FINITE DIFFERENCE SCHEMES FOR SCALAR CONSERVATION LAWS WITH SOURCE TERMS

HANS JOACHIM SCHROLL * AND RAGNAR WINTHER **

— IMA J. Num. Anal., 1996(16), pp 201-215. —

Abstract. Explicit and semi–implicit finite difference schemes approximating nonhomogenous scalar conservation laws are analyzed. Optimal error bounds independent of the stiffness of the underlying equation are presented.

Key words. Hyperbolic conservation law, source term, finite difference scheme, error estimate.

AMS(MOS) subject classifications. 35L65, 65M06, 65M15

1. Introduction. The purpose of the present paper is to study finite difference schemes applied to the Cauchy problem for nonhomogenous scalar conservation laws of the form:

$$u_t + f(u)_x = g(u). \tag{1.1}$$

From a numerical point of view, one obviously has to distinguish the stiff and the nonstiff case. According to Pember (1992), the conservation law (1.1) is stiff if the time scale introduced by the source term g is small compared to the characteristic speed f' and some appropriate length scale. The generic form of a stiff conservation law is

$$u_t + f(u)_x = \frac{1}{\delta}\, g(u), \tag{1.2}$$

where the relaxation time δ is a small parameter while the equation (1.1) corresponding to $\delta = 1$ is referred to as the nonstiff case.

The schemes that shall be analyzed are modifications of Godunov's method. In case of the nonstiff equation (1.1) we analyze the explicit method:

$$\frac{u_j^{n+1} - u_j^n}{\Delta t} + \frac{f(u^*(u_j^n, u_{j+1}^n)) - f(u^*(u_{j-1}^n, u_j^n))}{\Delta x} = g(u_j^n). \tag{1.3}$$

Here $u^*(u^-, u^+)$ is the solution of the homogenous Riemann problem

$$u_t + f(u)_x = 0, \quad u(x,0) = \begin{cases} u^-, & x < 0 \\ u^+, & 0 < x \end{cases}$$

evaluated along the ray $x/t = 0$.

* Numerische Mathematik, RWTH–Aachen, Templergraben 55, D-52056 Aachen.
This author has been supported by The Norwegian Research Council (NFR), program no. 100284/431.
** Department of Informatics, P. O. Box 1080 Blindern, University of Oslo, Norway.
This author has been supported by The Norwegian Research Council (NFR), program no. 100284/431 and STP.29643.
e-mail: schroll@igpm.rwth-aachen.de, ragnar@ifi.uio.no

As an approximation of the stiff equation (1.2), we discuss the semi–implicit method:

$$\frac{u_j^{n+1}-u_j^n}{\Delta t}+\frac{f(u^*(u_j^n,u_{j+1}^n))-f(u^*(u_{j-1}^n,u_j^n))}{\Delta x}=\frac{1}{\delta}\,g(u_j^{n+1}). \tag{1.4}$$

Note that for monotone fluxes, for example $f'>0$ the solution of the Riemann problem is given by the left state $u^*(u^-,u^+)=u^-$ and both schemes reduces to simple backward difference methods.

We are particularly interested in estimating the global error of both schemes in L^1. In the nonstiff case we will prove, that if the proper CFL-condition is satisfied, the L^1-error is $\mathcal{O}(\sqrt{\Delta t})$. This corresponds to a well known result for homogenous, scalar conservation laws, see eg. Kuznetsov (1976) or Lucier (1985). It should be mentioned here that it is not difficult to extend Lucier's argument to the nonhomogenous case and to obtain an error estimate for the real Godunov scheme applied to (1.1) or (1.2). The problem is that in this situation Godunov's method does not reduce to a computable scheme, since it requires the exact solution of the conservation law with the source term included on small time intervals. In the stiff case we shall establish under proper additional assumptions on g that the $\mathcal{O}(\sqrt{\Delta t})$ estimate holds independently of the relaxation time δ.

For a explicit scheme similar to (1.3) convergence towards the entropy solution of (1.1) was proved by Chalabi (1992). As it is based on Helly's theorem, his result does not give any information on the rate of that convergence.

A semi–implicit scheme of type (1.4) applied to a special stiff system has been analyzed by Schroll, Tveito & Winther (1994). The argument presented in the present paper is strongly related to the technique introduced by Schroll et al. (1994). We will refer to that paper for several details.

The present paper is organized as follows: In Section 2 we introduce the notation, give the exact assumptions on both models (1.1) and (1.2) and review properties of the entropy solution. Furthermore, the main results are stated. In Section 3 we prove the error estimate in the nonstiff case, and Section 4 is devoted to the semi–implicit scheme and the stiff case. Finally, Section 5 provides proofs for some auxiliary results.

2. Preliminaries and the main results. Throughout the paper, we assume the flux f and the source g to be given smooth functions satisfying

$$g(0)=0.$$

We shall consider solutions of the Cauchy problem with given initial data

$$u^0\in L^1(\mathcal{R})\cap BV(\mathcal{R},\mathcal{R}).$$

Here, BV denotes the subspace of L^1_{loc} consisting of functions with bounded total variation

$$TV(z):=\sup_{h\neq 0}\int_{\mathcal{R}}\frac{|z(x+h)-z(x)|}{|h|}dx.$$

Given the above assumptions, it is well known that there is a unique entropy solution of the Cauchy problem for (1.1) or (1.2). This solution is defined by the so–called Kruzkov inequality (Kruzkov 1970, definition 1):

$$
\begin{aligned}
&\int_0^T \int_{\mathcal{R}} [|u-k|\varphi_t + \sigma(u-k)(f(u)-f(k))\varphi_x]\, dx\, dt \\
&+\int_{\mathcal{R}} [|u^0-k|\varphi(x,0) - |u(x,T)-k|\varphi(x,T)]\, dx \\
&\geq -\frac{1}{\delta}\int_0^T \int_{\mathcal{R}} [\sigma(u-k)g(u)\varphi]\, dx\, dt, \\
&\text{for any } T>0,\ k\in\mathcal{S} \text{ and any } \varphi\in\mathcal{D}_+(T).
\end{aligned} \tag{2.1}
$$

Here, σ is the sign function and $\mathcal{D}_+(T)$ is the set of all nonnegative C^∞–functions with compact support in $\mathcal{R}\times[0,T]$.

Throughout this paper, the final time $T>0$ will be considered fixed. The solution of the nonstiff equation (1.1) is assumed to be located in the state space

$$\mathcal{S} = \{u\in\mathcal{R} : |u|\leq\alpha\}$$

for a suitable positive constant α. Furthermore, if

$$L_g = \sup_{u\in S} |g'(u)|$$

we assume that

$$\exp(L_g T)\|u^0\|_\infty \leq \alpha, \tag{2.2}$$

where $\|\cdot\|_\infty$ denotes the L^∞-norm. The inequality (2.2) will imply that the entropy solution u stays in the state space, i.e. $u(x,t)\in\mathcal{S}$ for $t\in[0,T]$. Also, since $g(0)=0$,

$$|g(u)|\leq L_g|u|, \quad \forall u\in\mathcal{S}. \tag{2.3}$$

The assumptions above also imply that there is a constant M, depending on $TV(u^0)$, such that

$$TV(u(\cdot,t))\leq M, \quad \forall t\in[0,T], \tag{2.4}$$

and

$$\|u(\cdot,t)-u(\cdot,\tau)\|_1 \leq M|t-\tau|, \quad \forall t,\tau\in[0,T]. \tag{2.5}$$

In fact, the TV–bound (2.4) and the L^1–continuity (2.5) follow from Lemma 3.1 below.

Next, we describe precisely the schemes which shall be analyzed. Choosing positive grid sizes Δt and Δx in t and x direction, we define cell boundaries $x_{j+1/2} := (j+1/2)\Delta x$

and corresponding computational cells $I_j := [x_{j-1/2}, x_{j+1/2})$. The initial data for the scheme is computed as cell averages of u^0

$$u_j^0 = \frac{1}{\Delta x}\int_{I_j} u^0(x)\, dx. \tag{2.6}$$

The discrete solution $\{u_j^n\}$ generated by either of the schemes (1.3) or (1.4) is considered as approximation of the average of the entropy solution in I_j at time level $t_n := n\Delta t$.

Due to the explicit treatment of the convective terms, it is reasonable to assume a CFL–condition

$$\mu |f'(u)| \leq 1, \quad \mu := \frac{\Delta t}{\Delta x} \tag{2.7}$$

for all relevant values of u.

Now, we are in the position to state the result in the nonstiff case:

THEOREM 2.1. *Assume that the conditions given above are satisfied. Let u be the entropy solution of (1.1) and $\{u_j^n\}$ the approximations generated by (1.3). Then there is a constant M,independent of Δt and Δx, such that*

$$\|u(\cdot, t_n) - u^n\|_1 \leq M\sqrt{\Delta t}, \quad 0 \leq t_n \leq T, \tag{2.8}$$

where u^n denotes the piecewise constant function representing u_j^n. □

For the stiff equation (1.2) we would like to have a similar result, where the estimate (2.8) holds independently of the relaxation time δ. Several other authors observed that such a result can not hold for general source terms, i.e. source terms that allow more than one equilibrium state (LeVeque & Yee 1990, Berkenbosch, Kaasschieter & ten Thije Boonkkamp 1994, Klingenstein 1994). Therefore, we need additional assumptions to deal with the stiff equation.

The assumption we propose is that the source g decreases strictly with respect to u. Therefore, zero is the only equilibrium. More precisely, the function g is assumed to satisfy

$$g(0) = 0, \qquad g'(u) \leq -\gamma < 0 \quad \forall u \in [0,1]. \tag{2.9}$$

If g satisfies this property, then the unit interval is a time invariant state space for the equation (1.2). Furthermore, it will be assumed below that the initial data is "close to equilibrium". We emphasize that the only stability condition that is assumed is the CFL–condition (2.7) with respect to the unit interval.

THEOREM 2.2. *In addition to the assumptions given in Theorem 2.1 assume that g satisfies (2.9). Furthermore, the initial data satisfies $u^0 \in BV(\mathcal{R}, [0,1])$ and there is a constant M_0 independent of $\delta > 0$, such that $\|g(u^0)\|_1 \leq M_0\delta$.*

Then the semi–implicit scheme (1.4) has a unique solution $u_j^n \in [0,1]$. Furthermore, if u is the entropy solution of (1.2) there is a constant M, depending on M_0, but independent of $\Delta t, \Delta x$ and δ, such that

$$\|u(\cdot, t_n) - u^n\|_1 \leq M\sqrt{\Delta t}, \quad 0 \leq t_n \leq T. \; \square \tag{2.10}$$

Here, we want to point out that due to the implicit treatment of the source term the error bound does not depend on δ. Furthermore, both error estimates of order $O(\sqrt{\Delta t})$ are optimal, since the same estimate is optimal for homogenous conservation laws, cf. Lucier (1985).

The outline of the arguments for both proofs is as follows: In order to compare the approximation generated by the scheme to the entropy solution, we define a comparison function $\bar{u}_\Delta$ on $\mathcal{R} \times \mathcal{R}^+$ which interpolates the data u_j^n in the sense that

$$\bar{u}_\Delta(y, t_n^+) = \lim_{\substack{\tau \to t_n \\ \tau > t_n}} \bar{u}_\Delta(y, \tau) = u_j^n \text{ for } y \in I_j. \tag{2.11}$$

This comparison function will be discontinuous at the discrete time levels t_n. Furthermore, a Kruzkov–type inequality for $\bar{u}_\Delta$ will be derived which can be seen as a weak entropy formulation of the scheme. Using this formulation and the original Kruzkov inequality (2.1), it is possible to compare $\bar{u}_\Delta$ and u.

3. The nonstiff case.

3.1. A–priori estimates. We begin the study of the explicit scheme (1.3) by collecting some properties of the discrete solutions.

LEMMA 3.1. *Given the conditions of Theorem 2.1, there is a constant M, independent of Δt and Δx, such that the following a–priori estimates hold for the solution of the explicit scheme (1.3):*

$$\begin{aligned}
&i) && u_j^n \in \mathcal{S}, \\
&ii) && \|u^n\|_1 \le M \\
&iii) && TV(u^n) \le M, \\
&iv) && \|u^n - u^m\|_1 \le M\Delta t|n-m|, \quad \forall 0 \le n, m \le N = T/\Delta t. \quad \square
\end{aligned}$$

Here, the total variation for grid functions is defined by

$$TV(u) := \sum_{j \in \mathcal{Z}} |u_j - u_{j-1}|$$

and the discrete L^1–norm is

$$\|u\|_1 := \Delta x \sum_{j \in \mathcal{Z}} |u_j|.$$

The proof of Lemma 3.1 is given in Section 5. The a–priori estimates of Lemma 3.1 carry over to the entropy solution of (1.1) by a proper application of Helly's theorem cf. Oleinik (1957) or Smoller (1983, chapter 16). This justifies our assumptions on the entropy solution.

3.2. The comparison function. Let us introduce the averaging operator P acting on $u \in L^1_{loc}$:

$$P(u)(x) := \frac{1}{\Delta x} \int_{I_j} u(z)\, dz \text{ for } x \in I_j.$$

Now, we define the comparison function $\bar{u}_\Delta$ iteratively. It will be convenient to use the arguments y and τ instead of x and t and to omit the subscript Δ.

First, we initialize

$$\bar{u}^0(y) = \bar{u}(y, 0^+) = P(u^0)(y).$$

Then, we iterate the following three steps for $n \in \mathcal{Z}_0^+$.

1) In (t_n, t_{n+1}) we solve

$$\bar{u}_\tau + f(\bar{u})_y = 0 \tag{3.1}$$

with initial data $\bar{u}(y, t_n^+)$.

2) At t_{n+1} we take cell averages

$$\bar{u}(y, t_{n+1}) = P(\bar{u}(\cdot, t_{n+1}^-))(y), \tag{3.2}$$

where $\bar{u}(\cdot, t_n^-) := \lim\limits_{\substack{\tau \to t_n \\ \tau < t_n}} \bar{u}(\cdot, \tau)$.

3) We solve an pseudo Euler step to define the data for step 1)

$$\bar{u}(y, t_{n+1}^+) = \bar{u}(y, t_{n+1}) + \Delta t g(\bar{u}(y, t_n^+)). \tag{3.3}$$

The interpolation property (2.11) follows by induction. The integral form of (3.1) on the grid block $I_j \times [t_n, t_{n+1})$ yields

$$\bar{u}(y, t_{n+1}) = u_j^n - \frac{1}{\Delta x} \int_{t_n}^{t_{n+1}} f(\bar{u}(y_{j+1/2}, \tau)) - f(\bar{u}(y_{j-1/2}, \tau))) \, d\tau.$$

In (t_n, t_{n+1}) $\bar{u}$ solves a series of homogenous Riemann problems. The CFL–condition (2.7) ensures that waves from neighboring Riemann problems do not interact and $\bar{u}(y_{j+1/2}, \tau) = u^*(u_j^n, u_{j+1}^n)$ is a constant for $\tau \in (t_n, t_{n+1})$. Therefore

$$\bar{u}(y, t_{n+1}) = u_j^n - \mu[f(u^*(u_j^n, u_{j+1}^n)) - f(u^*(u_{j-1}^n, u_j^n))]$$

and $\bar{u}(y, t_{n+1}^+) = u_j^{n+1}$, $y \in I_j$ is obvious.

3.3. The Kruzkov–type inequality. On the open time intervals (t_n, t_{n+1}), where $\bar{u}$ solves the homogenous equation (3.1), the usual Kruzkov inequality is valid:

$$\int_{t_n}^{t_{n+1}} \int_{\mathcal{R}} [|\bar{u} - k|\varphi_\tau + \sigma(\bar{u} - k)(f(\bar{u}) - f(k))\varphi_y] \, dy \, d\tau$$
$$+ \int_{\mathcal{R}} [|\bar{u}(y, t_n^+) - k|\varphi(y, t_n) - |\bar{u}(y, t_{n+1}^-) - k|\varphi(y, t_{n+1})] \, dy \geq 0$$

for all $k \in \mathcal{S}$ and any $\varphi \in \mathcal{D}_+(T)$.

We sum this inequality from $n = 0$ to $N - 1$ to obtain

$$\begin{aligned} &\int_0^T \int_{\mathcal{R}} [|\bar{u} - k|\varphi_\tau + \sigma(\bar{u} - k)(f(\bar{u}) - f(k))\varphi_y] \, dy \, d\tau \\ &+ \sum_{n=1}^{N} \int_{\mathcal{R}} [|\bar{u}(y, t_n^+) - k| - |\bar{u}(y, t_n^-) - k|]\varphi(y, t_n) \, dy \\ &+ \int_{\mathcal{R}} [|\bar{u}^0 - k|\varphi(y, 0) - |\bar{u}(y, t_N^+) - k|\varphi(y, T)] \, dy \geq 0. \end{aligned} \tag{3.4}$$

The discontinuity $\bar{u}(\cdot, t_n^-) \neq \bar{u}(\cdot, t_n^+)$ is governed by the projection (3.2) and the Euler step (3.3). Subtracting the constant k in (3.3) and multiplying by $\sigma(\bar{u}(y, t_{n+1}^+) - k)$, we find

$$|\bar{u}(y, t_n^+) - k| \leq |\bar{u}(y, t_n) - k| + \Delta t \; \sigma(\bar{u}(y, t_n^+) - k) \; g(\bar{u}(y, t_{n-1}^+)).$$

Finally, we insert the last inequality into (3.4) to obtain the Kruzkov–type inequality:

$$\begin{aligned}
&\int_0^T \int_{\mathcal{R}} [|\bar{u} - k|\varphi_\tau + \sigma(\bar{u} - k)(f(\bar{u}) - f(k))\varphi_y] \; dy \; d\tau \\
&+ \sum_{n=1}^{N} \int_{\mathcal{R}} [|\bar{u}(y, t_n) - k| - |\bar{u}(y, t_n^-) - k|]\varphi(y, t_n) \; dy \\
&+ \int_{\mathcal{R}} [|\bar{u}^0 - k|\varphi(y, 0) - |\bar{u}(y, t_N^+) - k|\varphi(y, T)] \; dy \\
&\geq -\Delta t \sum_{n=1}^{N} \int_{\mathcal{R}} [\sigma(\bar{u}(y, t_n^+) - k) g(\bar{u}(y, t_{n-1}^+))\varphi(y, t_n)] \; dy
\end{aligned} \tag{3.5}$$

for all $k \in \mathcal{S}$ and any $\varphi \in \mathcal{D}_+(T)$.

3.4. Proof of Theorem 2.1. The two inequalities (2.1) and (3.5) are combined using a mollifier function. This mollifier is defined by

$$\omega_\epsilon(x) := \frac{1}{\epsilon}\Omega(x/\epsilon),$$

where $\Omega : \mathcal{R} \to \mathcal{R}$ is smooth, nonnegative, symmetric with support in $[-1, 1]$ and $\int_{\mathcal{R}} \Omega(x) \; dx = 1$.

For arbitrary $T > 0$, we proceed by selecting the constant k and the smooth function φ. In the Kruzkov–type inequality (3.5) we select $k = u(x, t)$, $\varphi(y, \tau) := \omega_\epsilon(x - y)\omega_\epsilon(t - \tau)$ and integrate in x and t:

$$\begin{aligned}
-\int_0^T \int_{\mathcal{R}} \int_0^T \int_{\mathcal{R}} \quad & [|\bar{u} - u| \; \omega_\epsilon(x - y)\omega'_\epsilon(t - \tau) \\
& + \quad \sigma(\bar{u} - u)(f(\bar{u}) - f(u)) \; \omega'_\epsilon(x - y)\omega_\epsilon(t - \tau)] \; dy \; d\tau \; dx \; dt \\
+\int_0^T \int_{\mathcal{R}} \sum_{n=1}^{N} \int_{\mathcal{R}} & [|\bar{u}(y, t_n) - u| - |\bar{u}(y, t_n^-) - u|]\omega_\epsilon(x - y)\omega_\epsilon(t - t_n) \; dy \; dx \; dt \\
+\int_0^T \int_{\mathcal{R}} \int_{\mathcal{R}} & [|\bar{u}^0 - u|\omega_\epsilon(t) - |\bar{u}(y, t_N^+) - u|\omega_\epsilon(t - T)]\omega_\epsilon(x - y) \; dy \; dx \; dt \\
\geq -\int_0^T \int_{\mathcal{R}} \sum_{n=1}^{N} \int_{t_{n-1}}^{t_n} \int_{\mathcal{R}} & [\sigma(\bar{u}(y, t_n^+) - u) g(\bar{u}(y, t_{n-1}^+))\omega_\epsilon(x - y)\omega_\epsilon(t - t_n) \; dy \; d\tau \; dx \; dt.
\end{aligned}$$

For notational convenience, we suppress the argument when it is clear, i.e. $u = u(x, t)$ and $\bar{u} = \bar{u}(y, \tau)$.

Similarly, we insert $k = \bar{u}(y, t_n^+)$ and $\varphi(y,\tau) := \omega_\epsilon(x-y)\omega_\epsilon(t-\tau)$ into (2.1), integrate over $\mathcal{R} \times [t_{n-1}, t_n]$ with respect to y and τ and sum in n:

$$\sum_{n=1}^{N} \int_{t_{n-1}}^{t_n} \int_{\mathcal{R}} \int_0^T \int_{\mathcal{R}} [|u - \bar{u}(y,t_n^+)|\, \omega_\epsilon(x-y)\omega_\epsilon'(t-\tau)$$
$$+\sigma(u - \bar{u}(y,t_n^+))(f(u) - f(\bar{u}(y,t_n^+))\, \omega_\epsilon'(x-y)\omega_\epsilon(t-\tau)]\, dx\, dt\, dy\, d\tau$$
$$+\sum_{n=1}^{N} \int_{t_{n-1}}^{t_n} \int_{\mathcal{R}} \int_{\mathcal{R}} [|u^0 - \bar{u}(y,t_n^+)|\omega_\epsilon(\tau) - |u(x,T) - \bar{u}(y,t_n^+)|\omega_\epsilon(T-\tau)]$$
$$\times \omega_\epsilon(x-y)\, dx\, dy\, d\tau$$
$$\geq -\sum_{n=1}^{N} \int_{t_{n-1}}^{t_n} \int_{\mathcal{R}} \int_0^T \int_{\mathcal{R}} [\sigma(u - \bar{u}(y,t_n^+))g(u)\omega_\epsilon(x-y)\omega_\epsilon(t-\tau)\, dx\, dt\, dy\, d\tau.$$

By summing the last two inequalities and rearranging the terms we receive

$$L(\epsilon) \leq R(\epsilon) + F(\epsilon) + G(\epsilon) + H(\epsilon), \tag{3.6}$$

where:

$$L(\epsilon) \;:=\; \sum_{n=1}^{N} \int_{t_{n-1}}^{t_n} \int_{\mathcal{R}} \int_{\mathcal{R}} [|u(x,T) - \bar{u}(y,t_n^+)|\omega_\epsilon(x-y)\omega_\epsilon(T-\tau)]\, dy\, dx\, d\tau$$
$$+\; \int_0^T \int_{\mathcal{R}} \int_{\mathcal{R}} [|\bar{u}(y,t_N^+) - u|\, \omega_\epsilon(x-y)\omega_\epsilon(t-T)]\, dy\, dx\, dt,$$

$$R(\epsilon) := \sum_{n=1}^{N} \int_{t_{n-1}}^{t_n} \int_{\mathcal{R}} \int_{\mathcal{R}} [|u^0 - \bar{u}(y,t_n^+)|\, \omega_\epsilon(x-y)\omega_\epsilon(\tau)]\, dy\, dx\, d\tau$$
$$+\; \int_0^T \int_{\mathcal{R}} \int_{\mathcal{R}} [|\bar{u}^0 - u|\, \omega_\epsilon(x-y)\omega_\epsilon(t)]\, dy\, dx\, dt,$$

$$F(\epsilon) := \int_0^T \int_{\mathcal{R}} \sum_{n=1}^{N} \int_{t_{n-1}}^{t_n} \int_{\mathcal{R}} [|u - \bar{u}(y,t_n^+)| - |\bar{u} - u|]\, \omega_\epsilon(x-y)\omega_\epsilon'(t-\tau)$$
$$+\; [\sigma(u - \bar{u}(y,t_n^+))(f(u) - f(\bar{u}(y,t_n^+))) - \sigma(\bar{u} - u)(f(\bar{u}) - f(u))]$$
$$\times \omega_\epsilon'(x-y)\omega_\epsilon(t-\tau)\, dy\, d\tau\, dx\, dt,$$

$$G(\epsilon) := \int_0^T \int_{\mathcal{R}} \sum_{n=1}^{N} \int_{t_{n-1}}^{t_n} \int_{\mathcal{R}} \sigma(u - \bar{u}(y,t_n^+))[g(u)\, \omega_\epsilon(t-\tau) - g(\bar{u}(y,t_{n-1}^+))\, \omega_\epsilon(t-t_n)]$$
$$\times \omega_\epsilon(x-y)\, dy\, d\tau\, dx\, dt$$

and finally,

$$H(\epsilon) := \int_0^T \int_{\mathcal{R}} \sum_{n=1}^{N} \int_{\mathcal{R}} [|\bar{u}(y,t_n) - u| - |\bar{u}(y,t_n^-) - u|]\, \omega_\epsilon(x-y)\omega_\epsilon(t-t_n)\, dy\, dx\, dt.$$

These expressions are bounded in the following way:

LEMMA 3.2. *For any $T > 0$ there is a constant M independent of step sizes and ϵ such that:*

$$i)\quad |L(\epsilon) - \|u(\cdot,T) - \bar{u}(\cdot,t_N^+)\|_1| \le M\epsilon,$$

$$ii)\quad |R(\epsilon) - \|u^0 - \bar{u}(\cdot,0)\|_1| \le M(\epsilon + \Delta t),$$

$$iii)\quad |F(\epsilon)| \le M\frac{\Delta t}{\epsilon},$$

$$iv)\quad |G(\epsilon)| \le M\left(\epsilon + \Delta t + \frac{\Delta t}{\epsilon} + \Delta t \sum_{n=0}^{N-1} \|u(\cdot,t_n) - \bar{u}(\cdot,t_n^+)\|_1\right),$$

$$v)\quad |H(\epsilon)| \le M\frac{\Delta t}{\epsilon}. \quad \square$$

Except for *iii)* and *iv)*, all these estimates carry over from the corresponding estimates for the system cf. Schroll et al. (1994). We first apply these estimates to complete the proof of Theorem 2.1. From (3.6) we get the error recursion

$$e^N \le e^0 + M(\epsilon + \Delta t + \frac{\Delta t}{\epsilon}) + \Delta t M \sum_{n=0}^{N-1} e^n,$$

where $e^n := \|u(\cdot,t_n) - \bar{u}(\cdot,t_n^+)\|_1$. Here, we choose $\epsilon = \sqrt{\Delta t}$ and use a well known property of the projection operator

$$e^0 = \|u^0 - \bar{u}(\cdot,0)\|_1 = \|u^0 - P(u^0)\|_1 \le M\Delta t$$

to find

$$e^N \le \Delta t M \sum_{n=0}^{N-1} e^n + M\sqrt{\Delta t}.$$

Finally, the discrete Gronwall lemma completes the proof of Theorem 2.1. $\square$

3.5. Proof of Lemma 3.2. It remains to prove the bounds *iii)* and *iv)* in Lemma 3.2. Here, the a–priori estimates of Lemma 3.1 and the properties of the entropy solution are used. In order to prove *iii)* note that $\sigma(u-k)(f(u)-f(k))$ as a function of k is Lipschitz continous with a Lipschitz constant independent of u:

$$|\sigma(u-k)(f(u)-f(k)) - \sigma(u-q)(f(u)-f(q))| \le M|k-q|, \quad u,k,q \in \mathcal{S}.$$

Using this fact the estimate of F easily carries over from the corresponding estimate in Schroll et al. (1994).

Concerning the function G, we have:

$$\begin{aligned} |G(\epsilon)| \le \quad & \int_0^T \int_{\mathcal{R}} \sum_{n=1}^{N} \int_{t_{n-1}}^{t_n} \int_{\mathcal{R}} |g(u) - g(\bar{u}(y,t_{n-1}^+))|\, \omega_\epsilon(x-y)\omega_\epsilon(t-\tau)\; dy\; d\tau\; dx\; dt \\ + \quad & \int_0^T \sum_{n=1}^{N} \int_{t_{n-1}}^{t_n} \int_{\mathcal{R}} |g(\bar{u}(y,t_{n-1}^+))|\, |\omega_\epsilon(t-t_n) - \omega_\epsilon(t-\tau)|\; dy\; d\tau\; dt \\ =: \quad & G_1(\epsilon) + G_2(\epsilon). \end{aligned}$$

By using the TV-bound (2.4) and the L^1-continuity (2.5) of the entropy solution u we have

$$
\begin{aligned}
|G_1(\epsilon)| \leq\ & L_g\int_0^T\int_{\mathcal{R}}\sum_{n=1}^N\int_{t_{n-1}}^{t_n}\int_{\mathcal{R}}|u-\bar{u}(y,t_{n-1}^+)|\ \omega_\epsilon(x-y)\omega_\epsilon(t-\tau)\ dy\ d\tau\ dx\ dt \\
\leq\ & L_g\int_0^T\int_{\mathcal{R}}\int_{\mathcal{R}}|u(x,t)-u(y,t)|\ \omega_\epsilon(x-y)\ dy\ dx\ dt \\
& +L_g\int_0^T\sum_{n=1}^N\int_{t_{n-1}}^{t_n}\|u(\cdot,t)-u(\cdot,t_{n-1})\|_1\ \omega_\epsilon(t-\tau)\ d\tau\ dt \\
& +L_g\sum_{n=1}^N\int_{t_{n-1}}^{t_n}\|u(\cdot,t_{n-1})-u(\cdot,t_{n-1}^+)\|_1\ d\tau \\
\leq\ & M\ TV(u(\cdot,t))\ \epsilon+M\int_0^T\sum_{n=1}^N\int_{t_{n-1}}^{t_n}|t-t_{n-1}|\ \omega_\epsilon(t-\tau)\ d\tau\ dt+M\Delta t\sum_{n=1}^N e^n \\
\leq\ & M(\epsilon+\Delta t)+M\int_0^T\int_0^T|t-\tau|\ \omega_\epsilon(t-\tau)\ d\tau\ dt+M\Delta t\sum_{n=1}^N e^n \\
\leq\ & M(\epsilon+\Delta t)+M\Delta t\sum_{n=1}^N e^n.
\end{aligned}
$$

Furthermore, G_2 is bounded by

$$|G_2(\epsilon)|\leq L_g\int_0^T\sum_{n=1}^N\int_{t_{n-1}}^{t_n}\|\bar{u}(y,t_{n-1}^+)\|_1\ |\omega_\epsilon(t-t_n)-\omega_\epsilon(t-\tau)|\ d\tau\ dt.$$

Here, the a–priori estimate $ii)$ of Lemma 3.1 implies

$$|G_2(\epsilon)|\leq M\int_0^T\sum_{n=1}^N\int_{t_{n-1}}^{t_n}\int_{t_{n-1}}^{t_n}|\omega_\epsilon'(s-t)|\ ds\ d\tau\ dt\leq M\frac{\Delta t}{\epsilon}. \quad \Box$$

4. The stiff case. An interesting point in the arguments developed in the previous section is that they also apply to the stiff equation (1.2).

In the case of the semi–implicit scheme (1.4), we have to make sure that there is a unique solution in the state space $\mathcal{S}_\delta := [0,1]$. Furthermore, the generic constant M may not depend on δ. We have the following a–priori estimates:

LEMMA 4.1. *Given the conditions of Theorem 2.2, the semi–implicit scheme (1.4) has a unique solution satisfying the following estimates:*

$$
\begin{aligned}
&i) && u_j^n\in\mathcal{S}_\delta, \\
&ii) && TV(u^n)\leq TV(u^0)\leq M, \\
&iii) && \|g(u^n)\|_1\leq M\delta \\
&iv) && \|u^n-u^m\|_1\leq M\Delta t|n-m|, \quad \forall\ 0\leq n,m\leq N=T/\Delta t.
\end{aligned}
$$

Here, M is independent of step sizes and δ. $\Box$

Again, we delay the proof to Section 5 and point out that these a–priori estimates carry over to the entropy solution by Helly's theorem cf. Tveito & Winther (1994).

In order to prove Theorem 2.2, we define the comparison function as in Section 4, except the pseudo Euler step (3.3) must be replaced by

$$\bar u(y, t_{n+1}^+) = \bar u(y, t_{n+1}) + \frac{\Delta t}{\delta} g(\bar u(y, t_{n+1}^+)).$$

The corresponding Kruzkov–type inequality reads

$$\begin{aligned}
&\int_0^T \int_{\mathcal{R}} [|\bar u - k|\varphi_\tau + \sigma(\bar u - u)(f(\bar u) - f(k))\varphi_y]\; dy\; d\tau \\
&+ \sum_{n=1}^{N} \int_{\mathcal{R}} [|\bar u(y, t_n) - k| - |\bar u(y, t_n^-) - k|]\varphi(y, t_n)\; dy \\
&+ \int_{\mathcal{R}} [|\bar u^0 - k|\varphi(y, 0) - |\bar u(y, t_N^+) - k|\varphi(y, T)]\; dy \\
&\geq -\frac{\Delta t}{\delta} \sum_{n=1}^{N} \int_{\mathcal{R}} [\sigma(\bar u(y, t_n^+) - k) g(\bar u(y, t_n^+))\varphi(y, t_n)]\; dy
\end{aligned}$$

for all $k \in \mathcal{S}$ and any $\varphi \in \mathcal{D}_+(T)$.

The same arguments as in the proof of Theorem 2.1 lead to inequality (3.6) with the expressions as before except for G which now reads:

$$\begin{aligned}
G(\epsilon) :=\;& \frac{1}{\delta}\int_0^T \int_{\mathcal{R}} \sum_{n=1}^{N} \int_{t_{n-1}}^{t_n} \int_{\mathcal{R}} \sigma(u - \bar u(y, t_n^+))[g(u)\;\omega_\epsilon(t-\tau) - g(\bar u(y, t_n^+))\;\omega_\epsilon(t - t_n)] \\
&\qquad \times \omega_\epsilon(x - y)\; dy\; d\tau\; dx\; dt \\
=\;& \frac{1}{\delta}\int_0^T \int_{\mathcal{R}} \sum_{n=1}^{N} \int_{t_{n-1}}^{t_n} \int_{\mathcal{R}} \sigma(u - \bar u(y, t_n^+))[g(u) - g(\bar u(y, t_n^+))]\omega_\epsilon(x-y)\omega_\epsilon(t-\tau) \\
&\qquad dy\; d\tau\; dx\; dt \\
+\;& \frac{1}{\delta}\int_0^T \int_{\mathcal{R}} \sum_{n=1}^{N} \int_{t_{n-1}}^{t_n} \int_{\mathcal{R}} \sigma(u - \bar u(y, t_n^+))\; g(\bar u(y, t_n^+))\; [\omega_\epsilon(t-\tau) - \omega_\epsilon(t-t_n)] \\
&\qquad \times \omega_\epsilon(x - y)\; dy\; d\tau\; dx\; dt.
\end{aligned}$$

The first term in that sum is negative because of $g' < 0$ and therefore

$$G(\epsilon) \leq \frac{1}{\delta}\int_0^T \sum_{n=1}^{N} \int_{t_{n-1}}^{t_n} \|g(\bar u(\cdot, t_n^+))\|_1\; |\omega_\epsilon(t-\tau) - \omega_\epsilon(t-t_n)|\; d\tau\; dt.$$

Here, the stiffness parameter δ cancels by the a–priori estimate $iii)$ of Lemma 4.1

$$G(\epsilon) \leq M\frac{\Delta t}{\epsilon}.$$

Therefore, we deduce the error estimate

$$e^N \leq e^0 + M\left(\epsilon + \Delta t + \frac{\Delta t}{\epsilon}\right)$$

from (3.6). Choosing $\epsilon = \sqrt{\Delta t}$ concludes the proof of Theorem 2.2. $\square$

5. Appendix.

5.1. Proof of Lemma 3.1. The explicit scheme (1.3) reads

$$u_j^{n+1} = \tilde{u}_j^{n+1} + \Delta t g(u_j^n)$$

where $\tilde{u}_j^{n+1}$ is computed by one step of the original Godunov scheme for the homogenous conservation law with data u^n. Since Godunov's method does not introduce new extrema, we have

$$|\tilde{u}_j^{n+1}| \leq \max(|u_{j-1}^n|, |u_j^n|, |u_{j+1}^n|)$$

and by (2.3) it follows

$$|u_j^{n+1}| \leq \max(|u_{j-1}^n|, |u_j^n|, |u_{j+1}^n|) + \Delta t L_g |u_j^n|.$$

Therefore,

$$\|u^{n+1}\|_\infty \leq (1 + L_g \Delta t)\|u^n\|_\infty \leq \exp(L_g T)\|u^0\|_\infty,$$

confirming i).

Furthermore,

$$\|u^{n+1}\|_1 \leq \|\tilde{u}^{n+1}\|_1 + \Delta t L_g \|u^n\|_1.$$

Since Godunov's method is L^1–contracting, we derive that

$$\|u^{n+1}\|_1 \leq (1 + \Delta t L_g)\|u^n\|_1.$$

This implies ii).

Concerning the TV–bound iii), we observe that

$$|u_j^{n+1} - u_{j-1}^{n+1}| \leq |\tilde{u}_j^{n+1} - \tilde{u}_{j-1}^{n+1}| + \Delta t L_g |u_j^n - u_{j-1}^n|.$$

Again, since Godunov's method is total variation diminishing it follows

$$TV(u^{n+1}) \leq (1 + \Delta t L_g) TV(u^n) \leq \exp(L_g T) TV(u^0).$$

Finally, the time–Lipschitz continuity is a consequence of

$$\begin{aligned} |u_j^{n+1} - u_j^n| &\leq |\tilde{u}_j^{n+1} - u_j^n| + \Delta t L_g |u_j^n| \\ &\leq \mu |f'(\xi_j^n)| |u^*(u_j^n, u_{j+1}^n) - u^*(u_{j-1}^n, u_j^n)| + \Delta t L_g |u_j^n| \end{aligned}$$

where $f'(\xi_j^n) = (f(u^*(u_j^n, u_{j+1}^n)) - f(u^*(u_{j-1}^n, u_j^n)))/(u^*(u_j^n, u_{j+1}^n) - u^*(u_{j-1}^n, u_j^n))$. The CFL–condition implies

$$\|u^{n+1} - u^n\|_1 \leq \Delta x TV(u^n) + \Delta t L_g \|u^n\|_1 = \mathcal{O}(\Delta t).$$

Therefore iv) follows by the triangle inequality. $\square$

5.2. Proof of Lemma 4.1. The solution of scheme (1.4) is implicitly defined by

$$u_j^{n+1} - \frac{\Delta t}{\delta} g(u_j^{n+1}) = \tilde{u}_j^{n+1}, \tag{5.1}$$

where again $\tilde{u}_j^{n+1}$ is computed by one step of the original Godunov method applied to the homogenous equation. The left hand side as a function of u is strictly increasing, therefore, (5.1) has at most one solution. The right hand side is in $\mathcal{S}_\delta = [0, 1]$ for data in $\mathcal{S}_\delta$. Furthermore, since

$$\left(I - \frac{\Delta t}{\delta} g\right)(0) = 0 \quad \text{and} \quad \left(I - \frac{\Delta t}{\delta} g\right)(1) = 1 - \frac{\Delta t}{\delta} g(1) > 1$$

there is exactly one solution $u_j^{n+1} \in \mathcal{S}_\delta$.

Concerning the TV–bound, we have for the implicit scheme

$$(1 - \frac{\Delta t}{\delta} g'(\zeta_{j-1/2}^{n+1}))(u_j^{n+1} - u_{j-1}^{n+1}) = \tilde{u}_j^{n+1} - \tilde{u}_{j-1}^{n+1},$$

where $g'(\zeta_{j-1/2}^n) = (g(u_j^n) - g(u_{j-1}^n))/(u_j^n - u_{j-1}^n)$. Since $g' < 0$, it follows

$$|u_j^{n+1} - u_{j-1}^{n+1}| \le |\tilde{u}_j^{n+1} - \tilde{u}_{j-1}^{n+1}|$$

and since Godunov's scheme is TVD, the total variation is decreasing.

In order to prove iii), we observe

$$g(u_j^{n+1}) = g(u_j^n) - \mu g'(\eta_j^{n+1/2}) f'(\xi_j^n)(u^*(u_j^n, u_{j+1}^n) - u^*(u_{j-1}^n, u_j^n)) + \frac{\Delta t}{\delta} g'(\eta_j^{n+1/2}) g(u_j^{n+1}),$$

where $g'(\eta_j^{n+1/2}) = (g(u_j^{n+1}) - g(u_j^n))/(u_j^{n+1} - u_j^n)$. We multiply by $\sigma(g(u_j^{n+1}))$ and make use of (2.9) to find

$$(1 + \gamma \frac{\Delta t}{\delta})|g(u_j^{n+1})| \le |g(u_j^n)| + M|u^*(u_j^n, u_{j+1}^n) - u^*(u_{j-1}^n, u_j^n)|.$$

Consequently, by the TV-bound

$$(1 + \gamma \frac{\Delta t}{\delta})\|g(u^{n+1})\|_1 \le \|g(u^n)\|_1 + M\Delta t.$$

Hence, if $\tilde{M} = \max(M_0, M/\gamma)$ we obtain

$$\|g(u^n)\|_1 \le \tilde{M}\delta \quad 0 \le n \le N = T/\Delta t.$$

Finally, iv) is an easy consequence of ii) and iii) because of

$$\|u^{n+1} - u^n\|_1 \le M\Delta t TV(u^n) + \frac{\Delta t}{\delta}\|g(u^{n+1})\|_1 = \mathcal{O}(\Delta t). \quad \square$$

REFERENCES

Berkenbosch, A. C., Kaasschieter, E. F. & ten Thije Boonkkamp, J. H. M. (1994), 'The numerical wave speed for one–dimensional scalar hyperbolic conservation laws with source terms', Preprint, Dept. of Math. and Comp. Sci., Eindhoven Univ. of Technology.

Chalabi, A. (1992), 'Stable upwind schemes for hyperbolic conservation laws with source terms', *IMA J. Numer. Anal.* **12**, 217–241.

Klingenstein, P. (1994), 'Hyperbolic conservation laws with source terms: Errors of the shock location', Preprint, Seminar für Ang. Math., ETH Zürich, Switzerland.

Kruzkov, S. N. (1970), 'First order quasi linear equations with several space variables', *Math. USSR. Sb.* **10**, 217–243.

Kuznetsov, N. N. (1976), 'Accuracy of some approximate methods for computing the weak solutions of a first order quasi linear equation', *USSR. Comp. Math. and Math. Phys.* **16**, 105–119.

LeVeque, R. J. & Yee, H. C. (1990), 'A study of numerical methods for hyperbolic conservation laws with stiff source terms', *J. Comput. Phys.* **86**, 187–210.

Lucier, B. J. (1985), 'Error bounds for the methods of Glimm, Godunov and LeVeque', *SIAM J. Numer. Anal.* **22**, 1074–1081.

Oleinik, O. A. (1957), 'Discontinuous solutions of non-linear differential equations', *Amer. Math. Soc. Trans.* **26**(2), 95–172.

Pember, R. B. (1992), 'Numerical methods for hyperbolic conservation laws with stiff relaxation I.', Preprint, Center for Pure and Appl. Math., Univ. of California, Berkely, USA.

Schroll, H. J., Tveito, A. & Winther, R. (1994), 'An error bound for a finite difference scheme applied to a stiff system of conservation laws', Preprint, Dep. of Informatics, Univ. of Oslo, Norway, to appear in SIAM J. Numer. Anal.

Smoller, J. (1983), *Shock Waves and Reaction-Diffusion Equations*, Springer-Verlag, New York.

Tveito, A. & Winther, R. (1994), 'On the rate of convergence to equilibrium for a system of conservation laws including a relaxation term', Preprint, Dep. of Informatics, Univ. of Oslo, Norway, to appear in SIAM J. Math. Anal.