TA 81.

FORSCHUNG - AUSBILDUNG - WEITERBILDUNG

Bericht Nr. 104

IMAGE PROCESSING USING A WAVELET ALGORITHM FOR NONLINEAR DIFFUSION

J. Fröhlich[1], J. Weickert

[1]present address: Konrad-Zuse-Zentrum
Heilbronner Str. 10
10711 Berlin, Germany

UNIVERSITÄT KAISERSLAUTERN
Fachbereich Mathematik
Postfach 3049

D-67653 Kaiserslautern

März 1994

Image Processing Using a Wavelet Algorithm for Nonlinear Diffusion

J. Fröhlich [1], J. Weickert

Department of Mathematics, University of Kaiserslautern,
67663 Kaiserslautern, Germany

Abstract

The edge enhancement property of a nonlinear diffusion equation with a suitable expression for the diffusivity is an important feature for image processing. We present an algorithm to solve this equation in a wavelet basis and discuss its one dimensional version in some detail. Sample calculations demonstrate principle effects and treat in particular the case of highly noise perturbed signals. The results are discussed with respect to performance, efficiency, choice of parameters and are illustrated by a large number of figures. Finally, a comparison with a Fourier method and a finite volume method is performed.

[1] present adress: Konrad-Zuse-Zentrum, Heilbronner Str. 10, 10711 Berlin, Germany

1 Introduction

Since the work of Grossmann and Morlet [20] wavelets have found wide application in signal and image processing. These functions fit well in a Fourier transform-like approach without the drawbacks of trigonometric functions, however. The literature on this subject is rapidly evolving and we refer to review articles and books such as [7, 12, 32] and others on this topic.

An important feature of wavelets is the possibility to compress information with quite spectacular results [3, 13]. This induced researchers in the field of numerical analysis to use such functions as bases in their algorithms for the solution of partial differential equations (PDEs) [18, 22, 30]. The aim is to set up adaptive algorithms and/or effective preconditioners by using the inherent hierarchy and orthogonality of a discrete wavelet basis.

Another junction between image processing and the numerical solution of PDEs is the fact that applying any 'meaningful' image transform corresponds to the solution of a PDE [1, 2]. The idea is that the numerical solution of a suitable PDE may lead to a practical algorithm when used in such a way that the original image is taken as initial condition and the solution of the PDE after some temporal evolution is taken as the output. The convolution with a Gaussian kernel, for example, is nothing else than solving a linear diffusion equation [21]. This approach has been persued in recent years by considering nonlinear diffusion-type equations for image processing [6, 9, 10, 28, 29, 33, 34]. Indeed, the results of these authors demonstrate that this method outperforms the linear models as a preprocessing step for segmentation, as it blurs or shifts edges much less than the former.

Our motivation for the present work is the following. The cited algorithms for nonlinear diffusion equations are based on a low order finite difference discretization in space. On the other hand people often do filtering using wavelet coefficients for compression. An alternating combination of both seems to give good results [26]. So why not combine both and use the same basis functions for filtering and for solving the PDE ?

We contribute such an algorithm based on [22] for periodic functions. It is applied to the 1D case with periodic boundary conditions and permits the observation of typical phenomena. Note that the smoothing of a noisy signal in one dimension may be a much more delicate task than the two dimensional case. This is due to the increase of the number of neighbours in higher dimensions. Of course the treatment of purely two dimensional features such as the response near corners has to be investigated seperately. The detection of edges and segments, however, does carry over to this case, since near a straight edge, a 2D image behaves like a function of one variable.

In the whole text an image is represented by a real valued function $f(x)$ on $[0,1]$ corresponding to the intensity of light. In the discrete case $[0,1]$ is replaced by a set of equidistant grid points (pixels). Image processing then starts from $f(x)$ given at these points and aims to construct a function $g(x)$ that retains the required information. The definition of what this required information consists of is a major part of the process and may lead to substantially different methods.

This paper is organized as follows:

Sections 2 and 3 juxtapose the concepts of wavelet and nonlinear diffusion filtering. In section 4 the numerical algorithm which combines both is presented. After clarifying the role of the different parameters in section 5, it is used to process signals with jumps and high

noise level in section 6. The study is completed by a comparison with classical algorithms.

2 The wavelet transform and its direct use in image processing

Since this paper is meant to focus on the image processing aspect, we will only recall some essential features of wavelet approximation that will be important in the sequel. An exhaustive treatment of the subject can be found in [7, 12, 25].
The discrete wavelet approach resides on the concept of multiresolution. A *multiresolution* of $L^2(R)$ is a sequence of imbedded subspaces V_j with

$$V_j \subset V_{j+1} \qquad \forall j \in Z \tag{1}$$

$$\overline{\bigcup}_{j\in Z} V_j = L^2(R) \tag{2}$$

$$\bigcap_{j\in Z} V_j = \{0\} \tag{3}$$

$$f(x) \in V_j \Leftrightarrow f(2x) \in V_{j+1} \tag{4}$$

There exists a *scaling function* $\phi(x)$ serving to construct a basis in each V_j, via

$$V_j = \overline{span}\{\phi_{ji}\}_{i\in Z} \tag{5}$$

with

$$\phi_{ji}(x) = 2^{j/2}\phi(2^j x - i) \qquad j, i \in Z \tag{6}$$

In the classical case this basis is orthonormal, so that

$$< \phi_{ji}, \phi_{jk} >_R = \delta_{ik} \tag{7}$$

with $< f, g >_R = \int_{-\infty}^{+\infty} f(x)\overline{g}(x)\,dx$ being the usual $L^2(R)$ inner product. The main issue of the wavelet approach now is to work with the orthogonal complement spaces W_j defined by

$$V_{j+1} = V_j \oplus W_j \tag{8}$$

Based on the function $\phi(x)$ one can find a function $\psi(x)$, the so-called *mother wavelet*, of which the translates and dilates constitute orthonormal bases of the spaces W_j.

$$W_j = \overline{span}\{\psi_{ji}\}_{i\in Z} \tag{9}$$

generated by the *wavelets*

$$\psi_{ji}(x) = 2^{j/2}\psi(2^j x - i) \qquad j, i \in Z \tag{10}$$

Each function $f \in L^2(R)$ can now be expressed as

$$f(x) = \sum_{i \in Z} c_{j_0 i}\, \phi_{j_0 i}(x) + \sum_{j=j_0}^{\infty} \sum_{i \in Z} d_{ji}\, \psi_{ji}(x) \tag{11}$$

where

$$c_{ji} = < f, \phi_{ji} >_R \qquad d_{ji} = < f, \psi_{ji} >_R \tag{12}$$

The transition from $f(x)$ to the coefficients c_{ji} and d_{ji} is called the *discrete wavelet transform*. Of course, in numerical applications the sums in (11) are truncated which corresponds to the projection of f into a subspace of $V_J \subset L^2(R)$. The decomposition (11) is orthogonal, as, by construction,

$$< \psi_{ji}, \psi_{lk} >_R = \delta_{jl}\ \delta_{ik} \tag{13}$$

$$< \psi_{ji}, \phi_{lk} >_R = 0 \qquad j \geq l \tag{14}$$

in addition to (7).

An image is characterized by a limited extent of information, so that the question arises of how to treat the boundaries. The simplest way would be to extend the interval where the image is defined by completing with $f(x) = 0$ in the exteriour, and using the above multiresolution of the real line. This, however, introduces a strong singularity at the boundary which is often inconvenient but could be remedied by a more regular continuation. We do not adopt this approach here for the main reason that it would lead to complications with the PDE approach to be studied. Therefore the classical technique of periodization is applied which consists of imposing periodicity of the approximating function (and its existing derivatives through regularity). To this aim one can construct a multiresolution of 1-periodic functions, i.e. on the *torus* $T = R/Z$. It should be noted that the definition of a multiresolution on an interval without periodicity is principally opposed to the translational invariance of the bases in (5) and (9). Among the different constructions on the interval $[0, 1]$, the one of [8] seems to be the most promising at the moment.

Let us now describe how a multiresolution of $L^2(T)$ can be obtained from a multiresolution on $L^2(R)$ following [25, 31]. The key point is to map a function $f \in L^2(R)$ onto a function $\tilde{f} \in L^2(T)$ by the relation

$$\tilde{f}(x) = \sum_{k \in Z} f(x + k) \tag{15}$$

(periodicity will always be denoted by a tilde). Note that in Fourier space this relation reads

$$\hat{\tilde{f}}_k = \hat{f}(k) \qquad k \in Z \tag{16}$$

where

$$\hat{f}(\omega) = \int_{-\infty}^{+\infty} f(x)\ e^{-2\pi i\omega x} dx \tag{17}$$

defines the continuous Fourier transform in $L^2(R) \cap L^1(R)$ and

$$\hat{\tilde{f}}_k = \int_0^1 \tilde{f}(x)\ e^{-2\pi ikx} dx \tag{18}$$

the Fourier transform in $L^2(T)$. Applying this technique to ϕ_{ji} and ψ_{ji} one obtains [31]

$$\tilde{V}_j = \{\tilde{f} \mid \tilde{f}(x) = \sum_{i=0}^{2^j-1} c_{ji}\, \tilde{\phi}_{ji}(x)\} \qquad j \geq 0 \tag{19}$$

$$\tilde{W}_j = \{\tilde{f} \mid \tilde{f}(x) = \sum_{i=0}^{2^j-1} d_{ji}\, \tilde{\psi}_{ji}(x)\} \qquad j \geq 0 \tag{20}$$

where in particular

$$\tilde{V}_0 = \{\tilde{f} \mid \tilde{f}(x) = c_{00}\, \tilde{\phi}_{00}(x)\} \tag{21}$$

with $\tilde{\phi}_{00} \equiv 1$. The definition of a multiresolution in the periodic case carries over directly from the nonperiodic one described above. A function $\tilde{f}(x) \in L^2(T)$ is projected onto $\tilde{V}_J$ and decomposed into contributions from $\tilde{V}_0$ and $\tilde{W}_j, (j = 0, \ldots, J-1)$. For conciseness we set

$$\tilde{\psi}_{-10} = \tilde{\phi}_{00} \qquad\qquad d_{-10} = c_{00} \tag{22}$$

and obtain

$$\tilde{f}_J(x) = \sum_{j=-1}^{J-1} \sum_{i=0}^{max\{2^j-1,0\}} d_{ji}\, \tilde{\psi}_{ji}(x) \tag{23}$$

with

$$d_{ji} = < \tilde{f}, \tilde{\psi}_{ji} >_T \tag{24}$$

The decomposition (23) is orthogonal, since, by construction, the orthogonality carries over from (11) to the torus with

$$< f, g >_T = \int_0^1 f(x)\overline{g}(x) dx \tag{25}$$

Note, that in contrast to the nonperiodic case (11) the sums over i in (19), (20) and (23) do not spread all over Z and that there is a minimal j, namely $j_0 = 0$.
In the present study we apply real valued periodic spline wavelets of even order m and refer to the appendix for their expressions.
For someone who is not familiar with the wavelet framework it is sufficient to retain the following properties of the basis $\{\tilde{\psi}_{ji}\}$:

- hierarchy in scale (i.e. localization in frequency space)
- localization in space
- translation invariance
- orthogonality
- regularity, here determined by the order of the spline

Fig. 2.1 gives an illustration of these properties. The upper part displays the function, the lower part the decomposition in the wavelet coefficient space: each box represents one coefficient with i being counted from left to right and j from bottom to top, so that the boxes centers are located at the centers of the corresponding wavelet function. The grey values represent the amplitude of the coefficient according to a logarithmic scale.
The coefficients in (23) are practically calculated by first interpolating the values f_i of the function $f(x)$ at the $N = 2^J$ equidistant grid points $\{x_i = i/N\}_{i=0,\ldots,N-1}$ thus determining the coefficients $\{c_{Ji}\}_{i=0,\ldots,2^J-1}$. Subsequently, a recursive algorithm is applied to obtain $\{c_{ji}\}_{i=0,\ldots,2^j-1}$ and $\{d_{ji}\}_{i=0,\ldots,2^j-1}$, successively down to $j = 0$ [31]. This transformation requires $O(N \log N)$ operations due to the tree-structure of the multiresolution and the use of FFT for the fast evaluation of convolutions. Recomposition is done similarly by inverting each step.
In the remainder of this section we describe the standard way of applying wavelets in image processing in order to oppose it to the methods described later. It consists, generally speaking, of a scale analysis, i.e. a separation of the image into components on different scales, similar to Fourier analysis where one determines the amplitudes of frequencies. (Note that usually the term "frequency" is reserved for the amplitudes of a trigonometric decomposition, whereas "scale" is related to self-similar structures and therefore used in connection with wavelets.) The above decomposition is a fast method of obtaining the amplitudes d_{ji} of the wavelets $\tilde{\psi}_{ji}$ in (23). As these functions are concentrated in Fourier space around the frequency 2^j, the sum $\sum_i d_{ji}\tilde{\psi}_{ji}$ (j fixed) may be interpreted as the result of a band pass filtering. Furthermore, since the functions $\tilde{\psi}_{ji}$ have local character, the bandpass filtering retains information about localization as well. In fact, the wavelet transform is similar to windowed Fourier transform with the difference, however, that the window narrows with increasing frequency (see [11] or [12] for a thorough comparison between windowed Fourier transform and wavelet transform).
A first application of the wavelet transform is the pure analysis of a signal or image. Let us illustrate here, how the wavelet transform can be used for the detection of edges. Fig. 2.2a displays a function composed of different steps (it will serve as test case for most of the reported calculations) after projection on V_J with $J = 7$ and cubic spline wavelets. The obvious oscillations near the jumps (Gibbs phenomenon) are due to the projection. In all figures the true interpolating function is plotted, not only the values at the grid points. It is to be observed, that the presence of steps is related to fine scale coefficients. Indeed, this transform is well suited for the detection of singularities which can also be of other kind, e.g. related to the derivative of a function [24]. What, however, if random noise is added, e.g. $\mu(x) = A \cdot (\varepsilon(x) - \frac{1}{2})$ with $\varepsilon(x)$ being a random number in $[0,1]$ and A a positive amplitude? This has been done in Fig. 2.2b with $A = 0.1$, i.e. roughly 20% of the height of the steps. The peaks in fine scales are still present, but it is clear that the identification of steps will become impossible with this technique if its amplitude increases. Fig. 2.2c shows

the same function as before but with random noise of amplitude $A = 1.0$ beeing added. We shall describe a method that permits even in this case to obtain a function consisting of steps which is relatively close to the non perturbed one.
A second standard application of the wavelet transform is the compression of images. It permits to determine a smoothed function described by only few wavelet coefficients of which e.g. the L^2-norm is close to the one of the original signal. This approach can lead to very high compression rates [3, 13]. In Fig 2.3 the amplitudes of the signal in Fig. 2.2b have been set to zero if their absolutie value was below $\varepsilon = 0.01$. The simple threshold filter results in a compression ratio of one forth in this case. It can be observed that in the interiour of the steps the function is smoothed, but that, in contrast, near the edges oscillations are enhanced. In particular when the noise level is high this method is not useful since no distinction can be made between noise and signal (c.f. coefficients in Fig 2.2c).
In the last part of this section we describe the result of what happens when just the fine scale amplitudes are eliminated from the noisy signal of Fig 2.2c. This operation corresponds to the projection into subspaces $V_j (j < J)$ and may be interpreted as a linear low-pass filtering. From the graphs in Fig 2.4 we observe that the original curve is only poorly recognizable since the location of steps is too much blurred due to a lack of high frequencies in this part. (Fig. 2.4 has been obtained with $m = 6$ in order to compare it to latter results.)
We conclude that even if it would be possible to detect edges like in Fig. 2.2b, it is not straightforward and seems to be delicate to define ad hoc an appropriate smoothing procedure that at the same time takes into account the presence of edges. It has to be applied in coefficient space, but must treat wavelets near edges differently with respect to those in the interiour of segments. Indeed, in [24] such a method is described that tries to exploit the regularity depending convergence of wavelet sums. Its practical implementation, however, is rather complicated. One reason is the difficulty of tracking features through different scales, which is a classical problem in image processing [35], [1]. We therefore use a different strategy here in wedding the wavelet approach with the nonlinear diffusion process in order to obtain edge-preserving smoothing.
Let us indicate for completeness that there exists a different approach of using the wavelet transform in image processing and in particular for edge detection [23]. It consists roughly speaking of coding the given function f by the zeros of the functions $g_j = P_{W_j} f$, i.e. the projections of the functions f on the difference spaces W_j.

3 The nonlinear diffusion equation

In this section we briefly describe the concept of nonlinear diffusion filters for image processing and discuss the equations to be solved later on. For the present study we limit ourselves to the basic model of Perona and Malik [29] and its modification by [6].
In order to describe the nonlinear diffusion approach let us go back to the removal of noise for a moment. Noise which is related to high frequencies, i.e. small scales can be removed by a low-pass-type filtering, i.e. by smoothing. On the other hand, solutions of parabolic PDEs typically exhibit increasing smoothness with time. Take as an example the Cauchy problem for the linear diffusion equation. This equation may be interpreted as a process which gradually damps the spatial frequency proportional to the square of the wavenumber.

Remember also the equivalence of solving a linear diffusion equation and the convolution with a Gaussian [21]. Thus, linear diffusion is just another way of expressing smoothing without generating new details. The key point now is to introduce a nonlinearity in the diffusion process which is particularly designed for a given task. In our case we are concerned with smoothing while at the same time preserving or even enhancing edges. It has been discussed in the previous section that fine scales are required to correctly represent edges (Fig. 2.2a). A linear diffusion process leads to smoothing in the whole domain. This, however, is only required in the interior of a segment and not near the edges. Two steps are therefore to be designed. First the probability for the presence of an edge has to be estimated. Second, diffusion has to be carried out in such a way that it is large where no edge is present and small where this is the case. The latter step can be accounted for by a spatially varying diffusivity which is small near edges and large elsewhere. Whether there is an edge at a certain point or not can in a first approach be measured by simply taking the modulus of the gradient at this point.
The mechanism described so far is cast in mathematical form by the equation

$$\partial_t u = \nabla(\ D(|\nabla u|)\ \nabla u\ ) \tag{26}$$

for the scalar function $u(x,t)$ with $|.|$ being the Euclidean norm. The initial value is

$$u(x,0) = f(x) \tag{27}$$

where f is the given image. The boundary conditions have to be chosen in a suitable way. Homogeneous Neumann or periodic boundary conditions seem to be appropriate as they lead to conservation of the average grey value of the whole image. In our case we get rid of boundary conditions by imposing periodicity of u in space (which is convenient for wavelet or Fourier methods). Moreover we restrict ourselves to the one dimensional setting.
In order to be useful in the way described above, the diffusivity D should fulfil some additional requirements, e.g.

$$\begin{aligned} D(0) &= 1 \\ \lim_{s\to\infty} D(s) &= 0 \\ D &\in C^\infty[0,\infty[\\ D'(s) &\leq 0 \end{aligned} \tag{28}$$

One expression, introduced in [29], is

$$D(|\nabla u|) = \frac{1}{1+\left(\frac{|\nabla u|}{\lambda}\right)^2} \qquad \lambda \in R^{>0} \tag{29}$$

which will be applied here with a slight modification discussed below.
The model (26), (27), (29) contains one parameter only, the scalar λ that plays the role of a contrast parameter. To illustrate this, consider for simplicity a one dimensional form of (26). Introducing the flux

$$\phi(s) = sD(|s|) \tag{30}$$

yields

$$\partial_t u = \partial_x(\phi(\partial_x u)) = \phi'(\partial_x u)\ \partial_{xx} u \tag{31}$$

Plugging (29) into (30) leads to

$$\begin{aligned} \phi'(\partial_x u) &> 0 \qquad \text{if} \quad |\partial_x u| < \lambda \\ \phi'(\partial_x u) &< 0 \qquad \text{if} \quad |\partial_x u| > \lambda \end{aligned} \tag{32}$$

Hence, the sign of ϕ' in (31) may become negative (near strong edges) which locally leads to a backward diffusion problem. It can be shown by deriving (31) with respect to x that indeed the gradient $|\partial_x ui|$ at an edge increases in this case [29]. It is for this reason that the equation can no longer be classified in a standard way. Furthermore it is conjectured that it is not possible to prove existence and uniqueness for this initial boundary value problem with general initial data [6].

Note that the above behaviour is closely related to the speed of decrease towards zero in (29). If e.g. the square in the denominator is removed, only forward diffusion is possible and the initial boundary value problem is well-posed. Unfortunatly, the resulting equation is not very useful for our purposes, since it does no longer allow edge enhancement. Edges become smeared similiar to the linear diffusion case with constant diffusivity.

In addition to difficulties in theory a second drawback of (26),(27),(29) is of practical nature. The Perona-Malik filter interprets strong noise (which causes large gradients) as edges and therefore tends to preserve it.

As a remedy for these theoretical and practical problems, [6] suggest to introduce a presmoothing when evaluating the diffusivity. They replace (29) by

$$D(|\nabla S(u)|) = \frac{1}{1 + \left(\frac{|\nabla S(u)|}{\lambda}\right)^2} \qquad \lambda \in R^{>0} \tag{33}$$

with S being a smoothing operator described in the following sections (similiar regularizations have been proposed by [27] as well). This modification ensures better noise elimination and allows to prove existence and uniqueness of the solution for the corresponding initial value problem [6]. Following [6, 27] it can be shown that $u \equiv const.$ is the only steady solution of these equations with the above boundary conditions. Concerning the smoothing operator S it should be recalled that the spatial discretization which has to be introduced for the numerical solution of the nonlinear diffusion equation has an additional smoothing effect: for a fixed resolution the representable features cannot become arbitrarily fine. In some cases the regularizing effect of the spatial discretization is already sufficient to carry out the computations with (29), i.e. without explicit smoothing as in (33) (c.f. sections 5-7). Nitzberg and Shiota [27] investigated the loss of regularization by locally increasing the spatial resolution with an adaptive mesh. The solution then exhibits a strong irregular behaviour and tends to destabilize.

Finally, we note that due to the divergence form of (26) the integral of u is invariant for appropriate boundary conditions, in particular if u is required to be periodic. This means that the average grey level of a picture is unchanged. Such a property should be enherited by a numerical method for this problem, and we will see that this is the case for the applied spatial discretization. On the other hand, the range $R = \max\{u\} - \min\{u\}$ is not conserved and tends to zero for a convergent numerical method. An additional rescaling mechanism may in this case preserve the contrast and will be discussed later.

4 Solving the nonlinear diffusion equation in a wavelet basis

The method for the solution of (31), (33) on the torus T that we describe here is of *Petrov-Galerkin* type (also termed *method of weighted residuals* [5]) and has been developed in [22]. As we will see below, it requires the operator which is inverted in each step to be linear with constant coefficients. This can be achieved by means of an appropriate temporal discretization which is described first.
A straightforward implicit time scheme is excluded since the particular form of the diffusivity would lead to the inversion of a nonlinear system which is too costly. One can of course apply an explicit discretization of second order (Adams-Bashforth) which amounts to replacing

$$\partial_t u = \partial_x(D\,\partial_x u) \tag{34}$$

by

$$\frac{3u^{n+1} - 4u^n + u^{n-1}}{2\Delta t} = 2\partial_x(D\,\partial_x u)^n - \partial_x(D\,\partial_x u)^{n-1} \tag{35}$$

where the upper index denotes the time level.
However, as expected, the scheme suffers from severe stability restrictions. A simple remedy can in some cases be the addition of a stabilizing term on both sides of the equation. It is discretized in an implicit way on one side and in an explicit way on the other side without modifying the order of the scheme. Here we take in a natural way the Laplacian and obtain

$$\partial_t u - \nu\partial_{xx} u = \partial_x((D-\nu)\partial_x u) \tag{36}$$

leading to

$$\begin{aligned} \frac{3}{2\Delta t}u^{n+1} - \nu\partial_{xx}u^{n+1} &= \frac{4}{2\Delta t}u^n - \frac{1}{2\Delta t}u^{n-1} \\ &+ 2(\partial_x(D-\nu)\partial_x u)^n - (\partial_x(D-\nu)\partial_x u)^{n-1} \end{aligned} \tag{37}$$

which is still of second order in time. Such schemes have been applied successfully e.g. in computational fluid dynamics [16]. The scalar ν is chosen to optimize the numerical stability of the algorithm (c.f. below). Scheme (37) leads to solving

$$(\sigma - \nu\partial_{xx})u^{n+1} = g \tag{38}$$

for u^{n+1} in each time step ($\sigma = \frac{3}{2\Delta t}$) which will be described now.
The periodic function u^{n+1} is developed in a wavelet sum (23). The method of weighted residuals consists of inserting this sum into (38) and applying the scalar product with suitable test functions $\tilde{\theta}_{lm}$. One thus obtains the following equations for the unknown coefficients d_{ji}

$$\sum_{ji} d_{ji} < (\sigma - \nu\partial_{xx})\tilde{\psi}_{ji}, \tilde{\theta}_{lm} >_T \; = \; < g, \tilde{\theta}_{lm} >_T \qquad (39)$$

$$l = -1, \ldots, J-1$$
$$m = 0, \ldots, \max(2^j - 1, 0)$$

where the linearity of the operator $(\sigma - \nu\partial_{xx})$ and the scalar product have been used. The test functions $\tilde{\theta}_{lm}$ are in our case chosen so as to diagonalize the matrix on the left hand side of (39). This is accomplished when

$$(\sigma - \nu\partial_{xx})\tilde{\theta}_{lm} = \tilde{\psi}_{lm} \qquad (40)$$

$$l = -1, \ldots, J-1$$
$$m = 0, \ldots, \max(2^j - 1, 0)$$

Due to the orthogonality of the wavelets $\tilde{\psi}_{ji}$, equation (39) reads

$$d_{lm} = < g, \tilde{\theta}_{lm} >_T \qquad (41)$$

The functions $\tilde{\theta}_{lm}$ can be determined from (40) in Fourier space

$$(\hat{\tilde{\theta}_{lm}})_k = \frac{(\hat{\tilde{\psi}_{lm}})_k}{\sigma + \nu k^2} \qquad k \in Z \qquad (42)$$

and what remains is to find an efficient way of evaluating the scalar products in (41). [22] have constructed a tree-structured algorithm that works similar to the wavelet transform described in section 2 and resembles the non-standard form of an operator as defined by [4]. It uses the wavelet coefficients d_{ji} as well as the scaling coefficients c_{ji} of the function g to which filters are applied that are based on the scalar products

$$< \tilde{\theta}_{l,m}, \tilde{\psi}_{ji} >_T \quad \text{and} \quad < \tilde{\theta}_{l,m}, \tilde{\phi}_{ji} >_T \qquad (43)$$

These filters are calculated once and for all (relatively costly) which does not allow the coefficients σ and ν to change during the temporal evolution, hence, together with (42), the requirement of constant coefficients.
The fast evaluation of (41) is based on the observation that the scalar products in (43) vanish for $|l - j|$ being sufficiently large which can easily be justified by considering the product in Fourier space. The support of $(\hat{\tilde{\psi}_{ji}})_k$ extends all over Z with the amplitudes being centered around 2^j. They decrease due to the regularity of the wavelet, in particular proportional to $1/k^{m+1}$ for spline wavelets of order m. For numerical calculations one may assume a zero value for distances sufficiently large from 2^j, depending on the required precision. As to be

seen from (42) , the support of $(\hat{\tilde{\theta}}_{lm})_k$ has a similar behaviour (somewhat faster exponential decay for $k \to \infty$) so that the scalar products in (43) vanish if the numerical supports in Fourier space are disjoint. This is e.g. already the case for $|l - j| > 1$ when requiring a precision of 10^{-4} in the scalar products and using quintic spline wavelets (m=6) [22].
Another point to be mentioned is the evaluation of the nonlinear term $\partial_x((D - \nu)\partial_x u)$ starting from the wavelet coefficients of u. The outer derivative can be eliminated by partial integration (leading to the derivation of $\tilde{\theta}$ which is done in Fourier space). The remainder is evaluated in a pseudospectral-like way, i.e. by calculating its values on the grid points and reinterpolating. Recall that the interpolating step is done in $O(N \log N)$ operations due to the use of FFT.
An additional feature of this algorithm is that it permits local refinement of the spatial discretization in a straightforward manner by locally adding supplementary trial functions $\tilde{\psi}_{ji}$ and the corresponding test functions $\tilde{\theta}_{ji}$. This will not be used in the present study since in image processing it is not reasonable to work with a spatial resolution substantially below the pixel size. It may, however, be of interest to use this feature for coarsening the grid where this is possible in order to save computation time. We refer to [22] and [17], which also describe the present algorithm in more detail.
The remainder of this section is devoted to additional features that have been added to the cited algorithm. In what concerns the smoothing operator $S(u)$ mentioned in (33) for the calculation of the diffusivity, it is natural to employ the wavelet spectrum of u to this aim. A function u which is in $\tilde{V}_J$ due to the applied wavelet discretization is projected into $\tilde{V}_{J_S} \subset \tilde{V}_J$ $(J_S < J)$ by just cancelling the amplitudes d_{ji} with $j \geq J_S$. This type of smoothing is similar to a low pass Fourier filtering, very cheap and constitutes an advantageous property of the present representation of u.
Let us also indicate that from the continuous point of view the mean value of u does not change in the evolution governed by (31) with periodic boundary conditions. Since it is represented by the coefficient $d_{-1,0}$ of $\tilde{\psi}_{-1,0} = 1$ (the functions $\tilde{\psi}_{ji}$ have zero mean for $j \geq 0$), this coefficient is no longer an unknown but a constant equal to its initial value. The mean of the signal can thus be conserved exactly without effort.
Finally, we add some remarks on the beforementioned rescaling that has been applied in some cases. As indicated, the only steady state of the diffusion model alone, eqs. (26), (33), is the constant function representing the mean of the initial state. When analyzing an image in practice, it is often convenient to rescale the output when persuing the diffusion process for a long time. Otherwise, the range (global contrast) would decrease too strongly. When the solution more and more approaches the constant, the gradients decrease and the impact of the nonlinearity diminishes as well. Rescaling the function after each time step of the diffusion process, the gradients are increased and the nonlinearity maintains its importance. This leads to a modified mechanism that generates - by definition - non-trivial steady solutions even with pre-smoothing.
When devising such a process for noisy signals, it is somewhat problematic to use a procedure that conserves the values of the extrema since these can occur at very different points which may lead to oscillatory behaviour. It is therefore more convenient to conserve the L^2-norm, i.e. the energie of the signal where the above reasoning applies as well. From a computational point of view this can be realized very easily, since the wavelet coefficients directly represent the L^2-norm. The rescaling is therefore done in coefficient space,

an advantage of this type of representation of the signal. The quantity conserved in the corresponding calculations is thus

$$e = \int_0^1 (u - \overline{u})^2 dx \quad = \quad \sum_{j=0}^{J-1} \sum_{i=0}^{2^j-1} (d_{ji})^2 \tag{44}$$

where the mean $\overline{u}$ is substracted by disregarding $d_{-1,0}$ as described above. This procedure does not alter the average grey value, i.e. the arithmetic mean of the signal.
When employing this additional step we are aware of the fact that an underlying PDE cannot be formulated and that the procedure as a whole of course is no more a diffusion mechanism. The practical results however seem to justify its employment under certain conditions so that we applied the above rescaling for tests in some cases.
Let us note in an aside that non-trivial steady solutions can also be obtained by degenerated diffusivities that become in exactly zero above some critical value of the gradient. This leads to complete suppression of the diffusion process when, near an edge, the gradient exceeds this quantity. However this can only carry over to the discrete system if a finite volume type discretization is used. For a Galerkin scheme this is not guaranteed and can therefore not be used in the present case. Nontrivial steady states may also be obtained by the adaptive choice of $\lambda = \lambda(t)$ [29] or an additional reaction term [9, 10, 28].

5 Demonstration of principal effects

Before reporting on signal processing computations it is convenient to clarify in this section the role of the different parameters.
On one hand there are parameters for the numerical algorithm such as the number of grid points N, the regularity m of the spline wavelets and the stability parameter ν. On the other hand the continuous model (26), (33) is characterized by the reference gradient λ and the amount of pre-smoothing. Finally, the influence of rescaling in each time step will be considered.
In problems of signal or image processing the number of samples or grid points can usually not be chosen at will but is imposed by the application. All calculations reported in this study have been done with $N = 128$, corresponding to eight embedded subspaces V_j, $j = 0, ..., 7$.
The influence of the stabilizing factor ν is difficult to be quantified exactly, since the numerical stability of the scheme depends on the instantaneous values of the nonlinear term. Let us however note that for large λ, i.e. in the linear diffusion limit, the scheme becomes completely implicit and unconditionally L^2-stable with $\nu = 1$. A semi-implicit scheme like the present one may remain unconditionally stable provided the explicit contribution is sufficiently small. Even if the scheme becomes only conditionally stable, the critical time step is generally larger than for the corresponding explicit scheme (e.g. [15]). In all the cases where this was tested for the present nonlinear diffusion equation we found substantial stabilization, so that the value $\nu = 1$ was used throughoutly. For illustration we indicate that e.g. the calculation of Fig. 5.1b below could not be carried out with a time step of $\Delta t = 10^{-4}$ and $\nu = 0$, since meaningless high frequency oscillations occured.

Let us now turn to the choice of the diffusivity, which is characterized by the value of λ. For the continuous model it has been discussed that gradients larger than λ may be increased by the nonlinear mechanism whereas smaller gradients are damped. This carries over to the discrete system and is illustrated in Fig. 5.1. The initial condition is $u(x,0) = \sin(2\pi x - 0.7)$, chosen in order not to exhibit the same symmetries as the dyadic wavelet basis (as would be the case e.g. for $u = \sin(2\pi x)$). The different curves in one picture display the solution at the first 10 time steps and thereby give an impression of the temporal evolution. For $\lambda = 10$ it is to be observed that the diffusion process is approximately linear: the shape of the solution remains unchanged like it would be the case for $D = const.$. Setting $\lambda = 4$ one observes a first phase of nonlinear diffusion which is characterized by the fact that large slopes change only slightly (near the zero crossings of the second spatial derivative) while moderate slopes decrease. As a result the sine function is flattened around its extrema. After some time the linear contribution predominates and the solution recovers its sinusoidal shape. For $\lambda = 2$ the nonlinear effect becomes stronger as being displayed by the third part of the figure. From Fig. 5.1 it can also be observed how the exchange by diffusion is hindered when λ gets smaller: the difference in level between the right hand side and the left hand side of the solution equilibrates much slower. Let us note in an aside that the results of Fig. 5.1 remain almost unchanged when varying the degree of the spline wavelets. This has been verified for $m = 4$ and $m = 12$.

It should also be recalled that the formation of steps out of a very smooth function like discussed here is a severe test case. The extrema of the slope are not very pronounced so that the location of the steps to be created by the algorithm is subject to uncertainity due to numerical effects such as round off errors. As pointed out in section 3 we are locally faced with a backward diffusion process. This is an ill-posed problem and can lead to the typical numerical phenomena related to this situation. In our case we remark a tendency to oscillatory behaviour when decreasing the time step size. The calculations in Fig. 5.2 have been done under the same conditions as the one in Fig. 5.1c, but with $\Delta t = 10^{-3}$ and $\Delta t = 10^{-4}$, respectively. It should be noted that this effect is not due to the wavelet discretization in space which is applied here. It has been observed similiarly with a finite volume scheme (cf. section 7).

Independently of the above, the oscillations of course increase with the regularity of the basis functions. They become somewhat smaller in the cubic case ($m = 4$) and somewhat stronger for $m = 12$. Note, however, that the order $m = 6$ which has been used in Fig. 5.2 is already fairly high.

We shall now discuss different ways to face the described numerical problems. The simplest one would be to use large time steps and large values of λ only. However, this eliminaties the advantages of the nonlinear diffusion model if pushed too far.

As described in section 4, the diffusivity can also be determined for a filtered version of the instantaneous solution instead of the true one. This has a strong regularizing effect and makes the results less dependent on numerical parameters like the time step size or the order of the spline wavelets. Fig 5.3 displays the result under the same conditions as in Fig. 5.2a, but with $J_s = J - 2$, i.e. the the two finest scales of the solution do not enter in the calculation of the diffusivity. Observe for example the higher smoothness of the solution in the laps of time before the pronounced steps are formed.

Another option is the rescaling procedure described at the end of the previous section. In

addition to the local edge enhancement of nonlinear diffusion, this mechanism leads to a global increase of the gradients. It does not avoid ill-posedness, but it allows the choice of a larger λ and increases the lifetime of edges. Thus edge detection becomes easier. In Fig. 5.4a, rescaling depends on the extrema of the solution (compare to Fig. 5.1b where apart from rescaling the same parameters were used). Part b of Fig. 5.4 shows the evolution with $\Delta t = 10^{-3}$. The presmoothing from above can be used to diminish the oscillatory tendency and gives the result depicted in Fig. 5.4c. The alternative procedure of conserving the L^2-norm described in (44) has been applied in Fig. 5.5.

6 Computational results for the detection of steps

In this section we report calculations that aim to reconstruct a "step-function" from a perturbed signal (cf. Fig. 2.2).
Let us start with a low degree of noise and take the signal from Fig. 2.2b as initial condition for our iterative method. One observes that the step sized function is recovered without any problem after 10 iterations. No presmoothing or rescaling needs to be applied. The Gibbs phenomenon present in Fig. 2.2a is of course not recovered (recall that everywhere the entire interpolating function is depicted). The result of Fig. 6.1 is obviously better than the direct wavelet filtering either by amplitude or by scale.
The main part of this section will now be devoted to the treatment of highly noise pertubed signals. As been discussed in section 2 direct wavelet based filtering is not satisfactory. In particular supressing wavelet coefficients with amplitudes below a threshold is unreasonable since signal and noise can no longer be distinguished by a difference in the magnitude of the wavelet amplitudes. The following figures take as initial signal the data from Fig. 2.2c (using $m = 6$ here). The time step in all computations has been chosen for stability reasons mainly. It has been kept constant to permit easy comparison.
First of all it is to be seen that even with a very high pertubation the initial signal is recovered with satisfactory accuracy (cf. Fig. 6.2b). Furthermore, Fig. 6.2 gives an idea of the influence of presmoothing. It succeeds in suppressing fine-scale oscillations. Too much of it however does not lead to pronounced edges as to be seen in Fig. 6.2c. In the present case, $J_s = 5$ seems to be a good choice. Fig. 6.2c illustrates that when applying the nonlinear diffusion method one has to decide on when to stop the iterations. If carried out too long, the solution will be a constant, and all the information is lost. Criteria for estimating an appropriate final time could perhaps be developped form diffusion based arguments [6] or by calculating the variation $\int_0^1 |\partial_x u|\, dx$ of the solution. We do not go further in this direction here, but propose as a remedy the rescaling mechanism discussed in the previous section. The graphs of Fig. 6.3 have been obtained with this additional feature, the other parameters are the same as in Fig. 6.2. Since in each step the gradient of u is increased the diffusion mechanism is diminished. This leads to longer integration in time. The results of Fig. 6.3ab are more satisfactory than the corresponding ones before in that the edges are more pronounced and the downward jump in the right part is better detected. Fig. 6.3c shows that for extremely long calculations ("long" depending on the choice of parameters) the finite gradient of the numerical solution still induces diffusion leading together with the present rescaling to erroneous peak-like solutions. Note, however, that in this figure

already the second curve at time $t_2 = 10^{-3}$ gives a very good result and that the last one corresponds to $50t_2$.
An interesting question is how to choose the value of λ. We saw that if it is too large, steps do not form sufficiently, if it is too small, in particular when much noise is present, unwanted solutions may be created. Fig. 6.4 corresponds to Fig. 6.3.b with $\lambda = 1$. The number of iterations had to be augmented since the smaller λ, the lower the average diffusion. It can be seen that a satisfactory solution is obtained which exhibits nicely pronounced steps. If however λ is further decreased, the formation of finer structures has to be inhibited by stronger presmoothing.
In the last figure of this section a comparison is made for different degrees of regularity of the wavelet functions. Apart from the order m of the spline wavelets the parameters have been chosen the same as in Fig. 6.3. As expected, the oscillatory tendency of the solution is increased with m. On the other hand the steps become steeper, therefore the long time behaviour of the solution is better, since diffusion between adjacent segments is better suppressed.
The results presented in this section show that the nonlinear diffusion wavelet method leads to correct detection of steps in a highly pertubed signal. The choice of paramters depends on the signal, but some hints can be drawn from our experience. The basis functions should not be too regular to avoid oscillatory behaviour. A good choice for the order is $m = 6$, i.e. quintic spline wavelets. Some presmoothing should be applied (more if m is large). The experiments indicate that rescaling can counterbalance a large value of λ chosen not to get erroneous peaks.

7 Comparison with a Fourier and a finite volume method

In this section we complete the picture by two supplementary methods which cover the high regularity and the low regularity end, respectively.
The Fourier method is classical (c.f. [5] for spectral methods). For the present problem the same time scheme is chosen, i.e. (37). The solution of (38) can be obtained very efficiently in Fourier space. The evaluation of the right hand side g is done by the pseudospectral method consisting of evaluating derivatives in Fourier space and products with non-constant factors in physical space, i.e. on the grid points.
The Fourier method requires $O(N \log N)$ operations per time step with N being the number of grid points. This is the same order as for the wavelet method of sections 4 to 6.
Following the argumentation for the wavelet technique one may also obtain the average grey value conservation of the Fourier method.
For the finite volume technique we use again time scheme (37). The spatial approximation at a grid point x_i is accomplished by means of

$$\partial_x(D(|\partial_x v|)\,\partial_x u) \;\approx\; D(|\frac{v_{i+1}-v_i}{\Delta x}|)\,\frac{u_{i+1}-u_i}{(\Delta x)^2} \;-\; D(|\frac{v_i-v_{i-1}}{\Delta x}|)\,\frac{u_i-u_{i-1}}{(\Delta x)^2} \qquad (45)$$

with D from (33). It leads together with the standard discretization of $\partial_{xx} u$ to a second order scheme in space and time.

The values v_i in (45) denote the presmoothing $v = S(u)$ from (33). For this purpose it is in the present context convenient to apply a convolution of the periodized values u_i with a Gaussian kernel

$$K_\sigma(x) = \frac{1}{\sqrt{2\pi}\,\sigma}\; e^{(-\frac{x^2}{2\sigma^2})} \tag{46}$$

Since the solution $v(t)$ of the initial value problem

$$\begin{aligned} \partial_t u &= u_{xx} \\ u(0,x) &= f(x) \end{aligned} \tag{47}$$

with $f \in L^2(R)$ is given by the convolution

$$v(t,x) = K_{\sqrt{2t}}(x) * f(x) \tag{48}$$

a presmoothing may be carried out by discretizing the linear diffusion equation. We used one step of the unconditionally stable, implicit Euler scheme and obtained v_i as a solution of

$$\frac{v_i - u_i}{\tau} = \frac{v_{i+1} - 2v_i + v_{i-1}}{(\Delta x)^2} \tag{49}$$

with $\tau = \sigma^2/2$.
Hence, v_i^n and u_i^{n+1} with $i = 1, \ldots, N$ are each determined by a cyclic tridiagonal system which may be solved with linear effort by an adapted Gaussian algorithm (see e.g. [14]). Therefore the finite volume method requires only $O(N)$ operations.
It is easily verified that the preceeding scheme fulfills the redistribution property of [19] and thus preserves the average grey value of the original image.
The following figures compare the two methods of this section with the wavelet-based algorithm. Fig. 7.1 has been obtained with the initial condition of section 5, a small time step and no presmoothing. The Fourier method (Fig. 7.1a) develops small scale oscillations all over the domain. These are first of all generated by the Gibbs phenomenon near the edges. Due to the global nature of the trial functions, problems arising in the backward diffusion areas are spread out to those parts where forward diffusion should produce a nearly constant solution. This makes the use of the Fourier method as a numerical tool for the nonlinear diffusion equation without presmoothing very doubtful. If a rescaling procedure is added, the results get even worse.
The finite volume method (Fig. 7.1c) has the most local character of the three techniques, as it uses only information of a (3×1) mask. Consequently, smooth behaviour of the solution in the forward and steep behaviour in the backward diffusion region is very well separated. The wavelet method (Fig. 7.1b) benefits from the local character of its basis as well. The Gibbs phenomenon is hardly visible for spline wavelets of order six. The regularity of these functions leads to a smoother solution than for the finite volume method.
As already stated, practical calculations generally require the use of presmoothing for the diffusivity. With this additional feature the differences between the three methods vanish

almost completely, see Fig. 7.2. For comparison the same presmoothing by a "litteral" convolution with a Gaussian has been applied in these calculations.
Let us terminate this section by some remarks on stability and computational cost. The computational efficiency highly depends on implementational issues. Qualitatively it can be stated that the wavelet method is the slowest and the Fourier method the fastest one. The Fourier and wavelet method allow an easy and very fast implementation of presmoothing in wavelet or Fourier space. For the finite volume method this may be carried out in a second diffusion step. Advantages of the finite volume technique become evident for the discretization in the case of homogeneous Neumann boundary conditions, since the resulting tridiagonal systems can be solved about twice as fast as cyclic tridiagonal ones from periodization.
It has been observed that the stability of all three methods mainly depends on the time scheme. Since the same scheme has been used throughoutly, no significant stability differences appeared.

8 Conclusion

The present contribution discusses in detail the mechaniscm of nonlinear diffusion from a phenomenological and from a numerical point of view. Approximations of different spatial regularity have been considered focussing in particular on the use of a spline wavelet basis. We think that the presented technique of solving a nonlinear diffusion equation in a wavelet basis can be used as a complementary step for signal analysis with wavelets. It permits e.g. in a first step to eliminate fine scale noise without affecting the sharpness of edges that are present in the underlying signal. A subsequent wavelet analysis in the classical manner should then give a more pronounced result.

Acknowledgement. This research was supported by grants from DFG, ZPM and Stiftung Innovation des Landes Rheinland-Pfalz.

Appendix

The spline wavelets used in the present calculations are defined by the following formula that are taken from [31]. For the nonperiodic case of a spline multiresolution of $L^2(R)$ one has

$$\hat{S}(\xi) = \frac{\sin^m(\pi\xi)}{(\pi\xi)^m \, P_{\frac{m}{2}-1}(\sin^2(\pi\xi))} \tag{50}$$

$$\hat{\phi}(\xi) = \frac{\sin^m(\pi\xi)}{(\pi\xi)^m \sqrt{P_{m-1}(\sin^2(\pi\xi))}} \tag{51}$$

$$\hat{\psi}(\xi) = \frac{\sin^{2m}(\pi\xi/2)}{(\pi\xi/2)^m} \sqrt{\frac{P_{m-1}(\cos^2(\pi\xi/2))}{P_{m-1}(\sin^2(\pi\xi/2)) \, P_{m-1}(\sin^2(\pi\xi))}} \, e^{-i\xi\pi} \tag{52}$$

for the cardinal Lagrange spline, the scaling function, and the mother wavelet, respectively. The m-th order polynomial P_m

$$P_m(x) = \sum_{i=0}^{m} a_i^m \, x^i \tag{53}$$

is determined by

$$\frac{P_{m-1}(\sin^2 z)}{\sin^{2m} z} = \sum_{l \in Z} \frac{1}{(z + l\pi)^{2m}} \tag{54}$$

and can be calculated via

$$a_0^r = 1 \qquad\qquad r = 0, \ldots, m \tag{55}$$

$$a_i^r = \frac{1}{r(2r+1)} \left((r-i)(2r-2i+1)\, a_i^{r-1} - 2(r-i+1)^2 \, a_{i-1}^{r-1} \right) \qquad i = 1, \ldots, r \tag{56}$$

Periodization is obtained by sampling in frequency space as described in section 2. This defines the periodic multiresolution and gives the following expressions for the (discrete) Fourier transform of the interpolatory filter L, and the decomposition filters G, and H, respectively

$$(\widehat{L_J})_k = \frac{1}{2^{3J/2}} \hat{L}(\frac{k\pi}{2^J}) \tag{57}$$

$$(\widehat{H_j})_k = \frac{1}{2^j} \hat{H}(\frac{k\pi}{2^j}) \tag{58}$$

$$(\widehat{G_j})_k = \frac{1}{2^j} \hat{G}(\frac{k\pi}{2^j}) \tag{59}$$

with

$$\hat{L}(\omega) = \frac{\sqrt{P_{m-1}(\sin^2 \omega)}}{P_{\frac{m}{2}-1}(\sin^2 \omega)} \tag{60}$$

$$\hat{H}(\omega) = \sqrt{2} \cos^m(\omega) \sqrt{\frac{P_{m-1}(\sin^2 \omega)}{P_{m-1}(\sin^2(2\omega))}} \tag{61}$$

$$\hat{G}(\omega) = \sqrt{2} \sin^m(\omega) \sqrt{\frac{P_{m-1}(\cos^2 \omega)}{P_{m-1}(\sin^2(2\omega))}} \; e^{-2i\pi\omega} \tag{62}$$

References

[1] L. Alvarez, F. Guichard, P.L. Lions, J.M. Morel, *Axiomes et équations fondamentales du traitement d'images. (Analyse multiéchelle et E.D.P)*, C. R. Acad. Sci. Paris, t. 315, Série I, 135–138, 1992.

[2] L. Alvarez, F. Guichard, P.L. Lions, J.M. Morel, *Axioms and fundamental equations in image processing*, Preprint 9231, Université Paris IX – Dauphine, Place du Maréchal de Lattre de Tassigny, F-75775 Paris cedex 16, France, 1992 (to appear in Arch. Rat. Mech. Anal.).

[3] M. Antonini, M. Barlaud, P. Mattieu, I. Daubechies, *Image coding using wavelet transform*, IEEE Trans. Image Proc., vol. 1, 205–220, 1992.

[4] G. Beylkin, R. Coifman, V. Rokhlin, *Fast wavelet transforms and numerical algorithms I*, Comm. Pure Appl. Math., vol. 44, 141–183, 1991. Academic Press, New York, 1992.

[5] C. Canuto, M.Y. Hussaini, A. Quarteroni, T.A. Zang, *Spectral methods in fluid dynamics*, Springer-Verlag, 1988.

[6] F. Catté, P.L. Lions, J.M. Morel, T. Coll, *Image selective smoothing and edge detection by nonlinear diffusion*, SIAM J. Numer. Anal., vol. 29, 182–193, 1992.

[7] C.K. Chui, *An introduction to wavelets*, Academic Press, New York, 1992.

[8] A. Cohen, I. Daubechies, B. Jawerth, P. Vial, *Multiresolution analysis, wavelets and fast algorithms on the interval*, C. R. Acad. Sci. Paris, t. 316, Série I, 417–421, 1993.

[9] G.H. Cottet, *Diffusion approximation on neural networks and applications for image processing*, F. Hodnett (Ed.), Proc. Sixth European Conf. on Math. in Industry, 3–9, Teubner, Stuttgart 1992.

[10] G.H. Cottet, L. Germain, *Image processing through reaction combined with nonlinear diffusion*, Math. Comp., vol. 61, 659–673, 1993.

[11] I. Daubechies, *The wavelet transform, time frequency localization and signal analysis*, IEEE Trans. Inform. Theory, vol. 36, 961–1005, 1990.

[12] I. Daubechies, *Ten lectures on wavelets*, SIAM, Philadelphia, 1992.

[13] R. DeVore, B. Jawerth, B. Lucier, *Image compression through wavelet transform coding*, preprint, University of South Carolina.

[14] G. Engeln-Müllges, F. Reutter, *Formelsammlung zur Numerischen Mathematik*, Bibliographisches Institut, Mannheim, 1983.

[15] J. Fröhlich, *Comparison of solution algorithms for the variable coefficient heat equation applying a spectral collocation method*, preprint no. 212, Mathematics Department, University Nice, 1988.

[16] J. Fröhlich, R. Peyret, *Calculations of non-Boussinesq convection by a pseudospectral method*, Comp. Meth. Appl. Mech. Eng., vol. 90, 631–642, 1991.

[17] J. Fröhlich, K. Schneider, *One- and two-dimensional flame computations with an adaptive wavelet discretization*, to appear in Europ. J. Mechanics B.

[18] R. Glowinski, R.W. Lawton, M. Ravachol, E. Tennenbaum, *Wavelet solution of linear and nonlinear elliptic, parabolic, and hyperbolic problems in one space dimension*, R. Glowinski, A. Lichnewsky (Eds.), Proc. Ninth Int. Conf. on Comp. Meth. in Appl. Sci. and Eng., SIAM, Philadelphia, 1990.

[19] R. Gorenflo, *Conservative difference schemes for diffusion problems*, preprint no. 39, Mathematics Dept., Free University of Berlin, 1977.

[20] A. Grossmann, J. Morlet, *Decomposition of Hardy functions into square integrable wavelets of constant shape* , SIAM J. Math. Anal., vol. 15, 723–736, 1984.

[21] J.J. Koenderink, *The structure of images*, Biol. Cybern., vol. 50, 363–370, 1984.

[22] J. Liandrat, Ph. Tchamitchian, *Resolution of the 1D regularized Burgers equation using a spatial wavelet approximation algorithm and numerical results*, ICASE report, 1990.

[23] S. Mallat, *Zero-crossings of a wavelet transform*, IEEE Trans. Inf. Theory, vol. 37, 1019–1033, 1991.

[24] S. Mallat, W.L. Hwang, *Singularity detection and processing with wavelets*, technical report no. 549, New York University, Dept. of Computer Science, Courant Inst. of Mathematical Sciences, 251 Mercer St, New York 10012, USA, 1991.

[25] Y. Meyer, *Ondelettes et operateurs*, Herman, Paris, 1990.

[26] H. Neunzert, R. Roesch, private communication, 1991.

[27] M. Nitzberg, T. Shiota, *Nonlinear image filtering with edge and corner enhancement*, IEEE PAMI, vol. 14, 826–833, 1992.

[28] N. Nordström, *Biased anisotropic diffusion - A unified regularization and diffusion approach to edge detection*, O. Faugeras (Ed.), Computer Vision – ECCV 90, Lecture Notes in Computer Science, vol. 427, Springer, Berlin, 18–27, 1990.

[29] P. Perona, J. Malik, *Scale space and edge detection using anisotropic diffusion*, IEEE PAMI, vol. 12, 629–639, 1990.

[30] V. Perrier, *Ondelettes et simulation numerique*, thesis, University Paris IV, 1991.

[31] V. Perrier, C. Basdevant, *La décomposition en ondelettes périodiques, un outil pour l'analyse de champs inhomogènes. Théorie et algorithmes*, Rech. Aérosp., 53–67, 1989-3.

[32] O. Rioul, M. Vetterli, *Wavelets and signal processing*, IEEE SP Mag., 14–38, Oct. 1991.

[33] J. Weickert, *Anisotropic diffusion filters for image processing based quality control*, A. Fasano, M. Primicerio (Ed.), Proc. Seventh Europ. Conf. on Math. in Industry, Teubner, Stuttgart, 355–362, 1994.

[34] J. Weickert, *Anisotropic diffusion in image processing*, Ph.D. thesis, to be filed, 1994.

[35] A.P. Witkin, *Scale-space filtering*, Proc. Eighth Int. Joint. Conf. on Artificial Intell., Karlsruhe, 1019–1022, 1983.

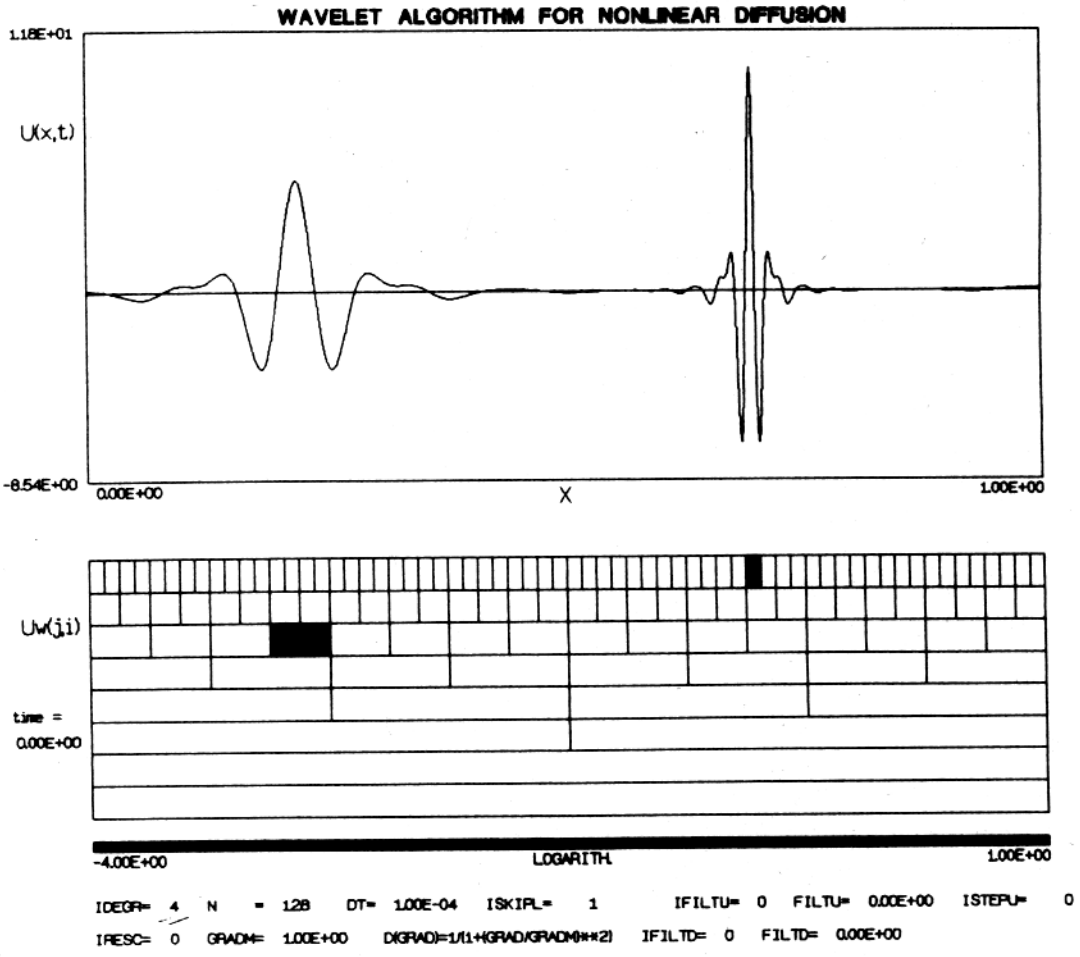

Fig. 2.1: Illustration of the wavelet basis: $\psi_{4,3}$ and $\psi_{6,44}$, $m = 4$.

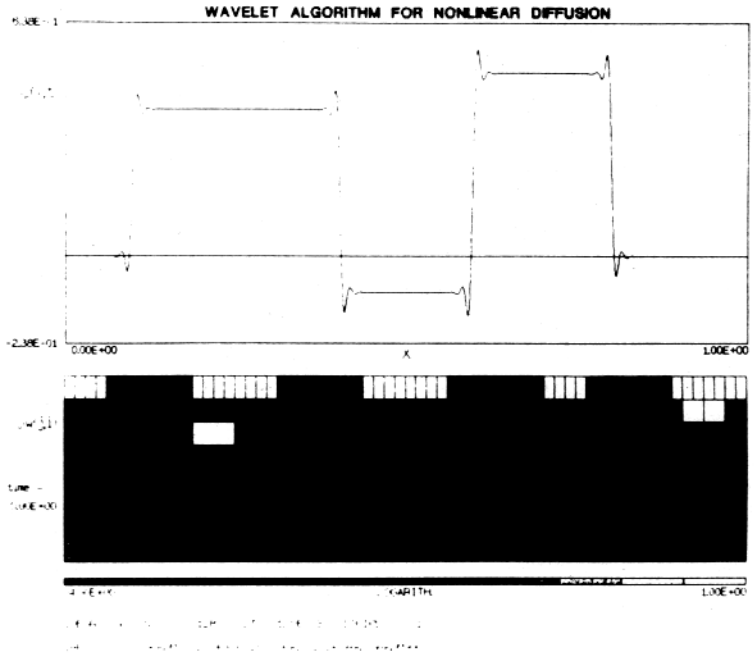

Fig. 2.2a: Interpolating spline function for step shaped data, no noise, $m = 4$.

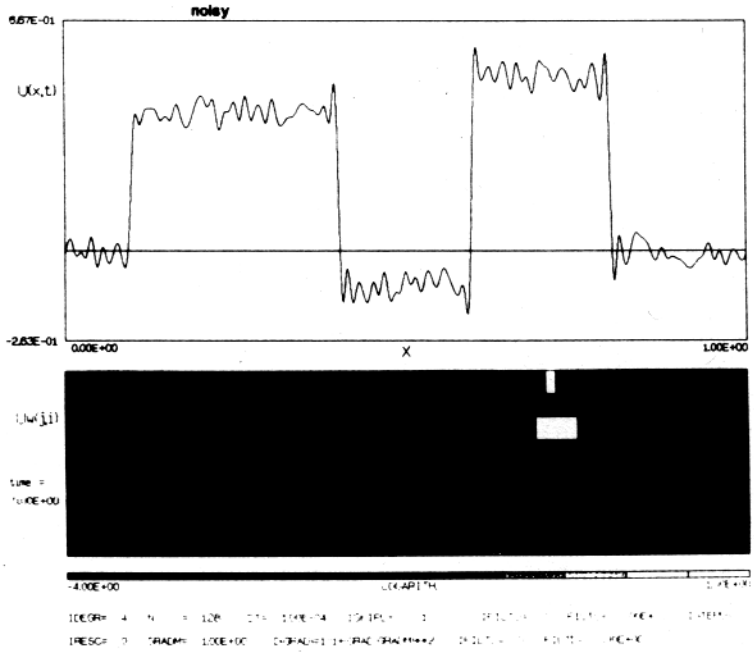

Fig. 2.2b: Interpolating spline function for step shaped data, noise ampl. $A = 0.1$, $m = 4$.

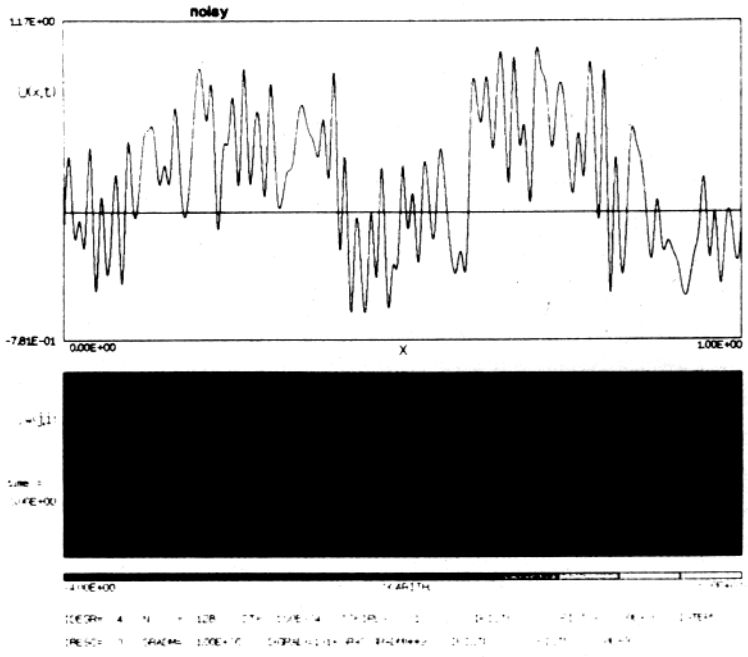

Fig. 2.2c: Interpolating spline function for step shaped data, noise ampl. $A = 1$, $m = 4$.

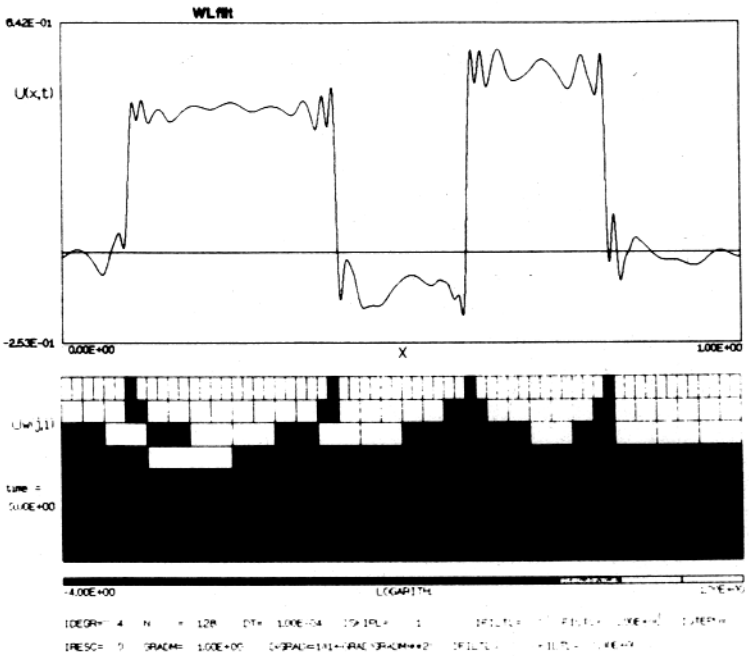

Fig. 2.3: Wavelet filtering of the function in Fig.2.2b. Only the coefficients d_{ji} with $|d_{ji}| > 10^{-2}$ have been retained.

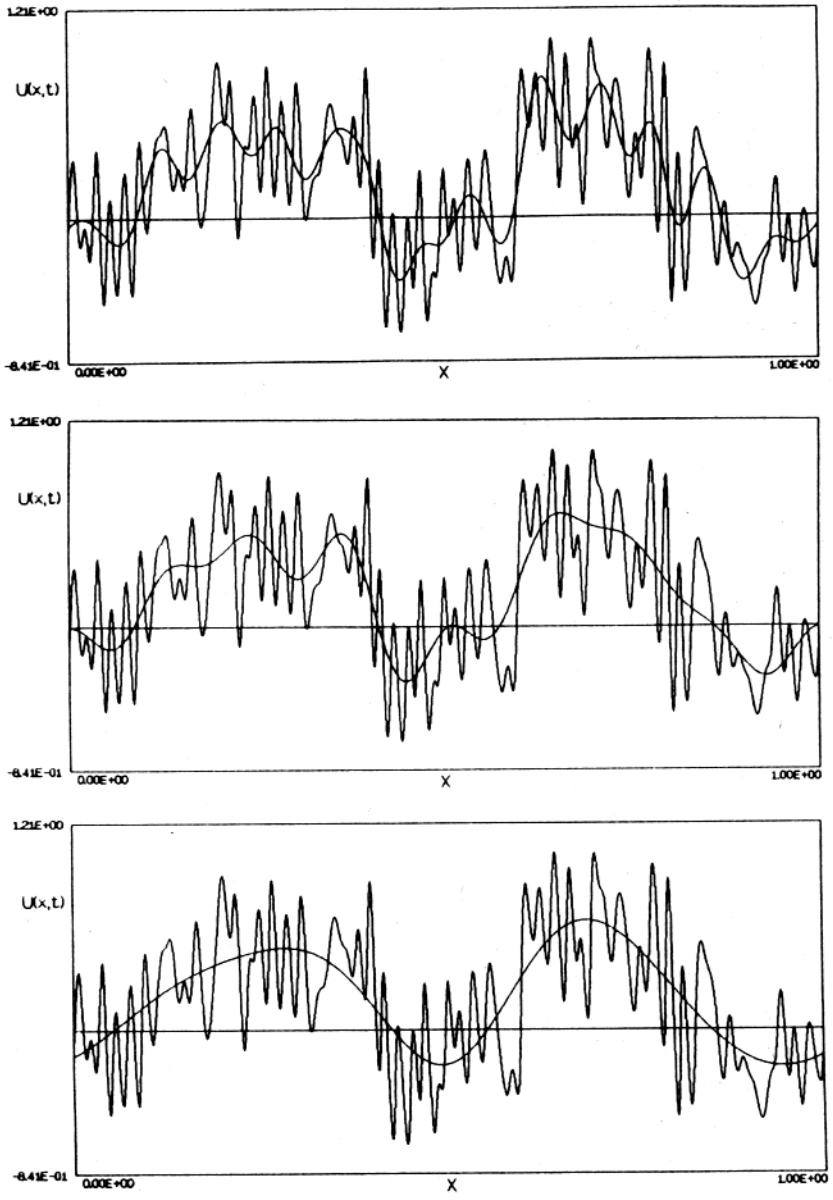

Fig. 2.4: Scalewise filtering in the wavelet basis: a) $J_s = 4$, b) $J_s = 3$, c) $J_s = 2$.

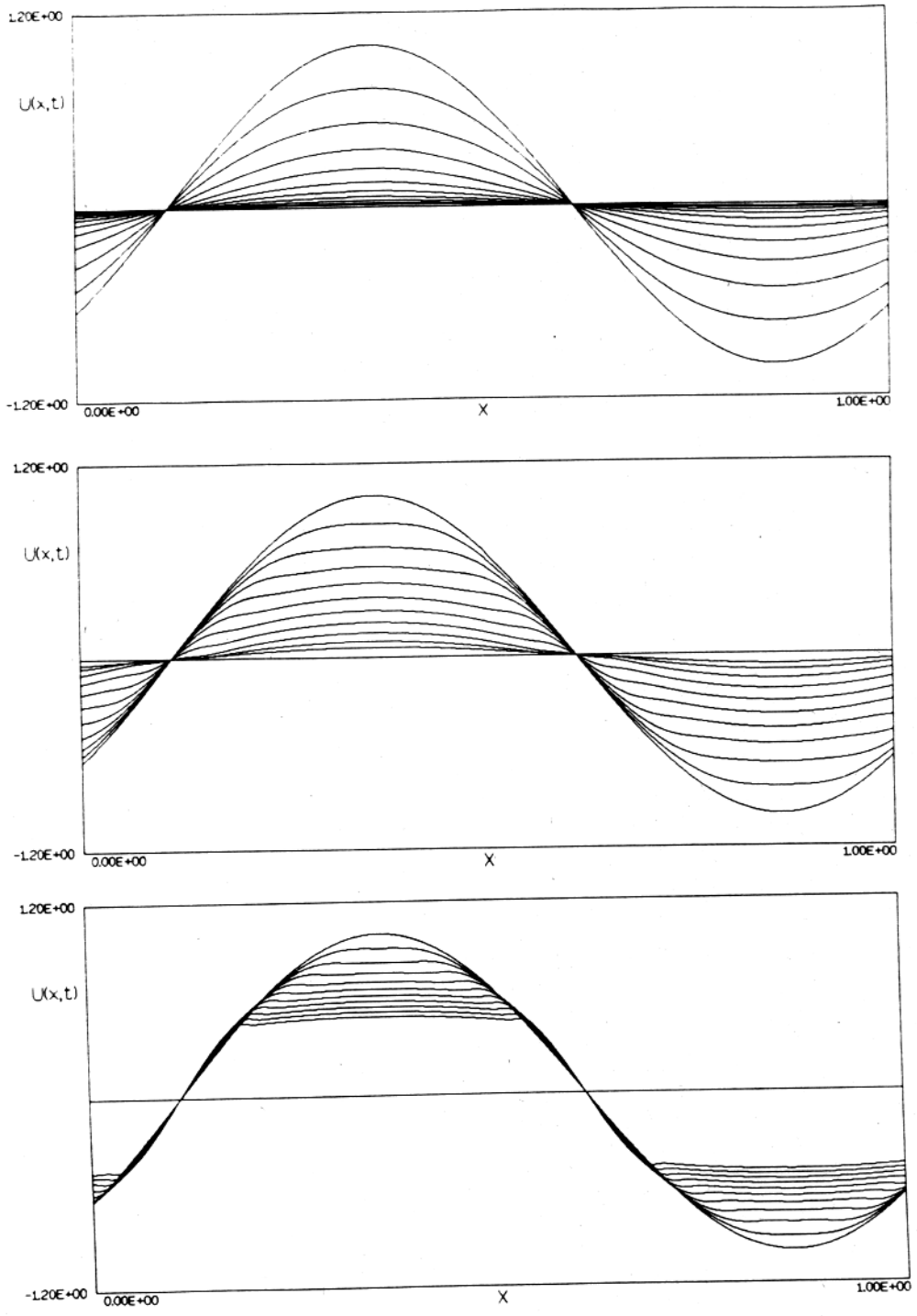

Fig. 5.1: Evolution of a sine function for different values of λ without presmoothing. Plot for $t = 0.01, 0.02, ...,\ m = 6,\ \Delta t = 10^{-2}$. a) $\lambda = 10$, b) $\lambda = 4$, c) $\lambda = 2$.

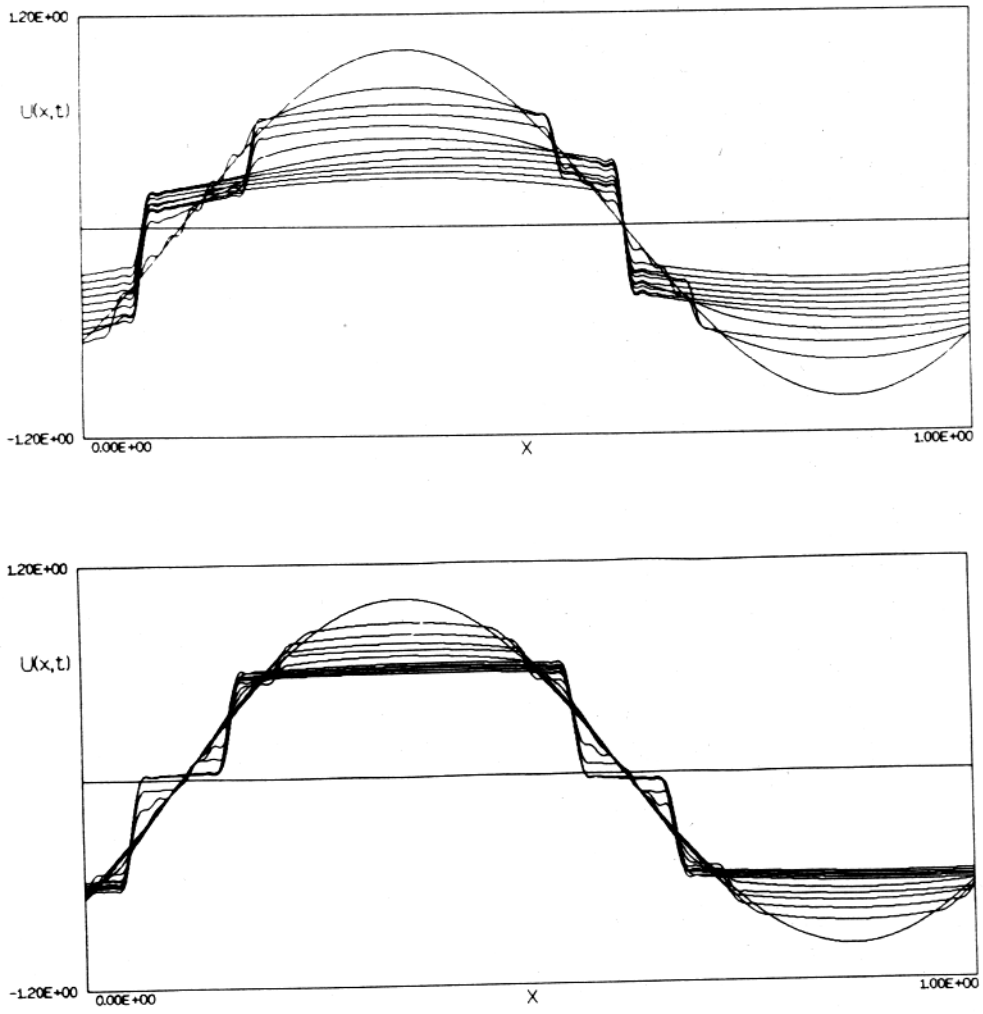

Fig. 5.2: Evolution of a sine function for different time steps, no presmoothing. Plot for $t = 0.01, 0.02, ...$, $m = 6$, $\lambda = 2$. a) $\Delta t = 10^{-3}$, b) $\Delta t = 10^{-4}$.

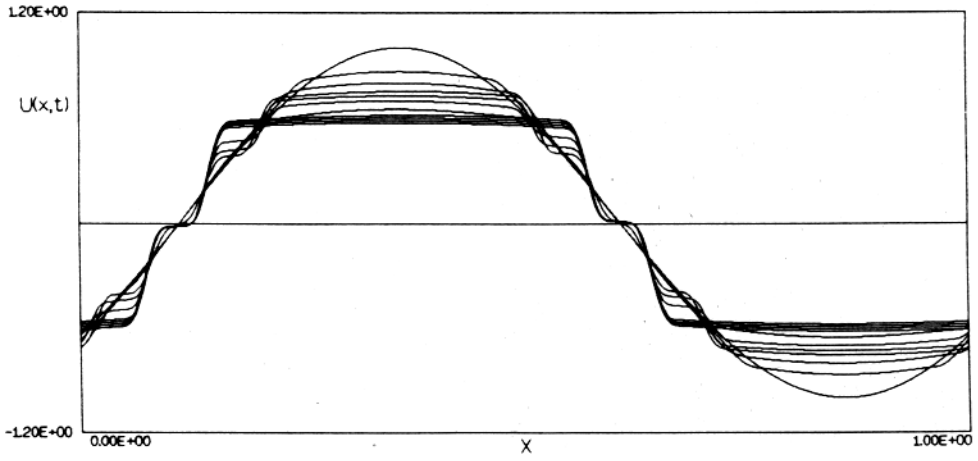

Fig. 5.3: The effect of presmoothing: parameters as Fig.5.2a, presmoothing with $J_s = 4$.

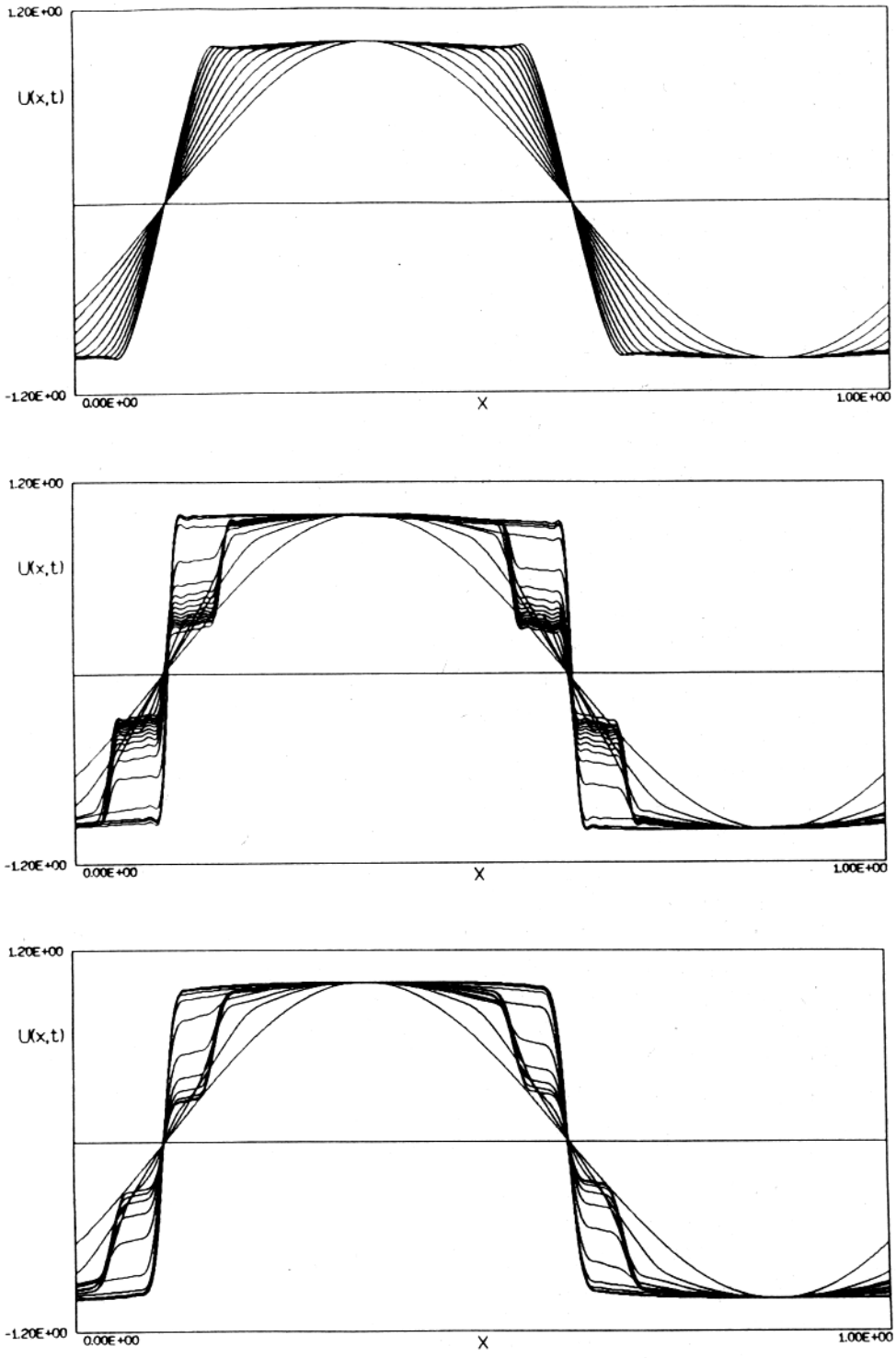

Fig. 5.4: Rescaling in each time step based on the maximum norm, $m = 6$, $\lambda = 4$
a) $\Delta t = 10^{-2}$, no presmoothing, b) $\Delta t = 10^{-3}$, no presmoothing, c) $\Delta t = 10^{-3}$, presmoothing with $J_s = 4$.

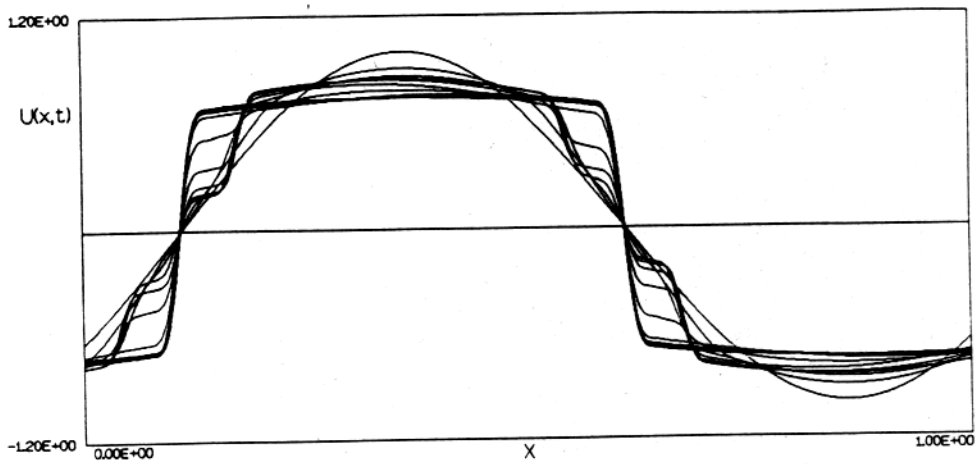

Fig. 5.5: Rescaling in each time step based on the energy norm, $m = 6$, $\lambda = 4$, $\Delta t = 10^{-3}$, presmoothing with $J_s = 4$.

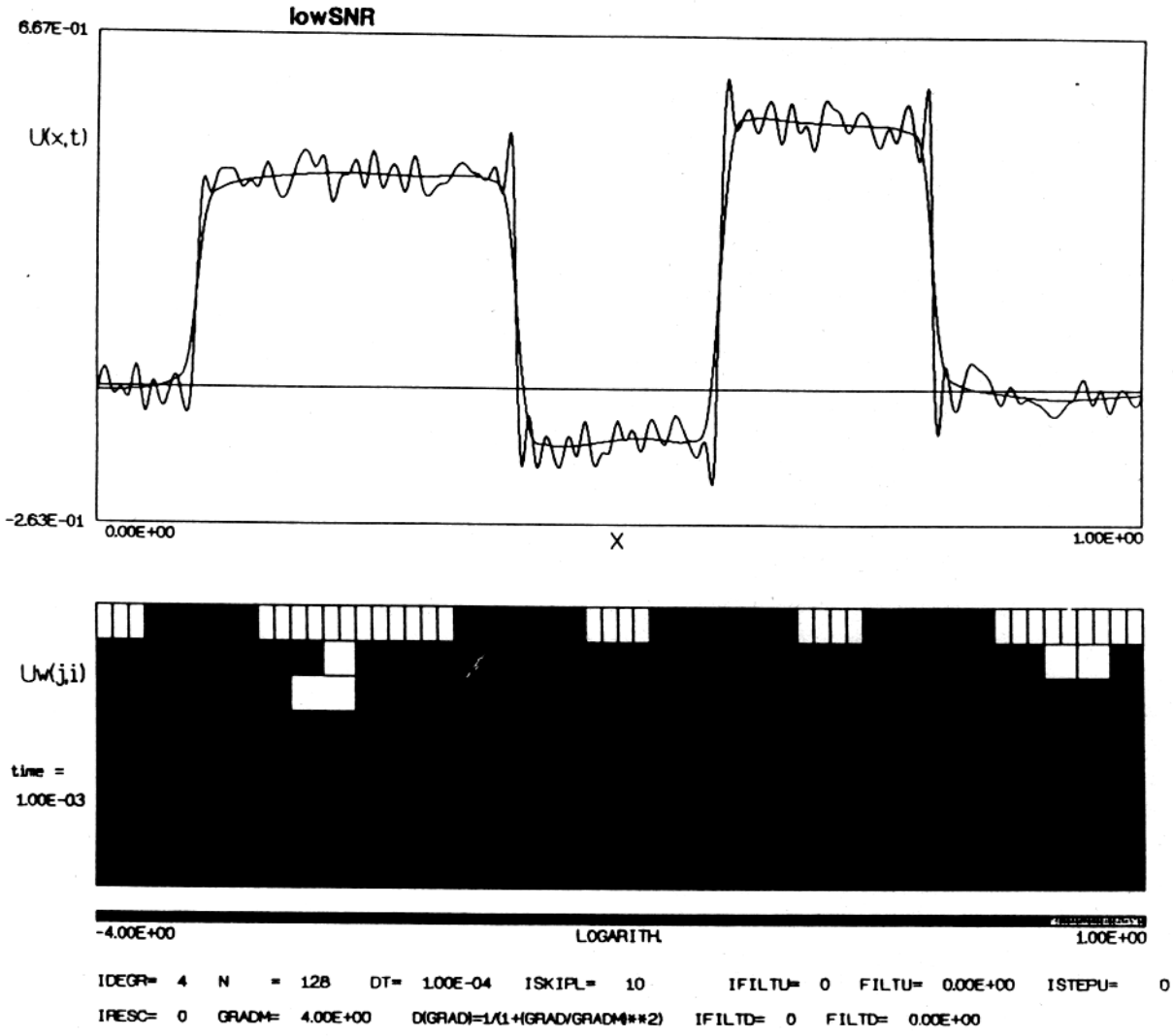

Fig. 6.1: Nonlinear diffusion for small noise amplitude, $m = 4$, $\lambda = 4$, $\Delta t = 10^{-4}$, no presmoothing.

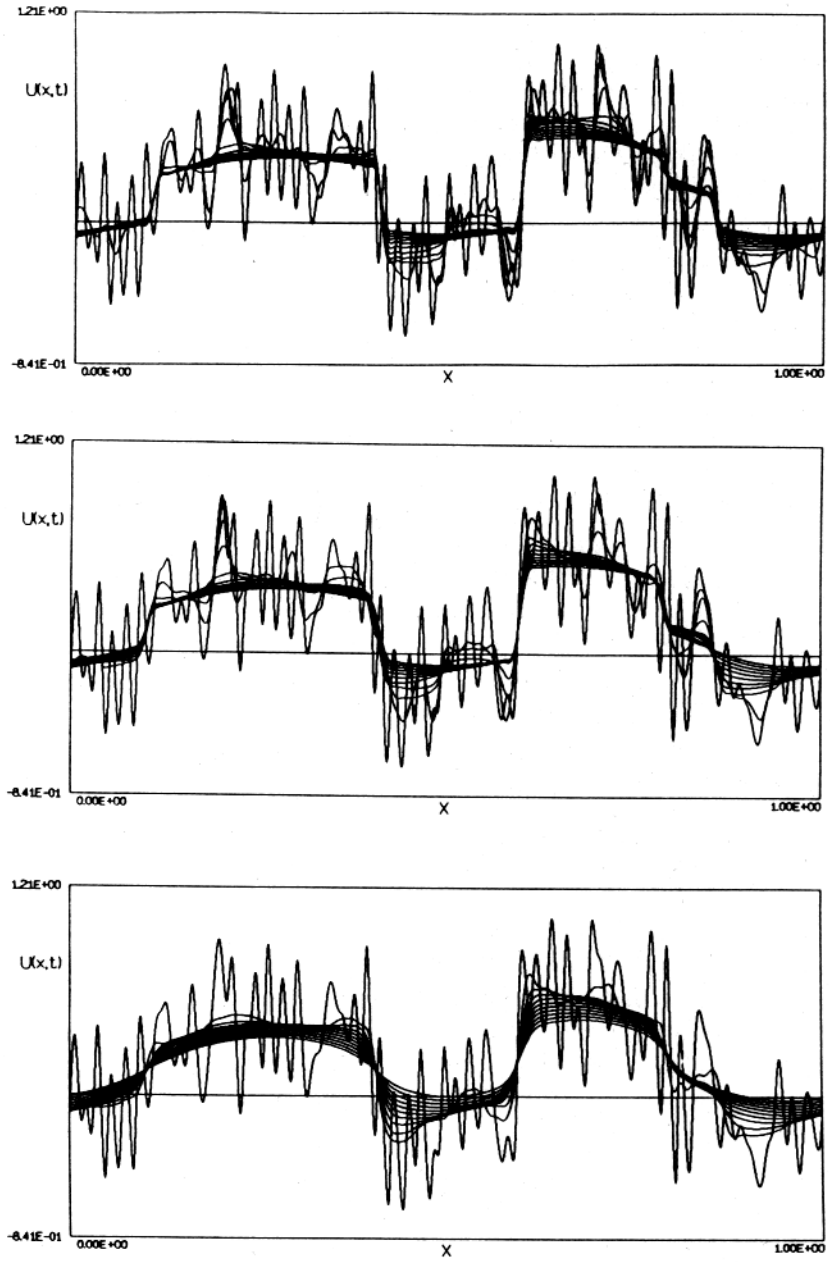

Fig. 6.2: The effect of presmoothing for large noise amplitude, $m = 6$, $\lambda = 4$, $\Delta t = 10^{-4}$, the function is plotted for $t = 0, 0.001, 0.002, \ldots$. a) no presmoothing, b) $J_s = 5$, c) $J_s = 4$.

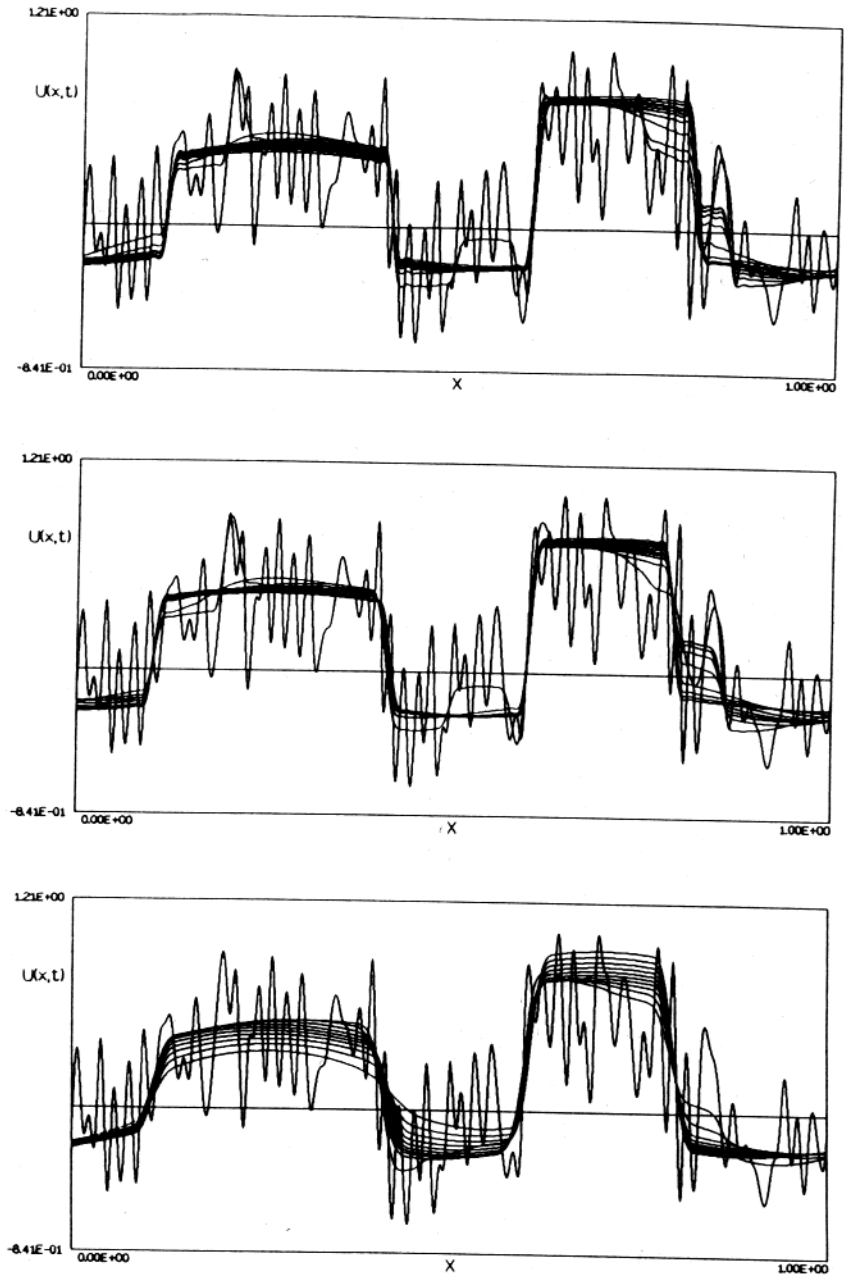

Fig. 6.3: Rescaling with the energy norm, $m = 6$, $\lambda = 4$, $\Delta t = 10^{-4}$, the function is plotted for $t = 0, 0.005, 0.01, \ldots$. a) no presmoothing, b) $J_s = 5$, c) $J_s = 4$.

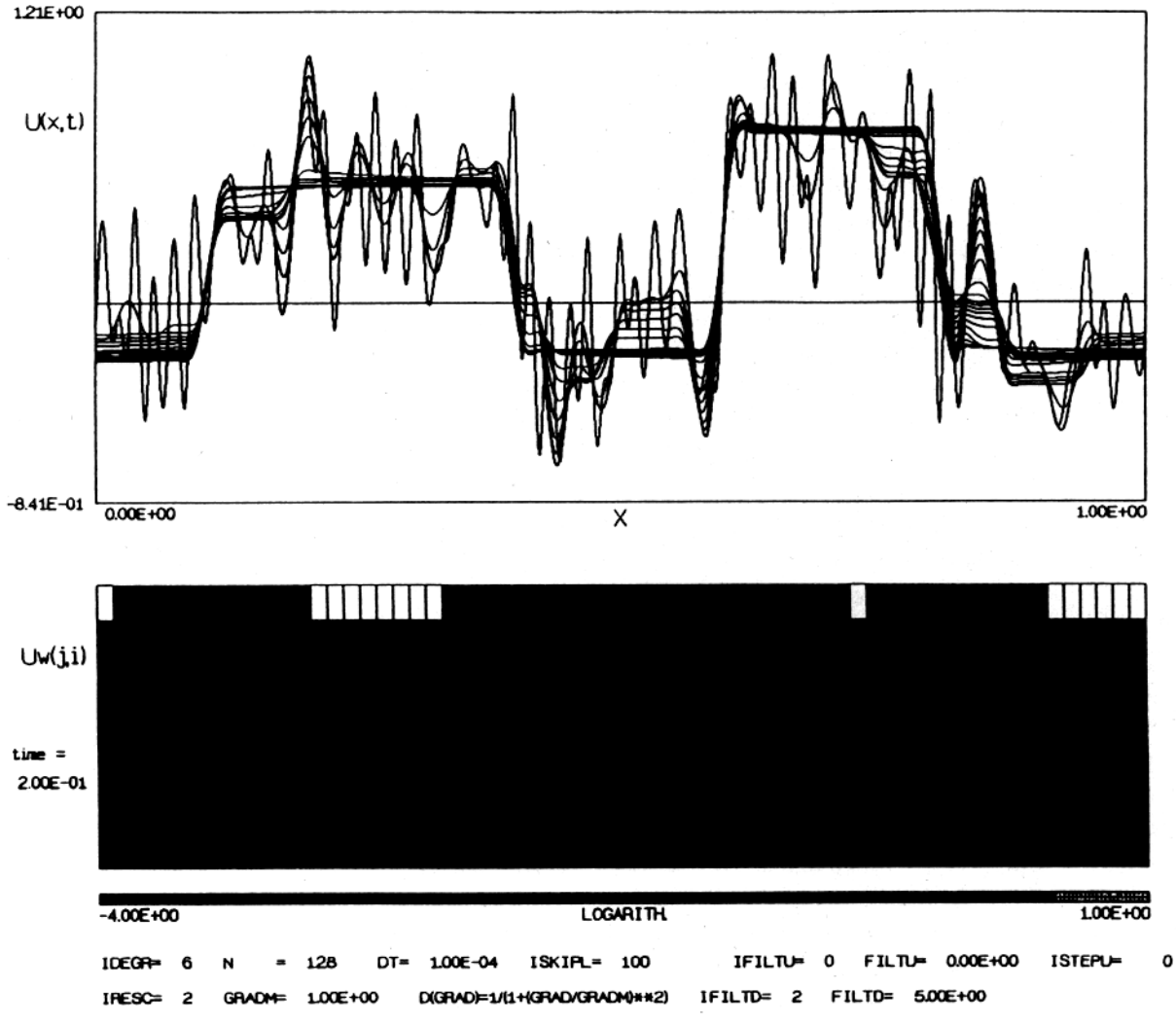

Fig. 6.4: A small value of λ while rescaling with the energy norm, $m = 6$, $\lambda = 1$, $\Delta t = 10^{-4}$, the function is plotted for $t = 0, 0.01, 0.02, \ldots$.

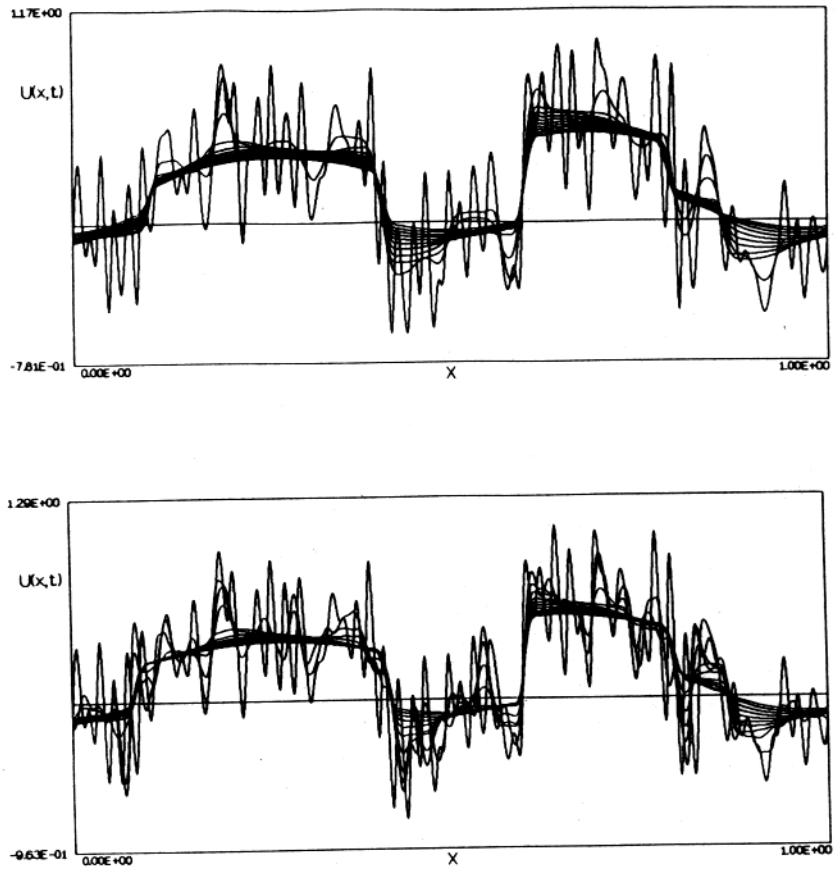

Fig. 6.5: Different regularity of the spline wavelet basis, $\lambda = 4$, $\Delta t = 10^{-4}$, $J_s = 5$, the function is plotted for $t = 0, 0.001, 0.002, \ldots$. a) $m = 4$, b) $m = 12$.

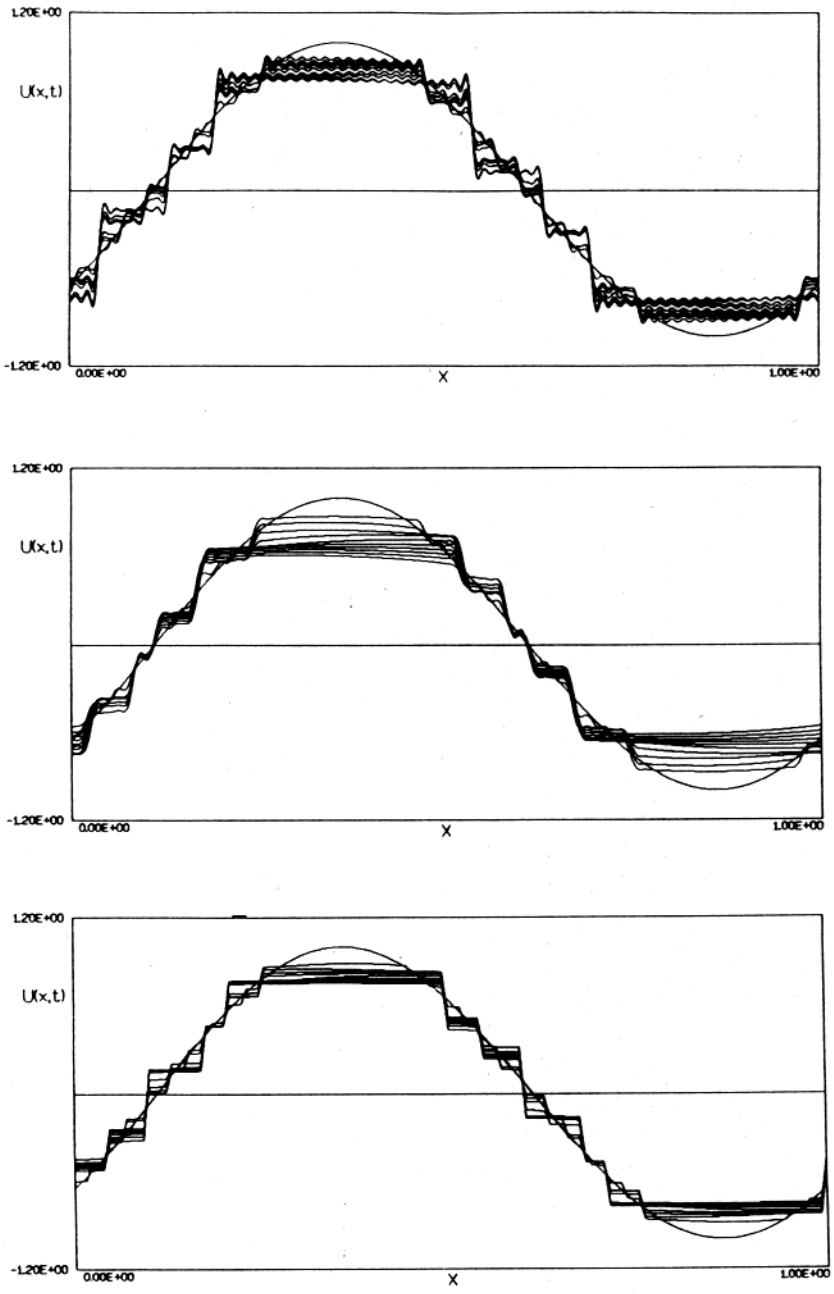

Fig. 7.1: Different algorithms applied to the sine function, $\lambda = 2$, $\Delta t = 10^{-4}$, no presmoothing, the function is plotted for $t = 0, 0.01, 0.02, \ldots$. a) Fourier method, b) wavelet method with $m = 6$, c) finite volume method.

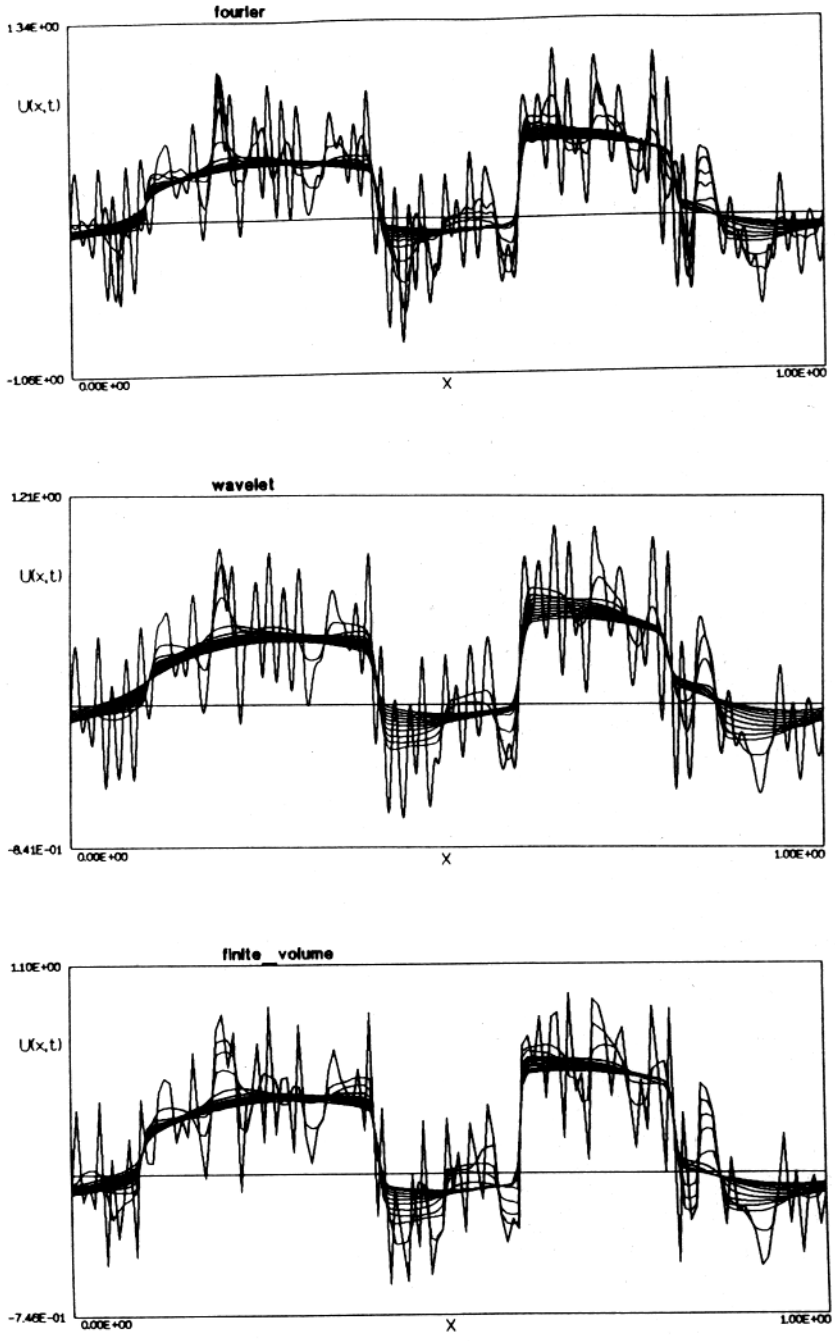

Fig. 7.2: Different algorithms applied to the noisy signal, $\lambda = 4$, $\Delta t = 5.10^{-5}$, presmoothing by Gaussian, the function is plotted for $t = 0, 0.001, 0.002, \ldots$. a) Fourier method, b) wavelet method with $m = 6$, c) finite volume method.