

Universität des Saarlandes



Fachrichtung 6.1 – Mathematik

Preprint Nr. 148

**Numerical Aspects of TV Flow**

Andrea Bürgel, Michael Breuß, Thomas Sonar, Thomas Brox and  
Joachim Weickert

Saarbrücken 2005



## Numerical Aspects of TV Flow

**Andrea Bürgel**

Computational Mathematics, TU Braunschweig  
Pockelsstraße 14  
38106 Braunschweig, Germany  
a.buergel@tu-bs.de

**Michael Breuß**

Computational Mathematics, TU Braunschweig  
Pockelsstraße 14  
38106 Braunschweig, Germany  
m.breuss@tu-bs.de

**Thomas Sonar**

Computational Mathematics, TU Braunschweig  
Pockelsstraße 14  
38106 Braunschweig, Germany  
t.sonar@tu-bs.de

**Thomas Brox**

Mathematical Image Analysis Group,  
Faculty of Mathematics and Computer Science,  
Saarland University, Building 27.1,  
66041 Saarbrücken, Germany  
brox@mia.uni-saarland.de

**Joachim Weickert**

Mathematical Image Analysis Group,  
Faculty of Mathematics and Computer Science,  
Saarland University, Building 27,  
66041 Saarbrücken, Germany  
weickert@mia.uni-saarland.de

Edited by  
FR 6.1 – Mathematik  
Universität des Saarlandes  
Postfach 15 11 50  
66041 Saarbrücken  
Germany

Fax: + 49 681 302 4443  
e-Mail: [preprint@math.uni-sb.de](mailto:preprint@math.uni-sb.de)  
WWW: <http://www.math.uni-sb.de/>

## Abstract

The singular diffusion equation called total variation (TV) flow plays an important role in image processing and appears to be suitable for reducing oscillations in other types of data. Due to its singularity for zero gradients, numerical discretizations have to be chosen with care. We discuss different ways to implement TV flow numerically, and we show that a number of discrete versions of this equation may introduce oscillations such that the scheme is in general not TV-decreasing. On the other hand, we show that TV flow may act self-stabilising: even if the total variation increases by the filtering process, the resulting oscillations remain bounded by a constant that is proportional to the ratio of mesh widths. For our analysis we restrict ourselves to the one-dimensional setting.

*Key Words:* TV flow, singular diffusion equation, finite difference methods, numerical stability  
*AMS subject classification:* 65M06, 65M12, 94A08

## Contents

<b>1</b>	<b>Introduction</b>	<b>2</b>
<b>2</b>	<b>Alternative formulations of TV flow</b>	<b>2</b>
2.1	Regularisation of the diffusivity . . . . .	2
2.2	Direct use of the chain rule . . . . .	3
2.3	The chain rule through regularisation . . . . .	4
<b>3</b>	<b>Numerical implementations</b>	<b>5</b>
3.1	Direct discretisation . . . . .	6
3.2	Discretisation of TV flow with regularised diffusivity . . . . .	9
3.3	Discretisation based on analytical solutions for the two gridpoints case . . . . .	10
3.4	Discretisation by means of the direct chain rule . . . . .	12
3.5	Explicit discretisation by means of the regularised chain rule . . . . .	12
3.6	Implicit discretisation by means of the regularised chain rule . . . . .	14
<b>4</b>	<b>Summary and Conclusions</b>	<b>19</b>

# 1 Introduction

One of the paradigms in total variation (TV) diminishing filters can be found in the equation

$$\partial_t u = \nabla \cdot \left( \frac{\nabla u}{|\nabla u|} \right). \quad (1)$$

This equation has gained much interest in the image processing community, where it is called *TV flow* or *TV diffusion*. It requires no additional parameters, it is well-posed [3, 5, 8], scale-invariant [2], it preserves the shape of some objects [5], and it leads to constant signals in *finite* time [4]. In [21] it has further been shown that in the space-discrete 1-D case, TV flow with initial data  $f$  is equivalent to the so-called *TV regularisation* [1, 6, 17], where one minimises the energy functional

$$E(u) := \int \left( (u - f)^2 + 2t|\partial_x u| \right) dx.$$

This variational method is well-known for its good denoising capabilities.

One obvious difficulty of TV flow arises when  $\partial_x u$  becomes zero, since this creates an unbounded diffusivity  $1/|\partial_x u|$ . Sophisticated numerical ways of dealing with such nonlinear singular diffusion equations have been developed in the past, see e.g. [7, 8, 21]. They are either based on regularizing strategies [8] or on the simulation of the analytic behaviour of TV flow [7, 21].

The goal of the present paper is to discuss a number of numerical problems that can arise with such discretisations, and to investigate if similar problems arise in the context of regularized versions of (1). For the sake of simplicity, we restrict ourselves thereby to the 1-D case. While the 1-D case is well-understood in the space-discrete case [20, 21], it will reveal many aspects that are also useful in a higher dimensional setting.

We show that several discrete versions of TV flow may introduce oscillations such that the scheme in general is not TV-decreasing. On the other hand, however, we show that TV flow has a remarkable self-stabilising property: even if the total variation increases, the resulting oscillations remain bounded by a constant that is proportional to the time step size.

The remainder of the paper is organised as follows. In Section 2 we derive a number of alternative formulations of the TV flow. These formulations inspire different numerical algorithms which are analysed and compared in Section 3. The paper is concluded with a summary in Section 4.

## 2 Alternative formulations of TV flow

We now investigate some alternative formulations for the one-dimensional TV flow equation

$$\partial_t u = \partial_x (\text{sign } \partial_x u). \quad (2)$$

These formulations will be exploited by some numerical implementation of TV flow later in Section 3.

### 2.1 Regularisation of the diffusivity

A common strategy to circumvent the degeneracy of the TV flow is to replace it by the regularisation

$$\partial_t u = \partial_x \left( \frac{\partial_x u}{\sqrt{\beta^2 + (\partial_x u)^2}} \right) \quad (3)$$

with some small and fixed regularisation parameter  $\beta > 0$ . In this way the diffusivity is bounded by  $\frac{1}{\beta}$ . However, this regularised equation is only an approximation of TV flow. The equivalence to TV regularisation, for instance, can no longer be expected to be true. Feng and Prohl have shown that given a domain  $\Omega \subset R^N$ , one can establish, for almost every time  $t$  in a given interval  $(0, T)$ , strong convergence of a weak solution  $u^\beta(t)$  to  $u(t)$  in the  $L_p$ -norm for any  $p \in [1, N/(N - 1))$  [8].

## 2.2 Direct use of the chain rule

We rewrite the right-hand side of Equation (2) as

$$\text{sign } w(x) = (H \circ w)(x) - (H \circ -w)(x)$$

where  $H$  denotes the Heaviside function. Let  $x^*$  be a single zero of  $w$  in an interval  $I^* := [a, b] := [x^* - \varepsilon, x^* + \varepsilon]$  with  $\varepsilon > 0$ . According to [13] we will view the mapping  $w$  as a change of coordinates, and hence we have to assume that  $w' \neq 0$  holds in the vicinity of  $x^*$ . Thus we assume that  $w'$  is of only one sign in  $I^*$ .

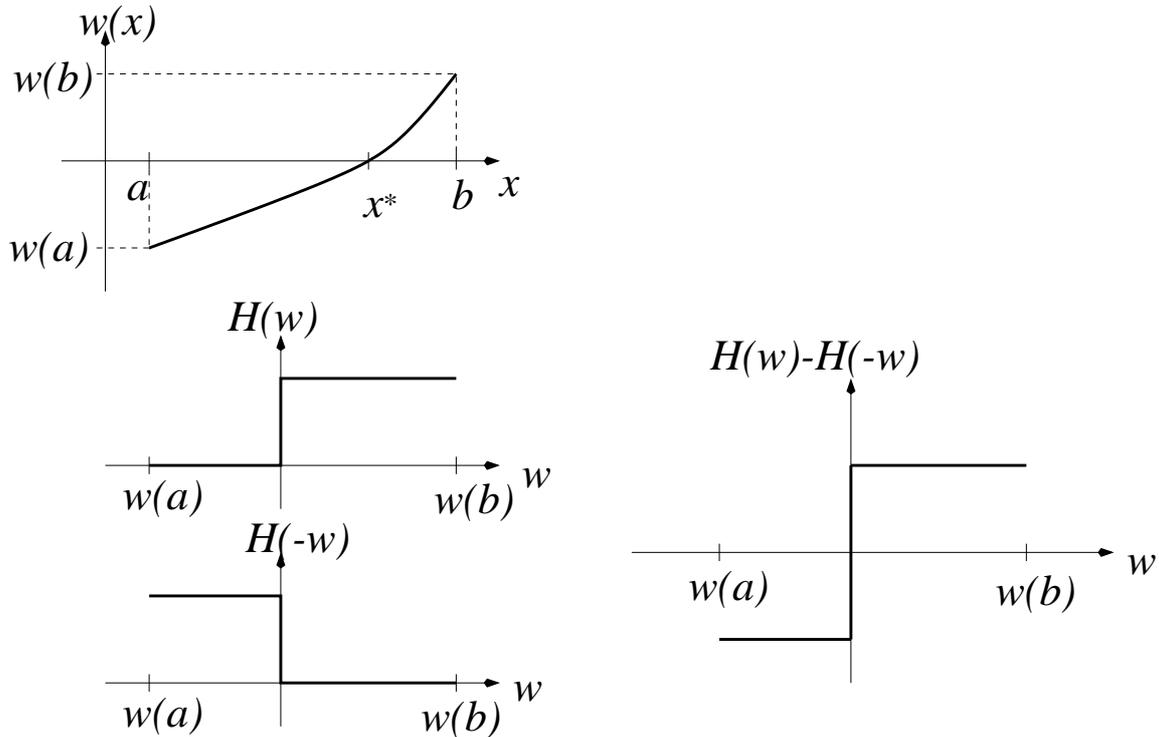


Figure 1: The case  $w' > 0$

(i) **Assume  $w'(x) > 0$  for all  $x \in I^*$ :** See Figure 1 for this situation. In order to express  $H \circ \pm w$  as a function of  $x$  we note that the difference  $H(x - x^*) - H(x^* - x)$  is exactly  $(H \circ w)(x) - (H \circ -w)(x)$  as can be seen from Figure 2. Hence, we end up with

$$\text{sign } w(x) = H(x - x^*) - H(x^* - x)$$

which now can be differentiated in the sense of distributions to yield

$$\partial_x \text{sign } w(x) = 1 \cdot \delta(x - x^*) - (-1) \cdot \delta(x^* - x) = 2\delta(x - x^*). \quad (4)$$

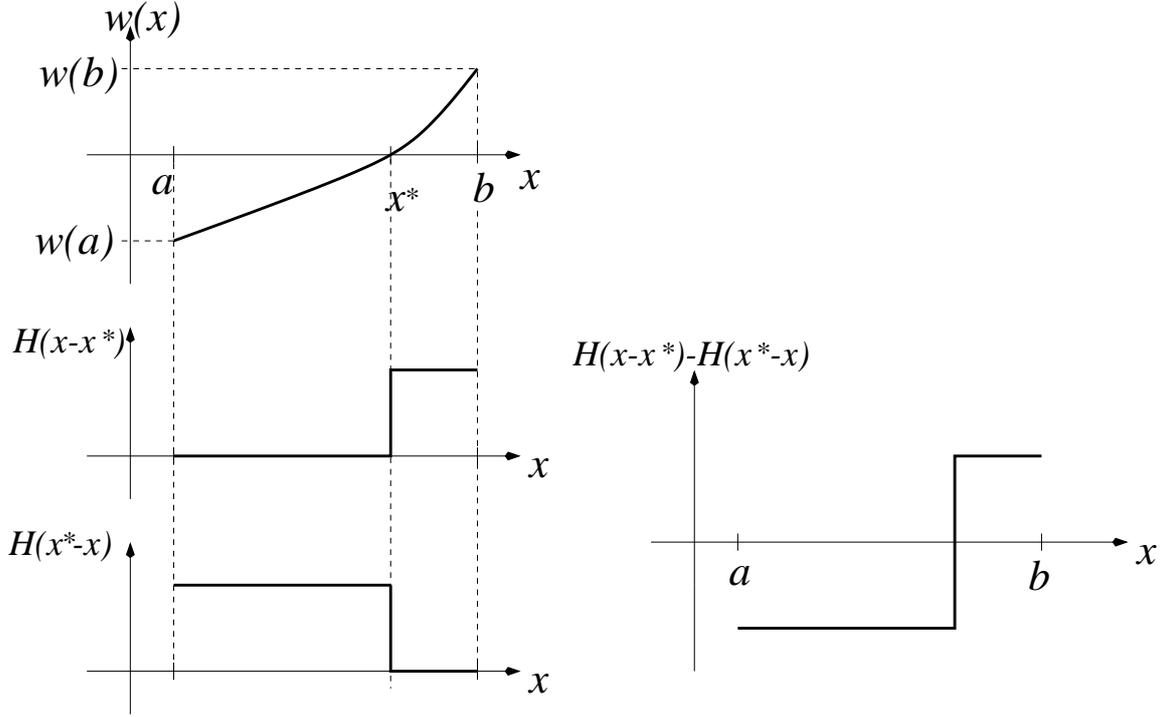


Figure 2: A useful sketch illuminating the computations for the case  $w' > 0$

(ii) **Assume  $w'(x) < 0$  for all  $x \in I^*$ :** Proceeding analogously to case (i) one obtains

$$\begin{aligned} \partial_x \text{sign } w(x) &= \partial_x (H(x^* - x) - H(x - x^*)) \\ &= (-1) \cdot \delta(x^* - x) - 1 \cdot \delta(x - x^*) = -2\delta(x - x^*). \end{aligned} \quad (5)$$

Including cases (i) and (ii) in one case we have shown

**Lemma 2.1** *Let  $w$  be a smooth function with a single zero  $x^*$  in the interval  $[x^* - \varepsilon, x^* + \varepsilon]$ ,  $\varepsilon > 0$  in which  $w(x) \neq 0$  holds. Then it holds*

$$\partial_x \text{sign } w(x) = \text{sign } w'(x^*) \cdot 2\delta(x - x^*) \quad (6)$$

*in the sense of distributions.*

Thus, TV flow may be rewritten as

$$\partial_t u = 2\delta_{(\partial_x u=0)} \cdot \partial_x^2 u,$$

which is to be understood in the sense of distributions.

### 2.3 The chain rule through regularisation

An alternative way to derive the chain rule (6) consists in regularising the sign function by means of

$$\widetilde{\text{sign}}^m w(x) := \frac{2}{\pi} \arctan(mw(x)), \quad m \in \mathbb{N}. \quad (7)$$

The derivative is given by

$$\partial_x \widetilde{\text{sign}}^m w(x) = \frac{2m}{\pi(1 + m^2 w(x)^2)} w'(x)$$

where the function  $\widetilde{\delta}^m(w(x)) := \frac{m}{\pi(1+m^2w(x)^2)}$  is a member of a delta sequence so that

$$\widetilde{\delta}^m(w(x)) = \frac{m}{\pi(1+m^2w(x)^2)} \xrightarrow{\mathcal{D}'} \delta(w(x))$$

and hence

$$\partial_x \widetilde{\text{sign}}^m w(x) \xrightarrow{\mathcal{D}'} 2\delta(w(x))w'(x).$$

If we again assume that  $w$  has an isolated zero at  $x^*$  and that  $w' \neq 0$  in  $I^*$  it follows from the transformation properties of the delta distribution (see [12], [13])

$$\delta(w(x)) = \frac{\delta(x - x^*)}{|w'(x^*)|}$$

that we arrive at

$$\partial_x \widetilde{\text{sign}}^m w(x) \xrightarrow{\mathcal{D}'} 2\delta(w(x))w'(x) = \frac{w'(x)}{|w'(x^*)|} \cdot 2\delta(x - x^*).$$

Since  $\delta(x - x^*) \neq 0$  only for  $x = x^*$  this is identical to

$$\partial_x \widetilde{\text{sign}}^m w(x) \xrightarrow{\mathcal{D}'} \frac{w'(x^*)}{|w'(x^*)|} \cdot 2\delta(x - x^*) = \text{sign } w'(x^*) \cdot 2\delta(x - x^*)$$

and this is again (6).

### 3 Numerical implementations

In order to construct discrete filters from (2) we have described several starting points in the last section, namely

1. the original TV flow equation (2):  $\partial_t u = \partial_x (\text{sign } \partial_x u)$ ,
2. its variant (3) with regularised diffusivity:  $\partial_t u = \partial_x \left( \frac{\partial_x u}{\sqrt{\beta^2 + (\partial_x u)^2}} \right)$ ,
3. the equation  $\partial_t u = 2 \text{sign} (\partial_x^2 u(x^*)) \cdot \delta(x - x^*)$  using (6), and
4. the regularised equation  $\partial_t u = \partial_x \left( \widetilde{\text{sign}}^m (\partial_x u) \right)$ , see (7).

Besides a discussion of the schemes arising by these formulations, we also compare the derived methods by means of an oscillatory test signal shown in Figure 3. The signal is the output of an oscillatory numerical method for approximating hyperbolic conservation laws, the so-called Lax-Wendroff scheme. It contains 200 discrete values with the equidistant grid-size  $\delta x = \pi/100$ .

We now turn to the discretisations of (1) discussed in the following. While we keep the spatial grid of the test signal we consider different time step sizes  $\Delta t$  for numerically approximating (1). Thus,  $\lambda := \Delta t / \Delta x$  is changed. The parameters in use are indicated in the appropriate fashion in the subsequent text and vary between  $\lambda = 0.00005$  and  $\lambda = 5$ .

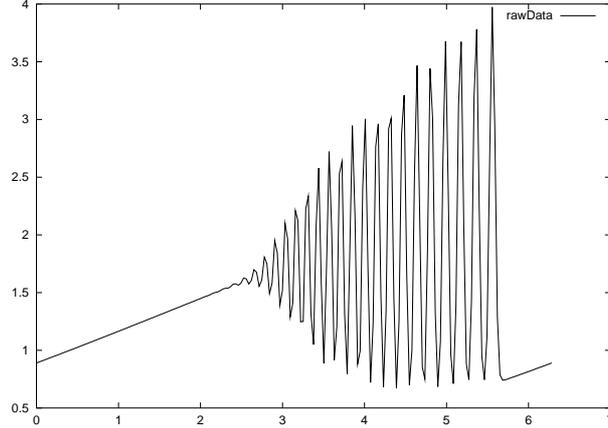


Figure 3: Signal to be filtered

### 3.1 Direct discretisation

We start with discretising equation (2) directly. On a grid

$$\mathbb{G} := \{(x, t) \mid x = j\Delta x, t = n\Delta t, j \in \mathbb{Z}, n \in \mathbb{N}, \Delta x > 0, \Delta t > 0\}$$

we denote by  $U_j^n$  an approximation to  $u(j\Delta x, n\Delta t)$ . For the sake of simplicity the time derivative will always be the simple forward difference

$$\partial_t u(j\Delta x, n\Delta t) = \frac{U_j^{n+1} - U_j^n}{\Delta t} + \mathcal{O}(\Delta t).$$

Discretising

$$\partial_x u(j\Delta x, n\Delta t) = \frac{U_{j+1}^n - U_{j-1}^n}{2\Delta x} + \mathcal{O}(\Delta x^2)$$

and replacing the outer derivative in (2) by the same difference formula we arrive at a discrete analogue of (2) in form of

$$U_j^{n+1} = U_j^n + \frac{\Delta t}{2\Delta x} (\text{sign}(U_{j+2}^n - U_j^n) - \text{sign}(U_j^n - U_{j-2}^n)).$$

This discrete analogon obviously suffers from an odd-even decoupling which can be completely removed by decreasing the width of the difference stencil, i.e.,

$$U_j^{n+1} = U_j^n + \frac{\Delta t}{\Delta x} (\text{sign}(U_{j+1}^n - U_j^n) - \text{sign}(U_j^n - U_{j-1}^n)), \quad (8)$$

defining the discrete operator  $C(\Delta t)$  via

$$U_j^{n+1} = C(\Delta t)U_j^n. \quad (9)$$

In order to show properties of the discrete operator  $C(\Delta t)$  we discuss the defined numerical default test case.

Figure 4 shows the numerical results after 10000 and 20000 time steps using  $\lambda = 0.0005$ , respectively. As is readily observed, the filter decreases the total variation of the input signal significantly. Note also that the whole signal shrinks and seems on its way to a (still far away) steady state constant. However, we also observe some high-frequent oscillations. The plot of

the total variation against the number of time steps displayed in Figure 5 (left) assures us that in fact the total variation, measured by  $\sum_{i=1}^{199} |U_{i+1}^n - U_i^n|$  on time level  $n$ , is diminished. However, a closer look at the total variation between the time steps 19900 and 20000 reveals that we still have oscillations left which are not decreased in a monotone fashion, see Figure 5 (right).

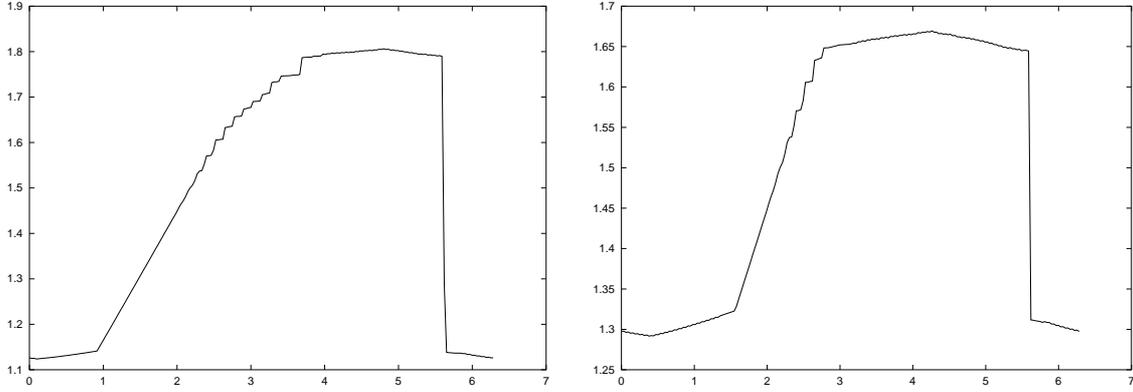


Figure 4: Numerical results using  $\lambda = 0.0005$  after (left) 10000 and (right) 20000 filtering steps, respectively

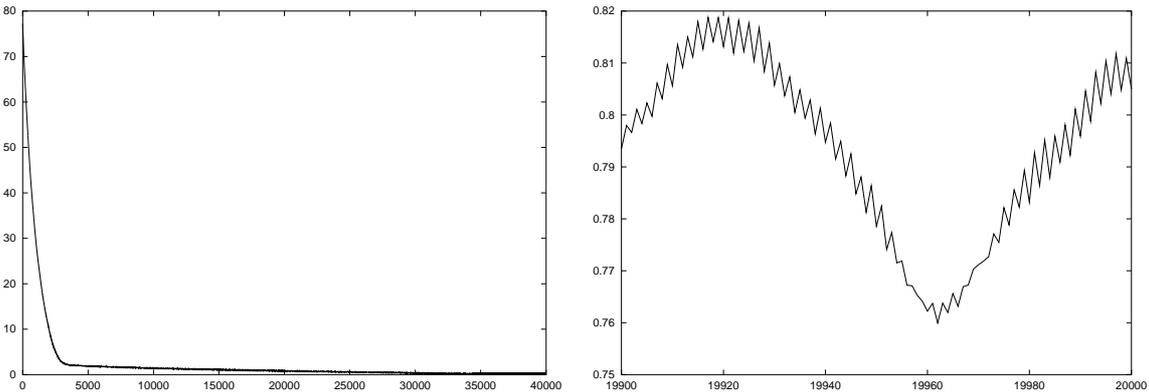


Figure 5: Total variation plotted against the number of time steps for  $\lambda = 0.0005$

An increase of the grid coefficient to  $\lambda = 0.5$  yields after 10000 time steps the numerical solution shown in Figure 6. The filtered signal is noisy everywhere and the total variation increases to a level where oscillations around a certain “mean total variation” take place. Note that the high-frequency oscillations in the filtered signal seem to be of order  $\lambda$ .

A closer look at the first 100 time steps as well as to the time steps between 7000 and 7010, see Figure 7, reveals the overall structure of the evolution of the total variation.

The grid coefficient  $\lambda$  can be further increased without restriction in sharp contrast to a severe CFL condition which one would expect for an explicit time discretisation. Figure 8 shows the filtered signal after 10000 time steps and the corresponding evolution of the total variation for  $\lambda = 5$ . The filtered signal shows huge oscillations while the total variation is increased far beyond the variation of the unfiltered signal!

The investigated numerical behaviour motivates us to formulate the following general assertion.

**Theorem 3.1** Consider the discrete operator  $C(\Delta t)$  as defined by (8), (9). Let  $U_j^n$  be a local

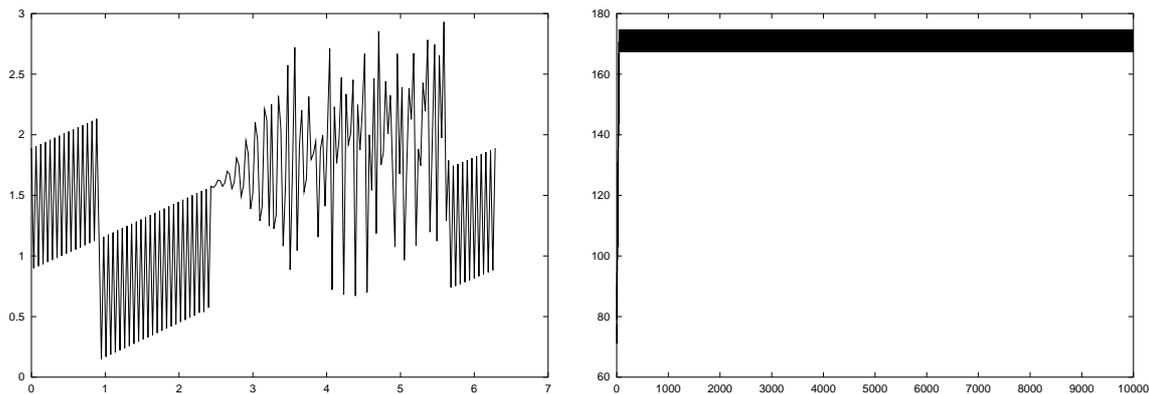


Figure 6: Filtered signal (left) and total variation vs. time steps (right) using  $\lambda = 0.5$

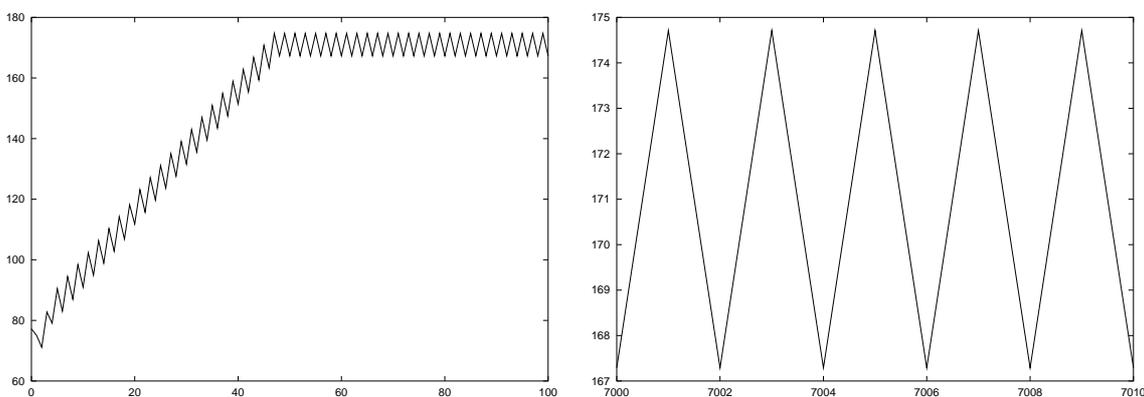


Figure 7: Total variation during the first 100 time steps (left) and between time steps 7000 and 7010 (right) for  $\lambda = 0.5$

*maximum (minimum) and let*

$$0 < \max\{|U_j^n - U_{j-1}^n|, |U_j^n - U_{j+1}^n|\} < \frac{\Delta t}{\Delta x}$$

*hold. Then  $U_j^{n+1}$  is a new local minimum (maximum) and the arising oscillation at the grid point  $x_j$  is of amplitude  $2\lambda = 2\Delta t/\Delta x$ .*

**Proof** Assume, without loss of generality, a maximum at  $x_j$ , i.e.,  $U_j^n > U_{j-1}^n$  and  $U_j^n > U_{j+1}^n$ . Then, according to (8),

$$\begin{aligned} U_j^{n+1} &= U_j^n + \frac{\Delta t}{\Delta x} \left( \underbrace{\text{sign}(U_{j+1}^n - U_j^n)}_{=-1} - \underbrace{\text{sign}(U_j^n - U_{j-1}^n)}_{=1} \right) \\ &= U_j^n - 2\frac{\Delta t}{\Delta x} = U_j^n - 2\lambda. \end{aligned}$$

Obviously, if  $\max\{|U_j^n - U_{j-1}^n|, |U_j^n - U_{j+1}^n|\} < 2\frac{\Delta t}{\Delta x}$  then the maximum is turned into a minimum. If we assume a minimum at  $x_j$  the same proof applies.  $\blacksquare$

While the above theorem is concerned with the oscillatory properties, we summarize the observed behaviour with respect to the total variation as follows.

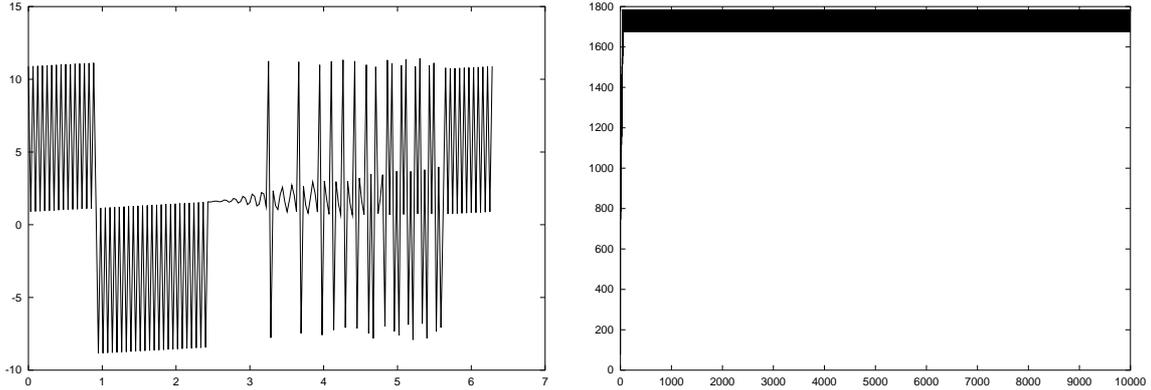


Figure 8: Filtered signal (left) and total variation vs. time steps (right) for  $\lambda = 5$

**Corollary 3.1** *The discrete operator  $C(\Delta t)$  as defined in (8), (9), is in contrast to its differential counterpart (2) TV-increasing or decreasing, depending on the choice of the grid coefficient  $\lambda$ . However, the total variation of the steady state is bounded in terms of  $\lambda$ , hence even in case of an explicit time stepping scheme with large time step sizes, oscillations will remain bounded: from (8), (9) we obtain*

$$TV(U^{n+1}) := \sum |U_{j+1}^{n+1} - U_j^{n+1}| \leq \left(1 + \frac{\Delta t}{\Delta x} \|C\|\right) TV(U^n),$$

where  $C$  is given by identifying (8) with

$$U_j^{n+1} = U_j^n + \frac{\Delta t}{\Delta x} C U_j^n,$$

yielding  $0 \leq \|C\| \leq 2$ .

We want to emphasize that, though the total variation can increase, the corollary is a stability result: the scheme satisfies the essence of general stability, i.e., a norm of interest is allowed to grow if the growth is uniformly bounded; compare e.g. [14].

To remedy the problem with nondecreasing total variation, and to obtain stability in a stronger sense, the discrete equation (8) might be applied with small  $\lambda$  until the lower bound (determined by  $\lambda$ ) of total variation is reached. Then one switches to a much simpler filter, say, the linear heat equation, in order to suppress the remaining oscillations. Interestingly, the strategy just described resembles a recently developed algorithm, see Steidl et al. [21], derived from carefully examining a two-gridpoints scenario. We will investigate this scheme later on.

### 3.2 Discretisation of TV flow with regularised diffusivity

Let us now consider a discretisation of the regularised TV flow

$$\partial_t = \partial_x \left( \frac{\partial_x u}{\sqrt{\beta^2 + (\partial_x u)^2}} \right). \quad (10)$$

An explicit finite difference scheme for this equation is given by

$$U_j^{n+1} = U_j^n + \frac{\Delta t}{\Delta x^2} \left( \frac{U_{j+1}^n - U_j^n}{\sqrt{\beta^2 + \left(\frac{U_{j+1}^n - U_j^n}{\Delta x}\right)^2}} - \frac{U_j^n - U_{j-1}^n}{\sqrt{\beta^2 + \left(\frac{U_j^n - U_{j-1}^n}{\Delta x}\right)^2}} \right). \quad (11)$$

It is not difficult to see that for

$$\frac{\Delta t}{\Delta x^2} \leq \frac{\beta}{2}$$

this discretisation is stable in a very strong sense: it satisfies a discrete maximum–minimum principle. Moreover, for

$$\frac{\Delta t}{\Delta x^2} \leq \frac{\beta}{4}$$

it is sign stable; cf [9]. In this case it acts TV diminishing.

However, these stability bounds show that, when we approach the original TV flow by using the limit for the numerical parameter  $\beta \rightarrow 0$ , the time step size of the explicit scheme has to become infinitesimally small. Thus it seems attractive to use implicit schemes instead. In this case we will see later on that the condition number of the resulting linear systems of equations may become unbounded.

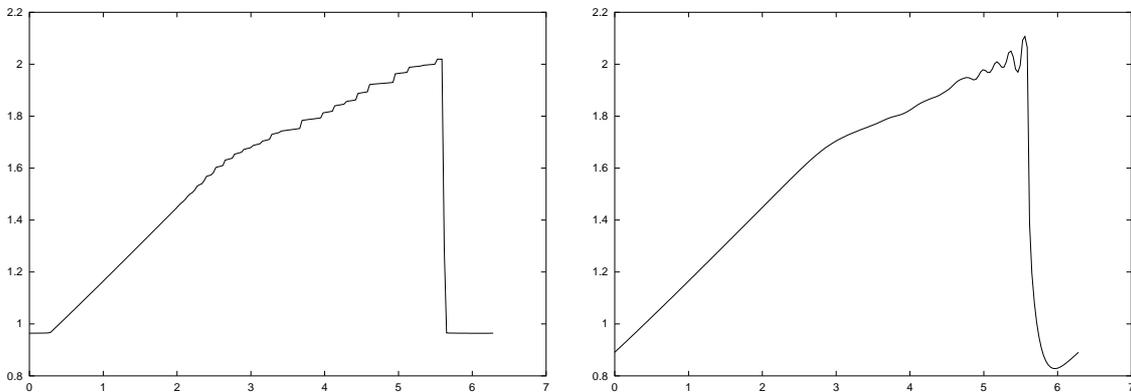


Figure 9: Signal filtered with  $\beta$ -regularised diffusivity and  $\beta = 0.0001$  (left) and signal filtered with  $\beta = 0.1$  (right),  $t = 1.75$ . Larger  $\beta$  introduce blurring effects and yield a slightly slower removal of oscillations.

Figure 9 shows the result obtained when  $\beta$ -regularized TV flow is applied to the test signal from Figure 3. On the left, a small  $\beta$  has been used, thus the result reflects the behaviour of unregularized TV flow. In order to ensure stability,  $\lambda$  had to be constrained to  $\lambda = 0.00005$ , so 35000 time steps were necessary to reach the total diffusion time of  $t = 1.75$ . On the right, the result for the same experiment with  $\beta = 0.1$  is depicted. This allows for an increase of  $\lambda$  to  $\lambda = 0.05$ , and consequently reduces the number of time steps to 35. However, one can see that this reduction in computational effort introduces a blurring effect that is not part of TV flow. Furthermore, one can observe that the bounded diffusivities ( $\beta$ -regularisation prevents unbounded diffusivities) lead to a slower removal of oscillations.

### 3.3 Discretisation based on analytical solutions for the two grid-points case

Recently an alternative to the previous discretisation has been introduced [21]. It is based on combining analytical solutions to the two gridpoints case in a suitable way, and it does not require any regularisation. Let us sketch the main idea.

One can split the 1-D space-discrete TV flow in some pixel  $i$  into an interaction between the two-pixel pairs  $(U_{i-1}, U_i)$  and  $(U_i, U_{i+1})$ . Each two-pixel scenario can be described by a dynamical system with homogeneous Neumann boundary conditions, where the singular

diffusivity lead to a discontinuous right hand side. It has a unique analytical solution that is not difficult to compute and that satisfies a maximum–minimum principle. Using such local analytical solutions as building blocks for a numerical scheme leads to

$$U_i^{k+1} = U_i^k + \frac{\Delta t}{\Delta x} \text{sign}(U_{i+1}^k - U_i^k) \min\left(1, \frac{\Delta x}{4\Delta t} |U_{i+1}^k - U_i^k|\right) - \frac{\Delta t}{\Delta x} \text{sign}(U_i^k - U_{i-1}^k) \min\left(1, \frac{\Delta x}{4\Delta t} |U_i^k - U_{i-1}^k|\right).$$

This scheme can be regarded as a stabilisation of the direct scheme (8) for the TV diffusion. Although it is an explicit scheme, it is unconditionally stable, since it is a convex combination of the analytical two-gridpoints interactions that satisfy a maximum–minimum principle. Therefore it also fulfills the extremum principle

$$\min_j F_j \leq U_i^{k+1} \leq \max_j F_j \quad \forall i.$$

In [21] it is shown that this two-gridpoint scheme is an  $O(\Delta t + \Delta x^2)$  approximation to the continuous TV diffusion for

$$\Delta t \leq \frac{\Delta x}{4} \min(|U_{i+1}^k - U_i^k|, |U_i^k - U_{i-1}^k|),$$

while for larger  $\Delta t$  it approximates a linear diffusion process. The latter happens in regions where the gradient is already close to zero. In this case, however, the visual differences between linear diffusion and TV diffusion are small.

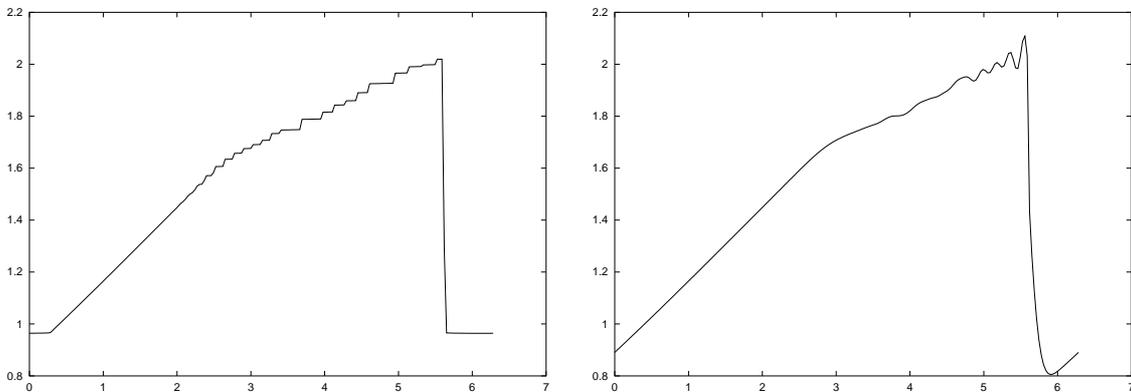


Figure 10: Signal filtered with two-gridpoint numerics and  $\lambda = 0.00005$  (left) and signal filtered with  $\lambda = 0.05$  (right),  $t = 1.75$ . Larger time steps introduce blurring effects and yield a slightly slower removal of oscillations.

Correspondingly to Figure 9, Figure 10 shows the result obtained with the two-gridpoints implementation of TV flow applied to the test signal from Figure 3. Here, instead of  $\beta$ , the time step size is the critical parameter that determines the accuracy. On the left,  $\lambda = 0.00005$  has been used, so like in the experiment with  $\beta$ -regularisation 35000 time steps were necessary to reach  $t = 1.75$ . Again, the corresponding result approximates TV flow very well, while the result on the right with  $\lambda = 0.05$  and 35 iterations reveals similar blurring effects as the result with correspondingly larger  $\beta$ .

This shows that both strategies to stabilize TV flow obviously work in a similar way. Both numerical schemes introduce homogeneous diffusion smoothing in order to prevent over- and

undershoots when the gradient tends to 0. While  $\beta$ -regularisation introduces the homogeneous type of smoothing by limiting the diffusivity at zero, the two-gridpoints numerics contains an averaging step that becomes relevant when  $\partial_x u$  is close to zero. In order to obtain solutions consistent with TV flow,  $\beta$  and the time step size have to be chosen sufficiently small.

### 3.4 Discretisation by means of the direct chain rule

Next we employ our chain rule (6). If  $x^*$  denotes a single zero of  $\partial_x u$ , then our filter equation on the differential level is given by

$$\partial_t u = 2\text{sign}(\partial_x^2 u(x^*)) \cdot \delta(x - x^*).$$

We detect zeros of  $\partial_x u$  on the discrete level by means of the finite difference product  $(U_{j+1}^n - U_j^n) \cdot (U_j^n - U_{j-1}^n)$  which will be non-positive in case of a zero. Hence, with the help of the Heaviside function and the finite difference operators introduced before, we obtain

$$\begin{aligned} & H(-(U_{j+1}^n - U_j^n) \cdot (U_j^n - U_{j-1}^n)) \\ &= \begin{cases} 1 & ; \text{ zero at } j \text{ in the discrete derivative,} \\ 0 & ; \text{ no zero at } j, \end{cases} \end{aligned}$$

which will be used as the detector. The second derivative is implemented by the central difference

$$\begin{aligned} & \partial_x^2 u(x^*) \Big|_{x=j\Delta x} \\ &= H(-(U_{j+1}^n - U_j^n) \cdot (U_j^n - U_{j-1}^n)) \cdot \left( \frac{U_{j+1}^n - 2U_j^n + U_{j-1}^n}{\Delta x^2} + \mathcal{O}(\Delta x^2) \right) \end{aligned}$$

while the Dirac functional at a zero  $x_j^*$  of the discrete  $u$ -derivative will be approximated by

$$\tilde{\delta}(x - x_j^*) := \begin{cases} \frac{1}{\Delta x} & ; x_j^* - \frac{\Delta x}{2} < x < x_j^* + \frac{\Delta x}{2}, \\ 0 & ; \text{ elsewhere,} \end{cases}$$

defining a valid delta sequence with  $\int \tilde{\delta}(x - x_j^*) dx = 1$ .

Not surprisingly, the numerical results computed by use of the resulting scheme are similar to the ones obtained by using the direct discretisation. Thus we do not present a corresponding figure. As in the latter case, the chain rule approach leads to oscillations that may increase the total variation.

### 3.5 Explicit discretisation by means of the regularised chain rule

Finally we consider the regularised equation

$$\partial_t u = \partial_x \left( \widetilde{\text{sign}}^m(\partial_x u) \right) = \frac{2}{\pi} \partial_x (\arctan(m\partial_x u)). \quad (12)$$

A discretised version can be obtained by applying the central difference operator  $(E - E^{-1})$  for the inner as well as for the outer derivative and then decreasing the width of the stencil in order to get rid of odd-even decoupling. This proceeding results in the method

$$U_j^{n+1} = U_j^n + \frac{2}{\pi} \lambda \left( \arctan(m(U_{j+1}^n - U_j^n)) - \arctan(m(U_j^n - U_{j-1}^n)) \right). \quad (13)$$

As can already be seen, the behaviour of the numerical solution can not be very much different from the behaviour of the numerical solutions we studied before. Since  $|\arctan(x)| < \frac{\pi}{2}$  the worst case in (13) occurs when the difference of the arctan-functions gives nearly  $\pm\pi$ . In this case we would get

$$U_i^{n+1} = U_i^n \pm 2\lambda$$

and again oscillations of size  $2\lambda$  would occur in the numerical solution.

In fact, computations with  $\lambda = 0.05$  and  $m = 1$  employing 10000 time steps show encouraging behaviour as shown in Figure 11. The filtered signal appears to be smooth. Increasing the

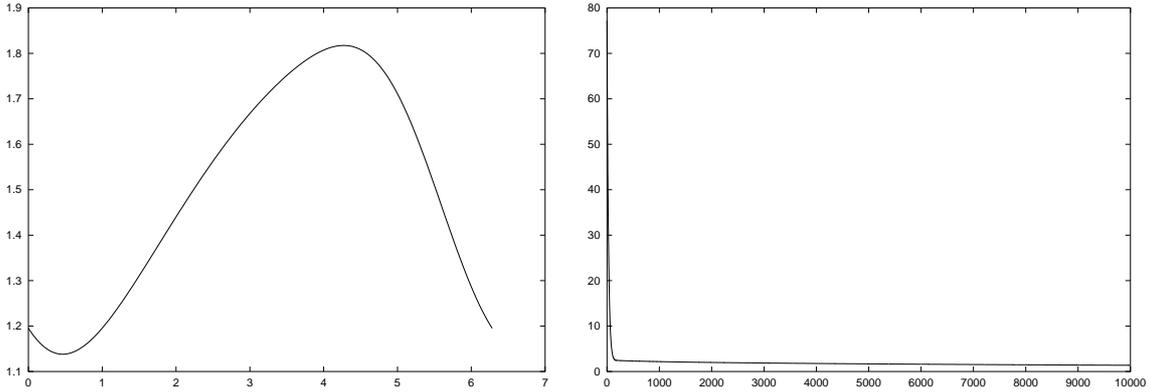


Figure 11: Filtered signal (left) and total variation vs. time steps (right) using  $\lambda = 0.05$  and  $m = 1$

regularisation parameter  $m$  to  $m = 100$  while retaining  $\lambda = 0.05$  results in the situation shown in Figure 12. Clearly, an increase of  $m$  drives the discrete equation towards the direct

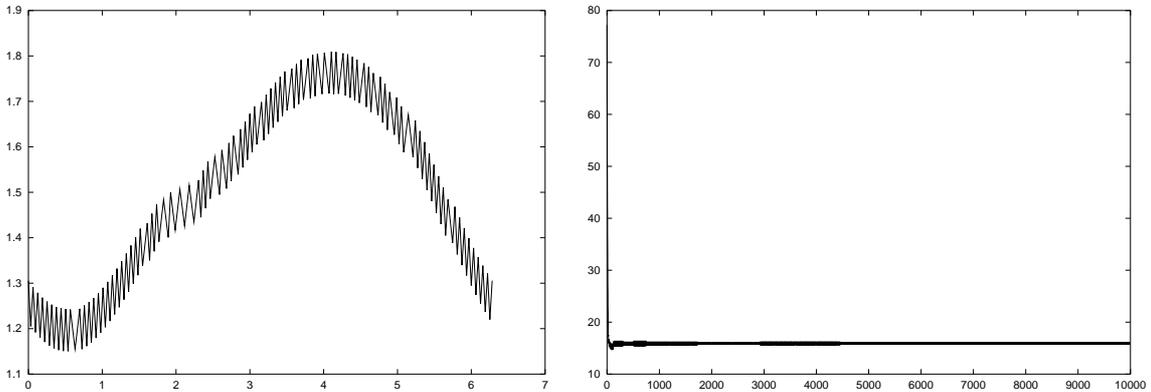


Figure 12: Filtered signal (left) and total variation vs. time steps (right) for  $\lambda = 0.05$  and  $m = 100$

discretisation (8) and oscillations occur exactly on the scale of  $\lambda$ . This is clearly seen by increasing  $\lambda$  to  $\lambda = 5$  while retaining  $m = 100$ , see Figure 13. The numerical solution is now completely spoiled by oscillations of height  $\lambda$ .

The discussed regularisation leads to a reasonable filter only if the regularisation parameter is small (and hence the amount of regularisation is large). The same is true if the derivative in (12) is carried out before discretisation. One then gets a regularisation of the Delta distribution which works nicely, but only in case of very small regularisation parameters.

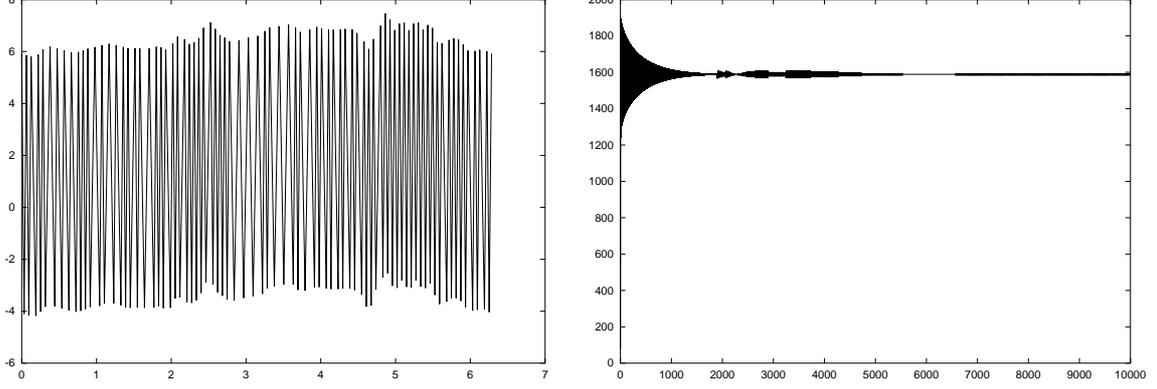


Figure 13: Filtered signal (left) and total variation vs. time steps (right) using  $\lambda = 5$  and  $m = 100$

### 3.6 Implicit discretisation by means of the regularised chain rule

Let us now consider the regularised equation (12) in an implicit setting:

$$\begin{aligned} U_j^{n+1} &= U_j^n + \lambda \left[ \widetilde{\text{sign}}^m (U_{j+1}^{n+1} - U_j^{n+1}) - \widetilde{\text{sign}}^m (U_j^{n+1} - U_{j-1}^{n+1}) \right] \\ &:= H(U_{j-1}, U_j, U_{j+1}, U_j^n)^{n+1}. \end{aligned} \quad (14)$$

Employing (7) we now compute the relevant partial derivatives of  $H$  we need for the linearisation of (14) around the time level  $n\Delta t$ . We obtain

$$\begin{aligned} &\frac{\partial}{\partial U_{j-1}} H(U_{j-1}, U_j, U_{j+1}, U_j^n) \\ &= \frac{\partial}{\partial U_{j-1}} \left[ U_j^n + \lambda \left( \frac{2}{\pi} \arctan(m(U_{j+1} - U_j)) - \frac{2}{\pi} \arctan(m(U_j - U_{j-1})) \right) \right] \\ &= -\frac{2\lambda}{\pi} \cdot \frac{1}{1 + (m(U_j - U_{j-1}))^2} \cdot (-m) \\ &= m \frac{2\lambda}{\pi} \cdot \frac{1}{1 + m^2(U_j - U_{j-1})^2}, \end{aligned} \quad (15)$$

$$\begin{aligned} &\frac{\partial}{\partial U_j} H(U_{j-1}, U_j, U_{j+1}, U_j^n) \\ &= \frac{\partial}{\partial U_j} \left[ U_j^n + \lambda \left( \frac{2}{\pi} \arctan(m(U_{j+1} - U_j)) - \frac{2}{\pi} \arctan(m(U_j - U_{j-1})) \right) \right] \\ &= \frac{2\lambda}{\pi} \cdot \frac{1}{1 + (m(U_{j+1} - U_j))^2} \cdot (-m) - \frac{2\lambda}{\pi} \cdot \frac{1}{1 + (m(U_j - U_{j-1}))^2} \cdot m \\ &= \left( -m \frac{2\lambda}{\pi} \right) \cdot \left( \frac{1}{1 + m^2(U_{j+1} - U_j)^2} + \frac{1}{1 + m^2(U_j - U_{j-1})^2} \right), \end{aligned} \quad (16)$$

and

$$\begin{aligned}
& \frac{\partial}{\partial U_{j+1}} H(U_{j-1}, U_j, U_{j+1}, U_j^n) \\
&= \frac{\partial}{\partial U_{j+1}} \left[ U_j^n + \lambda \left( \frac{2}{\pi} \arctan(m(U_{j+1} - U_j)) - \frac{2}{\pi} \arctan(m(U_j - U_{j-1})) \right) \right] \\
&= \frac{2\lambda}{\pi} \cdot \frac{1}{1 + (m(U_{j+1} - U_j))^2} \cdot m \\
&= m \frac{2\lambda}{\pi} \cdot \frac{1}{1 + m^2(U_{j+1} - U_j)^2}. \tag{17}
\end{aligned}$$

Using also the abbreviation  $\Delta^t U_i^n := U_i^{n+1} - U_i^n$  a linearisation around the data in  $t^n$  reads

$$\begin{aligned}
U_j^{n+1} &= H(U_{j-1}, U_j, U_{j+1}, U_j^n)^{n+1} \\
&= H(U_{j-1}, U_j, U_{j+1}, U_j^n)^n \\
&\quad + \sum_{i \in J} \left\{ \left[ \frac{\partial}{\partial U_i} H(U_{j-1}, U_j, U_{j+1}, U_j^n) \right] \right|^n \Delta^t U_i^n + \mathcal{O}([\Delta^t U_i^n]^2) \right\},
\end{aligned}$$

where  $J := \{j-1, j, j+1\}$ . Thus, neglecting the truncation error, we obtain after a few trivial manipulations using the Kronecker symbol a simple implicit first order time stepping method as

$$\begin{aligned}
& \sum_{i \in J} \left( \delta_{ij} - \left[ \frac{\partial}{\partial U_i} H(U_{j-1}, U_j, U_{j+1}, U_j^n) \right] \right)^n \Delta^t U_i^n \\
&= \lambda \left[ \widetilde{\text{sign}}^m(U_{j+1}^n - U_j^n) - \widetilde{\text{sign}}^m(U_j^n - U_{j-1}^n) \right] \quad \forall j. \tag{18}
\end{aligned}$$

We denote for further use by  $A$  the matrix defined via the left hand side of (18). Let us describe some properties of the matrix  $A$  by the following Lemma.

**Lemma 3.1** *The described matrix  $A$  is a real, symmetric matrix. Furthermore, it is strictly diagonal dominant and positive definite. The eigenvalues of  $A$  are all real and positive. Furthermore,  $A$  is irreducible.*

**Proof** By construction, each individual line of the matrix  $A$  contains components given in (15), (16) and (17).

For an arbitrary index  $k$ ,

$$(U_{(k-1)+1} - U_{k-1})^2 = (U_k - U_{k-1})^2$$

holds, and thus  $A$  is symmetric.

Since  $A$  is symmetric and contains only real numbers, it exhibits only real eigenvalues.

By (16) and (18), it is evident that the diagonal entries  $a_{jj}$  satisfy  $a_{jj} > 0$  for all  $j$ . Because of the special structure imposed by the entries defined by (15), (16), (17) and (18), the matrix  $A$  is obviously strictly diagonal dominant.

Because  $A$  is symmetric, strictly diagonal dominant and since it features  $a_{jj} > 0 \forall j$ , it is also positive definite (see [11], page 52).

Since the stencil includes the nodes with the indices

$$j + i, i \in \{-1, 0, +1\}$$

and because  $A$  is obviously regular, it is also irreducible. ■

Let us remark, that the irreducibility property we have shown seems at first glance redundant; however, it is to be read as the property that all computational nodes are coupled which is important for the propagation of numerical errors (which is, of course, not a priori bad since it could imply a smoothing effect).

For numerical purposes, the investigated matrix  $A$  seems to have at first glance good structural properties by this Lemma. However, we have already observed in earlier sections that the approximated process does in general give quite irregular solutions. Thus, we investigate the condition of  $A$ .

In order to give a thorough discussion we investigate heuristically which situations of interest may arise. The crucial number which determines the magnitude of the entries within  $A$  is given by

$$\gamma(m, \varepsilon) := m \frac{2\lambda}{\pi} \cdot \frac{1}{1 + m^2 \varepsilon^2}. \quad (19)$$

Here,  $\varepsilon$  stands for a change within the local data as in (15), (16) or (17). Since only  $\varepsilon^2$  is really of importance, we consider  $\varepsilon$  to be a positive number, i.e., we investigate absolute data variations.

As can easily be observed, the magnitude of  $m^2 \varepsilon^2$  is of special interest.

The case

$$m^2 \varepsilon^2 \ll 1, \text{ i.e. } \varepsilon \ll \frac{1}{m}$$

leads to

$$m \frac{2\lambda}{\pi} \cdot \frac{1}{1 + m^2 \varepsilon^2} \approx m \frac{2\lambda}{\pi}.$$

Note that this relation is exact if  $\varepsilon = 0$ . Thus, in this case the entries within the corresponding line of  $A$  are given by

$$1 + 2m \frac{2\lambda}{\pi} \quad (\text{diagonal entry}), \quad -m \frac{2\lambda}{\pi} \quad (\text{off-diagonal entries}). \quad (20)$$

We observe that in the case of "small" data variations, the absolute magnitude of the entries grows with  $m$  to infinity – of course, provided  $\varepsilon \ll 1/m$  holds. However, variations within the data can be assumed to be arbitrarily small in practice: for example, very small oscillations are usually induced by use of a high order scheme approximating conservation laws away from shocks.

Note that in the case  $\varepsilon = 0$ , the matrix entries also show the described behaviour.

The case

$$\varepsilon^2 \approx \frac{1}{m}$$

leads for  $m$  large enough to

$$m \frac{2\lambda}{\pi} \cdot \frac{1}{1 + m^2 \varepsilon^2} \approx m \frac{2\lambda}{\pi} \cdot \frac{1}{1 + m} \approx m \frac{2\lambda}{\pi} \cdot \frac{1}{m} = \frac{2\lambda}{\pi}, \quad (21)$$

thus, already for "medium" given data variations (in the described sense), the matrix entries have a reasonable size. The situation is even more appealing for

$$\varepsilon \gg \frac{1}{m}, \text{ i.e., especially for } 1 \geq \varepsilon^2 = \alpha \gg 0.$$

In this case, we can estimate for  $m$  large enough

$$m \frac{2\lambda}{\pi} \cdot \frac{1}{1 + m^2 \varepsilon^2} \approx m \frac{2\lambda}{\pi} \cdot \frac{1}{m^2 \alpha} = \frac{1}{m} \frac{2\lambda}{\alpha \pi}. \quad (22)$$

For  $m \rightarrow \infty$ , we observe by (22) convergence of the corresponding diagonal entry of  $A$  to 1, while the off-diagonal entries converge with exactly the same rate to 0.

To conclude, the "ideal" situation for the described method is for an arbitrary index  $k$  qualitatively equivalent to

$$\dots, U_{k-1} = -1, U_k = 1, U_{k+1} = -1, \dots,$$

together with the choice of a large  $m$ , since in this case  $A \rightarrow I$  as  $m \rightarrow \infty$ . On the other hand, the matrix entries can grow very large for small data variations and large  $m$ .

We can use some results of this discussion to get an impression of the distribution of eigenvalues by the Gerschgorin circles. Let us remember, that all eigenvalues can be found in the union of the circles

$$Z_i := \left\{ z : z \in \mathbf{C}, |z - a_{ii}| \leq \sum_{j, j \neq i} |a_{ij}| \right\}. \quad (23)$$

Using the abbreviations

$$\varepsilon_{j+} := U_{j+1} - U_j \quad \text{and} \quad \varepsilon_{j-} := U_j - U_{j-1}$$

together with the function  $\gamma$  from (19), the circles  $Z_i$  are described in the case of an inner discretisation point via

$$|z - (1 + \gamma(m, \varepsilon_{j+}) + \gamma(m, \varepsilon_{j-}))| \leq \gamma(m, \varepsilon_{j+}) + \gamma(m, \varepsilon_{j-}). \quad (24)$$

Now we use Lemma 3.1 to allow only positive real numbers as eigenvalues. Then by (24), the eigenvalues can only be found in the following real intervals:

$$[1.0, 1.0 + 2\gamma(m, \varepsilon_{j+}) + 2\gamma(m, \varepsilon_{j-})] \quad \forall j. \quad (25)$$

By the discussion above, it is clear that for small variations  $\varepsilon_{j+}$  or  $\varepsilon_{j-}$ , the right end of the intervals in (25) can be arbitrarily large.

We want to give a heuristic conclusion at this point. Consider e.g. initial data resulting from a shock approximation using a high order method with data variations of varying orders of magnitude. Then, the size of the Gerschgorin circles (25) suggests the following conjecture: *The eigenvalues corresponding to a node near the oscillating shock approximation are located near 1. Eigenvalues corresponding to small oscillations naturally arising away from the shock vicinity may be very large, while eigenvalues corresponding to the viscous part of the approximation are distributed near 1. This results in a large spectrum of eigenvalues of different orders of magnitude which may result in numerical difficulties.*

We finish our discussion with the following assertion.

**Lemma 3.2** *The condition number of  $A$  is not bounded for  $m \rightarrow \infty$ .*

**Proof** Since  $A$  is symmetric, each of its eigenvectors is also an eigenvector of  $A^T$ . Thus,  $A$  is a normal matrix [18].

Within the proof, we will use the Rayleigh quotients

$$\eta(x) = \frac{(Ax, x)}{(x, x)} \quad (26)$$

which are defined for all non-zero vectors  $x \in \mathbb{C}^n$ . The set of all possible Rayleigh quotients as  $x$  is chosen arbitrarily from  $\mathbb{C}^n$  is the well-known *field of values* of  $A$ . Then, since  $A$  is a normal matrix, the field of values of  $A$  is the convex hull of its spectrum [18].

Since  $A$  is normal and all eigenvalues of  $A$  are positive, it holds (see e.g. [16])

$$\text{cond}_2(A) = \frac{\lambda_n}{\lambda_1}. \quad (27)$$

Thereby,  $\text{cond}_2(A)$  denotes the condition number with respect to the 2-norm, while  $\lambda_n$  and  $\lambda_1$  denote the largest and smallest eigenvalue, respectively.

We now construct two special eigenvectors together with the corresponding eigenvalues, by which we want to estimate the condition number  $\text{cond}_2(A)$  from below.

Let us investigate a large oscillation, e.g. around an inner node with number  $k$  as it may be given in the numerical approximation of a conservation law in the vicinity of a shock. For  $m \rightarrow \infty$ , this results in

$$a_{kk} \rightarrow 1, \quad a_{k,k-1} \rightarrow 0 \quad \text{and} \quad a_{k,k+1} \rightarrow 0.$$

By the symmetry of  $A$ , this also means

$$a_{k-1,k} \rightarrow 0 \quad \text{and} \quad a_{k+1,k} \rightarrow 0.$$

We especially notice, that in the limit  $m \rightarrow \infty$  all entries in the  $k$ -th column vector of  $A$  are 0 with the exception of the diagonal entry which is 1. Using as a test vector the  $k$ -th unit vector, i.e.

$$e^k := (0, \dots, 0, 1, 0, \dots, 0) \quad \text{with 1 at the } k\text{-th position,}$$

the Rayleigh quotient gives

$$\eta(e^k) = \frac{(Ae^k, e^k)}{(e^k, e^k)} = \frac{(e^k, e^k)}{(e^k, e^k)} = 1. \quad (28)$$

By the Gerschgorin intervals (25) and since  $A$  is normal,  $\eta(e^k)$  is exactly the lower bound of the convex hull containing all eigenvalues of  $A$ .

Let us now investigate an "arbitrarily small" oscillation, i.e. we consider the oscillation as being of size 0. Furthermore, we discuss the for many cases important situation of non-periodic boundary conditions.

We may locate such locally flat data within the first two (or, alternatively, last two) discretisation points. Let us fix the index  $\ell$  as  $\ell = 1$ , and we consider the corresponding unit vector

$$e^1 = (1, 0, \dots, 0) \quad \text{with 1 at the first position.}$$

Then the corresponding Rayleigh quotient gives

$$\eta(e^1) = \frac{(Ae^1, e^1)}{(e^1, e^1)} = \left(1 + 2m \frac{\lambda}{\pi}\right) \frac{(e^1, e^1)}{(e^1, e^1)} - m \frac{\lambda}{\pi} \frac{(e^1, e^1)}{(e^1, e^1)} = 1 + m \frac{\lambda}{\pi}, \quad (29)$$

i.e., in the limit  $m \rightarrow \infty$  holds  $\eta(e^1) = \infty$ . Since  $A$  is normal,  $\eta(e^1)$  is less or equal to the upper bound of the convex hull of all eigenvalues of  $A$ . Thus, using (28) and (29), by (27)

$$\lim_{m \rightarrow \infty} \text{cond}_2(A) = \lim_{m \rightarrow \infty} \frac{\lambda_n}{\lambda_1} = \infty$$

follows. ■

The results of numerical experiments are similar to the ones for the case of the explicit method if the regularisation parameter  $m$  is large, confirming the theoretical results derived above. A notable difference in the numerical behaviour compared to the explicit algorithm we have observed is, that for fixed  $\lambda$  one can choose a larger regularisation parameter  $m$  so that we obtain non-oscillatory results; however, the factor by which  $m$  can be enlarged is not impressive in our experiments. The matrix obtained by an analogous implicit discretisation of the  $\beta$ -regularisation model discussed in Section 3.2 has qualitatively the same properties, so one cannot expect a different numerical behaviour in that case. In contrast to the implicit scheme that we investigate here, let us note that semi-implicit discretisations investigated in [23] for the case of the  $\beta$ -regularisation (10) are absolutely stable in the infinity norm.

## 4 Summary and Conclusions

In this paper we have investigated a number of numerical schemes for TV flow. We have seen that on the one hand one has to be careful with discretisations of this singular differential equation: many numerical schemes that appear straightforward may introduce oscillations when the spatial gradient tends to zero. As a consequence of this instability, the discrete total variation of the solution may even increase. Explicit discretisations of a regularised TV flow with bounded diffusivity may not show these problems if very small time step sizes are used, while implicit discretisations may lead to ill-conditioned problems. Thus, it is very difficult to find efficient and stable schemes for TV flow.

On the other hand, it is also difficult to find unstable schemes. Many discretisations of TV flow act self-stabilising in the following sense: they introduce oscillations which lead to large gradients. Increasing the gradient magnitude decreases the diffusivity such that the scheme becomes stable again until the gradient becomes too small. In our paper we have analysed this situation in detail and have shown that the oscillations remain bounded by a constant that depends on the time step size.

Knowing the properties of several numerical schemes, TV flow can be applied for removing signal oscillations. In particular, we think of employing TV flow as filtering step for improving solutions of hyperbolic conservation laws based on the concept presented in [10, 22]. The sound integration of TV flow into such numerical schemes is part of our future research.

## Acknowledgements

The authors gratefully acknowledge the financial support of their work by the *Deutsche Forschungsgemeinschaft (DFG)* under the grants No. SO 363/9-1 and WE 2602/1-2.

## References

- [1] R. Acar and C. R. Vogel, Analysis of bounded variation penalty methods for ill-posed problems, *Inverse Problems*, Vol. 10 (1994), 1217–1229.
- [2] L. Alvarez, F. Guichard, P.-L. Lions and J.-M. Morel, Axioms and fundamental equations in image processing, *Archive for Rational Mechanics and Analysis*, Vol. 123 (1993), 199–257.
- [3] F. Andreu, C. Ballester, V. Caselles and J. M. Mazón, Minimising total variation flow, *Differential and Integral Equations*, Vol. 14, No. 3 (2001), 321–360.
- [4] F. Andreu, V. Caselles, J. I. Diaz and J. M. Mazón, Qualitative properties of the total variation flow, *Journal of Functional Analysis*, Vol. 188, No. 2 (2002), 516–547.
- [5] G. Bellettini, V. Caselles and M. Novaga, The total variation flow in  $R^N$ , *Journal of Differential Equations*, Vol. 184, No. 2 (2002), 475–525.
- [6] T. F. Chan, S. Osher, and J. Shen, The digital TV filter and nonlinear denoising, *IEEE Transactions on Image Processing*, Vol. 10, No. 2 (2001), 231–241.
- [7] F. Dibos and G. Koepfler, Global total variation minimisation, *SIAM Journal on Numerical Analysis*, Vol. 37, No. 2 (2000), 646–664.
- [8] X. Feng and A. Prohl, Analysis of total variation flow and its finite element approximations, *Mathematical Modelling and Numerical Analysis*, Vol. 37, No. 3, (2003), 533–556.
- [9] K. Glashoff and H. Kreth, Vorzeichenstabile Differenzenverfahren für parabolische Anfangsrandwertaufgaben, *Numerische Mathematik*, Vol. 35 (1980), 343–354.
- [10] T. Grahs, A. Meister and T. Sonar, Image processing for numerical approximations of conservation laws: nonlinear anisotropic artificial dissipation, *SIAM Journal on Scientific Computing*, Vol. 23 (2002), 1439–1455.
- [11] W. Hackbusch, *Elliptic Differential Equations*, Springer Verlag, 1992.
- [12] D.S. Jones, *The theory of generalised functions*, Cambridge University Press 2nd ed., 1982.
- [13] R.P. Kanwal, *Generalised Functions: Theory and Technique*, Birkhäuser Verlag, 2nd ed., 1998.
- [14] R.J. LeVeque, *Numerical Methods for Conservation Laws*, Birkhäuser Verlag, 2nd ed., 1992.
- [15] R.J. LeVeque, *Finite Volume Methods for Hyperbolic Problems*, Cambridge University Press, 2002.
- [16] A. Meister, *Numerik linearer Gleichungssysteme*, Vieweg Verlag, 1999.
- [17] L. I. Rudin, S. Osher and E. Fatemi, Nonlinear total variation based noise removal algorithms, *Physica D*, Vol. 60 (1992), 259–268.
- [18] Y. Saad, *Iterative Methods for Sparse Linear Systems*, PVS Publ. Comp., 1996.

- [19] G.A. Sod, *Numerical Methods in Fluid Dynamics*, Cambridge University Press, 1985.
- [20] D. M. Strong, *Adaptive Total Variation Minimizing Image Restoration*, Ph.D. thesis, Department of Mathematics, University of California, Los Angeles, CA, 1997.
- [21] G. Steidl, J. Weickert, T. Brox, P. Mrázek and M. Welk, On the equivalence of soft wavelet shrinkage, total variation diffusion, total variation regularisation, and SIDEs, *SIAM Journal on Numerical Analysis*, Vol. 42, No. 2 (2004), 686–713.
- [22] G. Wei, Shock capturing by anisotropic diffusion oscillation reduction, *Computer Physics Communications*, Vol. 144 (2002), 317–342.
- [23] J. Weickert, *Anisotropic Diffusion in Image Processing*, B.G. Teubner, Stuttgart, 1998.