ADAPTIVE SOLUTION OF ONE–DIMENSIONAL SCALAR CONSERVATION LAWS WITH CONVEX FLUX

Folkmar A. Bornemann

Fachbereich Mathematik, Freie Universität Berlin, Arnimallee 2–6, D-14195 Berlin, and
Konrad–Zuse–Zentrum Berlin, Heilbronner Str. 10, D-10711 Berlin, Germany
bornemann@sc.zib-berlin.de

Summary

A new adaptive approach for one-dimensional scalar conservation laws with convex flux is proposed. The initial data are approximated on an adaptive grid by a problem dependent, monotone interpolation procedure in such a way, that the multivalued problem of characteristic transport can be easily and explicitly solved. The unique entropy solution is chosen by means of a selection criterion due to HOPF and LAX. For arbitrary times, the solution is represented by an adaptive monotone spline interpolation. The spatial approximation is controlled by local L^1-error estimates. As a distinctive feature of the approach, there is no discretization in time in the traditional sense. The method is monotone on fixed grids. Numerical examples are included, to demonstrate the predicted behavior.

Key Words. method of characteristics, adaptive grids, monotone interpolation, L^1–error estimates

AMS(MOS) subject classification. 65M15, 65M25, 65M50

Introduction

The construction of an adaptive algorithm for time-dependent hyperbolic conservation laws has to face a well known problem: Explicit direct discretization in time will yield a CFL-restriction of the time step, which will be prohibitively small when shocks are resolved by an adaptive space mesh. On the other hand, an implicit direct discretization in time introduces much numerical viscosity, since the velocity of information transport is modeled the wrong way. Thus, the time derivative should not be discretized directly at all.

Instead, one should try to attack the *evolution operator* of the problem, $\mathcal{E}_t : X \to X$, which maps admissible initial data to the solution at time t. An abstract framework of promising methods is given by the so called PERU-schemes of K.W. MORTON [11], which we state in a slightly more general way: Replace the evolution operator $\mathcal{E}_t$ by the sequence $\hat{\mathcal{E}}_t = \mathcal{P}\mathcal{E}_t\mathcal{R}$,

$$X \xrightarrow{\mathcal{R}} \hat{X} \xrightarrow{\mathcal{E}_t} X \xrightarrow{\mathcal{P}} X.$$

Here, $\mathcal{R}$ is a *(re)construction* operator, which maps the initial data in a more desirable space $\hat{X} \subset X$. The operator $\mathcal{P}$ *(re)presents* the result of the transport on $\hat{X}$ as a more appropriate element of the space X. Moreover, the exact transport $\mathcal{E}_t$ on the space $\hat{X}$ might be replaced by some simplification $\tilde{\mathcal{E}}_t$.

An example of this rather general approach is provided by Godunov's method and its generalizations. Here, $X = L^1$ and $\hat{X} \subset X$ consists of the piecewise constant functions on some mesh Δ. The reconstruction and representation operator are just the cell average projection onto $\hat{X}$, $\mathcal{R} = \mathcal{P}$. We have to consider three cases, depending on the Courant number $C = C(t, \Delta)$:

- $C < 1/2$. The evaluation of $\mathcal{E}_t$ on $\hat{X}$ is just the solution of non-interacting Riemann problems (classical Godunov scheme).
- $C \leq 1$. Here one can use the weak formulation of conservation laws in order to evaluate at least the product $\mathcal{P}\mathcal{E}_t$ on $\hat{X}$ by means of Riemann solvers.
- $C > 1$. There are simplifications $\tilde{\mathcal{E}}_t$ of the exact evolution operator, which approximate the now interacting Riemann problems and which are TVD. LEVEQUE [9] handles the interactions linearly in the large time-step Godunov's method. BRENIER [2] uses the transport-collapse operator, which is a certain averaging operator of the multivalued solution of characteristic transport. Both introduce errors in time.

The question which will be addressed is, whether there are any spaces $\hat{X}$ which are more powerful than piecewise constants in terms of approximation properties and which allow (at least for the most simple problems) the computation of the exact evolution operator $\mathcal{E}_t$ for arbitrary times t.

The answer given in this article is as follows: $\hat{X}$ can be chosen as the range of a certain interpolation operator $\mathcal{R}$, which depends on the right hand side of the problem. The space $\hat{X}$ will be as powerful for approximation as piecewise linears. The interpolation operator $\mathcal{R}$ is constructed in a way, that the computation of the multivalued solution of characteristic transport of initial data from $\hat{X}$ is extremely simple. Rather than averaging the multivalued solution like BRENIER [2] with the transport-collapse operator, we select the single valued entropy solution by means of the Lax-Hopf formula [6, 7]. The final representation of the solution is given by some interpolation operator $\mathcal{P}$, which could be any monotone spline interpolation. The space meshes will be constructed by an adaptive interpolation procedure, which is guided by local L^1-error estimates.

Up to now, the presented algorithm strongly relies on convexity properties of the underlying problem, so that the problem class which can be immediately handled seems to be quite restricted. However, there are strong indications that our approach will be at least applicable to the class of those systems of conservation laws, which are equivalent to Hamilton-Jacobi equations with convex Hamiltonian.

Theoretical Preparations

We are concerned with the solution of scalar conservation laws

$$u_t + f(u)_x = 0, \qquad u(\cdot, 0) = u_0, \tag{1}$$

where $u(\cdot, t)$ is a function on $\mathbb{R}$. Our general *assumptions* on the flux f will be

- $f : \mathbb{R} \to \mathbb{R}$ strictly convex and C^1.

Thus, the derivative $\alpha \equiv f' : \mathbb{R} \to]\alpha_-, \alpha_+[$ is one–one, onto and nondecreasing. The inverse of α will be denoted by $\beta \equiv \alpha^{-1} :]\alpha_-, \alpha_+[\to \mathbb{R}$.

A crucial role will be played by the *Legendre–Fenchel dual* $f^* :]\alpha_-, \alpha_+[\to \mathbb{R}$ of f, defined as

$$\begin{aligned} f^*(z) &= \sup_v (vz - f(v)) \\ &= u\alpha(u) - f(u) \qquad \text{with } u = \beta(z). \end{aligned}$$

Convex analysis (e.g., TIKHOMIROV [14]) states that

$$(f^*)' = \beta.$$

Our whole approach heavily relies on the following characterization of that weak solution, which satisfies the entropy condition.

THEOREM 1. (LAX [7]). *Let* $u_0 \in L^1(\mathbb{R})$ *and* $U_0(y) = \int_{-\infty}^{y} u_0(\xi)d\xi$. *For* $x \in \mathbb{R}$, $t > 0$ *define*

$$\mathcal{E}_t u_0(x) = \beta\left(\frac{x-y}{t}\right),$$

where $y = y(x,t) \in]x - t\alpha_-, x - t\alpha_+[$ *is any value which minimizes*

$$U_0(y) + tf^*\left(\frac{x-y}{t}\right) = \min! \,. \tag{2}$$

Then $\mathcal{E}_t u_0 \in L^1(\mathbb{R})$ *is the unique entropy solution at time* t *of the conservation law under consideration.*

If there exist several different values y*, which minimize (2) for a given* x*, then* x *is the position of a shock discontinuity. The limits* $\mathcal{E}_t u_0(x-0)$, $\mathcal{E}_t u_0(x+0)$ *exist, and*

$$\mathcal{E}_t u_0(x-0) \geq \beta\left(\frac{x-y}{t}\right) \geq \mathcal{E}_t u_0(x+0)$$

holds for every such y*.*

Remark. For $u \in L^\infty(\mathbb{R})$, $u_- \leq u_0 \leq u_+$ a.e., the solution only depends on f restricted to the interval $[u_-, u_+]$.

Remark. This Theorem has quite some history. HOPF [6] stated it for the inviscid Burgers equation $u_t + uu_x = 0$. He obtained the result in the limit $\mu \to 0$ of his explicit solution (i.e., the Cole–Hopf transformation to the heat equation) of the viscid Burgers equation $u_t + uu_x = \mu u_{xx}$. Later LAX [7] generalized the result to arbitrary convex fluxes f. A nice interpretation as Bellman's approach to the Hamilton–Jacobi equation $v_t + f(v_x) = 0$ can be found in LAX [8] or CONWAY and HOPF [3]. In fact, $U(x,t) = \int_{-\infty}^{x} u(\xi,t)d\xi$ satisfies the Hamilton–Jacobi equation and

$$U(x,t) = \min_y \left(U_0(y) + tf^*\left(\frac{x-y}{t}\right)\right),$$

if f is adjusted to $f(0) = 0$. This formula is actually connected with the modern notion of *viscosity solutions* of more general Hamilton–Jacobi equations, we refer to the book of P.-L. LIONS [10]. In [4] this connection is used to propose discretizations of Hamilton-Jacobi equations for the numerical solution of conservation laws.

Our approach uses the fact that, for certain u_0, the set of values y which possibly minimize (2) can be considerably restricted. A first step in that direction is the following

COROLLARY 2. *Assumptions as in Theorem 1. Additionally let u_0 be continuous. Any value $y \in]x-t\alpha_-, x-t\alpha_+[$ which minimizes (2) satisfies*

$$y + t\alpha(u_0(y)) = x, \tag{3}$$

and allows to set $\mathcal{E}_t u_0(x) = u_0(y)$.

Proof. Differentiating relation (2) yields, by continuity of u_0,

$$u_0(y) - \beta\left(\frac{x-y}{t}\right) = 0.$$

This implies both assertions. □

The nonlinear equation (3) is just the one which allows to construct, for smooth u_0 and small t, by means of the implicit function theorem, a classical solution of the conservation law. For larger t, equation (3) does not have a unique solution y for whole intervals of x. Thus, the minimum condition (2) can be understood as a *selection principle* for the right value of y.

The following stability result allows us to change u_0 slightly, in order to obtain simpler problems of the kind (3).

THEOREM 3. (KEYFITZ [12]). *The entropy solutions of (1) form a nonlinear L^1–contractive semigroup $\mathcal{E}_t$. Thus, for $u_0, v_0 \in L^1(\mathbb{R})$, the corresponding entropy solutions satisfy the estimate*

$$\|\mathcal{E}_t u_0 - \mathcal{E}_t v_0\|_{L^1} \leq \|u_0 - v_0\|_{L^1} \tag{4}$$

for all $t \geq 0$. A proof may also be found in LAX [8].

For our further development we need some *notation*:

- The convex hull of two points u_0, u_1 will be denoted by $[u_0, u_1]$.
- For $u \in [u_0, u_1]$ the *barycentric coordinate* of u is denoted by $\lambda_{u_0,u_1}(u)$ and satisfies

$$u = (1-\lambda_{u_0,u_1}(u))u_0 + \lambda_{u_0,u_1}(u)u_1.$$

- Let $(y_0, u_0), (y_1, u_1)$ be two points in $\mathbb{R}^2$. The *β–interpolant* of these points is given as $\mu_{u_0,u_1} : [y_0, y_1] \to [u_0, u_1]$ by

$$\mu_{u_0,u_1}\Big((1-\lambda)y_0 + \lambda y_1\Big) = \beta\Big((1-\lambda)\alpha(u_0) + \lambda\alpha(u_1)\Big), \qquad \lambda \in [0,1].$$

Our assumptions imply, that this is a monoton connection of the two points.

These β–interpolants have the very nice property that (3) may be solved uniquely, as we show now.

LEMMA 4. *For given $t > 0$ define*

$$\varphi(y) = y + t\alpha(\mu_{u_0,u_1}(y)),$$

which maps $[y_0, y_1]$ onto $[\varphi(y_0), \varphi(y_1)]$. If $\varphi(y_0) \neq \varphi(y_1)$, the equation $x = \varphi(y)$ is uniquely solved by the value y given as

$$\lambda_{y_0,y_1}(y) = \lambda_{\varphi(y_0),\varphi(y_1)}(x).$$

Proof. Simply note that

$$\varphi(y) = (1-\lambda_{y_0,y_1}(y))(y_0 + t\alpha(u_0)) + \lambda_{y_0,y_1}(y)(y_1 + t\alpha(u_1)),$$

and that $\varphi(y_i) = (y_i + t\alpha(u_i))$ for $i = 0, 1$. □

Since integrals are involved in (2), it helps a lot that integrals of β–interpolants can be computed explicitly:

LEMMA 5. *We have, for* $y \in [y_0, y_1]$, *that*

$$\int_{y_0}^{y} \mu_{u_0,u_1}(\eta)d\eta = \begin{cases} \dfrac{y_1 - y_0}{\alpha(u_1) - \alpha(u_0)} f^* \Big|_{\alpha(u_0)}^{\alpha(\mu_{u_0,u_1}(y))} & \text{if } u_0 \neq u_1, \\ (y - y_0)u_0 & \text{if } u_0 = u_1. \end{cases}$$

Proof. Let $u_0 \neq u_1$. The substitution $\zeta = (1-\lambda_{y_0,y_1}(\eta))\alpha(u_0) + \lambda_{y_0,y_1}(\eta)\alpha(u_1)$ gives

$$\int_{y_0}^{y} \mu_{u_0,u_1}(\eta)d\eta = \frac{y_1 - y_0}{\alpha(u_1) - \alpha(u_0)} \int_{\alpha(u_0)}^{\alpha(\mu_{u_0,u_1}(y))} \beta(\zeta)d\zeta.$$

Thus, the assertion follows from $(f^*)' = \beta$. □

Now we try to approximate u_0 by its piecewise β–interpolant, for which we have seen that problems (3) and (2) turn out to be fairly simple. For that purpose, let u_0 be a piecewise continuous function with $\operatorname{supp} u_0 \subset\subset]a, b[$. Let $\Delta : a = y_0 < y_1 < \ldots < y_n = b$ be a subdivision of that interval, with mesh–size parameter

$$h = \max_{1 \le j \le n} (y_j - y_{j-1}).$$

Denote the subintervals by $I_j = [y_{j-1}, y_j]$, $j = 1, 2, \ldots, n$. The piecewise β–interpolant $\mathcal{R}_{f,\Delta}u_0$ is now defined as

$$\mathcal{R}_{f,\Delta}u_0(y) = \begin{cases} \mu_{u_0(y_{j-1}),u_0(y_j)}(y) & \text{for } y \in I_j, j = 1, 2, \ldots, n, \\ 0 & \text{elsewhere.} \end{cases}$$

Obviously we have $\mathcal{R}_{f,\Delta}u_0 \in C^0$ and $\mathcal{R}_{f,\Delta}u_0(y_j) = u_0(y_j)$, $j = 1, 2, \ldots, n$.

LEMMA 6. *Let u be a piecewise continuous function on $[a, b]$. Then the interpolation operator $\mathcal{R}_{f,\Delta}$ satisfies*

$$\|u - \mathcal{R}_{f,\Delta}u\|_{L^1[a,b]} \to 0 \quad \text{for } h \to 0.$$

Moreover, if the function u is piecewise C^2, and we make the assumptions on the flux f that $f \in C^3$ with

$$M = \|f'''/f''^3\|_{L^\infty[u_-,u_+]} \|f''\|^2_{L^\infty[u_-,u_+]} < \infty,$$

where $u_- \le u(x) \le u_+$ for all $x \in [a, b]$, then there is a constant $c = c(u, M)$, which gives us the estimates

$$\|u - \mathcal{R}_{f,\Delta}u\|_{L^1[a,b]} \le ch, \quad \text{and} \quad \|u - \mathcal{R}_{f,\Delta}u\|_{L^1(I_c)} \le ch^2.$$

Here $I_c = \bigcup_{j \notin J_u} I_j$ *with* $J_u = \left\{ 1 \le j \le n \;\middle|\; \left(u|_{I_j} \right) \notin C^2 \right\}$.

Proof. We proof the second, smooth part. The first part follows by usual density arguments. Take any $j \notin J_u$, and denote the linear interpolation operator at the nodes y_{j-1}, y_j by $\mathcal{I}_\Delta$. Since by construction

$$\mathcal{I}_\Delta \mathcal{R}_{f,\Delta} = \mathcal{I}_\Delta$$

we estimate by the usual error expression for linear interpolation

$$\begin{aligned} \|u - \mathcal{R}_{f,\Delta} u\|_{L^\infty(I_j)} &\le \|u - \mathcal{I}_\Delta u\|_{L^\infty(I_j)} + \|\mathcal{R}_{f,\Delta} u - \mathcal{I}_\Delta \mathcal{R}_{f,\Delta} u\|_{L^\infty(I_j)} \\ &\le \frac{h^2}{8} \left(\|u''\|_{L^\infty[a,b]} + \|\mu''_{u(y_{j-1}),u(y_j)}\|_{L^\infty(I_j)} \right). \end{aligned}$$

Now we compute, for $y \in I_j$, that

$$\begin{aligned} \mu''_{u(y_{j-1}),u(y_j)}(y) &= -\frac{f'''\left(\mu_{u(y_{j-1}),u(y_j)}(y)\right)}{f''\left(\mu_{u(y_{j-1}),u(y_j)}(y)\right)^3} \left(\frac{\alpha(u(y_j)) - \alpha(u(y_{j-1}))}{u(y_j) - u(y_{j-1})} \cdot \frac{u(y_j) - u(y_{j-1})}{y_j - y_{j-1}} \right)^2 \\ &= -\frac{f'''\left(\mu_{u(y_{j-1}),u(y_j)}(y)\right)}{f''\left(\mu_{u(y_{j-1}),u(y_j)}(y)\right)^3} f''(\eta)^2\, u'(\zeta)^2 \end{aligned}$$

for some $\eta \in [u(y_{j-1}), u(y_j)]$, $\zeta \in I_j$. Hence, it is $\|\mu''_{u(y_{j-1}),u(y_j)}\|_{L^\infty(I_j)} \le M \|u'\|^2_{L^\infty[a,b]}$. For $j \in J_u$, we simply estimate

$$\|u - \mathcal{R}_{f,\Delta} u\|_{L^1(I_j)} \le 2\|u\|_{L^\infty[a,b]} h.$$

Finally, we observe that $\# J_u \le \nu$ as $h \to 0$, because we assumed that u is piecewise C^2. Thus, we obtain

$$\|u - \mathcal{R}_{f,\Delta} u\|_{L^1[a,b]} \le \frac{b-a}{8} \left(\|u''\|_{L^\infty[a,b]} + M\|u'\|^2_{L^\infty[a,b]} \right) h^2 + 2\nu \|u\|_{L^\infty[a,b]} h$$

and

$$\|u - \mathcal{R}_{f,\Delta} u\|_{L^1(I_c)} \le \frac{b-a}{8} \left(\|u''\|_{L^\infty[a,b]} + M\|u'\|^2_{L^\infty[a,b]} \right) h^2.$$

□

Note that the same result holds for the piecewise linear interpolation operator $\mathcal{I}_\Delta$, as introduced in the proof.

Remark. The value of the constant M is invariant against transformations $f \mapsto \gamma f$ with $\gamma > 0$.

Another important property of $\mathcal{R}_{f,\Delta}$ is monotonicity. This is a fairly simple consequence of the assumed monotonicity of α, β.

LEMMA 7. *Let* u, v *be piecewise continuous. The pointwise inequality* $u \le v$ *implies that pointwise* $\mathcal{R}_{f,\Delta} u \le \mathcal{R}_{f,\Delta} v$. *The same holds for the linear interpolation operator* $\mathcal{I}_\Delta$.

The Algorithm

Our algorithm solves the following problem: Given a conservation law (1), a piecewise continuous initial u_0, compute for an accuracy TOL and a time $t > 0$ an approximation $\hat{\mathcal{E}}_t u_0$ to the solution $\mathcal{E}_t u_0$ such that

$$\|\mathcal{E}_t u_0 - \hat{\mathcal{E}}_t u_0\|_{L^1} \leq \mathsf{TOL}\,. \tag{5}$$

The algorithm can be stated very roughly as

A. *Construct* by an adaptive interpolation procedure a mesh Δ_0, such that the piecewise β-interpolant $\mathcal{R}_{f,\Delta_0} u_0$ of u_0 fulfills

$$\|u_0 - \mathcal{R}_{f,\Delta_0} u_0\|_{L^1} \leq \mathsf{TOL}\,/2.$$

B_x. *Propagate* the function $\mathcal{R}_{f,\Delta_0} u_0$ with the evolution operator $\mathcal{E}_t$ to the time t, such that $\mathcal{E}_t \mathcal{R}_{f,\Delta_0} u_0(x)$ is evaluable.

C. *Represent* by an adaptive interpolation procedure (with the help of step B_x) the function $\mathcal{E}_t \mathcal{R}_{f,\Delta_0} u_0$ on a mesh Δ_t by its piecewise β-interpolant (resp. piecewise linear interpolant) $\hat{\mathcal{E}}_t u_0 = \mathcal{R}_{f,\Delta_t} \mathcal{E}_t \mathcal{R}_{f,\Delta_0} u_0$ (resp. $\hat{\mathcal{E}}_t u_0 = \mathcal{I}_{\Delta_t} \mathcal{E}_t \mathcal{R}_{f,\Delta_0} u_0$), such that

$$\|\mathcal{E}_t \mathcal{R}_{f,\Delta_0} u_0 - \hat{\mathcal{E}}_t u_0\|_{L^1} \leq \mathsf{TOL}\,/2.$$

If these steps can be achieved, Theorem 3 guarantees for the accuracy requirement (5).

The choice of the interpolant in Step C, i.e., $\mathcal{I}_{\Delta_t}$ or $\mathcal{R}_{f,\Delta_t}$, is not really important. In fact, any adaptive monotone spline interpolation, which controls the L^1–approximation error, could be used to represent the solution for fixed times. The β–interpolant should be taken, if we intend to use the solution at a particular time as new initial data for another computation.

We now describe each step more closely. Note that Steps A and C are quite similar tasks.

Step A. Here, the choice of an appropriate mesh Δ_0 is the essential problem. This will be done in an *adaptive* way, starting with a coarse mesh Δ^0. The main loop reads as:

```
while  (estimated L^1-error > TOL /2)
   {
         Δ^{k+1} = refine(Δ^k);
         k = k + 1;
   }
```

Let the kth mesh be $\Delta^k : a = y_0^k < y_1^k < \ldots < y_{n_k}^k = b$. For the following, we will suppress the index k. The L^1–error of the piecewise β-interpolation on the mesh Δ is given as

$$\epsilon = \|u_0 - \mathcal{R}_{f,\Delta} u_0\|_{L^1} = \sum_{j=1}^{n} \epsilon_j,$$

with

$$\epsilon_j = \int_{y_{j-1}}^{y_j} \left| u_0(\xi) - \mu_{u_0(y_{j-1}),u_0(y_j)}(\xi) \right| d\xi.$$

The local error ϵ_j will be estimated by a trapezoidal sum, introducing the midpoint of I_j. Thus, noting the interpolation property, we obtain the local estimate

$$\epsilon_j \approx \eta_j = \frac{(y_j - y_{j-1})}{2} \left| u_0\left(\frac{y_{j-1} + y_j}{2}\right) - \mu_{u_0(y_{j-1}),u_0(y_j)}\left(\frac{y_{j-1} + y_j}{2}\right) \right|.$$

Observe that

$$\mu_{u_0(y_{j-1}),u_0(y_j)}\left(\frac{y_{j-1}+y_j}{2}\right)=\beta\left(\frac{1}{2}\left(\alpha(u_0(y_{j-1}))\ \ +\alpha(u_0(y_j))\right)\right).$$

In the case that I_j is bisected, we note that $u_0((y_{j-1}+y_j)/2)$ has already been computed for η_j. Thus, we can readily assign this value to the new node.

The proposed error estimate is sensible for accuracy and complexity reasons. The global estimate is finally given as

$$\eta=\sum_{j=1}^{n}\eta_j.$$

For actual refinement we need some *refinement strategy*, which uses the local information provided by the indicators η_j. We have implemented a strategy based on local extrapolation, which was introduced by BABUŠKA and RHEINBOLDT [1] for elliptic problems.

The actual implementation of the mesh refinement can easily be done by means of packages designed for finite element computations using tree data structures, e.g., ROITZSCH [13].

After the adaptive refinement we are provided with the final mesh Δ_0, an error estimate $\eta<\mathsf{TOL}/2$, and each node y_j carries the interpolation information $u_0(y_j)$. For purposes of Step B_x, we should additionally store in each node the integral information

$$\int_{-\infty}^{y_j}\mathcal{R}_{f,\Delta_0}u_0(\xi)d\xi,$$

which can be computed by successive application of Lemma 5.

Step C. The piecewise β-interpolant (piecewise linear interpolant) $\hat{\mathcal{E}}_t u_0$ of the propagated function $\mathcal{E}_t\mathcal{R}_{f,\Delta_0}u_0$ and the mesh Δ_t are computed in a similar fashion as $\mathcal{R}_{f,\Delta_0}u_0$ and Δ_0. This approximation procedure only demands the possibility of evaluating the expression $\mathcal{E}_t\mathcal{R}_{f,\Delta_0}u_0$ in certain points x.

Step B_x. How do we compute $\mathcal{E}_t\mathcal{R}_{f,\Delta_0}u_0(x)$ for a given x? This question will be addressed now. In preparation of any evaluation, the following values are computed

$$\varphi(y_j)=y_j+t\alpha(\mathcal{R}_{f,\Delta_0}u_0(y_j))$$

for $j=0,1,\ldots,n$. These are the positions of the characteristic transport of the mesh points y_j of Δ_0. Given x, we first determine the set J_x of indices j, such that

$$x=y+t\alpha(\mathcal{R}_{f,\Delta_0}u_0(y)) \tag{6}$$

possesses a solution $y\in I_j$. By construction of $\mathcal{R}_{f,\Delta_0}u_0$ and Lemma 4, this set is exactly given by

$$J_x=\{1\le j\le n\mid x\in[\varphi(y_{j-1}),\varphi(y_j)]\}.$$

For $j\in J_x$ with $\varphi(y_{j-1})\neq\varphi(y_j)$, we compute the barycentric coordinate

$$\lambda_j=\lambda_{\varphi(y_{j-1}),\varphi(y_j)}(x),$$

which is, by Lemma 4, also the barycentric coordinate of the unique solution $\bar{y}_j \in I_j$, i.e., $\bar{y}_j = (1-\lambda_j)y_{j-1} + \lambda_j y_j$. In the exceptional case $j \in J_x$, $\varphi(y_{j-1}) = \varphi(y_j)$, all values $y_{j-1} \le y \le y_j$ satisfy (6). Thus, the value of the expression in (2) remains constant on the whole interval $[y_{j-1}, y_j]$. Hence, we may take $\bar{y}_j = y_{j-1}$ as representative candidate for the minimizing value of (2). Summarizing, our construction of $\mathcal{R}_{f,\Delta_0}u_0$ allows us to compute a set of critical points of (2) with cardinality $\#J_x$, in which a minimizing value is included.

In view of Theorem 1 and its Corollary we choose the smallest value $\bar{y}_\ell$ among those $\bar{y}_j, j \in J_x$, which minimize (2). Our desired value of $\mathcal{E}_t\mathcal{R}_{f,\Delta_0}u_0(x)$ is given by

$$\mathcal{E}_t\mathcal{R}_{f,\Delta_0}u_0(x) = \mathcal{R}_{f,\Delta_0}u_0(\bar{y}_\ell) = \beta\left((1-\lambda_\ell)\alpha(u_0(y_{\ell-1})) + \lambda_\ell\alpha(u_0(y_\ell))\right).$$

For the evaluation of (2) it is necessary to rely on Lemma 5.

If there are several values $\bar{y}_k$ which minimize (2) among the $\bar{y}_j, j \in J_x$, we are allowed, due to Theorem 1, to take any of them: In this case, x is exactly the position of a shock of $\mathcal{E}_t\mathcal{R}_{f,\Delta_0}u_0$. All minimizing $\bar{y}_k$ produce values for $\mathcal{E}_t\mathcal{R}_{f,\Delta_0}u_0(x)$ which are between the left and the right shock value. In fact, since we choose the smallest $\bar{y}_k$, which minimizes (2), we can be more specific. We obtain

$$\mathcal{E}_t\mathcal{R}_{f,\Delta_0}u_0(x) = \mathcal{E}_t\mathcal{R}_{f,\Delta_0}u_0(x-0) \tag{7}$$

for any shock position x. Note that the specification $\bar{y}_j = y_{j-1}$ in the case $\varphi(y_{j-1}) = \varphi(y_j)$ also served this purpose: It guarantees, that the smallest minimizing value of the $\bar{y}_j$ is really the smallest value of all minimizing values for (2).

A simple implication of the monotonicity property of $\mathcal{R}_{f,\Delta}$ and $\mathcal{I}_\Delta$ (Lemma 7), together with the monotonicity of the semigroup, is the monotonicity of our algorithm as long as we fix the meshes Δ_0, Δ_t.

LEMMA 8. *Let u_0, v_0 be piecewise continuous functions, such that pointwise $u_0 \le v_0$. If we use for both functions the same meshes Δ_0, Δ_t, we obtain that pointwise $\hat{\mathcal{E}}_t u_0 \le \hat{\mathcal{E}}_t v_0$.*

Proof. Care should be taken, if x is a shock position of both $\mathcal{E}_t\mathcal{R}_{f,\Delta_0}u_0$ and $\mathcal{E}_t\mathcal{R}_{f,\Delta_0}v_0$. Here, one has to rely on (7). If (7) wouldn't hold, one would have to exclude a neighborhood of x. □

In order to run our algorithm, we need procedures for evaluating f, α and β. If β is not given analytically, we may compute it by Newton's method.

Numerical Examples

Important: Since we can evaluate $\mathcal{E}_t\mathcal{R}_{f,\Delta_0}u_0(x)$ *exactly* for any (x,t), $t > 0$, it should be clear, that the time-steps of the examples have been solely introduced for graphical reasons. They are completely arbitrary and independent, and we work for all times with the same $\mathcal{R}_{f,\Delta_0}u_0$. Thus, *there is no discretization in time!* Once more, we remark, that any adaptive monotone spline interpolation, which controls the L^1-approximation error, could be used to represent the function $\mathcal{E}_t\mathcal{R}_{f,\Delta_0}u_0$. For simplicity, we have chosen piecewise linear interpolation whenever displacing our approximate solution graphically.

Example 1. Here, we consider the nonlinear conservation law

$$u_t + \left(\frac{u^4}{4}\right)_x = 0$$

with initial data

$$u_0(x) = \begin{cases} 1 & \text{for } 0 \le x \le 1, \\ 0 & \text{elsewhere.} \end{cases}$$

The inverse of the flux derivative, $\beta = (\cdot)^{1/3}$, is quite different from a linear function, giving the β-interpolant a distinguishable shape.

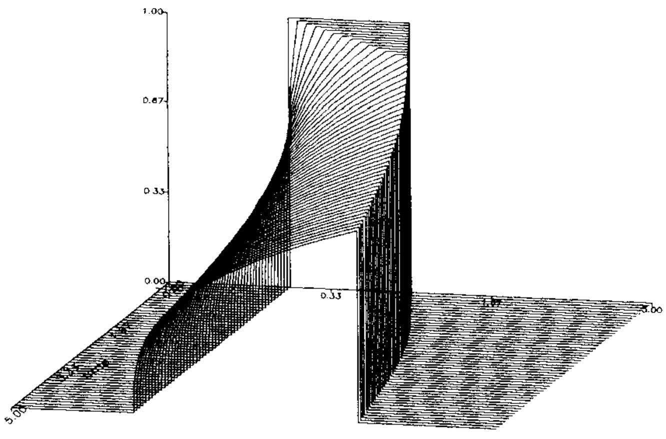

FIG. 1. *Example 1. Evolution of the computed solution.*

The exact solution is given by

$$u(x,t) = \begin{cases} \begin{cases} \left(\frac{x}{t}\right)^{1/3} & 0 \le x \le t \\ 1 & t \le x \le 1 + t/4 \\ 0 & \text{elsewhere} \end{cases} & \text{for } 0 < t \le 4/3, \\ \begin{cases} \left(\frac{x}{t}\right)^{1/3} & 0 \le x \le \left(\frac{4}{3}\right)^{3/4} t^{1/4} \\ 0 & \text{elsewhere} \end{cases} & \text{for } t > 4/3. \end{cases}$$

The computed solution for $0 \le t \le 5$, using a time-step $\tau = 0.1$ *for graphical reasons*, can be seen in Fig. 1. If not stated otherwise, we choose as accuracy $\mathsf{TOL} = 10^{-4}$. The solution was represented by the adaptive linear interpolation of Step C.

The solution for the particular time $t = 1.0$ is shown in Fig. 2, represented by the adaptive linear interpolation. We observe that, as a result of our construction (7), there is no mesh point with a value between the left and the right shock value.

The development of the interpolation mesh in time, here with time-step $\tau = 0.05$, can be seen in Fig. 3. We can observe nicely, how the rarefaction wave runs into the shock and slows it down.

Using the adaptive β-interpolation to represent the solution, we get much less mesh points. This is precisely what should have been expected for this example: Rarefaction waves are exactly represented by the β-interpolant of the left and right value. For the same accuracy as above, the corresponding mesh ($\tau = 0.05$) is shown in Fig. 4.

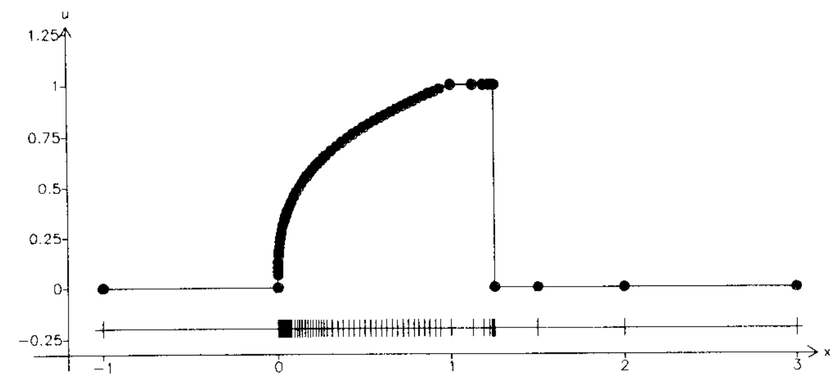

FIG. 2. *Example 1. Solution for $t = 1.0$, represented by adaptive linear interpolation.*

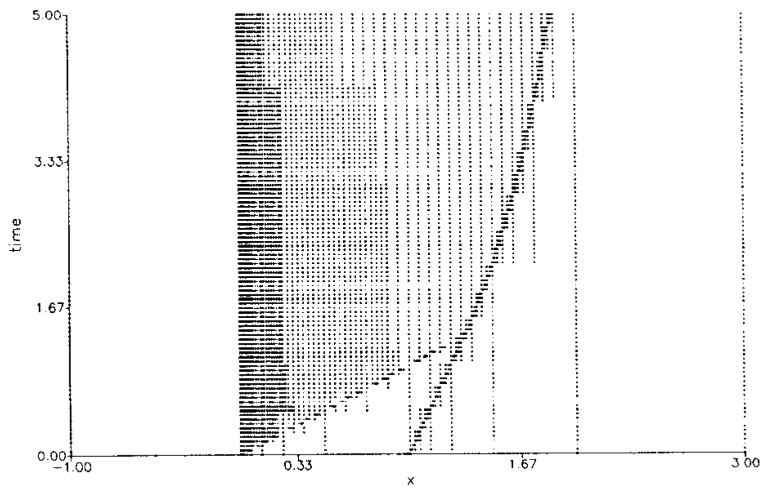

FIG. 3. *Example 1. Grid, using adaptive linear interpolation.*

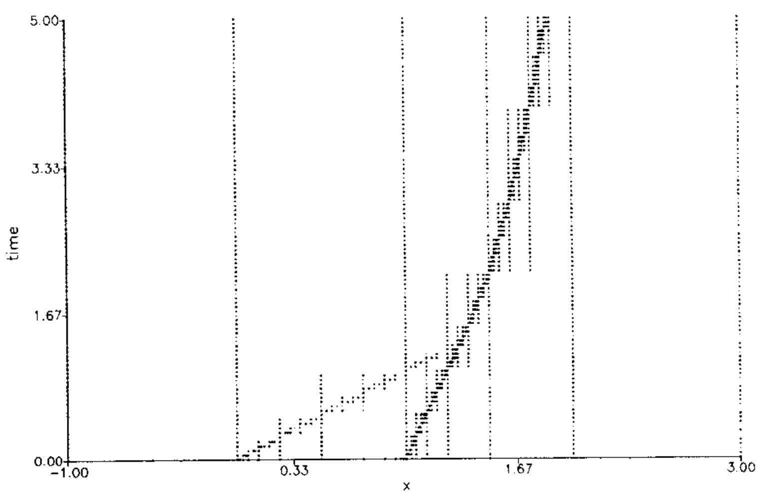

FIG. 4. *Example 1. Grid, using adaptive β-interpolation.*

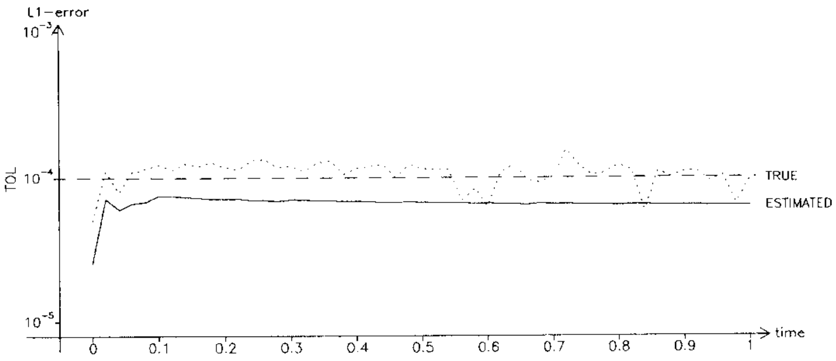

FIG. 5. *Example 1. Evolution of the error, true (...) and estimated (—).*

The quality of our error estimator can be seen in Fig. 5. We observe a slight error underestimation. Our estimated error η at time t is the *sum* of the estimated β–interpolation error η_0 of the initial data and of the estimated linear interpolation error η_t of $\hat{u}$,

$$\eta = \eta_0 + \eta_t .$$

Compared with the true L^1–error ϵ, we obtained for all of our experiments (i.e., $0 \le t \le 5$, $\tau = 0.05$, $\mathsf{TOL} = 10^{-1}, \ldots, 10^{-8}$, linear as well as β–interpolation of the solutions), that

$$0.33 \le \frac{\eta}{\epsilon} \le 1.97 \, .$$

Finally, we show in Fig. 6 the dependence of the CPU–time (in seconds) on the accuracy TOL, for the case, that we represent the solution at each time by the adaptive linear interpolation. The comparison has been made using 100 time–steps of size $\tau = 0.05$ for each accuracy. The dotted line in the double–logarithmic scale has slope $-1/2$. We observe, that asymptotically

$$\text{CPU-time} \propto \mathsf{TOL}^{-1/2} \, . \tag{8}$$

This is an optimal result, since, for the set $\mathcal{S}_{2,n}$ of piecewise linear functions with no more

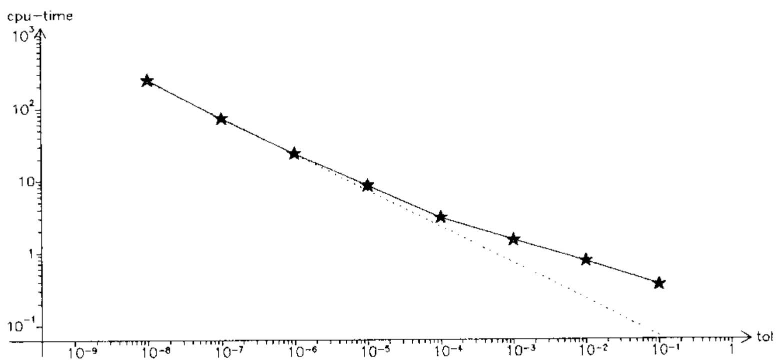

FIG. 6. *Example 1. Computing time vs.* TOL.

than n breaks in the first derivative, we obtain

$$\mathrm{dist}(\mathcal{E}_t u_0, \mathcal{S}_{2,n}) = \mathcal{O}\left(n^{-2}\right),$$

a result, which can be found in DE BOOR [5, Theorem III.2]. Thus, the behavior (8) shows two things: First, that our mesh was chosen nearly optimal, second, that we realized our algorithm with an complexity of $\mathcal{O}(\#\text{nodes})$.

Example 2. The problem of this example is given by the inviscid Burgers equation

$$u_t + \left(\frac{u^2}{2}\right)_x = 0$$

with initial data

$$u_0(x) = \begin{cases} 2.4 + \sin(\pi(x-0.5)) & \text{for } 0.5 \le x \le 2.5, \\ 2.4 & \text{elsewhere.} \end{cases}$$

This initial data does not have a compact support, but we can obviously modify our algorithm to handle this kind of problems.

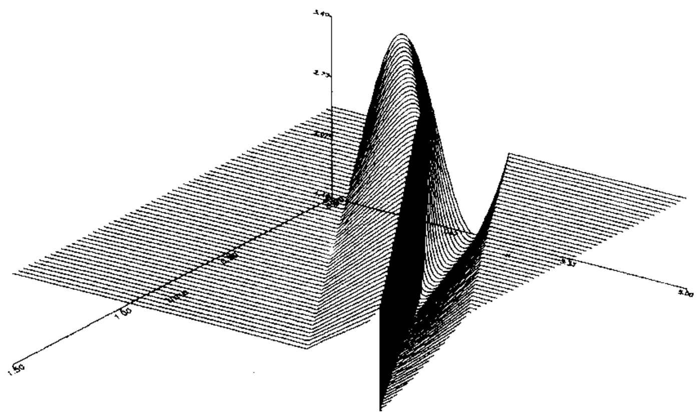

FIG. 7. *Example 2. Evolution of the computed solution.*

The continuous initial u_0 develops a shock at time $t = 1/\pi \approx 0.318$. The computed solution can be seen in Fig. 7. It was computed with accuracy $\mathsf{TOL} = 10^{-4}$ in the time interval $[0, 1.5]$, using a time-step $\tau = 0.025$.

The corresponding mesh is shown in Fig. 8. The number of mesh points varies between 330 at the beginning and 11 at the end.

Figs. 9 and 10 show a zoom into the computed solution represented by piecewise linear interpolation just before and just after the shock formation. In both cases we have taken the position $x_s = 1.5 + 2.4t$, and have shown the computed solution in the interval $[x_s - 0.01, x_s + 0.01]$.

Note that in this problem there is no difference between piecewise linear and β-interpolation: $\mathcal{I}_\Delta = \mathcal{R}_{f,\Delta}$.

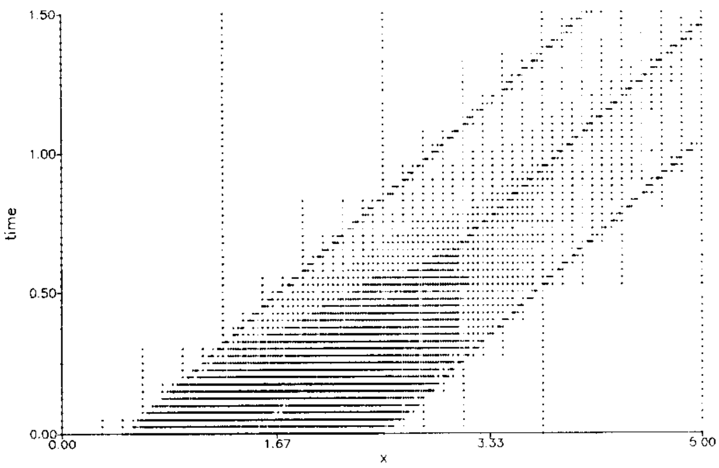

FIG. 8. *Example 2. Evolution of the mesh.*

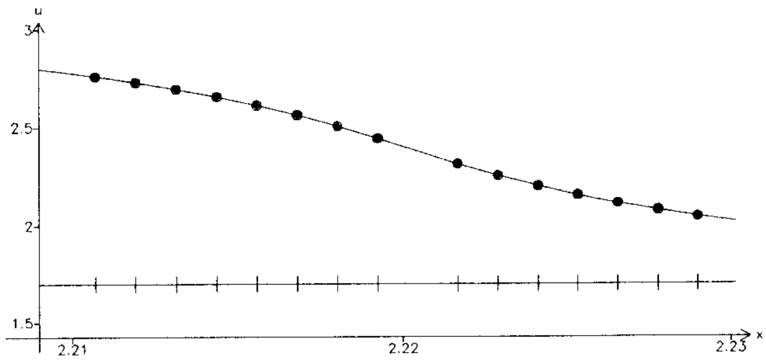

FIG. 9. *Example 2. Zoom into solution, just before the shock (t = 0.3).*

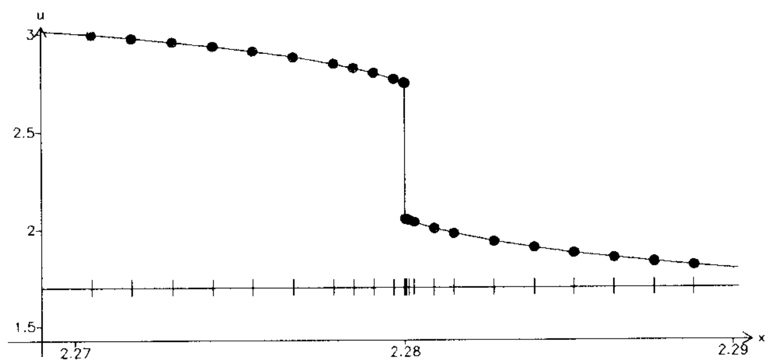

FIG. 10. *Example 2. Zoom into solution, just after the shock (t = 0.325).*

REFERENCES

[1] I. BABUŠKA AND W. C. RHEINBOLDT, *Error estimates for adaptive finite element computations*, SIAM J. Numer. Anal., 15 (1978), pp. 736–754.

[2] Y. BRENIER, *Averaged multivalued solutions for scalar conservation laws*, SIAM J. Numer. Anal., 21 (1984), pp. 1013–1037.

[3] E. CONWAY AND E. HOPF, *Hamilton's theory and generalized solutions of the Hamilton–Jacobi equation*, J. Math. Mech., 13 (1964), pp. 939–986.

[4] L. CORRIAS, M. FALCONE, AND R. NATALINI, *Numerical schemes for conservation laws via Hamilton-Jacobi equations*, Istituti per le applicazioni del calcolo, Roma, 1992.

[5] C. DE BOOR, *A Practical Guide to Splines*, Springer–Verlag, New York, Heidelberg, Berlin, 1978.

[6] E. HOPF, *The partial differential equation* $u_t + uu_x = \mu u_{xx}$, Comm. Pure Appl. Math., 3 (1950), pp. 201–230.

[7] P. D. LAX, *Hyperbolic systems of conservation laws II*, Comm. Pure Appl. Math., 7 (1957), pp. 537–566.

[8] ——, *Hyperbolic Systems of Conservation Laws and the Mathematical Theory of Shock Waves*, Conf. Board Math. Sci. 11, SIAM, Philadelphia, 1973.

[9] R. J. LEVEQUE, *A large time step generalization of Godunov's method for systems of conservation laws*, SIAM J. Numer. Anal., 22 (1985), pp. 1051–1073.

[10] P. L. LIONS, *Generalized Solutions of Hamilton–Jacobi Equations*, Research Notes in Mathematics 69, Pitman, Boston, London, Melbourne, 1982.

[11] K. W. MORTON, *Private communication.*

[12] B. L. QUINN-KEYFITZ, *Solutions with shocks, an example of an* L_1*-contractive semigroup*, Comm. Pure Appl. Math., 24 (1971), pp. 125–132.

[13] R. ROITZSCH, *KASKADE Programmer's Manual*, Konrad–Zuse–Zentrum, Berlin, 1989. Technical Report TR89–5.

[14] V. M. TIKHOMIROV, *Convex analysis*, in Analysis II, Encyclopaedia of Mathematical Sciences Vol. 14, R. Gamkrelidze, ed., Springer–Verlag, Berlin, Heidelberg, New York, 1990.