t\leq T} |J_i(t)|\,.$$

*In particular,* $m > hz(h)$. *Let*

$$M := \max\left\{\sup_{l_{k-1}+1\leq j\leq l_k - 1} |D^+ v_j(0)|, \sup_{l_k+1\leq j\leq l_{k+1}-1} |D^+ v_j(0)|\right\}.$$

*Then there exists a* $\delta_J > hz(h)$, *depending only on* $m$ *and* $M$ *such that if* $|J_k(t_0)| \leq \delta_J$ *for some* $t_0 < T$, *then* $|J_k(t_0)|$ *is nonincreasing on* $t \in [0, T]$.

PROOF: The argument is the same as that of Lemma 2.10 with minor modifications. Without loss of generality, take $D^+ v_{l_k}(0) > 0$. Choose $\delta_J > hz(h)$ so small that $R_h(\delta_J/h) > \max\{R_h(m/h), R_h(M)\}$. Let $t_0 := \inf\{t \geq 0 : J_k(t) = \delta_J\}$. We must show: $\partial_t J_k(t) < 0$ for $t \in [t_0, T)$. Assume not, and let $T_* := \inf\{t \in [t_0, T) : \partial_t J_k(t) \geq 0\}$. The hypothesis and the definitions of $t_0$ and $T_*$ imply

$$R(D^+ v_{l_k}(T_*)) = \sup_{(j,t)\in\Gamma^{T_*}_{l_{k-1},l_k}} R(D^+ v_j(T_*)) = \sup_{(j,t)\in\Gamma^{T_*}_{l_k,l_{k+1}}} R(D^+ v_j(T_*))\,.$$

By the strict maximum principle (Proposition 3.2),

$$D^+ v_{l_k-1}(T_*),\, D^+ v_{l_k+1}(T_*) < R_*^{-1}(D^+ v_{l_k}(T_*))$$

so that $\dot v_{l_k-1}(T_*) > 0$ and $\dot v_{l_k+1}(T_*) < 0$. That means $\partial_t J_k(T_*) < 0$, a contradiction. □

LEMMA 3.5 *Let $\{v_j(t)\}$ and $m$, $M$, $\delta_J$, and $T$ be as in Lemma 3.4. Assume $t_0 \in [0, T]$ is such that $|J_k(t_0)| < \delta_J$. Then*

$$T \leq t_0 + \frac{2K(l_k - l_{k-1})h}{R_h(|J_k(t_0)|/h) - R_h(m/h)}$$

*where $K = \sup_j |v_j(0)|$.*

PROOF: Without loss of generality, assume that $J_k(0) > 0$. Lemma 3.4 implies that $J_k'(0) < 0$ for all $t \in [t_0, T]$. Therefore

$$\begin{aligned}
&\frac{d}{dt}\sum_{j=l_{k-1}+1}^{l_k-1} h|K - v_j(t)| \\
&\quad= \frac{d}{dt}\sum_{j=l_{k-1}+1}^{l_k-1} h(K - v_j(t)) = -\sum_{j=l_{k-1}+1}^{l_k-1} hD^-\big(R_h(D^+ v_j(t))\big) \\
&\quad= R_h(D^+ v_{l_{k-1}}) - R_h(D^+ v_{l_k}) \leq R_h\left(\frac{m}{h}\right) - R_h\left(J_k\frac{(t_0)}{h}\right).
\end{aligned}$$

Integrating in $t$ over $[t_0, T]$, we find

$$\begin{aligned}
0 &\leq \sum_{j=l_{k-1}+1}^{l_k-1} h|K - v_j(T)| \\
&\leq 2K(l_k - l_{k-1})h + \left(R_h\left(\frac{m}{h}\right) - R_h\left(J_k\frac{(t_0)}{h}\right)\right)(T - t_0)\,,
\end{aligned}$$

which gives the desired inequality. □

**Energy Estimates**

The energy estimates we obtained for the proposed continuum limit have discrete versions. We start with the analogue of Proposition 2.6 that holds for the less singular boundary conditions.

PROPOSITION 3.6 (Steepest Descent) *Let* $\{v_j(t)\}_{j=1}^m$ *be the solution generated by the scheme* (3.1), *with the constitutive function* $g_k(\xi) = (1+\xi/k)^{(\beta-1)}$ *for* $\beta \in (0, \frac{1}{2})$. *Define the energy*

$$\mathbf{E}_v^h(t) = \sum_{j=0}^{m-1} h\Phi_{k,\beta}\big((D^+v_j)^2\big) \tag{3.3}$$

*where*

$$\Phi_{k,\beta}(\xi) = \frac{k}{\beta}\left(\left(1+\frac{\xi}{k}\right)^{\beta} - 1\right) \quad \textit{and} \quad k = h^{\frac{2\beta-1}{1-\beta}}. \tag{3.4}$$

*Then*

$$\frac{1}{2}\frac{d}{dt}\mathbf{E}_v^h(t) = -\sum_{j=0}^{m} h(\dot{v}_j)^2. \tag{3.5}$$

PROOF: In complete analogy with the proof of Proposition 2.6, we compute

$$\begin{aligned}\frac{d}{dt}\mathbf{E}_v^h(t) &= \sum_{j=0}^{m-1} 2hR(D^+v_j)D^+\dot{v}_j \\ &= -\sum_{j=1}^{m-1} 2hD^-(R(D^+v_j))\dot{v}_j - 2R(D^+v_0)\dot{v}_0 + 2R(D^+v_{m-1})\dot{v}_m \\ &= \sum_{j=0}^{m} 2h(\dot{v}_j)^2 \quad \text{since } R(\xi) = \xi\Phi'_{k,\beta}(\xi^2) \text{ and } \dot{v}_j = D^-(R(D^+v_j)),\end{aligned}$$

which is what we wanted. □

*Remark.* Assumptions on the initial data $\phi^h$ imply that $\sup_{h>0} \mathbf{E}_{v^h}^h(0) < \infty$.

Next we obtain Hölder continuity in time and the $L^2$-bound for difference quotients still in the $\beta \in (0, \frac{1}{2})$ case.

COROLLARY 3.7 (Hölder Continuity, $\beta \in (0, \frac{1}{2})$) *Let* $\{v_j(t)\}_{j=1}^m$ *be as in Proposition* 3.6. *Then*

$$v \in C^{1/2}\big([0,\infty); L^2(G_h)\big) \quad \textit{and} \quad v \in C^{1/4}\big([0,\infty); L^\infty(G_h)\big).$$

*Furthermore, the Hölder constants involved depend only on the energy and jump locations of the initial condition.*

PROOF: Take $0 \le t_1 \le t_2$. By the Hölder inequality,

$$\begin{aligned}\sum_{j=0}^{m} h|v_j(t_2) - v_j(t_1)|^2 &= \sum_{j=0}^{m} h\left(\int_{t_1}^{t_2} \dot{v}_j\, dt\right)^2 \\ &\le |t_2 - t_1| \sum_{j=0}^{m} h \int_{t_1}^{t_2} (\dot{v}_j)^2\, dt \\ &= |t_2 - t_1| \int_{t_1}^{t_2} -\frac{1}{2}\frac{d}{dt}\mathbf{E}_v^h(t)dt \\ &= |t_2 - t_1|\frac{1}{2}\big(\mathbf{E}_v^h(t_1) - \mathbf{E}_v^h(t_2)\big) \le |t_2 - t_1|\mathbf{E}_v^h(t_1)\end{aligned}$$

which is the definition of $C^{1/2}$ Hölder continuity in time with values in $L^2(G_h)$. Just as in the continuum case (i.e., as in Corollary 2.7) Hölder continuity with values in $L^\infty(G_h)$ now follows from an interpolation lemma (discrete analogue of Lemma 4.1) once we notice that

$$\sup_{|\xi| \le z(k)} \frac{\xi^2}{\Phi_{k,\beta}(\xi^2)} := C(\beta) < \infty$$

and therefore

$$\sum_{j \notin S(v)} h|D^+ v_j(t)|^2 \le C(\beta)\mathbf{E}_v^h(t)\,. \tag{3.6}$$

That concludes the proof. □

*Note.* The whole point of Corollary 3.7 is that the Hölder constants do not depend on discretization size $h$, provided that the approximations to the initial condition remain bounded in energy as $h \to 0^+$.

Turning now to the more singular case of $\beta = 0$, we have first:

PROPOSITION 3.8 ($L^p$-bound for Difference Quotients) *Let $\{v_j(t)\}_{j=0}^m$ be generated from the initial condition $\{\phi_j\}_{j=0}^m$ by the scheme given in* (3.1) *with the constitutive function $g_k(\xi) = 1/(1+\xi/k)$ where the parameter $k$ is subject to the scaling $k = 1/h$ as usual. Then for $p \in (1, 2)$,*

$$\sup_{t \ge 0} \sum_{j \notin S(v(t))} h|D^+ v_j|^p \le \sum_{j \notin S(\phi)} h|D^+ \phi_j|^p + C(p) < \infty\,.$$

*Furthermore, the constant $C(p)$ depends only on the initial jump heights $J_k(0)$ in addition to $p$.*

PROOF: For convenience, we will sum from $j = -1$ to $j = m$ with the understanding that $D^+ v_1 = D^+ v_m = 0$; hence $j \notin S(v)$ means $j \in \{-1, 0, 1, \dots, m\} -$

$S(v)$. The calculation is completely analogous to the one in the proof of Proposition 2.5. Fix $\varepsilon > 0$ and let $F_\varepsilon(x) = (x^2 + \varepsilon^2)^{p/2}$. We compute

$$\frac{d}{dt} \sum_{j \notin S(v)} h((D^+ v_j)^2 + \varepsilon^2)^{p/2} = \sum_{j \notin S(v)} hF'(D^+ v_j) D^+ \dot{v}_j = \mathrm{I} + \mathrm{II} + \mathrm{III} \tag{3.7}$$

where

$$\mathrm{I} = -\sum_{j=0}^{l_1 - 1} hD^- F'(D^+ v_j)\dot{v}_j - \sum_{k=1}^{n} \sum_{j=l_k+2}^{l_{k+1}-1} hD^- F'(D^+ v_j)\dot{v}_j - \sum_{j=l_n+2}^{m} hD^- F'(D^+ v_j)\dot{v}_j \,,$$

$$\mathrm{II} = \sum_{k=1}^{n} \left( \dot{v}_{l_k} F'(D^+ v_{l_k - 1}) - \dot{v}_{l_k+1} F'(D^+ v_{l_k+1}) \right),$$

$$\mathrm{III} = -\dot{v}_{-1} F'(D^+ v_{-1}) + \dot{v}_{m+1} F'(D^+ u_m)\,.$$

We start with term I. We have

$$D^- F'(D^+ v_j)\dot{v}_j = F''(\xi_j) R'(\eta_j)(D^- D^+ v_j)^2$$

where $\xi_j$ and $\eta_j$ are between $D^+ v_{j-1}$ and $D^+ v_j$. Whenever $j \notin S(v) \cup \{S(v)+1\}$, we have that $|D^+ v_{j-1}|, |D^+ v_j| \in (-z(h), z(h))$; therefore for such $j$ we also have $R'(\eta_j) > 0$. Since $F''(\xi) > 0$ for all $\xi$, we see at once that $\mathrm{I} < 0$.

Term III is easily seen to be 0 by the remarks made at the beginning of the proof concerning $D^+ v_{-1}$ and $D^+ v_m$.

Turning our attention to term II, we recall the definition of $R_*^{-1}(x)$ given in (1.8). We make the observation that for any increasing function $f(x)$ and $a, b \in \mathbb{R}$ we have

$$F'(f(ha+b))a \geq F'(f(b))a\,. \tag{3.8}$$

Apply (3.8) with $a = \dot{v}_{l_k}$, $b = R(D^+ v_{l_k - 1})$, and $f(x) = R_*^{-1}(x)$. Noting that

$$ha + b = R(D^+ v_{l_k}) \quad \text{and} \quad R_*^{-1}(R(D^+ v_{l_k})) = \frac{1}{v_{l_k+1} - v_{l_k}} = \frac{1}{J_k(t)}\,,$$

we get

$$\dot{v}_{l_k} F'(D^+ v_{l_k - 1}) \leq F'\left(\frac{1}{J_k(t)}\right)\dot{v}_{l_k}\,.$$

Then we apply (3.8), this time with $a = -\dot{v}_{l_k+1}$ and $b = R(D^+ v_{l_k+1})$. Noting that we again have $ha + b = R(D^+ v_{l_k})$, we get in this case

$$\dot{v}_{l_k+1} F'(D^+ v_{l_k+1}) \geq F'\left(\frac{1}{J_k(t)}\right)\dot{v}_{l_k+1}\,.$$

That means

$$\mathrm{II} \le -\sum_{k=1}^{n} F'\left(\frac{1}{J_k(t)}\right)\frac{d}{dt}J_k(t) \tag{3.9}$$

so that

$$\frac{d}{dt}\sum_{j\notin S(v)} h\big((D^+v_j)^2+\varepsilon^2\big)^{p/2} \le -\sum_{k=1}^{n} F'\left(\frac{1}{J_k(t)}\right)\frac{d}{dt}J_k(t)\,.$$

Integrating this inequality in $t$ over $[t_1, t_2]$, we find

$$\sum_{j\notin S(v)} h\big((D^+v_j)^2+\varepsilon^2\big)^{p/2}\Big|_{t_2} \le \sum_{j\notin S(v)} h\big((D^+v_j)^2+\varepsilon^2\big)^{p/2}\Big|_{t_1} - \sum_{k=1}^{n}\int_{t_1}^{t_2} F'\left(\frac{1}{J_k(t)}\right)\frac{d}{dt}J_k(t)dt\,.$$

After making the change of variables $y = J_k(t)$ and sending $\varepsilon \to 0^+$, this expression becomes

$$\sum_{j\notin S(v)} h|D^+v_j(t_2)|^p \le \sum_{j\notin S(v)} h|D^+v_j(t_1)|^p - \sum_{k=1}^{n}\int_{J_k(t_1)}^{J_k(t_2)} \frac{dy}{|y|^{(p-2)/2}y}\,.$$

The integrand on the right-hand side is integrable over any bounded interval of time provided that $p \in (1,2)$. That proves the proposition. ☐

COROLLARY 3.9 (Hölder Continuity, $\beta = 0$) *Let $\{v_j(t)\}_{j=0}^m$ be the solution generated by scheme* (3.1)*, this time with $\beta = 0$. Assume that the jump set $S(v)$ is constant over the interval of time $[T_1, T_2]$. Let $\hat{G} = \{1, 2, \dots, m\} - S(v(T_1)) \cup \{S(v(T_1)) + 1\}$. Then for any $\nu \in [0, \frac{1}{2})$ and $\mu \in [0, \frac{1}{4})$ we have*

$$v \in C^\nu\big([T_1, T_2]; L^2\big(\{x_j : j \in \hat{G}\}\big)\big) \quad \textit{and} \quad v \in C^\mu\big([T_1, T_2]; L^\infty\big(\{x_j : j \in \hat{G}\}\big)\big)\,.$$

*Moreover, the Hölder constants involved can be bounded by the discrete $W^{1,p}(\hat{G})$-norm and jump heights and locations of the initial data.*

PROOF: Integrate equation (3.7) in time over $[t_1, t_2] \subseteq [T_1, T_2]$. Using the notation in Proposition 3.8, we get

$$\begin{aligned}\int_{t_1}^{t_2} \mathrm{I}\,dt &= -\int_{t_1}^{t_2}\sum_{j\in\hat{G}} hD^-F'(D^+v_j)\dot{v}_j\,dt \\ &= \sum_{j\notin S(v)} hF(D^+v_j(t))\Big|_{t_1}^{t_2} - \int_{t_1}^{t_2}\mathrm{II}\,dt\,.\end{aligned}$$

Recalling estimate (3.9) for the term II, the equation above turns into

$$\int_{t_1}^{t_2} \sum_{j\in\hat{G}} hD^- F'(D^+ v_j)\dot{v}_j\,dt \leq \sum_{j\notin S(v)} hF(D^+ v_j)\Big|_{t_1} - \sum_{k=1}^{n}\int_{t_1}^{t_2} F'\left(\frac{1}{J_k}\right)\frac{d}{dt}J_k(t)dt\,.$$

Notice that

$$D^- F'(D^+ v_j)\dot{v}_j = F''(\xi_j)(D^- D^+ v_j)\dot{v}_j \geq F''(\xi_j)(\dot{v}_j)^2$$

where $\xi_j$ is between $D^+ v_{j-1}$ and $D^+ v_j$; this is because $\dot{v}_j = R'(\eta_j)D^- D^+ v_j$ with $\eta_j$ between $D^+ v_{j-1}$ and $D^+ v_j$, and because $0 \leq R'(\eta_j) \leq 1$ for $|\eta_j| \leq z(h)$. Consequently, we obtain the inequality

$$\int_{t_1}^{t_2} \sum_{j\in\hat{G}} hF''(\xi_j)(\dot{v}_j)^2\,dt \leq \sum_{j\notin S(v)} hF(D^+ v_j(t_1)) - \sum_{k=1}^{n}\int_{t_1}^{t_2} F'\left(\frac{1}{J_k(t)}\right)\frac{d}{dt}J_k(t)dt\,.$$

Here the right-hand side can be bounded in terms of the $W^{1,p}(\hat{G})$-norm and the jump heights of $v$ at $t = T_1$. Once we note the trivial fact $|\xi_j| \leq |D^+ v_{j-1}| + |D^+ v_j|$, it is possible to proceed exactly as we did in Corollary 2.7, relying on Proposition 3.8 when we need $L^p$-bounds on difference quotients. One gets, in particular,

$$\int_{t_1}^{t_2} \sum_{j\in\hat{G}} h|\dot{v}_j|^p\,dt \leq C|t_2 - t_1|^{\frac{2-p}{2}} \tag{3.10}$$

where the constant depends only on the $W^{1,p}(\hat{G})$-norm of the initial condition and its jump heights, as it should. From here onwards, the argument is again the same as that of Corollary 2.7, using the discrete analogue of Lemma 4.1. □

*Remark.* Corollary 3.9 allows us to get uniform-in-time estimates on $\|v(t_2) - v(t_1)\|_{L^\infty}$ on the entire grid $G_h$, and not just on $\{x_j : j \in \hat{G}\}$. Indeed, the Hölder continuity result of the corollary allows us to estimate the contribution to $\|v(t_2) - v(t_1)\|_{L^2}$ from the smaller grid. But the contribution to this norm from $G_h - \{x_j : j \in \hat{G}\}$ is order $h$ by virtue of maximum principles. Therefore, by taking $h$ small enough, $\|v(t_2) - v(t_1)\|_{L^2}$ can be estimated on the full grid $G_h$. Then, by interpolation, we can turn that into an estimate of $\|v(t_2) - v(t_1)\|_{L^\infty}$ on $G_h$.

The following technical lemma gives us an $L^\infty$-bound on the difference quotients generated by the numerical scheme; the bound depends only on the energy of the initial data.

LEMMA 3.10 *Given initial data $\phi$ and a sequence $\{\phi^h\}$ of numerical approximations subject to the usual assumptions, for any large enough $N$ there exists a* $\delta = \delta(N) > 0$ *with the following property*:

*For any* $t \geq 0$ *there exists* $t_0 \in [t, t+\delta]$ *such that*

$$\sup_{j \notin S(v^h(t_0))} |D^+ v_j^h(t_0)| < N$$

*where* $v^h(t)$ *is the solution generated by scheme* (3.1). *Furthermore, it can be arranged that* $\delta(N) \to 0^+$ *as* $N \to \infty$.

PROOF: We concentrate first on the case $\beta > 0$. By the energy identity (3.5) we have

$$\sum_j h(D^- R(D^+ v_j))^2 = -\frac{1}{2}\frac{d}{dt}\mathbf{E}_v^h(0)\,.$$

Integration of this equality in $t$ over $[t, t+\delta]$ gives

$$\int_t^{t+\delta} \sum_j h(D^- R(D^+ v_j))^2\, dt \leq \mathbf{E}_v^h(0)\,. \tag{3.11}$$

Therefore, there exists $t_0 \in [t, t+\delta]$ such that

$$\sum_{j=0}^{m} h(D^- R(D^+ v_j))^2 \leq \frac{\mathbf{E}_v^h(0)}{\delta}\,. \tag{3.12}$$

Second, if we combine inequality (3.6) with the fact that $|R(\xi)| \leq |\xi|$ for all $|\xi|$, we find

$$\sum_{j \notin S(v(t))} hR(D^+ v_j)^2 \leq C(\beta)\mathbf{E}_v^h(0)\,. \tag{3.13}$$

Now apply the discrete analogue of the interpolation inequality (Lemma 4.1) on the domain $\{x_j : j \notin S(v(t_0))\}$ with $f = R(D^+ v_j)$ and $\theta = p = 2$. Estimates (3.12) and (3.13) yield

$$\sup_{j \notin S(v(t_0))} |R(D^+ v_j(t_0))| \leq C_1\sqrt{\mathbf{E}_v^h(0)} + \frac{C_2}{\delta^{1/4}}\sqrt{\mathbf{E}_v^h(0)}\,.$$

But if $|\xi| < z(h)$, then $|\xi| \leq \theta^{-1}(\beta)|R(\xi)|$ where $\theta(\beta) > 0$ (see (1.9)). With that, we get the same inequality as the last one, this time for $D^+ v_j$:

$$\sup_{j \notin S(v(t_0))} |D^+ v_j(t_0)| \leq C_1\sqrt{\mathbf{E}_v^h(0)} + \frac{C_2}{\delta^{1/4}}\sqrt{\mathbf{E}_v^h(0)}$$

but with different constants $C_1$ and $C_2$. That implies the conclusion of the lemma for the case $\beta \in (0, \frac{1}{2})$. For the case $\beta = 0$, we make the following modification: Inequality (3.10) gives

$$\int_t^{t+\delta} \sum_{j \in \hat{G}} h|\dot{v}_j|^p\, dt = \int_t^{t+\delta} \sum_{j \in \hat{G}} h|D^- R(D^+ v_j)|^p\, dt \leq C|\delta|^{\frac{2-p}{2}}\,.$$

Once we replace inequality (3.11) with the one above, the same argument carries through. □

### 3.3 Convergence

Let $u$ be the solution to the continuum problem described by the PDE system (1.12) and the prescription given in Section 2.4, with initial data $\phi$. Let $S(\phi) = \{p_1, p_2, \ldots, p_n\}$ be ordered, and let $J_k(t) = u(p_k^+, t) - u(p_k^-, t)$ be the associated jump heights.

We assume throughout that we have access to a sequence $\phi_h : G_h \to \mathbb{R}$ of discrete approximations to $\phi$ that satisfy the assumptions on approximate initial data listed in Section 3.1.

The purpose of this section is to prove, under suitable conditions, the convergence of the discrete solutions $v^h(t) : G_h \times [0, \infty) \to \mathbb{R}$ generated by scheme (3.1) from the initial conditions $\phi^h$ to the proposed continuum limit.

Our proof has two components: a convergence argument, with a rate, that is valid on any interval of time during which the jump heights of the continuum solution are bounded away from 0, and an argument that controls the behavior of the numerical and continuum solutions once a jump height becomes smaller than a critical value (denoted $\delta_J$ in what follows). The essential property of both the continuum and the discrete solutions that makes this possible is the following: Once the jump height at a discontinuity point becomes small enough, quenching is imminent. Moreover, how much the solutions get modified in such a small interval of time can be controlled.

PROPOSITION 3.11 *Let $T > 0$ be such that*

$$\delta := \min_k \inf_{t \in [0,T]} |J_k(t)| > 0$$

*so that the jump set $S(u)$ of $u$ does not change for $t \in [0, T]$ and consequently the evolution of $u$ is described by* (1.12)*. Let*

$$M := \sum_{j=1}^{n+1} \sup_{t \in [0,T]} \|u\|_{C^3([p_{j-1}, p_j])} < \infty .$$

*Then*

$$\lim_{h \to 0} \sup_{t \in [0,T]} \sup_{j=0,1,\ldots,m} |u_j(t) - v_j(t)| = 0 .$$

The proof of Proposition 3.11 will of course need to concern itself with showing the compatibility of jump sets $S(u(t))$ and $S(v(t))$; note, however, that we already know by virtue of Proposition 3.2 that $S(v)$ is at most a decreasing function of time, by hypothesis that $S(u(t))$ is constant, and by assumptions on initial data that $S(u(0))$ and $S(v(0))$ agree. We will need the following lemma, which makes a *consistency* statement:

LEMMA 3.12 *Let $E_j = u_{xx}(x_j) - D^- R(D^+ u_j)$. Then*

$$|E_j| \leq \begin{cases} CMh^{(1-2\beta)/(1-\beta)} & \text{for } j \notin S(v) \cup \{S(v) + 1\} \\ CM & \text{otherwise} \end{cases}$$

*where $C$ is a constant that is independent of $h$ and $j$.*

PROOF: For $j \notin S(v) \cup \{S(v)+1\}$, the segment $[x_{j-1}, x_{j+1}]$ lies in one of the intervals on which $u$ is $C^3$. Therefore, for such $j$ we have

$$D^- D^+ u_j = u_{xx}(x_j) + O(h)\,.$$

Also, we have $R'(\xi) = 1 + O(h^{(1-2\beta)/(1-\beta)})$ for $|\xi| \le M$ and $D^- R(D^+ u_j) = R'(\xi_j) D^- D^+ u_j$. Hence we get, as claimed,

$$D^- R(D^+ u_j) = u_{xx}(x_j) + O\big(h^{(1-2\beta)/(1-\beta)}\big)\,.$$

For $j \in S(v) \cup \{S(v)+1\}$, we shall only consider $j \in S(v)$ since the case of $j \in \{S(v)+1\}$ is completely analogous. First note that there exists a $p_k \in S(u)$ such that $p_k \in [x_j, x_{j+1}]$ and $u_{j+1} - u_j = J_k(t) + O(h)$. The function $R_{\beta,h}(x/h)$ is Lipschitz in $x$, uniformly for $h > 0$; in other words, for any $a, b \in \mathbb{R}$ with $|a|, |b| \ge \delta > 0$ and $h > 0$ sufficiently small,

$$\left| R_{\beta,h}\left(\frac{a}{h}\right) - R_{\beta,h}\left(\frac{b}{h}\right) \right| \le C(\beta,\delta)|b-a| \tag{3.14}$$

where $C(\beta,\delta) = C(\beta)\delta^{2\beta-2}$. Furthermore, for $J$ bounded away from 0, we have $R(J/h) = J|J|^{2\beta-2} + O(h^{1/(1-\beta)})$. Applying (3.14) with $a = u_{j+1} - u_j$ and $b = J_k$ we find

$$\begin{aligned} R(D^+ u_j) &= R\left(\frac{J_k(t)}{h}\right) + O(h) && \text{(since } h(D^+ u_j) - J_k(t) = O(h)) \\ &= J_k(t)|J_k(t)|^{2\beta-2} + O(h) && \text{if } J_k(t) \text{ is bounded away from } 0 \\ &= u_x(x_j, t) + O(h) \end{aligned}$$

where we employed the boundary condition $u_x(p_k, t) = J_k(t)|J_k(t)|^{2\beta-2}$ and also the trivial fact $u_x(x_j, t) = u_x(p_k, t) + O(h)$ at the last step. We thus get

$$R(D^+ u_j) = R(D^+ u_{j-1}) + O(h)\,.$$

Since $u_{xx}$ is bounded, we are done. □

PROOF OF PROPOSITION 3.11: Consider $h > 0$ small enough so that $z(h) > 2M$. Let $\hat{S} = S(v) \cup \{S(v)+1\}$ and set

$$I(t) := \sum_{j=0}^{m} h(u_j - v_j)^2\,.$$

We compute

$$\begin{aligned}\frac{1}{2}\frac{d}{dt}&I(t)\\ &= \sum_{j=0}^{m} h(u_j - v_j)\big(D^- R(D^+ u_j) - D^- R(D^+ v_j) + u_{xx}(x_j) - D^- R(D^+ u_j)\big)\\ &= \mathrm{A} + \mathrm{B} + \mathrm{C}\end{aligned}$$

where

$$\mathrm{A} = -\sum_{j \notin S(v)} hD^+(u_j - v_j)\big(R(D^+ u_j) - R(D^+ v_j)\big),$$

$$\mathrm{B} = -\sum_{k=1}^{n} \big((u_{l_k+1} - v_{l_k+1}) - (u_{l_k} - v_{l_k})\big)\big(R(D^+ u_{l_k}) - R(D^+ v_{l_k})\big),$$

$$\mathrm{C} = \sum_{j=0}^{m} h(u_j - v_j)E_j\,.$$

We first examine term A. To that end, note that $j \notin S(v)$ implies

$$|D^+ v_j| < z(h) \quad \text{and} \quad |D^+ u_j| \le \frac{z(h)}{2}$$

and that means

$$(R(D^+ u_j) - R(D^+ v_j))(D^+ u_j - D^+ v_j) \ge \theta (D^+ u_j - D^+ v_j)^2$$

for some $\theta > 0$ by (1.9). From this observation we get

$$\mathrm{A} \le -\theta \sum_{j \notin S(v)} h(D^+(u_j - v_j))^2\,. \tag{3.15}$$

To estimate the second term, we can use the trace theorem on each of the summands; let us call them $\mathrm{B}_k$ so that $\mathrm{B} = \sum \mathrm{B}_k$. We have

$$\begin{aligned}\big|R(D^+ u_{l_k}) - R(D^+ v_{l_k})\big| &\le \frac{C(\beta)}{\delta^{2-2\beta}}\big|(u_{l_k+1} - u_{l_k}) - (v_{l_k+1} - v_{l_k})\big|\\ &\le \frac{C(\beta)}{\delta^{2-2\beta}}\big(|u_{l_k+1} - v_{l_k+1}| + |u_{l_k} - v_{l_k}|\big)\end{aligned}$$

where we used (3.14) again. That leads to

$$|\mathrm{B}_k| \le \frac{C(\beta)}{\delta^{2-2\beta}}\big(|u_{l_k+1} - v_{l_k+1}|^2 + |u_{l_k} - v_{l_k}|^2\big)\,.$$

By the discrete trace theorem we have

$$|\mathrm{B}| \le \varepsilon \sum_{j \notin S(v)} h(D^+(u_j - v_j))^2 + C(\varepsilon)\sum_{j=0}^{m} h(u_j - v_j)^2\,. \tag{3.16}$$

If we choose $\varepsilon \in (0, \theta)$, we can absorb the first term on the right-hand side into A.

Turning our attention to term C, we first write it in the form

$$\mathrm{C} = -\sum_{j\in\hat{S}} h(u_j - v_j)E_j - \sum_{j\notin\hat{S}} h(u_j - v_j)E_j \,.$$

Then by virtue of Lemma 3.12 we can estimate

$$\begin{aligned}\left|\sum_{j\notin\hat{S}} h(u_j - v_j)E_j\right| &\le \sum_{j=0}^{m} h(u_j - v_j)^2 + \sum_{j\notin\hat{S}} hE_j^2 \\ &\le I(t) + C(M)h^{\frac{2-4\beta}{1-\beta}} \\ \left|\sum_{j\in\hat{S}} h(u_j - v_j)E_j\right| &\le 2nC(M)Nh\end{aligned} \tag{3.17}$$

where $N = \sup_{j,t} |u_j - v_j|$, which is finite by Propositions 2.4 and 3.1 (maximum principles). Putting together our estimates (3.15), (3.16), and (3.17), we end up with a differential inequality for $I(t)$ of the following kind:

$$I'(t) \le c_1 I(t) + c_2 h^{\mu} \tag{3.18}$$

where $\mu := \min\{1, (2-4\beta)/(1-\beta)\}$. Integrating (3.18), we find that

$$I(t) \le (I(0) + c_2 h^{\mu} t)e^{c_1 t} \,. \tag{3.19}$$

This result bounds the $L^2$-norm of $u - v$ under the assumption that the jump sets of $u$ and $v$ remain compatible (needed in Lemma 3.12). This condition, although true initially by hypothesis, cannot be verified for any definite interval of time by an $L^2$-estimate on $u-v$, since jump set information is not stable under perturbations small in $L^2$. We therefore improve (3.19) to an $L^\infty$-estimate, which is sufficient to ensure compatibility of jump sets. This is accomplished once again via interpolation. Indeed, letting $f_j = u_j - v_j$, for $\beta = 0$ Proposition 3.8 and for $\beta \in (0, \frac{1}{2})$ inequality (3.6) imply $D^+ f_j$ is bounded in the $L^p$-norm for $p \in (1, 2)$. Applying the discrete analogue of Lemma 4.1 with $p = q = r = \frac{3}{2}$, we obtain convergence in $L^\infty$, and that completes the proof of the proposition. □

THEOREM 3.13 *Let $T_1, T_2, \dots, T_n$ be the quenching times, in order, of the proposed limit with each $T_i$ distinct. Given $T \ge 0$, $\varepsilon > 0$, and $\delta > 0$ there exists $h_* > 0$ such that for all $h \in (0, h_*)$ we have*

(i) $\sup_{t\in[0,T]} \sup_{0\le j\le m(h)} |u(x_j, t) - v_j^h(t)| < \varepsilon$ *and*
(ii) $S(u(t))$ *and* $S(v^h(t))$ *are compatible for* $t \in [0, T] - \bigcup [T_i - \delta, T_i + \delta]$.

PROOF: It is sufficient to prove the claim for the first quenching time, which we assume takes place at $x = p_k$. The general statement then follows by induction. More precisely, the following statement will be proved: For all $\varepsilon_* > 0$ and $\delta_* > 0$, there exists an $h_* > 0$ such that if $h \in (0, h_*)$, then

(1) $\sup_{0\le t\le t_*} \sup_{0\le j\le m(h)} |u(x_j, t) - v_j^h(t)| < \varepsilon_*$ and

(2) $S(u(t))$ and $S(v^h(t))$ are compatible for $t \in [0, T_1 - \delta_*] \cup \{t_*\}$ for some $t_* \in (T_1, T_1 + \delta_*]$.

The proof relies on two parameters involving the proposed limit, namely,

$$m := \min_{i \neq k} \inf_{0 \leq t \leq T_1} |J_i(t)| \quad \text{and} \quad \delta_t := \min_{i=2,3,\ldots,n} T_i - T_{i-1} .$$

By hypothesis, $m > 0$ and $\delta_t > 0$.

*Step* 1. The Hölder continuity properties expressed in Propositions 2.7 and 3.7 give uniform-in-time estimates on how fast the continuum and discrete solutions change. As a consequence, for any given $\varepsilon > 0$ we can choose a $\delta \in (0, T_1)$ small enough so that

$$\begin{aligned} \sup_{i=1,2,\ldots,n} \|u_i(\cdot, t_1) - u_i(\cdot, t_2)\|_{L^\infty} &\leq \varepsilon \quad \text{if } |t_1 - t_2| \leq \delta , \\ \sup_{h>0} \|v_j^h(t_1) - v_j^h(t_2)\|_{L^\infty} &\leq \varepsilon \quad \text{if } |t_1 - t_2| \leq \delta . \end{aligned} \tag{3.20}$$

Since it is enough to prove the claim for only sufficiently small $\delta_*$, we can thus assume that (3.20) is satisfied with $\varepsilon = \min\{m/8, \varepsilon_*/4\}$ and $\delta = 2\delta_*$.

*Step* 2. Regularity properties of the proposed limit yield bounds on the spatial derivatives for positive time; we state it in two forms:

$$\sup_{1 \leq i \leq n} \|u_i(\cdot, \delta)\|_{C^3} \leq N(\delta) \quad \text{and} \quad \sup_{1 \leq i \leq n} \|\partial_x u_i(\cdot, \delta_t/2)\|_{L^\infty} \leq \frac{C}{\delta_t}$$

where the constant $C$ depends only on the $W^{1,p}$-norm of the initial condition (for the appropriate choice of $p > 1$). The constant $N$ may depend on $\delta$, as indicated ($\delta$ will be fixed in step 4). Furthermore, by Lemma 3.10, for all $h > 0$ there exists a $\delta_h \in [0, \delta_t/2]$ such that

$$\sup_{j \notin S(v^h)} |D^+ v_j^h(\delta_h)| \leq \frac{2C}{\delta_t} . \tag{3.21}$$

*Step* 3. Now, $m$ and $\delta_t$ determine a critical jump height $\delta_J$: For any $\delta_J > 0$ satisfying

$$\delta_J^{2\beta-1} \geq \frac{4C}{\delta_t} \quad \text{and} \quad \delta_J < \frac{m}{2} ,$$

Lemma 2.10 implies that if $|J_k(t_0)| \leq \delta_J$ for some $t_0 \in [\delta_t/2, T_1]$, then $|J_k(t_0)|$ is monotone decreasing on $[t_0, T_1]$. Moreover, according to Lemma 2.11, if we choose $\delta_J$ to be small enough so that

$$\frac{2K(p_k - p_{k-1})}{\delta_J^{2\beta-1} - (m/2)^{2\beta-1}} < \delta_*$$

where $K = \sup_{i,x} |u_i(x, 0)|$, then

$$|T_1 - t_0| < \delta_* ,$$

which is an upper bound on the quenching time of the proposed limit in terms of the jump height at the $k^{\text{th}}$ discontinuity. Let

$$t_0 := \inf\left\{t \geq 0 : |J_k(t)| < \frac{\delta_J}{2}\right\}.$$

*Step* 4. From now on let $\varepsilon := \min\{\varepsilon_*/20, \delta_J/20\}$. There exists a $\delta > 0$ such that (3.20) is satisfied for this choice of $\varepsilon$. Take a discrete initial condition $v^h(0)$ with $\sup_j |u(x_j, 0) - v_j^h(0)| \leq \varepsilon$. Then, $\sup_j |u(x_j, \delta) - v_j^h(\delta)| \leq 3\varepsilon = \min\{3\varepsilon_*/20, 3\delta_J/20\}$.

*Step* 5. The proposed limit is (piecewise) $C^3$ for $t \in [\delta, t_0]$, and its jump heights are bounded away from 0 on this interval of time. We can therefore apply the convergence argument Proposition 3.11 on $[\delta, t_0]$: For all $h > 0$ small enough (depending on $\varepsilon$), we have

$$\sup_{0\leq j\leq m(h)} \sup_{0\leq t\leq t_0} |u(x_j, t) - v_j^h(t)| < 4\varepsilon = \min\left\{\frac{\varepsilon_*}{5}, \frac{\delta_J}{5}\right\}. \tag{3.22}$$

In particular,

$$\frac{1}{10}\delta_J \leq |J_k^h(t_0)| \leq \frac{9}{10}\delta_J \quad \text{and} \quad \min_{i\neq k} \inf_{t\in[0,t_0]} |J_i^h(t)| \geq \frac{4}{5}m\,. \tag{3.23}$$

As a consequence, $S(u(t))$ and $S(v^h(t))$ are compatible for $t \in [0, t_0]$. In other words, neither the proposed limit nor the discrete solutions quench during this interval of time.

*Step* 6. The choice of $\delta_*$ made in step 1 implies that the inequalities in (3.23) and (3.22) can be improved to

$$\sup_{0\leq j\leq m(h)} \sup_{0\leq t\leq t_0+\delta_*} |u(x_j, t) - v_j^h(t)| < \varepsilon_* \tag{3.24}$$

and

$$\min_{i\neq k} \inf_{t\in[0,t_0+2\delta_*]} |J_i^h(t)| \geq \frac{m}{2}\,. \tag{3.25}$$

Inequalities (3.21), (3.23), and (3.25) now allow us to apply Lemma 3.4: The discrete jump height $|J_k^h(t)|$ is decreasing for $t \geq t_0$. We can therefore apply Lemma 3.5 on $[t_0, \min\{t_0 + \delta_*, T_1^h\}]$. Since

$$\frac{(l_k - l_{k-1})h}{R_h(\frac{9}{10h}\delta_J) - R_h(\frac{1}{2h}m)} < \frac{(p_k - p_{k-1})}{\delta_J^{2\beta-1} - (\frac{m}{2})^{2\beta-1}}$$

for $h$ small enough, and given the choice of $\delta_J$ in step 3, we conclude that $T_1^h < t_0 + \delta_*$. Moreover, $S(u(t))$ and $S(v^h(t))$ are constant, again by the choice of $\delta_*$, for $t \in [\max\{T_1, T_1^h\}, t_0 + \delta_*]$. Hence, $S(u(t))$ and $S(v^h(t))$ are compatible on $[0, t_0] \cup \{t_0 + \delta_*\}$. That proves the second assertion of the claim with $t_* = t_0 + \delta_*$, and inequality (3.24) proves the first assertion. □

## 4 Technical Lemmas

LEMMA 4.1 (Interpolation Lemma) *Let $f(x) \in C^1([a,b])$. Then*

$$\|f\|_{L^\infty} \leq C\left(\|f\|_{L^r} + \|f\|_{L^p}^{\theta} \|f'\|_{L^q}^{1-\theta}\right)$$

*where $p, q \geq 1$, $\theta \in (0,1)$, $\theta/p + (1-\theta)(1-q)/q = 0$, and $r \geq 1$. The constant $C$ depends only on the size of the interval and the choice of the exponents.*

PROOF: Let $\overline{f} := \|f\|_{L^1}/(b-a)$. By continuity, there is $x_0 \in [a,b]$ such that $f(x_0) = \overline{f}$. For any $\theta \in (0,1)$,

$$\begin{aligned} |f(x)|^{\frac{1}{1-\theta}} - |\overline{f}|^{\frac{1}{1-\theta}} &= |f(x)|^{\frac{1}{1-\theta}} - |f(x_0)|^{\frac{1}{1-\theta}} \\ &= \int_{x_0}^{x} \left(|f(s)|^{\frac{1}{1-\theta}}\right)' ds \\ &\leq C \int_{x_0}^{x} |f(s)|^{\frac{\theta}{1-\theta}} |f'(s)|\, ds \\ &\leq C\left(\int |f(s)|^{\frac{\theta q}{(1-\theta)(q-1)}}\, ds\right)^{\frac{q-1}{q}} \left(\int |f'(s)|^q\, ds\right)^{\frac{1}{q}} \end{aligned}$$

where we applied the Hölder inequality in the last step. Let $p := \theta q/((1-\theta)(q-1))$. Then the inequality above reads

$$\|f\|_{L^\infty}^{1/(1-\theta)} \leq C\left(|\overline{f}|^{1/(1-\theta)} + \|f\|_{L^p}^{\theta/(1-\theta)} \|f'\|_{L^q}\right),$$

which means

$$\|f\|_{L^\infty} \leq C\left(|\overline{f}| + \|f\|_{L^p}^{\theta} \|f'\|_{L^q}^{1-\theta}\right) = C\left(\|f\|_{L^1} + \|f\|_{L^p}^{\theta} \|f'\|_{L^q}^{1-\theta}\right).$$

Since we are on a bounded interval, $\|f\|_{L^1}$ is controlled by $\|f\|_{L^r}$ for any $r \geq 1$. □

Lemma 4.1 has an obvious discrete analogue.

LEMMA 4.2 (Parabolic Hopf Lemma) *Let $E = (p,q) \times (0,T)$, and for $t_0 \in (0,T)$ let $E_{t_0} := \{(x,t) \in E : t \leq t_0\}$. Suppose $u$ satisfies the uniformly parabolic inequality $au_{xx} - u_t \geq 0$ where $a$ is bounded. Suppose that $u$ is continuously differentiable at the boundary point $(x_0, t_0) \in \{p, q\} \times (0,T)$, that $u(x_0, t_0) = M$, and that $u(x,t) < M$ for all $(x,t) \in E_{t_0}$. If $\partial_\nu$ denotes any outward directional derivative from $E_{t_0}$ at $(x_0, t_0)$, then $\partial_\nu u > 0$ at $(x_0, t_0)$.*

PROOF: See [18] for the proof of a more general statement. □

**Acknowledgment.** The author thanks his advisor Robert V. Kohn, who suggested the approach pursued in this paper and provided guidance throughout its development.

# Bibliography


[1] Alvarez, L.; Lions, P.-L.; Morel, J.-M. Image selective smoothing and edge detection by nonlinear diffusion. II. *SIAM J. Numer. Anal.* **29** (1992), no. 3, 845–866.

[2] Catté, F.; Lions, P.-L.; Morel, J.-M.; Coll, T. Image selective smoothing and edge detection by nonlinear diffusion. *SIAM J. Numer. Anal.* **29** (1992), no. 1, 182–193.

[3] Chambolle, A. Image segmentation by variational methods: Mumford and Shah functional and the discrete approximations. *SIAM J. Appl. Math.* **55** (1995), no. 3, 827–863.

[4] Chambolle, A.; Dal Maso, G. Discrete approximation of the Mumford-Shah functional in dimension two. *M2AN Math. Model. Numer. Anal.* **33** (1999), no. 4, 651–672.

[5] Friedman, A. *Partial differential equations of parabolic type.* Prentice-Hall, Englewood Cliffs, N.J., 1964.

[6] Gobbino, M. Finite difference approximation of the Mumford-Shah functional. *Comm. Pure Appl. Math.* **51** (1998), no. 2, 197–228.

[7] Gobbino, M. Gradient flow for the one-dimensional Mumford-Shah functional. *Ann. Scuola Norm. Sup. Pisa Cl. Sci. (4)* **27** (1998), no. 1, 145–193.

[8] Kawohl, B.; Kutev, N. Maximum and comparison principle for one-dimensional anisotropic diffusion. *Math. Ann.* **311** (1998), no. 1, 107–123.

[9] Keller, J. B.; Lowengrub, J. S. Asymptotic and numerical results for blowing-up solutions to semi-linear heat equations. *Singularities in fluids, plasmas, and optics*, 111–130. NATO ASI Series. Kluwer, 1992.

[10] Kichenassamy, S. The Perona-Malik paradox. *SIAM J. Appl. Math.* **57** (1997), no. 5, 1328–1342.

[11] Kichenassamy, S.; Kumar, A.; Olver, P.; Tannenbaum, A.; Yezzi, A., Jr. Conformal curvature flows: from phase transition to active vision. *Arch. Rational Mech. Anal.* **134** (1996), no. 3, 275–301.

[12] Koenderink, J. J. The structure of images. *Biol. Cybernet.* **50** (1984), 363–370.

[13] Ladyženskaja, O. A.; Solonnikov, V. A.; Ural′ceva, N. N. *Linear and quasilinear equations of parabolic type.* Translated from the Russian by S. Smith. Translations of Mathematical Monographs, 23. American Mathematical Society, Providence, R.I., 1967.

[14] Morel, J.-M.; Solimini, S. *Variational methods in image segmentation. With seven image processing experiments.* Progress in Nonlinear Differential Equations and Their Applications, 14. Birkhäuser, Boston, 1995.

[15] Perona, P.; Malik, J. Scale space and edge detection using anisotropic diffusion. *Proc. IEEE Computer Soc. Workshop on Computer Vision* (1987), 16–22.

[16] Perona, P.; Malik, J. Scale space and edge detection using anisotropic diffusion. Dept. of EECS Technical Report, University of California, Berkeley, 1988.

[17] Perona, P.; Shiota, T.; Malik, J. Anisotropic diffusion. *Geometry-driven diffusion in computer vision*, 73–92. Computational Imaging and Vision, 1. Kluwer, Dordrecht, 1994.

[18] Protter, M. H.; Weinberger, H. F. *Maximum principles in differential equations.* Corrected reprint of the 1967 original. Springer, New York, 1984.

[19] Yuo, Y.; Xu, W.; Tannenbaum, A.; Kaveh, M. Behavioral analysis of anisotropic diffusion in image processing. *IEEE Trans. Image Process.* **5** (1996), no. 11, 1539–1553.

SELIM ESEDOGLU
University of Minnesota
Institute for Mathematics and Its Applications
207 Church Street S.E.
Minneapolis, MN 55455
E-mail: esedoglu@ima.umn.edu

Received January 2001.