Discrete dynamics for convex and non-convex smoothing functionals in PDE based image restoration

C.M.Elliott[†], B.Gawron[*], S.Maier-Paape[*], E.S.Van Vleck[‡]

[†]*Department of Mathematics*
University of Sussex
Falmer
East Sussex BN1 9QH, UK

[*]*Institut für Mathematik*
RWTH Aachen
52062 Aachen, Germany

[‡]*Department of Mathematics*
University of Kansas
Lawrence, KS 66045, USA

16th June 2003

Abstract

In this article we introduce a new model that generalizes the model of Perona-Malik and the total variation model. One main issue is to show that the discrete solutions of this new model conserve some properties of the solutions of the continuous model, in particular convergence of the iterative scheme to a critical point. We also implement this new model by a numerical algorithm with two different energy-based stopping methods. Computations are performed for a test picture.

Keywords: Perona-Malik, total variation model.

1 Introduction

Along with stochastic modeling and wavelets, the theory of PDEs is an important issue in modern image processing. A detailed outline is given in Aubert and Kornprobst [1]. Our goal is to develop a model for image processing, that generalizes Perona-Malik (PM) (see [9]) and the total variation model (TV) (cf. Rudin, Osher and Fatemi [10] or Osher and Fedkiw [8]). Besides giving a PDE for denoising purposes we analyse the global dynamics of a fully discrete model and discuss the problem of a well suited stopping time. This is important, because most PDE based models (including the two mentioned above) yield the best denoising results when the evolution time is halted after a while. In the definition of this stopping time our main emphasis is laid on the observation that no use of knowledge of the original picture should be made. This is crucial for practical reasons, because in most applications only a perturbed image of the real image is known.
Consider the following PDE on the domain $\Omega \subset \mathbb{R}^2$:

$$\begin{aligned} -\varepsilon \Delta u_t + u_t &= \nabla(g(|\nabla u|^2)\nabla u) \quad \text{in } \Omega \text{ for } t > 0 \\ \partial_\nu u &= 0 \quad \text{on } \partial\Omega \text{ for } t > 0 \\ u(t)_{|t=0} &= u_{noi} \quad \text{in } \Omega, \end{aligned} \tag{1}$$

where $g : [0, \infty) \to [0, \infty)$ with

$$g(s) = \frac{1}{(1 + \frac{s}{\gamma})^\alpha},$$

and $\varepsilon \geq 0$ and $\alpha, \gamma > 0$ are parameters and ν is the outward unit normal on $\partial\Omega$.

The initial condition u_{noi} is the noisy data (e.g. a picture) which is supposed to be denoised by the evolution of (1) without losing characteristic features like edges, for instance. Note that $\varepsilon = 0$ and $\alpha = 1$ in (1) delivers the (PM) model, whereas $\varepsilon = 0$ and $\alpha = 0.5$, $\gamma = 1$ represents some kind of a regularized version of the (TV) model, which is given by $g(s) = \frac{1}{\sqrt{s}}$.

The comparison of the models induced by different parameters is done as follows. We choose some target data u_{tar}, a function on $\Omega = [0,1]^2$ showing some geometric figures (cf. Figure 1).

As initial data for (1) we used (Figure 2):

$$u_{noi} = u_{tar} + "noise". \tag{2}$$

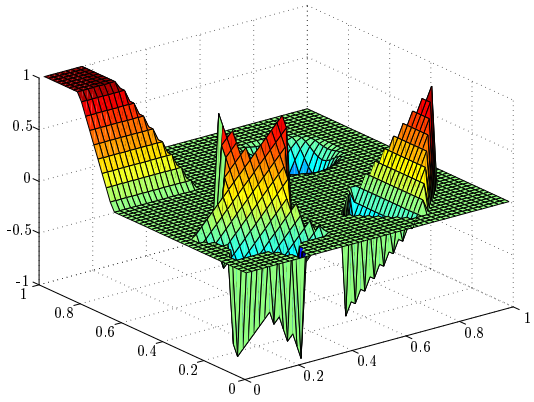

Figure 1: u_{tar}

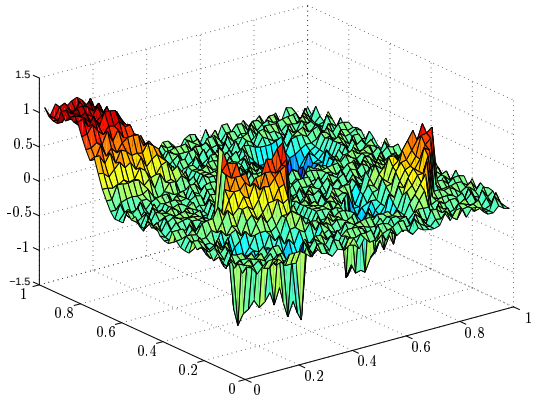

Figure 2: u_{noi}

Besides finding good parameters, an important issue will be to find a good stopping criterion, i.e., some stopping time t^* at which the denoising procedure is halted. Unfortunately (except for some special cases, like the heat equation model [3]) there is no satisfactory answer. In case u_{tar} is known (as in our case), one could choose as t^* that time t when $\|u(t) - u_{tar}\|_{L^2}$ stops decreasing. The closer $u(t^*)$ gets to u_{tar} for some stopping time $t^* > 0$, the better the reconstruction of the original picture works.
For practical applications, however, u_{tar} is usually not known. Therefore we need different stopping criteria for the denoising procedure which take this into account. We define the function $H(s)$ by $H'(s) = \frac{1}{2}g(s)$, i.e.,

$$H(s) = \frac{1}{2}\ \gamma\ log(1 + \frac{s}{\gamma}) \qquad \text{for } \alpha = 1 \tag{3}$$

and otherwise

$$H(s) = \frac{1}{2(1-\alpha)}\ \gamma\ [(1 + \frac{s}{\gamma})^{1-\alpha} - 1]. \tag{4}$$

We see that $H(0) = 0$ and $H'(s) > 0$ which implies $H(s) \geq 0$. From $g'(s) < 0$ we deduce that H is non-convex. With the additional parameter $\lambda > 0$ we construct an energy functional given by

$$G(u_{noi}, u) := \frac{\lambda}{2}\|u_{noi} - u\|_{L^2}^2 + \int_\Omega H(|\nabla u|^2)\ dx. \tag{5}$$

The first term measures the fidelity of the image u with the original noisy image u_{noi} whereas the second term measures the smoothness of the image u. In order to vary the balance of these two measures one has the scalar λ. Based on this energy functional we tested two different methods :

Method (1):

Set $\lambda = \lambda_1$. Solve (1) and stop when $E(u_{noi}, t) := G(u_{noi}, u(t))$ achieves a first local minimum in t. Note that the integral of H in (5) is a Liapunov functional for the evolution of (1). Therefore our stopping criterion balances the smoothing evolution of (1) represented by the term $\int_\Omega H(|\nabla u|^2)\ dx$ with a fidelity term $\frac{\lambda_1}{2}\|u_{noi} - u\|_{L^2}^2$ that punishes u going too far away from u_{noi}. In our numerical implementation we will see that for any discrete solution of (1) $u \to 0$ for $t \to \infty$ (cf. Corollary 4.2) and therefore we get

$$E(u_{noi}, t) \to \frac{\lambda_1}{2}\|u_{noi}\|_{L^2}^2 \quad \text{for} \quad t \to \infty.$$

Furthermore we will prove that $E(u_{noi},\ .\ )$ is decreasing for small t (cf. Corollary 3.4). If u_{noi} is sufficiently smooth we can choose λ_1 large enough to achieve

$$E(u_{noi}, 0) = \int_\Omega H(|\nabla u_{noi}|^2)\ dx \leq \frac{\lambda_1}{2}\|u_{noi}\|_{L^2}^2$$

and a positive finite minimum of $E(u_{noi},\ .\ )$ is guaranteed yielding a well defined stopping criterion.

In our second method we depart from the evolution (1) and consider gradient flow for the energy (5):

Method (2):

Instead of the PDE (1) we consider:

$$\begin{aligned} -\varepsilon\Delta u_t + u_t &= \nabla(g(|\nabla u|^2)\nabla u) + \lambda_2(u_{noi} - u) \quad \text{in } \Omega \text{ for } t > 0 \\ \partial_\nu u &= 0 \quad \text{on } \partial\Omega \text{ for } t > 0 \\ u(t)_{|t=0} &= u_{noi} \quad \text{in } \Omega, \end{aligned} \tag{6}$$

where $\lambda_2 > 0$. This evolution is H^1 gradient flow for the energy functional $G(u_{noi}, u)$ with $\lambda = \lambda_2$.

We will show that the discrete solution u of (6) converges as $t \to \infty$ (cf. Theorem 4.1) and we stop when u reaches a small neighborhood of a steady state, i.e., the evolution is halted when

$$\frac{\|u(t^* + \Delta t) - u(t^*)\|_{L^2}}{\Delta t} \leq tol$$

for some fixed small tolerance $tol > 0$ and at some discrete time point $t^* > 0$. One easily verifies that the PDE (6) is a gradient flow for the Liapunov functional $E(u_{noi}, \cdot)$. After an introduction of some notation in Section 2 we will see in Section 3 that our discretization preserves that property, i.e. a discrete version of (6) (with $\lambda_1 = \lambda_2$) is still a Liapunov functional for our discrete model. Furthermore in the discrete model (as well as in the continuous one) both methods have some natural invariance properties like conservation of the average value or gray-level shift invariance (cf. again Section 3). They will be useful to prove the convergence theorem in Section 4 and the existence of a discrete Liapunov functional. At last, in Section 5 we report on some computations using the discretization from above and both stopping methods.

One may ask why we chose this special family of functions $g = g_{\alpha,\gamma}$. There are some properties the evolution should have which lead to corresponding restrictions for g. For instance the condition $g(0) = 1$ implies behavior almost like the heat equation in case $|\nabla u|$ is small. In other words it leads to isotropic smoothing in flat regions. In points nearby a steep edge the evolution locally should be stopped which yields the restriction $\lim_{s\to\infty} g(s) = 0$. In a more precise examination it turns out that the function $b(s) = g(s) + 2sg'(s)$ plays an important role (cf. e.g. [1], Theorem 3.3.6 and [11], Section 3). The signs of $g(|\nabla u|^2)$ and $b(|\nabla u|^2)$ determine if our equation is locally forward parabolic or not. If both are positive we have local parabolicity which leads to a smoothing model. But it is absolutely reasonable to consider a smoothing-enhancing model, i.e. a model where edges are smoothed out or enhanced depending on the steepness of the edge. This leads to a replacement of the condition $b > 0$ by $b(s) > 0$ for $s \leq s_0$ and $b(s) < 0$ for $s > s_0$ for some threshold $s_0 > 0$. If in the considered point the condition $|\nabla u|^2 > s_0$ is valid, we get a locally backward

parabolic equation in direction of sharpest increase, which leads to sharpening of significant edges. With our special choice of g this condition only depends on the parameter α. We get a smoothing model for $\alpha \leq \frac{1}{2}$ and a smoothing-enhancing one for $\alpha > \frac{1}{2}$. A more detailed discussion is given in [11], Sections 2-4.

Finally we mention some classical existence results to this equation. For $\varepsilon > 0$ equation (6) (which is including (1)) has a classical solution on a maximal time of existence $[0, T)$ in an appropriate Banach space (compare [11], Section 12).
For $\varepsilon = 0$ we can generalize the problem by considering $u \in BV(\Omega)$. With the help of the notion of maximal operators one can prove the existence of a unique solution $u = u(t) : [0, \infty) \to L^2(\Omega)$ of the generalized problem (1) in case of $u_{noi} \in BV(\Omega)$, see [1], Theorem 3.3.1.

In summary, our major contributions are: Our choice of g in (1) generalizes the PDE ansatz of both the PM and TV model. We also extended these models by a viscous ε regularization. This regularization leads to a well posed PDE problem although g yields a non-convex variational problem. We focused on the analysis of the discrete dynamics which in effect yields a well-posed discrete scheme for the forward parabolic and the forward-backward parabolic case (even if the regularization parameter ε is zero). Finally we discussed the two different stopping criteria and made some numerical simulations.

Acknowledgement

This paper was completed whilst CME participated in the 2003 Programme "Computational Challanges in PDEs" at the Isaac Newton Institute, Cambridge, UK.

2 Preparations

For the spatial discretization we replace $\Omega = [0, 1]^2$ by a discrete grid of d^2 points

$$\Omega_d = \{z = (x, y) \in \Omega \mid dx - \frac{1}{2},\ dy - \frac{1}{2} \in \mathbb{Z}\}$$

and order these grid points $z_1, ..., z_{d^2}$ row after row.
For any function $v : [0, 1]^2 \to \mathbb{R}$ we name the vector $v \in \mathbb{R}^{d^2}$ on the d^2 grid points as its discrete counterpart:

$$v = (v_1, v_2, ..., v_{d^2})$$

with

$$v_1 = v(z_1),\ v_2 = v(z_2),\ v_3 = v(z_3),$$

This method has been extensively studied in [6], Chapter 4.

2.1 Notation

Let $v, w \in \mathbb{R}^{d^2}$ be arbitrary vectors (corresponding to functions $v, w : [0,1]^2 \to \mathbb{R}$). We will frequently use the following notation:

$diag(v) \in \mathbb{R}^{d^2 \times d^2}$ is the diagonal matrix

$$diag(v) = \begin{pmatrix} v_1 & 0 & \cdots & 0 \\ 0 & v_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & v_{d^2} \end{pmatrix},$$

$v \circ w \in \mathbb{R}^{d^2}$ is the vector valued product $\{v_i w_i\}_{i=1,\ldots,d^2}$. In fact, the point wise multiplication of the functions v, w may be simulated by $v \circ w = diag(v)w$. For $v \circ v$ we also write v^2.

For any function $f : \mathbb{R} \to \mathbb{R}$ we define the vector $f(v) \in \mathbb{R}^{d^2}$ by

$$f(v) = \{f(v_i)\}_{i=1,\ldots,d^2}.$$

We use the shortcut $\mathbb{1}$ for the vector $(1, 1, ..., 1) \in \mathbb{R}^{d^2}$. It represents the constant function $u = 1$. With the usual inner product in $\mathbb{R}^{d^2}$ $(\ ,\ )$ we obtain that

$$\frac{1}{d^2}(\mathbb{1}, v)$$

is the discrete analogue for $\int\limits_\Omega v(x)\mathrm{d}x$. Hence the L^2 product $\int\limits_\Omega v(x)w(x)\mathrm{d}x$ changes to

$$\frac{1}{d^2}(\mathbb{1}, v \circ w) = \frac{1}{d^2}(v, w).$$

The discrete p-norm ($p \geq 1$) is defined by

$$\|v\|_p = \sqrt[p]{\frac{1}{d^2}(\mathbb{1}, |v|^p)}.$$

Especially for $p = 2$ we have that $\|v\|_2^2 = \frac{1}{d^2}(\mathbb{1}, |v|^2) = \frac{1}{d^2}(|v|, |v|) = \frac{1}{d^2}(v, v)$.
In the case $p = \infty$ we set $\|v\|_\infty = \max\limits_{i=1,\ldots,d^2} |v_i|$.

2.2 Discrete derivatives and the discrete Laplacian

Our goal is to descretize the Laplacian Δ and the term $\nabla(\bar{g}(|\nabla u|^2)\nabla u)$, where $\bar{g}$ is defined by

$$\bar{g} = g - 1.$$

To approximate derivatives at a grid point we have to incorporate some grid points in its neighborhood which leads to problems at marginal grid points. For any given functions v, w on Ω_d the general idea is to extend them beyond the boundary on a greater grid Ω_{d+2} regarding the Neumann boundary conditions, to apply the usual derivative operators and to restrict then the result on Ω_d again. This restriction is realized by a linear operator

$$R : \mathbb{R}^{(d+2)^2} \to \mathbb{R}^{d^2}.$$

It removes all values of a given function v on the marginal grid points of the extended grid. We call Rv the **restriction** of v.
For $d = 3$ we have for instance:

$$R(v_1, ..., v_{25}) = (v_7, v_8, v_9, v_{12}, v_{13}, v_{14}, v_{17}, v_{18}, v_{19}).$$

Inversely for every linear operator

$$E : \mathbb{R}^{d^2} \to \mathbb{R}^{(d+2)^2}$$

with

$$REw = w$$

we call Ew an **extension** of $w \in \mathbb{R}^{d^2}$.

In our problem we have to regard the Neumann boundary condition

$$\partial_\nu v = 0 \quad \text{on} \quad \partial\Omega.$$

The derivative in outer normal direction in a point of the boundary may be approximated by the difference $\frac{1}{d}(v_c - v_m)$ where $v_m = v(z_m)$ is the value at a marginal point z_m (those with distance $\frac{1}{2d}$ to the boundary) and $v_c = v(z_c)$ is its continuation beyond the boundary. If we choose $v_c = v_m$ we match the Neumann boundary condition.
In other words missing neighbors are constructed by reflection of marginal points at the boundary. This special extension of v is realized by a linear operator

$$E_0 : \mathbb{R}^{d^2} \to \mathbb{R}^{(d+2)^2},$$

which we call the **Neumann extension**. By "usual derivative operators" we mean the matrices

$$D_x^s,\ D_x^d,\ D_y^l,\ D_y^u \in \mathbb{R}^{(d+2)^2 \times (d+2)^2} :$$

$$D_x^s = \frac{1}{d}\begin{pmatrix} 1 & 0 & \cdots & 0 \\ -1 & 1 & \cdots & 0 \\ \vdots & \ddots & \ddots & \vdots \\ 0 & \cdots & -1 & 1 \end{pmatrix}, \qquad D_x^d = \frac{1}{d}\begin{pmatrix} -1 & 1 & \cdots & 0 \\ \vdots & \ddots & \ddots & \vdots \\ 0 & \cdots & -1 & 1 \\ 0 & \cdots & 0 & -1 \end{pmatrix},$$

$$D_y^l = \frac{1}{d}\begin{pmatrix} I & -I & \cdots & 0 \\ \vdots & \ddots & \ddots & \vdots \\ 0 & \ddots & I & -I \\ 0 & \cdots & 0 & I \end{pmatrix}, \qquad D_y^u = \frac{1}{d}\begin{pmatrix} -I & 0 & \cdots & 0 \\ I & -I & \cdots & 0 \\ \vdots & \ddots & \ddots & \vdots \\ 0 & \cdots & I & -I \end{pmatrix},$$

where I is the $(d+2) \times (d+2)$ unit matrix. They are the sinistral, dexter, lower and upper derivative operators respectively. Confer for instance [11], Subsection 7.2, for more details.

For functions v with Neumann boundary conditions restricted to Ω_d we can approximate $\partial_x v \approx RD_x^s E_0 v$. We see that $D_x := RD_x^s E_0 \in \mathbb{R}^{d^2 \times d^2}$ is an endomorphism on $\mathbb{R}^{d^2}$. In the sequel we use D_x as **(sinistral) approximation of** ∂_x. An easy calculation gives:

$$D_x = \frac{1}{d}\begin{pmatrix} M & 0 & \cdots & 0 \\ 0 & M & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & M \end{pmatrix} \tag{7}$$

with

$$M = \begin{pmatrix} 0 & 0 & \cdots & 0 \\ -1 & 1 & \cdots & 0 \\ \vdots & \ddots & \ddots & \vdots \\ 0 & \cdots & -1 & 1 \end{pmatrix} \in \mathbb{R}^{d \times d}.$$

Similarly, for any given functions v, w on Ω_d (w with Neumann boundary conditions) we have approximately $\partial_x(v\ \partial_x w) \approx RD_x^d(Ev \circ D_x^s E_0 w)$. Hence for the discretization of $\partial_x(v\ \partial_x w)$ we use

$$RD_x^d(Ev \circ D_x^s E_0 w) = -D_x^\top diag(v) D_x w.$$

We call $-D_x^\top diag(v) D_x w \in \mathbb{R}^{d^2}$ the **discretization** of $\partial_x(v\ \partial_x w)$ on Ω_d.
In the same manner we receive as discrete version of ∂_y for functions with Neumann boundary conditions restricted to Ω_d the matrix $D_y := RD_y^l E_0 \in \mathbb{R}^{d^2 \times d^2}$:

$$D_y = \frac{1}{d}\begin{pmatrix} I & -I & \cdots & 0 \\ \vdots & \ddots & \ddots & \vdots \\ 0 & \cdots & I & -I \\ 0 & \cdots & 0 & 0 \end{pmatrix}. \tag{8}$$

Here I is the $d \times d$ unit matrix. The **discretization** of $\partial_y(v\ \partial_y w)$ on Ω_d is:

$$\partial_y(v\ \partial_y w) \approx RD_y^u(Ev \circ D_y^l E_0 w) = -D_y^\top diag(v) D_y w.$$

Now we are able to construct the discrete Laplacian subject to Neumann boundary conditions. We can write Δ in the form

$$\Delta w = \partial_x^2 w + \partial_y^2 w = \partial_x(v\ \partial_x w) + \partial_y(v\ \partial_y w),$$

with $v = 1$. Hence the discrete Laplacian $\Delta_d \in \mathbb{R}^{d^2 \times d^2}$ is given by

$$\begin{aligned} \Delta_d &= -D_x^\top diag(\mathbb{1}) D_x - D_y^\top diag(\mathbb{1}) D_y \\ &= -(D_x^\top D_x + D_y^\top D_y). \end{aligned}$$

For the discretization of $|\nabla v|^2$ we define a vector valued symmetric bilinear form $B : \mathbb{R}^{d^2} \times \mathbb{R}^{d^2} \to \mathbb{R}^{d^2}$ by

$$B(v, w) = (D_x v) \circ (D_x w) + (D_y v) \circ (D_y w).$$

Especially $B(v, v)$ may be seen as the discrete version of $|\nabla v|^2$. This is justified by the approximation:

$$\begin{aligned} |\nabla v|^2 &\approx R[(D_x^s E_0 v)^2 + (D_y^l E_0 v)^2] \\ &= (RD_x^s E_0 v)^2 + (RD_y^l E_0 v)^2 \\ &= (D_x v)^2 + (D_y v)^2 \\ &= B(v, v) \end{aligned}$$

Using this method and with the definition of the symmetric $d^2 \times d^2$ matrix

$$D(w) := -D_x^\top diag[\bar{g}(B(w, w))] D_x - D_y^\top diag[\bar{g}(B(w, w))] D_y \tag{9}$$

we see, that $\nabla(\bar{g}(|\nabla w|^2)\nabla w)$ is discretized by

$$D(w)w.$$

The following identities for $v, w \in \mathbb{R}^{d^2}$ essentially accord to integration by parts. They will be needed later on.

$$\begin{aligned} (\mathbb{1}, B(v, w)) &= (\mathbb{1}, D_x v \circ D_x w + D_y v \circ D_y w) \\ &= (w, D_x^\top D_x v + D_y^\top D_y v) \\ &= (w, -\Delta_d v) \end{aligned} \tag{10}$$

$$
\begin{aligned}
(\mathbb{1}, \bar{g}(B(v,v)) \circ B(v,w)) &= (\mathbb{1}, \bar{g}(B(v,v)) \circ [D_x v \circ D_x w + D_y v \circ D_y w]) \\
&= (w, D_x^\top [\bar{g}(B(v,v)) \circ (D_x v)] \\
&\qquad + D_y^\top [\bar{g}(B(v,v)) \circ (D_y v)]) \\
&\overset{(9)}{=} (w, -D(v)v) \qquad (11)
\end{aligned}
$$

$$
\begin{aligned}
(\mathbb{1}, g(B(v,v)) \circ B(v,w)) &= (\mathbb{1}, B(v,w) + \bar{g}(B(v,v)) \circ B(v,w)) \\
&= (w, -\Delta_d v - D(v)v) \qquad (12)
\end{aligned}
$$

We will also need:

$$
B(v,v) \circ B(w,w) - B(v,w)^2 = [(D_x v) \circ (D_y w) - (D_x w) \circ (D_y v)]^2. \quad (13)
$$

Further we deduce from $D_x \mathbb{1} = D_y \mathbb{1} = 0$:

$$
\Delta_d \mathbb{1} = 0 \qquad \text{and} \qquad D(v)\mathbb{1} = 0. \qquad (14)
$$

For some technical reasons we will need that the restriction of $-\Delta_d$ on the subspace $W = \{v \in \mathbb{R}^{d^2} \,|\, (\mathbb{1}, v) = 0\}$ consisting of functions with zero mass is strictly positive. To that end we note that the discrete Laplacian $\Delta_d : \mathbb{R}^{d^2} \to \mathbb{R}^{d^2}$ with Neumann boundary conditions has the eigenvectors e_{mn}, $m, n = 0, ..., d-1$, given by

$$
(e_{mn})_j = \cos(m\pi x_j) \cos(n\pi y_j), \, j = 1, ..., d^2 \qquad (15)
$$

where (x_j, y_j) are the coordinates of a grid point $z_j \in \Omega_d$. The corresponding eigenvalues are

$$
\mu_{mn} = -\frac{4}{d^2} \left(\sin^2 \left(\frac{m\pi}{2d} \right) + \sin^2 \left(\frac{n\pi}{2d} \right) \right). \qquad (16)
$$

Thus we have $\mu_{mn} = 0$ for $m = n = 0$ and otherwise $\mu_{mn} < 0$.

Lemma 2.1 *For $v \in W$ and some constant $c_1 > 0$ we have:*

$$
(v, -\Delta_d v) \geq c_1 (v, v). \qquad (17)
$$

Proof: $-\Delta_d$ is symmetric, so we have an orthonormal basis of eigenvectors $\{e_1, ..., e_{d^2}\}$. We denote the corresponding eigenvalues according to (16) as $\mu_i, i = 1, ..., d^2$.

$$
-\Delta_d e_i = \mu_i e_i, \quad i = 1, \ldots, d^2.
$$

Ordering them to $\mu_1 \leq \mu_2 \leq \ldots \leq \mu_{d^2}$ we get: $\mu_1 = 0$, $e_1 = \mathbb{1}$ and $0 < \mu_2 \leq \mu_i$ for $2 \leq i$.

Suppose that $v = \sum_{i=1}^{d^2} a_i e_i \in W$. Now we immediately obtain:

$$0 = (\mathbb{1}, v) = \sum_{i=1}^{d^2} a_i (e_1, e_i) = a_1$$

and hence

$$(v, -\Delta_d v) = \left(\sum_{i=2}^{d^2} a_i e_i, \sum_{i=2}^{d^2} a_i \mu_i e_i \right) \geq \mu_2 (v, v).$$

□

3 Numerical Algorithm, Invariance Properties and Liapunov Functional

Both methods are implemented numerically by the following algorithm:
We use the time discretization $t_0 = 0, t_{n+1} = t_n + \Delta t^n$. The width of the time steps Δt^n can be constant or regulated in an adaptive way. However in the following we shall assume that $\{\Delta t^n\}$ is bounded.
Then we start with the discretized initial function $u^0 = u(0) = u_{noi} \in \mathbb{R}^{d^2}$ and compute iteratively the solutions at later times $u^1 = u(t_1), u^2 = u(t_2), \ldots$.
The iterative scheme is given by:

$$C_n u^{n+1} = A u^n + \Delta t^n [D(u^n) u^n + \lambda_2 u^0] \tag{18}$$

with

$$A = I - \varepsilon \Delta_d \in \mathbb{R}^{d^2} \times \mathbb{R}^{d^2}$$

and

$$C_n = A + \Delta t^n (\lambda_2 I - \Delta_d) \in \mathbb{R}^{d^2} \times \mathbb{R}^{d^2}.$$

This is a well posed discrete scheme since C_n is symmetric positive definite and in particular bijective for all $\Delta t^n > 0$ and $\varepsilon \geq 0$.
To see this we write C_n as $C_n = \zeta_1 I - \zeta_2 \Delta_d$ with $\zeta_1, \zeta_2 > 0$. In terms of the proof of Lemma 2.1 we get

$$(C_n v, v) = \sum_{i=1}^{d^2} a_i^2 (\zeta_1 + \zeta_2 \mu_i) > 0 \text{ for all } v \neq 0.$$

The symmetry of C_n follows from the symmetry of Δ_d and I.

To compute the solution u^{n+1} at a later time t_{n+1} we have to solve a linear system of dimension d^2. This can be done by using the conjugate gradient scheme (c.f. [7]) or a fast fourier transformation. Note that this algorithm is used for **Method (1)** and **Method (2)**. If we consider the evolutions (1) and (6) we see that (6) includes (1), if we set $\lambda_2 = 0$. Corresponding to that we set $\lambda_2 = 0$ in the algorithm, if we work with **Method (1)**. For better understanding we give a formal derivation of this algorithm:
First we substitute in the original evolution

$$\begin{aligned} -\varepsilon\Delta u_t + u_t &= \nabla(g(|\nabla u|^2)\nabla u) + \lambda_2(u(0) - u) \\ &= \nabla(\bar{g}(|\nabla u|^2)\nabla u) + \Delta u + \lambda_2(u(0) - u) \end{aligned}$$

the function u by u^n or u^{n+1} and the operators by their discrete counterparts, respectively. The time derivative is approximated by $u_t \approx \frac{u^{n+1}-u^n}{\Delta t^n}$:

$$(-\varepsilon\Delta_d + I)\frac{u^{n+1} - u^n}{\Delta t^n} = D(u^n)u^n + \Delta_d u^{n+1} + \lambda_2(u^0 - u^{n+1})$$

If we order the terms with u^{n+1} to the left hand side this is equivalent to (18).

With this discretization of the evolution it is easy to see that the average value of u is conserved. Furthermore we have gray-level shift invariance. To be more precise:

Lemma 3.1 *Given the iterative scheme* (18) *the equation*

$$(\mathbb{1}, u^n) = \left(\mathbb{1}, u^0\right)$$

holds for all n.

Lemma 3.2 *Let* u^0 *generate the sequence* (u^n) *via* (18) *and let* $K \in \mathbb{R}$ *be given, then the modified start vector* $\tilde{u}^0 = u^0 + K\ \mathbb{1}$ *generates the sequence*
$(\tilde{u}^n) = (u^n + K\ \mathbb{1})$.
(*provided equal time steps* Δt^n)

A formal calculation shows that the right hand side of (5) with λ_2 instead of λ gives a Liapunov functional for (6). This leads to the conjecture that the iterative scheme (18) has the discrete Liapunov functional $J : \mathbb{R}^{d^2} \to \mathbb{R}$,

$$J(u) = \frac{\lambda_2}{2}\|u^0 - u\|_2^2 + \frac{1}{d^2}\Big(\mathbb{1}, H(B(u, u))\Big). \tag{19}$$

Theorem 3.3 *For the sequence* (u^n) *generated from* (18) *and some constant* $c_2 > 0$ *the inequality*

$$J(u^{n+1}) + \frac{c_2}{\Delta t^n}\|u^{n+1} - u^n\|_2^2 \leq J(u^n) \text{ holds for all } n \in \mathbb{N}_0. \tag{20}$$

In particular J *is a discrete Liapunov functional for* **Method (2)**. *Setting* $\lambda_2 = 0$ *it is also a discrete Liapunov functional for* **Method (1)**.

Proof: If we use Taylor expansion of J around u^n and the shortcut $h = u^{n+1} - u^n$, we receive a $\xi \in \{su^n + (1-s)u^{n+1} \mid s \in [0,1]\} \subset \mathbb{R}^{d^2}$ with:

$$\begin{aligned}
&d^2(J(u^{n+1}) - J(u^n)) \\
&= \left(\mathbb{1}, \lambda_2(u^n - u^0) \circ h + g(B(u^n, u^n)) \circ B(u^n, h)\right) \\
&\qquad + \frac{1}{2}\left(\mathbb{1}, \lambda_2 h^2 + 2g'(B(\xi,\xi)) \circ B(\xi, h)^2 + g(B(\xi,\xi)) \circ B(h,h)\right) \\
&\overset{g' \leq 0}{\leq} \left(\mathbb{1}, \lambda_2\left(\frac{u^{n+1} + u_n}{2} - u^0\right) \circ h + g(B(u^n, u^n)) \circ B(u^n, h)\right) \\
&\qquad + \left(\mathbb{1}, \frac{1}{2} g(B(\xi,\xi)) \circ B(h,h)\right) \\
&\overset{g - 1 = \bar{g} \leq 0}{\leq} \left(\mathbb{1}, \lambda_2\left(\frac{u^{n+1} + u^n}{2} - u^0\right) \circ h + g(B(u^n, u^n)) \circ B(u^n, h) + \frac{1}{2} B(h,h)\right) \\
&\overset{(12),(10),(18)}{=} -\left(\frac{\lambda_2}{2} + \frac{1}{\Delta t^n}\right)(h,h) - \left(\frac{1}{2} + \frac{\varepsilon}{\Delta t^n}\right)(h, -\Delta_d h) \\
&\leq -\frac{1}{\Delta t^n}(h,h).
\end{aligned}$$

since $-\Delta_d$ is positive. □

In particular we receive for $\lambda_2 = 0$ and $n = 0$ from (20):

$$\left(\mathbb{1}, H(B(u^1, u^1))\right) - \left(\mathbb{1}, H(B(u^0, u^0))\right) \leq -\frac{1}{\Delta t^0}\left(u^1 - u^0, u^1 - u^0\right). \tag{21}$$

This results can be applied to **Method (1)**:
If we discretize the functional (5), we obtain $K : \mathbb{R}^{d^2} \to \mathbb{R}$,

$$K(u) := \frac{\lambda_1}{2}\|u^0 - u\|_2^2 + \frac{1}{d^2}\left(\mathbb{1}, H(B(u,u))\right).$$

The following corollary gives conditions under which **Method 1** is justified:

Corollary 3.4 *If the sequence* (u^n) *is generated from* (18) *with* $\lambda_2 = 0$ *and the first time step* Δt^0 *fulfills* $\Delta t^0 \leq \frac{2}{\lambda_1}$, *then*

$$K(u^1) \leq K(u^0).$$

In particular the discretized version of functional (5) *of* **Method 1** *decreases at the beginning of the iteration.*

Proof: We set $h = u^1 - u^0$:

$$d^2\ (K(u^1) - K(u^0)) \overset{(21)}{\leq} \left(\frac{\lambda_1}{2} - \frac{1}{\Delta t^0}\right)(h,h) \leq 0,$$

if we choose $\Delta t^0 \leq \frac{2}{\lambda_1}$. □

4 Convergence

Due to the conservation of the average value (Lemma 3.1) and the gray-level shift invariance (Lemma 3.2) we can reduce our observations to the subspace $W = \{u \in \mathbb{R}^{d^2} | (\mathbb{1}, u) = 0\}$ of functions with mass zero.

Theorem 4.1 *Assume* $\alpha \in [0, \frac{1}{2}]$ *and* $u^0 \in W$ *and let arbitrary positive constants* $0 < c < C$ *be given. If* $c < \Delta t^n < C$ *for all* n *then the sequence* (u^n) *constructed by* (18) *(with* $\varepsilon \geq 0$*) converges to an unique critical point* u *sufficing the condition:*

$$-\Delta_d u = D(u)u + \lambda_2(u^0 - u). \tag{22}$$

Proof: The concept of this proof is adapted from [2] and [5]. We define $p = 1 - \alpha$ and $\tilde{p} = 2p$. Then $p \in [\frac{1}{2}, 1]$ and $\tilde{p} \in [1, 2]$.

It is easy to see that for nonnegative numbers a, b the following inequality holds:

$$a^p - b^p \leq |a - b|^p.$$

We deduce from Hölders inequality:

$$a^p + b^p \leq 2^{1-p}(a + b)^p. \tag{23}$$

This yields for nonnegative numbers a, b, c, d:

$$\begin{aligned}(a^p - b^p)^2 + (c^p - d^p)^2 &\leq |a - b|^{2p} + |c - d|^{2p} \\ &\leq 2^{1-p}[(a - b)^2 + (c - d)^2]^p. \end{aligned} \tag{24}$$

The bilinear form $B(u,u) = (D_x u) \circ (D_x u) + (D_y u) \circ (D_y u)$ is a vector with components of the form

$$(u_i - u_j)^2 + (u_k - u_l)^2$$

for some indices $i, j, k, l \in 1, ..., d^2$. This follows from the simple structure of the matrices D_x and D_y (compare Subsection 2.2). We infer:

$$\begin{aligned}
\frac{1}{d^2}(\mathbb{1}, B(|u|^p, |u|^p)) &= \frac{1}{d^2} \sum [(|u_i|^p - |u_j|^p)^2 + (|u_k|^p - |u_l|^p)^2] \\
&\overset{(24)}{\leq} 2^{1-p} \frac{1}{d^2} \sum [(|u_i| - |u_j|)^2 + (|u_k| - |u_l|)^2]^p \\
&\leq 2^{1-p} \frac{1}{d^2} \sum [(u_i - u_j)^2 + (u_k - u_l)^2]^p \\
&= 2^{1-p} \frac{1}{d^2} (\mathbb{1}, B(u,u)^p). \qquad (25)
\end{aligned}$$

Finally we obtain a discrete Poincare inequality:
There is a constant $c_1 > 0$ so that for all $\tilde{p} \in [1,2]$ and $u \in W$ we have:

$$\begin{aligned}
c_1 \|u\|_{\tilde{p}}^{\tilde{p}} &= c_1 \frac{1}{d^2} (\mathbb{1}, |u|^{2p}) \\
&= c_1 \frac{1}{d^2} (|u|^p, |u|^p) \\
&\overset{(17)}{\leq} \frac{1}{d^2} (|u|^p, -\Delta_d |u|^p) \\
&\overset{(10)}{=} \frac{1}{d^2} (\mathbb{1}, B(|u|^p, |u|^p)) \\
&\overset{(25)}{\leq} 2^{1-p} \frac{1}{d^2} (\mathbb{1}, B(u,u)^p). \qquad (26)
\end{aligned}$$

We have $\alpha \in [0, \frac{1}{2}]$ and therefore (see (4)):

$$\begin{aligned}
H(s) &= \frac{1}{2(1-\alpha)} \gamma \left[\left(1 + \frac{s}{\gamma} \right)^{1-\alpha} - 1 \right] \\
&= \frac{1}{2p} \gamma \left[\left(1 + \frac{s}{\gamma} \right)^{p} - 1 \right] \\
&\overset{(23)}{\geq} \frac{1}{2p} \gamma \left[\frac{1 + (\frac{s}{\gamma})^p}{2^{1-p}} - 1 \right] \\
&= c_3 s^p - k \qquad (27)
\end{aligned}$$

for $s \geq 0$ and some constants $c_3 > 0,\ k \in \mathbb{R}$.

From Theorem 3.3 we deduce that the sequence u^n is bounded in the $\tilde{p}$-norm:

$$\begin{aligned} J(u^0) &\geq J(u^n) \\ &\geq \frac{1}{d^2}(\mathbb{1}, H(B(u^n, u^n))) \\ &\overset{(27)}{\geq} c_3 \frac{1}{d^2}(\mathbb{1}, B(u^n, u^n)^p) - k \\ &\overset{(26)}{\geq} 2^{p-1} c_1 c_3 \|u^n\|_{\tilde{p}}^{\tilde{p}} - k. \end{aligned} \tag{28}$$

Repeated use of (20) leads to:

$$J(u^n) + c_2 \sum_{k=1}^{n} \frac{\|u^k - u^{k-1}\|_2^2}{\Delta t^{k-1}} \leq J(u^0).$$

Because $J(u) \geq 0$ the sum converges for $n \to \infty$, which implies

$$\frac{\|u^{n+1} - u^n\|_2^2}{\Delta t^n} \overset{n\to\infty}{\longrightarrow} 0.$$

Recalling $\Delta t^n < C$ for all n we obtain:

$$\|u^{n+1} - u^n\|_2 \overset{n\to\infty}{\longrightarrow} 0. \tag{29}$$

Due to $c < \Delta t^n$:

$$\frac{\|u^{n+1} - u^n\|_2}{\Delta t^n} \overset{n\to\infty}{\longrightarrow} 0. \tag{30}$$

Therefore we obtain for some positive constants c_5, c_6:

$$\begin{aligned} &\|D(u^n)u^n - D(u^{n-1})u^{n-1}\|_2 \\ &\overset{(18)}{=} \|\frac{1}{\Delta t^n} A(u^{n+1} - u^n) + (\lambda_2 I - \Delta_d)u^{n+1} \\ &\qquad\qquad - \frac{1}{\Delta t^{n-1}} A(u^n - u^{n-1}) - (\lambda_2 I - \Delta_d)u^n\|_2 \\ &\leq c_5 \frac{\|u^{n+1} - u^n\|_2}{\Delta t^n} + c_5 \frac{\|u^n - u^{n-1}\|_2}{\Delta t^{n-1}} + c_6 \|u^{n+1} - u^n\|_2. \end{aligned}$$

From (29) and (30) we infer:

$$\|D(u^n)u^n - D(u^{n-1})u^{n-1}\|_2 \overset{n\to\infty}{\longrightarrow} 0. \tag{31}$$

The sequence u^n is bounded in the $\tilde{p}$ norm. Then a convergent subsequence (u^{n_k}) exists with

$$\|u^{n_k} - u\|_{\tilde{p}} \stackrel{k\to\infty}{\longrightarrow} 0.$$

We write (18) in detail for u^{n_k-1}:

$$Au^{n_k} + \Delta t^{n_k-1}(\lambda_2 u^{n_k} - \Delta_d u^{n_k}) = Au^{n_k-1} + \Delta t^{n_k-1}(D(u^{n_k-1})u^{n_k-1} + \lambda_2 u^0)$$

or

$$\begin{aligned} &-\Delta_d u^{n_k} \\ &= \frac{1}{\Delta t^{n_k-1}} A(u^{n_k-1} - u^{n_k}) + D(u^{n_k})u^{n_k} \\ &\qquad +[D(u^{n_k-1})u^{n_k-1} - D(u^{n_k})u^{n_k}] + \lambda_2(u^0 - u^{n_k}). \end{aligned}$$

Since we have $1 \le \tilde{p} \le 2$ the $\tilde{p}$ norm is dominated by the 2 norm. This follows from Hölders inequality:

$$\begin{aligned} \|v\|_{\tilde{p}}^{\tilde{p}} &= \frac{1}{d^2}(\mathbb{1}, |v|^{\tilde{p}}) \\ &= \frac{1}{d^2}\sum_{i=1}^{d^2} |v_i|^{2p} \\ &\le \frac{1}{d^2}(\sum_{i=1}^{d^2} |v_i|^2)^p (\sum_{i=1}^{d^2} 1)^{1-p} \\ &= [\frac{1}{d^2}(\mathbb{1}, |v|^2)]^p \quad [\frac{1}{d^2}(\sum_{i=1}^{d^2} 1)]^{1-p} \\ &= \|v\|_2^{\tilde{p}}. \end{aligned} \tag{32}$$

Using (30), (31) and passing to the limit $k \to \infty$ in the $\tilde{p}$ norm we see that u fulfills (22):

$$-\Delta_d u = D(u)u + \lambda_2(u^0 - u). \tag{33}$$

Finally we show that (22) has a unique solution: Let u_1, u_2 be solutions of (22). We consider the function $f : \mathbb{R}^{d^2} \to \mathbb{R}$ given by:

$$f(u) = (\Delta_d u + D(u)u - \lambda_2 u, h)$$

with $h = u_1 - u_2$.

From (12) follows that

$$f(u) = -(\mathbb{1}, g(B(u,u)) \circ B(u,h) + \lambda_2 u \circ h).$$

For the derivative we obtain:

$$\begin{aligned} Df(u)\langle v\rangle = -(\mathbb{1}, & 2g'(B(u,u)) \circ B(u,v) \circ B(u,h) \\ & + g(B(u,u)) \circ B(v,h) + \lambda_2 v \circ h). \end{aligned}$$

Due to the mean value theorem there exists a vector $\xi \in \mathbb{R}^{d^2}$ with:

$$\begin{aligned} 0 &= f(u_1) - f(u_2) \\ &= Df(\xi)\langle h\rangle \\ &= -(\mathbb{1}, 2g'(B(\xi,\xi)) \circ B(\xi,h)^2 + g(B(\xi,\xi)) \circ B(h,h) + \lambda_2 h^2) \\ &\leq -(\mathbb{1}, 2g'(B(\xi,\xi)) \circ B(\xi,h)^2 + g(B(\xi,\xi)) \circ B(h,h)) \leq \ldots \end{aligned}$$

using $g'(s) \leq 0$ and (13)

$$\begin{aligned} \ldots &\leq -(\mathbb{1}, 2g'(B(\xi,\xi)) \circ B(\xi,\xi) \circ B(h,h) + g(B(\xi,\xi)) \circ B(h,h)) \\ &= -(B(h,h), 2g'(B(\xi,\xi)) \circ B(\xi,\xi) + g(B(\xi,\xi))) \\ &\leq 0. \end{aligned}$$

Note that the function $b(s) = g(s) + 2g'(s)s$ is strictly positive for $\alpha \leq \frac{1}{2}$.

Hence it follows that $B(h,h) = 0$:

$$\begin{aligned} 0 &= (\mathbb{1}, B(h,h)) \\ &\overset{(10)}{=} (h, -\Delta_d h) \\ &\overset{(17)}{\geq} c_1(h,h). \end{aligned}$$

We obtain $h = 0$.

Altogether we have shown that every convergent subsequence converges to the same critical point. This, in turn, implies that the whole sequence converges. □

Corollary 4.2 *For $\lambda_2 = 0$, i.e. for* **Method (1)** *the critical point u from Theorem 4.1 is the constant function with mass zero $u = 0$.*

Proof: For $\lambda_2 = 0$ (22) can be written as

$$(\Delta_d + D(u))u = 0.$$

We deduce from (12):

$$(u, -(\Delta_d + D(u))u) = (\mathbb{1}, g(B(u,u)) \circ B(u,u)) = 0.$$

Since g is positive and $B(u,u) \geq 0$ in every component it follows

$$(\mathbb{1}, B(u,u)) = 0$$

and like in the proof of Theorem 4.1 $u = 0$. □

If we allow that the constants in the estimates may depend on the grid dimension d, we get boundedness of the sequence u^n for all $\alpha \geq 0$:

Theorem 4.3 *Under the premises of Theorem 4.1 but with $\alpha \in (\frac{1}{2}, \infty)$ the sequence (u^n) is bounded and has therefore a convergent subsequence. Every convergent subsequence converges to a critical point u, i.e. a solution of (22). If these critical points are isolated, the whole sequence converges to one of them.*

Remark 4.4 *In general the solutions of (22) are nonunique.*

Proof of Theorem 4.3: This proof uses some ideas from Elliott [4], see also [5]. For any vector $v \in \mathbb{R}^{d^2}$ we have the estimates $\frac{1}{d^2}\|v\|_\infty \leq \|v\|_1 \leq \|v\|_\infty$. From Theorem 3.3 we deduce that

$$d^2 J(u^0) \geq d^2 J(u^n) \geq (\mathbb{1}, H(B(u^n, u^n))) = d^2 \|H(B(u^n, u^n))\|_1 \geq \|H(B(u^n, u^n))\|_\infty.$$

Since the function $H(s)$ is non-negative and monotone increasing we have $\|H(B(u^n,u^n))\|_\infty = H(\|B(u^n,u^n)\|_\infty)$. This implies:

$$C_0 := H^{-1}(d^2 J(u^0)) \geq \|B(u^n,u^n)\|_\infty \geq \|B(u^n,u^n)\|_1.$$

With $\|B(u^n,u^n)\|_1 = \frac{1}{d^2}(\mathbb{1}, B(u^n,u^n))$ we finally get:

$$\begin{aligned} d^2 C_0 &\geq (\mathbb{1}, B(u^n,u^n)) \\ &\overset{(10)}{=} (u^n, -\Delta_d u^n) \\ &\overset{(17)}{\geq} c_1(u^n,u^n). \\ &= d^2 c_1 \|u^n\|_2^2, \end{aligned}$$

where we used $u^n \in W$. Hence (u^n) is bounded. We claim next that every convergent subsequence (u^{n_k}) of (u^n) converges to a limit point u that fulfills (22). To see this we can adopt the proof of Theorem 4.1 numbers (29)-(33) setting $\tilde{p} = 2$.

Let P be the set of critical points. Assuming that all these critical points are isolated we have some $\delta > 0$ with

$$B_{3\delta}(u_i) \cap B_{3\delta}(u_j) = \emptyset \text{ for all } u_i, u_j \in P, i \neq j.$$

Let $\sigma(u^0) \subset P$ be the set of limit points of (u^n). Suppose $\hat{u} \in \sigma(u^0)$ and let (u^{n_q}) be a subsequence of (u^n) that converges to $\hat{u}$. We will show now that then also (u^{n_q+1}) converges to $\hat{u}$ and consequently the whole sequence, i.e. $\sigma(u^0)$ consists of the singleton $\hat{u}$.
As in the proof of Theorem 4.1 repeated use of (20) and $\Delta t^n < C$ for all n yields:

$$\|u^{n+1} - u^n\|_2 \overset{n\to\infty}{\longrightarrow} 0.$$

Therefore we may assume that $\|u^{n+1} - u^n\|_2 \leq \delta$ for all n. Hence $(u^{n_q+1}) \subset B_{2\delta}(\hat{u})$ for all $q > q_0(\delta)$:

$$\|\hat{u} - u^{n_q+1}\|_2 \leq \|\hat{u} - u^{n_q}\|_2 + \|u^{n_q} - u^{n_q+1}\|_2 \leq 2\delta.$$

Suppose (u^{n_q+1}) does not converge to $\hat{u}$. Then we get for some $\eta > 0$ a subsequence $(u^{n_k}) \subset B_{2\delta}(\hat{u}) \backslash B_\eta(\hat{u})$ of (u^{n_q+1}). This subsequence in turn possesses a convergent subsequence with limit point $u^* \in \overline{B_{2\delta}(\hat{u})} \backslash B_\eta(\hat{u}) \subset B_{3\delta}(\hat{u})$ but not equal to $\hat{u}$, which is a contradiction to $\hat{u}$ being isolated.

□

5 Computations

In our computations we used finite differences on a 50 x 50 grid for the spatial discretization of (18). The linear system in (18) is solved by the CGS-Method (iterative scheme).

We tested the following parameter constellations:

	α	γ	method
(a)	0.5	0.01	almost TV
(b)	1	1	Perona-Malik
(c)	0.5	1	regularized TV
(d)	2	1	unnamed

For example (1)(a) means: Use Method (1) and the parameters $\alpha = 0.5$, $\gamma = 0.01$. The free parameters are ε and λ_1 for Method (1)(respectively λ_2 for Method (2)).

We note that Method (a) is used to mimic the (unregularized) (TV)-model ($g(s^2) = \frac{1}{|s|}$). Method (d) has not been given a name in the literature so far as well as all models with $\varepsilon > 0$ regularization.

In order to compare the Methods (1)(a)-(2)(d) among each other, we use our knowledge of u_{tar}. The closer u gets to u_{tar} at the stopping time t^* , the better the method works. We investigated the following three norms:

$$\|u(t^*) - u_{tar}\|_1, \|u(t^*) - u_{tar}\|_2 \text{ , and } \|u(t^*) - u_{tar}\|_\infty.$$

After the computations we saw that for Methods (a)-(c), the best choice of the parameter $\lambda_i, i = 1, 2$, is somewhere intermediate (not too small and not too large), whereas Method (d) seems to prefer λ_i small. The range of $\varepsilon > 0$ (we tested $\varepsilon = 10^{-4}$ and 10^{-3}), where our numerical scheme converged showed no significant differences in the result compared to $\varepsilon = 0$. For ε larger than 10^{-3}, further computations have to be carried out to be able to judge any significant effect.
The $\|.\|_2$ results seem to be the most natural to compare the different methods. We observed that the optimal $\|.\|_2$ value for each of the methods (1)-(2), (a)-(d) is within a small range from 0.03 to 0.07. Therefore, from this point of view none of the methods seems to be significantly better than the others. However, knowing the best choice of λ_1/λ_2 is a critical issue, since the ratio of the worst over the best $\|.\|_2$ result reaches almost the factor 10. Therefore (d) could be the method of choice for a real application, because here the results are uniformly quite good. Also comparing the data as a whole, Method (2) is slightly better than Method (1). The whole data of our computations can be found in [11], Section 13.

In Figure 3 we show the optimal result $(2(a), \varepsilon = 0.001, \lambda_2 = 100)$.
It looks kind of rough. This is especially true when compared with Figure 4 (Method 2, $\alpha = 1, \gamma = 100, \varepsilon = 0, \lambda_2 = 1000$).

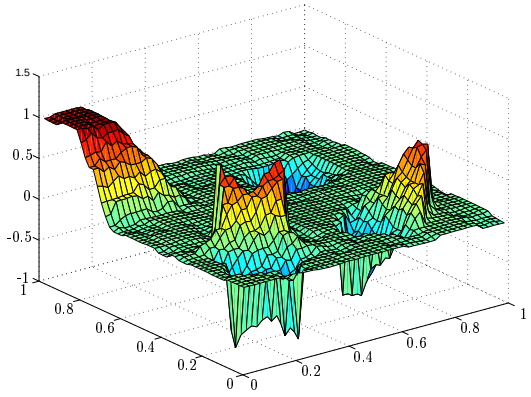

Figure 3: u_{opt}

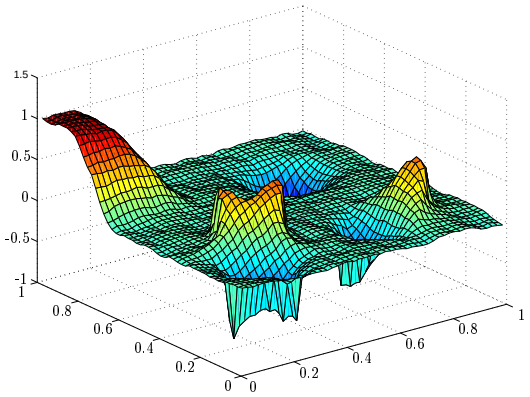

Figure 4: u_{smooth}

In Figure 4 the image got smoothed out much better, but the result is further away from the target u_{tar} as we can see when looking at the difference

$$u(t^*) - u_{tar}$$

in Figure 5 and 6.
We have $\|u_{opt} - u_{tar}\|_2 = 0.0329$ and $\|u_{smooth} - u_{tar}\|_2 = 0.0395$.

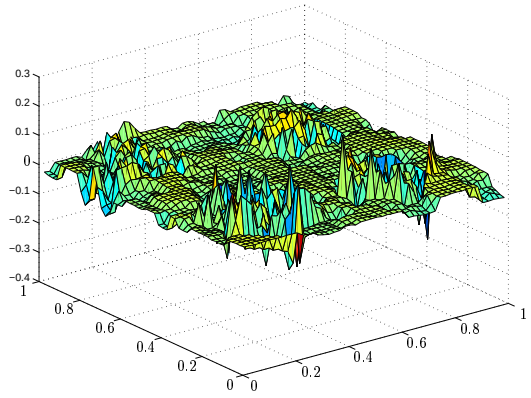

Figure 5: $u_{opt} - u_{tar}$

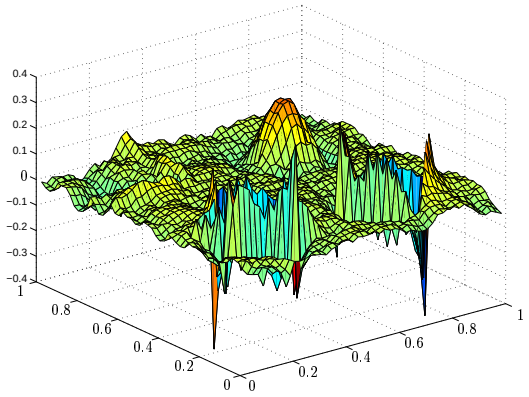

Figure 6: $u_{smooth} - u_{tar}$

References

[1] G. Aubert, P. Kornprobst. *Mathematical Problems in Image Processing.* Springer, 2002.

[2] J.W. Barrett, C.M. Elliott. Finite element approximation of a free boundary problem arising in the theory of liquid drops and plasma physics. *RAIRO, Modelisation Math. Anal. Numer. 25, No.2, 213-252, 1991.*

[3] C. Dolcetta, R. Ferretti. Optimal stopping time formulation of adaptive image filtering. *Appl. Math. Optimization 43, No.3, 245-258, 2001.*

[4] C.M. Elliott. The Cahn-Hilliard model for the kinetics of phase separation. *'Mathematical Models for Phase Change Problems'. ed. J. F. Rodrigues, International Series of Numerical Mathematics 88, Birkhäuser Verlag, 35-73, 1989.*

[5] C.M. Elliott, A.M. Stuart. The global dynamics of discrete semilinear parabolic equations. *SIAM, J. Numer. Anal. 30, No.6, 1622-1663, 1993.*

[6] W. Hackbusch. *Theorie und Numerik elliptischer Differentialgleichungen.* Zweite Auflage, Teubner, 1996.

[7] A. Meister. *Numerik linearer Gleichungssysteme.* Vieweg, 1999.

[8] S. Osher, R. Fedkiw. *Level Set Methods and Dynamic Implicit Surfaces.* Springer, 2003.

[9] P. Perona, J. Malik. *Scale-space and edge detection using anisotropic diffusion.* IEEE Transaction on Pattern Analysis and Machine Intelligence, July 1990. Springer, 1988.

[10] L. Rudin, S. Osher, E. Fatemi. Nonlinear Total Variation Based Noise Removal Algorithms. *Physica D 60, No.1-4, 259-268, 1992.*

[11] B. Gawron. Numerical study of image processing models. *Diplomarbeit, Universität Augsburg, 2002.*