Discrete cell entropy inequalities for scalar conservation laws

Thorsten Grahs, Thomas Sonar

Institute for Analysis
TU Braunschweig
Pockelsstr. 14
D-38106 Braunschweig
Germany

Received: date / Revised version: date

Summary We analyze semi-discrete three-point finite difference schemes for the numerical solution of scalar conservation laws and in particular their behaviour on discrete entropy inequalities. Special emphasize is given to the choice of the discretization of the numerical entropy flux. We show that our simple choice is consistent with a classical approach due to Crandall and Majda. Starting from this discretization of the entropy inequality we derive cell entropy inequalities which enables us to determine the dissipation model of the corresponding numerical method. In the following we show that this is consistent with the dissipation coefficient for the Roe scheme for scalar equations. We show that the necessary entropy fix for the Roe scheme proposed by Harten and Hyman is also necessary by the derivation from cell entropy inequalities.

1 Introduction

We are lookingfor approximate solutions of the scalar conservation law in one space dimension, i.e.

$$\partial_t u + \partial_x f(u) = 0 \tag{1}$$

where $u : \mathbb{R} \times \mathbb{R}^+ \to \mathbb{R}$ and the flux f is assumed to be differentiable.

Independently of the smoothness of the initial conditions discontinuities develop in general within finite time so that weak solutions

u defined by

$$\int_\omega (u\partial_t \Phi + f(u)\partial_x \Phi)\ dx\, dt = 0$$

have to be considered. Here, $\omega \subset \mathbb{R}^+ \times \mathbb{R}^2$, $\Phi \in C_0^1(\omega)$ and $u \in L^1 \cap L^\infty(\omega)$. The additional entropy inequality for the conservation law is

$$\partial_t \mathsf{u}(u) + \partial_x \mathsf{f}(u) \leq 0 \tag{2}$$

with the entropy $\mathsf{u}(u)$ and the corresponding entropy flux $\mathsf{f}(u)$ satisfying the compatability condition

$$\mathsf{f}'(u) = \mathsf{u}'(u) f'(u). \tag{3}$$

The conservation law (1) is approximated by a 3-point scheme in conservative form

$$u^i = \mathcal{H}(u_{i-1}, u_i, u_{i+1}))\ =\ u_i - \lambda\left[F(u_{i+1}, u_i) - F(u_i, u_{i-1})\right] \tag{4}$$

with the numerical flux function

$$F(u_{i+1}, u_i)\ =\ F_{i+1/2} = \tfrac{1}{2}[f(u_{i+1}) + f(u_i)] - Q(u_{i+1}, u_i)(u_{i+1} - u_i).$$

The numerical dissipation coefficient $Q(u_{i+1}, u_i) =: Q_{i+1/2}$ is yet to be determined. Tadmor [21] showed that every 3-point scheme of the form (4) can be cast in this form with an unique $Q_{i+1/2}$. Classical choices for Q are

$$\begin{aligned} Q^{LF}_{i+1/2} &= \tfrac{1}{2\lambda} && \text{Lax-Friedrichs scheme [9]},\\ Q^{Roe}_{i+1/2} &= \tfrac{1}{2}\left|a_{i+1/2}\right| && \text{Roe scheme [18]},\\ Q^{LW}_{i+1/2} &= \tfrac{1}{2}\lambda a^2_{i+1/2} && \text{Lax-Wendroff scheme [11].} \end{aligned}$$

with

$$a_{i+1/2} = \begin{cases} \frac{f_{i+1} - f_i}{u_{i+1} - u_i} & \text{for } u_{i+1} \neq u_i \\ f'_{i+1/2} & \text{else} \end{cases}.$$

Remark 1 The Roe scheme is the generalization to nonlinear systems for a scheme proposed by Courant-Isaacson-Rees [1] for quasi-linear Systems. The conservative formulation is due to Murman [14].

2 Entropy inequality and three-point schemes

Physically relevant solutionssolution of the conservation law (1) fulfil in addition the entropy inequality (2). This property should also be mirrored on the discrete level, i.e the discrete solution obey a discrete entropy inequality. Following Lax [10] we give the

Definition 1 *A numerical scheme of the form (4) is consistent with the entropy inequality (2) if there exists a continuous function*

$$\mathsf{F} : \mathbb{R}^2 \to \mathbb{R}$$

which

i) is consistent with the entropy flux f

$$\mathsf{F}(u, \dots, u) = \mathsf{f}(u),$$

ii) and for which the numerical solution satisfies a discrete entropy inequality

$$\begin{aligned} \mathsf{u}(u^i) - \mathsf{u}(u_i) &= \mathsf{u}^i - \mathsf{u}_i \\ &\leq \lambda \left[\mathsf{F}(u_{i-k+1}, \dots, u_{i+k}) - \mathsf{F}(u_{i-k}, \dots, u_{i+k-1})\right] \end{aligned}$$

with numerical entropy flux F.

The fundamental relevance of schemes obeying a discrete entropy inequality is described by Harten, Hyman and Lax [7], stating that a weak solution of a conservative scheme consistent with any entropy condition is the unique entropy solution of (1).

For monotone schemes they derive the following

Theorem 1 *A consistent monotone scheme is consistent with any entropy condition.*

The proof was given in [7] while a different approach was presented by Crandall and Majda [2].

But like for monotonicity (or TVD-property) and second order accuracy we have a similar relation between entropy and second order accuracy which results in the following result by Schonbeck [19]:

Theorem 2 *Assume a three-point difference scheme of the form (4) with C^2 numerical flux which is consistent with an entropy condition. Then the scheme is at most first order accurate.*

2.1 E-schemes

Osher [15] and Osher and Chakravarty [16] developed a theory concerning schemes automatically obeying an entropy condition. Osher and Tadmor [17] summerized and enlarged this theory concerning entropy satisfying schemes, called E-schemes. We start with the definition of E-schemes, given by Osher [15].

Definition 2 *A consistent conservative scheme is called an* E-scheme *if its numerical flux satisfies*

$$\mathsf{sgn}(u_{i+1}-u_i)[F_{i+1/2}-f(u)] \le 0, \quad \forall u \in [u_i, u_{i+1}].$$

According to Theorem 1 concerning monotone schemes, the following proposition is quite clear:

Proposition 1 *A 3-point monotone scheme is an E-scheme.*

This follows since $F(u,v)$ is non-decreasing in u and non-increasing in v, one recognizes

$$\begin{array}{lllll} F(u,v) & \le & F(u,w) & \le & F(w,w) \quad \text{if } u \le v \le w, \\ F(u,v) & \ge & F(w,v) & \ge & F(w,w) \quad \text{if } u \ge v \ge w, \end{array}$$

to obtain immediately

$$\mathsf{sgn}(v-u)[F(u,v)-F(w,w)] \le 0, \quad \forall w \in [u,v].$$

E-schemes can also be characterized by their dissipation coefficients. Tadmor [22] showed that they are consistent with any entropy condition if their dissipation coefficient is bounded in a crtain way:

Lemma 1 *Assume that the CFL-like condition*

$$\lambda \max_i \left| \frac{f_{i+1}-f_i}{u_{i+1}-u_i} \right| \le 1$$

holds. The E-fluxes are characterized by

$$\begin{array}{ll} F_{i+1/2} \le F^G_{i+1/2} & \mathit{if} u_i < u_{i+1}, \\ F_{i+1/2} \ge F^G_{i+1/2} & \mathit{if} u_i > u_{i+1}, \end{array}$$

where F^G is the Godunov-flux [3].

Proof For the Godunov-flux we are able to write under CFL ≤ 1

$$\begin{array}{ll} \min\limits_{u\in[u_{i-1},u_i]} f(u) & \text{if}\, u_i < u_{i+1}, \\ \min\limits_{u\in[u_i,u_{i-1}]} f(u) & \text{if}\, u_i > u_{i+1}, \end{array}$$

This proves that the numerical flux of Godunov's method is the limit for the flux of an E-scheme.

Tadmor [22] defined the modified Lax-Friedrichs scheme which posses only half of the dissipation as the original Lax-Friedrichs scheme, i.e.

$$Q^{mLF}_{i+1/2} = \tfrac{1}{4\lambda}.$$

With this definition he is able to proof

Theorem 3 *Under the CFL condition*

$$\lambda \max|a(u)| \leq 1/2,$$

an E-scheme with numerical dissipation coefficient $Q^E_{i+1/2}$*, satisfying*

$$Q^G_{i+1/2} \leq Q^E_{i+1/2} \leq Q^{mLF}_{i+1/2},$$

is consistent with any entropy condition.

But unfortunately like before Tadmor [21] showed the limitation for E-schemes:

Proposition 2 *An E-scheme with differentiable numerical flux is at most first order accurate.*

Remark 2 Tadmor [23] developed even a theory of second-order schemes obbying an entropy inequality. But unfortunately these schemes are not no means entropy stable [20].

3 Discrete entropy inequalities

As already mentioned, there are some degrees of freedom in choosing numerical entropy flux because the only requirement is the compatibility condition (3). So we are faced with the problem that every numerical entropy flux of the form

$$\mathsf{F}(u,v) := \tfrac{1}{2}[\mathsf{f}(u)+\mathsf{f}(v)] - \mathsf{Q}(u,v), \quad \mathsf{Q}(u,u)=0\ \forall u\in\Omega$$

is a consistent choice.

3.1 Numerical entropy flux

Lax [10] was the first who proved entropy stability for the Lax-Friedrichs scheme showing that the scheme applied to a system of conservation laws satisfies a discrete entropy inequality with numerical entropy flux chosen as

$$\mathsf{F}(u_{i+1}, u_i) = \tfrac{1}{2}[\mathsf{f}(u_{i+1}) + \mathsf{f}(u_i)] - \tfrac{1}{2\lambda}[\mathsf{u}(u_{i+1}) - \mathsf{u}(u_i)].$$

Exploiting this assumption Sonar [20] proposed the following

Definition 3 *A numerical entropy flux is called* Lax-consistent *if it can be written in the form*

$$\begin{aligned}\mathsf{F}_{i+1/2} &= \mathsf{F}(u_{i+1}, U_i)\\ &= \tfrac{1}{2}[\mathsf{f}(u_{i+1}) + \mathsf{f}(u_i)] - \mathsf{Q}(u_{i+1}, u_i)[\mathsf{u}(u_{i+1}) - \mathsf{u}(u_i)].\end{aligned}$$

with numerical entropy dissipation coefficient Q.

Remark 3 Sonar proposed a numerical entropy flux of the form

$$\begin{aligned}\mathsf{F}_{i+1/2} &= \mathsf{F}(u_{i+1}, u_i)\\ &= \tfrac{1}{2}[\mathsf{f}(u_{i+1}) + \mathsf{f}(u_i)] - \mathsf{Q}(u_{i+1}, u_i)[\mathsf{u}(u_{i+1}) - \mathsf{u}(u_i)].\end{aligned}$$

where Q stems from Q by replacing the u, f and the corresponding derivates by u, f. Here he has to show c-consistency (compatability-consistency) for the scheme. This is

$$\begin{aligned}\partial_u \mathsf{F}(u, v)|_{u=v} &= \mathsf{f}'(u) - \mathsf{u}'(u)\tilde{Q}(u, u)\\ &= \mathsf{u}'(u) f'(u) - \mathsf{u}'(u)\tilde{Q}(u, u)\\ &= \mathsf{u}'(u)\partial_u f(u, v)|_{u=v}\end{aligned}$$

which is nothing more than $\tilde{Q}(u, u) = Q(u, u)$.

We will give here a much simpler form of the numerical entropy flux, being still Lax-c-consistent.

Definition 4 *The numerical entropy flux*

$$\mathsf{F}(u, v) := \tfrac{1}{2}[\mathsf{f}(u) + \mathsf{f}(v)] - Q(u, v)[\mathsf{u}(u) - \mathsf{u}(v)] \tag{5}$$

is called the corresponding numerical entropy flux *to a numerical scheme of the form*

$$F(u, v) := \tfrac{1}{2}[f(u) + f(c)] - Q(u, v)[u - v].$$

Lemma 2 *The numerical entropy flux of the form (5) is Lax-c-consistent.*

Proof Lax-consistency follows directly from the definition of the numerical enropy flux (5). The c-consistency is trivial since the entropy flux is chosen such that the dissipation coefficients are equal for the numerical flux and the numerical entropy flux.

In the remainder of the section we like to show that the choice of the numerical entropy flux seems to be reasonable, i.e. is consistent with the choice given by Crandall and Majda [2]. For a three-point scheme this is

$$\begin{aligned}\mathsf{F}^{CM}_{i+1/2} = & F(\max(u_{i+1},k),\max(u_i,k)) \\ & -F(\min(u_{i+1},k),\min(u_i,k)), \quad k\in\mathbb{R}.\end{aligned} \tag{6}$$

Theorem 4 *The numerical entropy flux of the form*

$$\mathsf{F}_{i+1/2} = \frac{1}{2}(\mathsf{f}_{i+1}+\mathsf{f}_i) - Q_{i+1/2}(\mathsf{u}_{i+1}-\mathsf{u}_i) \tag{7}$$

with a dissipation coefficient Q is consistent with the Crandall-Majda entropy flux (6)

Proof Krutzkov [8] proved that it is enough to consider the family of entropy pairs $\mathsf{u}(u) = |u-k|$, $\mathsf{f}(u) = \mathrm{sgn}(u-k)(f(u)-f(k))$, $k\in\mathbb{R}$ to show that a numerical scheme obeys an entropy condition for all entropy pairs. Thus rewrite the entropy flux with $\max(u_{i+1/i},k) =$

$u_{i+1/i}\top k, \min(u_{i+1/i}, k) = u_{i+1/i}\bot k$ as

$$
\begin{aligned}
\mathsf{F}_{i+1/2} &= F(u_{i+1}\top k, u_i\top k) - F(u_{i+1}\bot k, u_i\bot k)\\
&= \frac{1}{2}[f(u_{i+1}\top k) + f(u_i\top k)]\\
&\quad -Q(u_{i+1}\top k, u_i\top k)(u_{i+1}\top k - u_i\top k)\\
&\quad -\frac{1}{2}[f(u_{i+1}\bot k) + f(u_i\bot k)]\\
&\quad +Q(u_{i+1}\bot k, u_i\bot k)(u_{i+1}\bot k - u_i\bot k)\\
&= \frac{1}{2}[f(u_{i+1\top k}) - f(u_{i+1}\bot k)] + \frac{1}{2}[f(u_i\top k) - f(u_i\bot k)]\\
&\quad -Q(u_{i+1}\top k, u_i\top k)(u_{i+1}\top k - u_i\top k)\\
&\quad +Q(u_{i+1}\bot k, u_i\bot k)(u_{i+1}\bot k - u_i\bot k)\\
&= \frac{1}{2}\mathrm{sgn}(u_{i+1} - k)(f(u_{i+1}) - f(k)) + \frac{1}{2}\mathrm{sgn}(u_i - k)(f(u_i) - f(k))\\
&\quad -Q(u_{i+1}\top k, u_i\top k)(u_{i+1}\top k - u_i\top k)\\
&\quad +Q(u_{i+1}\bot k, u_i\bot k)(u_{i+1}\bot k - u_i\bot k)\\
&= \frac{1}{2}[\mathsf{f}(u_{i+1}) + \mathsf{f}(u_i)]\\
&\quad -Q(u_{i+1}\top k, u_i\top k)(u_{i+1}\top k - u_i\top k)\\
&\quad +Q(u_{i+1}\bot k, u_i\bot k)(u_{i+1}\bot k - u_i\bot k)
\end{aligned}
$$

Hence, we have to prove that

$$
\begin{aligned}
Q(u_{i+1}\top k, u_i\top k)(u_{i+1}\top k - u_i\top k) - Q(u_{i+1}\bot k, u_i\bot k)(u_{i+1}\bot k - u_i\bot k)\\
= Q(u_{i+1}, u_i)(u_{i+1}, u_i)
\end{aligned}
$$

holds. We analyze the following cases:

i) $k \geq u_{i+1}, u_i$:
$\Longrightarrow$ $(u_{i/i+1}\top k) = k,\ (u_{i/i+1}\bot k) = u_{i/i+1}$:

$$
\begin{aligned}
&-Q(k,k)[k-k] + Q(U_{i+1}, u_i)[u_{i+1} - u_i]\\
&= Q(u_{i+1}, u_i)[u_{i+1} - u_i + k - k]\\
&= Q(u_{i+1}, u_i)[(u_{i+1} - k) - (u_i - k)]\\
&= Q(u_{i+1}, u_i)[\mathsf{sgn}(u_{i+1} - k)|u_{i+1} - k| - \mathsf{sgn}(u_i - k)|u_i - k|]\\
&= -Q(u_{i+1}, u_i)[\mathsf{u}_{i+1} - \mathsf{u}_i]
\end{aligned}
$$

ii) $k \leq u_{i+1}, u_i$:
$\Longrightarrow$ $(u_{i/i+1}\top k) = u_{i/i+1},\ (u_{i/i+1}\bot k) = k$:

$$
\begin{aligned}
&-Q(u_{i+1},u_i)[u_{i+1}-u_i]+Q(k,k)[k-k] \\
&\;= -Q(u_{i+1},u_i)[u_{i+1}-u_i+k-k] \\
&\;= -Q(u_{i+1},u_i)[(u_{i+1}-k)-(u_i-k)] \\
&\;= -Q(u_{i+1},u_i)[\mathsf{sgn}(u_{i+1}-k)|u_{i+1}-k|-\mathsf{sgn}(u_i-k)|u_i-k|] \\
&\;= -Q(u_{i+1},u_i)[\mathsf{u}_{i+1}-\mathsf{u}_i]
\end{aligned}
$$

iii) $u_{i+1} > k > u_i$:
$\Longrightarrow\ (u_{i+1}\top k) = u_{i+1},\ (u_i\top k) = k,\ (u_{i+1}\bot k) = k,\ (u_i\bot k) = u_i$:

$$
\begin{aligned}
&Q(u_{i+1},k)[u_{i+1}-k]-Q(k,u_i)[k-u_i] \\
&\;= Q(u_{i+1},k)[u_{i+1}-k]+Q(k,u_i)[u_i-k] \\
&\;= Q(u_{i+1},k)|u_{i+1}-k|-Q(k,u_i)|u_i-k| \\
&\;= Q(u_{i+1},k)\mathsf{u}_{i+1}-Q(k,u_i)\mathsf{u}_i
\end{aligned}
$$

iv) $u_{i+1} < k < u_i$:
$\Longrightarrow\ (u_{i+1}\top k) = k,\ (u_i\top k) = u_i,\ (u_{i+1}\bot k) = u_{i+1},\ (u_i\bot k) = k$:

$$
\begin{aligned}
&Q(k,u_i)[k-u_i]-Q(u_{i+1},k)[u_{i+1}-k] \\
&\;= -Q(k,u_i)[u_i-k]-Q(u_{i+1},k)[u_{i+1}-k] \\
&\;= -Q(k,u_i)|u_i-k|+Q(u_{i+1},k)|u_{i+1}-k| \\
&\;= -Q(k,u_i)\mathsf{u}_i+Q(u_{i+1},k)\mathsf{u}_{i+1}
\end{aligned}
$$

With the Taylor-expansions

$$
\begin{aligned}
Q(u_{i+1},k) &= Q(u_{i+1},u_i)+\partial_{u_i}Q(u_{i+1},u_i)(k-u_i)+Q(h^2) \\
&= Q(u_{i+1},u_i)+\partial_{u_i}Q(u_{i+1},u_i)\alpha(u_{i+1}-u_i)+O(h^2)
\end{aligned}
$$

$$
\begin{aligned}
Q(k,u_i) &= Q(u_{i+1},u_i)+\partial_{u_{i+1}}Q(u_{i+1},u_i)(u_{i+1}-k)+O(h^2) \\
&= Q(u_{i+1},u_i)+\partial_{u_{i+1}}Q(u_{i+1},u_i)(1-\alpha)(u_{i+1}-u_i)+O(h^2)
\end{aligned}
$$

we derive for the last case with $k = \alpha u_i + (1-\alpha)u_{i+1}, \quad \alpha \in (0,1)$

$$
\begin{aligned}
&Q(k,u_i)[k-u_i] - Q(u_{i+1},k)[u_{i+1}-k] \\
&= -Q(k,u_i)\mathsf{u}_i + Q(u_{i+1},k)\mathsf{u}_{i+1} \\
&= Q(u_{i+1},u_i)[\mathsf{u}_{i+1}-\mathsf{u}_i] \\
&\quad +\partial_{u_i}Q\alpha(u_{i+1}-u_i)\mathsf{u}_{i+1} - \partial_{u_{i+1}}Q(1-\alpha)(u_{i+1}-u_i)\mathsf{u}_i + O(h^2) \\
&= Q(u_{i+1},u_i)[\mathsf{u}_{i+1}-\mathsf{u}_i] \\
&\quad +\partial_{u_i}Q\alpha(u_{i+1}-u_i)\mathsf{sgn}(u_{i+1}-k)(u_{i+1}-\alpha u_{i+1}-(1-\alpha)u_i) \\
&\quad -\partial_{u_{i+1}}Q(1-\alpha)(u_{i+1}-u_i)\mathsf{sgn}(u_i-k)(u_i-\alpha u_{i+1}-(1-\alpha)u_i) + O(h^2) \\
&= Q(u_{i+1},u_i)[\mathsf{u}_{i+1}-\mathsf{u}_i] \\
&\quad -\left[\partial_{u_i}Q\alpha(1-\alpha)(u_{i+1}-u_i)^2 + \partial_{u_{i+1}}Q\alpha(1-\alpha)(u_{i+1}-u_i)^2\right] + O(h^2) \\
&= Q(u_{i+1},u_i)[\mathsf{u}_{i+1}-\mathsf{u}_i] \\
&\quad -\tfrac{1}{4}h^2 u'_{i+1/2}\alpha(1-\alpha)\left[\partial_{u_i}Q + \partial_{u_{i+1}}Q\right] + O(h^3)
\end{aligned}
$$

Since the function $f(\alpha) = \alpha(1-\alpha)$ posseses its maximum at $\alpha = 1/4$ we get

$$
\begin{aligned}
&|-Q(u_{i+1},k)[u_{i+1}-k] + Q(k,u_i)[k-u_i]| - |Q(u_{i+1},u_i)[\mathsf{u}_{i+1}-\mathsf{u}_i]| \\
&\qquad \le \frac{1}{4}h^2|(\partial_{u_{i+1}}+\partial_{u_i})Q(u_{i+1},u_i)|(u'_{i+1/2})^2 + O(h^3).
\end{aligned}
$$

which proofs the Theorem.

Remark 4 Numerical entropy fluxes of the form (5) were already used in [4] for the computation of numerical entropy inequalities. There they served as an entropy indicator to distinguish between regions with discontinuities and smooth ones. The results were quite satisfying.

3.2 Discrete cell entropy inequality

We start from the fact that we can compute the discrete entropy inequality for each cell. We follow Merriam [12,13] and start from the semi-discrete form of the entropy inequality (2)

$$\partial_t\, \mathsf{u}(u)_i + \tfrac{1}{\Delta x}[\mathsf{F}_{i+1/2,j} - \mathsf{F}_{i-1/2,j}] \le 0.$$

Since we assume piecewise constant distribution of u over the cell C_i, i.e.

$$\mathsf{u}_i = \mathsf{u}(\overline{u_i}) \;=\; \mathsf{u}(u_i),$$

with

$$\overline{u}_i = \frac{1}{h} \int\limits_{C_i} u(\xi, t^n)\, d\xi,$$

the chain rule is valid,

$$\partial_t \mathsf{u}(u_i) = \mathsf{u}' \partial_t(u_i) \ = \ \mathsf{u}'(u_t)_i. \tag{8}$$

Substituting this into the semi discrete cell inequality (8), using the numerical approximation for the time derivative, i.e. the difference of the numerical fluxes, yields the discrete cell entropy inequality for the cell C_i:

$$E_i := -\mathsf{u}_i'[F_{i+1/2} - F_{i-1/2}] + [\mathsf{F}_{i+1/2} - \mathsf{F}_{i-1/2}] \leq 0. \tag{9}$$

Such equations using $u_{i+1/2}$ and $u_{i-1/2}$ as interpolants at cell boundaries were extensively studied by Merriam [13]. He splitted (9) into two parts $E_{i+1/2}, E_{i-1/2}$ related to the corresponding cell boundaries, i.e.

$$\begin{aligned} E_i &= E_{i+1/2} + E_{i-1/2} \\ &= \underbrace{[-\mathsf{u}_i' F_{i+1/2} + \mathsf{F}_{i+1/2}]}_{E_{i+1/2}} + \underbrace{[\mathsf{u}_i' F_{i-1/2} - \mathsf{F}_{i-1/2}]}_{E_{i-1/2}} \ \leq \ 0. \end{aligned}$$

In order to satisfy the discrete entropy inequality, i.e. $E_i \leq 0$, one can follow the more restrictive approach to match the cell boundary inequalities

$$\begin{aligned} E_{i+1/2} &\leq 0, \\ E_{i-1/2} &\leq 0. \end{aligned} \tag{10}$$

This approach is easier to handle since it decouples the discrete cell entropy inequality (9) into two inequalities for each cell boundary.

With the proposed corresponding numerical entropy flux (5) above, we are able to write the according cell boundary entropy inequalities (10) as

$$\begin{aligned} E_{i+1/2} = -\mathsf{u}_i \left[\tfrac{1}{2}(f_{i+1} + f_i) - Q_{i+1/2}(u_{i+1} - u_i)\right] \\ +\tfrac{1}{2}(\mathsf{f}_{i+1} - \mathsf{f}_i) - Q_{i+1/2}(\mathsf{u}_{i+1} - \mathsf{u}_i) \end{aligned}$$

resp.

$$\begin{aligned} E_{i-1/2} = \mathsf{u}_i \left[\tfrac{1}{2}(f_i + f_{i-1}) - Q_{i-1/2}(u_i - u_{i-1})\right] \\ -\tfrac{1}{2}(\mathsf{f}_i - \mathsf{f}_{i-1}) - Q_{i-1/2}(\mathsf{u}_i - \mathsf{u}_{i-1}). \end{aligned}$$

The interesting result and the reward for the special choice of the numerical entropy flux given above is the following

Lemma 3 *A scheme of the form (4) with corresponding numerical entropy flux (5) satisfies the discrete cell entropy inequality (9) if one choose the numerical dissipation coefficient* $Q_{i+1/2}$ *to be*

$$Q^E_{i+1/2} = Q^E(u_{i+1}, u_i) := \max(Q^L_{i+1/2}, Q^R_{i+1/2}),$$

with

$$Q^R_{i+1/2} \geq -\frac{\frac{1}{2}[(\mathsf{f}_{i+1} - \mathsf{f}_i) - \mathsf{u}'_{i+1}(f_{i+1} - f_i)]}{[(\mathsf{u}_{i+1} - \mathsf{u}_i) - \mathsf{u}'_{i+1}(u_{i+1} - u_i)]}, \tag{11}$$

$$Q^L_{i+1/2} \geq \frac{\frac{1}{2}[(\mathsf{f}_{i+1} - \mathsf{f}_i) - \mathsf{u}'_i(f_{i+1} - f_i)]}{[(\mathsf{u}_{i+1} - \mathsf{u}_i) - \mathsf{u}'_i(u_{i+1} - u_i)]}. \tag{12}$$

Proof We start from the restrictive demand (10) and examine the inequality for $E_{i+1/2} \leq 0$, i.e.

$$\begin{aligned} E_{i+1/2} &= [-\mathsf{u}'_i F_{i+1/2} + \mathsf{F}_{i+1/2}] \\ &= -\mathsf{u}'_i \left[\tfrac{1}{2}(f_{i+1} + f_i) - Q_{i+1/2}(u_{i+1} - u_i)\right] \\ &\quad + \tfrac{1}{2}(\mathsf{f}_{i+1} - \mathsf{f}_i) - Q_{i+1/2}(\mathsf{u}_{i+1} - \mathsf{u}_i) \leq 0. \end{aligned}$$

If we rearrange this inequality in order to derive an inequality for the numerical dissipation coefficient $Q_{i+1/2}$, we end up with (11).

If we do the same for the cell inequality of the neigbouring cell C_{i+1} and consider E^L_{i+1}, we derive (12). The maximum satisfies both cell boundary entropy inequalities concerning the boundary $C_{i+1/2}$. Doing the same for $C_{i-1/2}$ reveals the desired result.

Remark 5 Having in mind that we have used the quite restrictive form (10) there seemed to be hope to follow the path where we admit the violation of (10) and require only the discrete cell inequality (9). Obviously more degrees of freedom coincide with a much more complex analysis since this leads to a coupled system for the dissipation coefficients.

One can think of the use of linear programming minimizing the entropy over all cells under the requirement that (9) is satisfied in each cell. From the theoretical point of view this may be an interesting scheme even if in practice this will lead to a very slow algorithm.

Corollary 1 *Up to second order the numerical dissipation coefficient* $Q^E_{i+1/2}$ *coincide with the dissipation coefficient of the Roe scheme[18] or Murman-Courant-Isaacson-Rees scheme [14, 1], i.e.*

$$Q^E_{i+/12} = \left|f'(u_{i+1/2}\right| + O(h^2)$$

Proof First (12) is considered. If we use Taylor-series expansion the expression writes

$$\frac{1}{2}\left[\frac{h[\mathsf{f}'(u_{i+1/2})u'_{i+1/2}-\mathsf{u}'(u_i)f'(u_{i+1/2})u'_{i+1/2}]+O(h^2)}{h[\mathsf{u}'(u_{i+1/2})u'_{i+1/2}-\mathsf{u}'(u_i)u'_{i+1/2}]+O(h^2)}\right]$$

Since the entropy flux obeys the compatability relation (3), we write

$$\frac{1}{2}\left[\frac{h[\mathsf{u}'(u_{i+1/2})-\mathsf{u}'(u_i)]f'(u_{i+1/2})u'_{i+1/2}+O(h^2)}{h[\mathsf{u}'(u_{i+1/2})-\mathsf{u}'(u_i)]u'_{i+1/2}+O(h^2)}\right].$$

Thus we derive

$$\frac{1}{2}\frac{f'(u_{i+1/2})+O(h^2)}{1+O(h^2)}=\tfrac{1}{2}f'(u_{i+1/2})+O(h^2).$$

On the other hand, if we examine (11) using Taylor-series expansion the expression writes as

$$-\frac{1}{2}\left[\frac{h[\mathsf{f}'(u_{i+1/2})u'_{i+1/2}-\mathsf{u}'(u_{i+1})f'(u_{i+1/2})u'_{i+1/2}]+O(h^2)}{h[\mathsf{u}'(u_{i+1/2})u'_{i+1/2}-\mathsf{u}'(u_{i+1})u'_{i+1/2}]+O(h^2)}\right].$$

Similar to the case above, we derive

$$-\frac{1}{2}\left[\frac{h[\mathsf{u}'(u_{i+1/2})-\mathsf{u}'(u_{i+1})]f'(u_{i+1/2})u'_{i+1/2}+O(h^2)}{h[\mathsf{u}'(u_{i+1/2})-\mathsf{u}'(u_{i+1})]u'_{i+1/2}+O(h^2)}\right],$$

and end up with

$$-\frac{1}{2}\frac{f'(u_{i+1/2})+O(h^2)}{1+O(h^2)}=-\tfrac{1}{2}f'(u_{i+1/2})+O(h^2)$$

Since the derived cell boundary entropy inequality satisfying dissipation coefficient is defined as the maximum over the both expression (11,12) and consequently the equation results in

$$Q^E_{i+1/2}=\tfrac{1}{2}|f'_{i+1/2}|+O(h^2)$$

which is a second order approximation of

$$Q^{Roe}_{i+1/2}=\frac{1}{2}\left|\frac{f(u_{i+1})-f(u_i)}{u_{i+1}-u_i}\right|.$$

If we look more in detail at the expressions (11),(12) there are some interesting consequences. If we start with (12), i.e.

$$\frac{\frac{1}{2}[(\mathsf{f}_{i+1}-\mathsf{f}_i)-\mathsf{u}'_i(f_{i+1}-f_i)]}{[(\mathsf{u}_{i+1}-\mathsf{u}_i)-\mathsf{u}'_i(u_{i+1}-u_i)]},$$

we first consider the denominator. From the convexity of $\mathsf{u}(u)$ immediately follows

$$(\mathsf{u}_{i+1}-\mathsf{u}_i)-\mathsf{u}'_i(u_{i+1}-u_i)\geq 0. \tag{13}$$

Using the Rankine-Hugoniot condition for the entropy inequality, i.e.

$$s[\mathsf{u}(u)]\geq[\mathsf{f}(u)]$$

with $s=(f_r-f_l)/(u_r-u_l)$ the entropy condition reads as

$$\mathsf{f}(u_r)-\mathsf{f}(u_l)-\left(\frac{\mathsf{u}(u_r)-\mathsf{u}(u_l)}{u_r-u_l}\right)(f(u_r)-f(u_l)\leq 0. \tag{14}$$

Inserting the Taylor-approximation for $\mathsf{u}(u_{i+1})$, i.e. the numerator of (12) with $u_r=u_{i+1}$, $u_l=u_i$:

$$\mathsf{f}(u_{i+1})-\mathsf{f}(u_i)-\mathsf{u}'_i(f(u_{i+1})-f(u_i)\leq 0.$$

If we assume

$$\mathsf{u}'_i(f_{i+1}-f_i)\geq\frac{\mathsf{u}_{i+1}-\mathsf{u}_i}{u_{i+1}-u_i}(f_{i+1}-f_i)$$

– which depends on the signs of $f_{i+1}-f_i$ and $u_{i+1}-u_i$ and holds for $\mathsf{sgn}(\Delta f_{i+1/2}),\mathsf{sgn}(\Delta u_{i+1/2})$ both positive or both negative – with the convexity requirement (13) it follows immediately by switching signs and adding the difference of the entropy fluxes, that

$$\begin{aligned}&\mathsf{f}(u_r)-\mathsf{f}(u_l)-\mathsf{u}'_i(f(u_r)-f(u_l))\\&\quad\leq\mathsf{f}(u_r)-\mathsf{f}(u_l)-\left(\frac{\mathsf{u}(u_r)-\mathsf{u}(u_l)}{u_r-u_l}\right)(f(u_r)-f(u_l))\end{aligned}$$

Since we want to assure the discrete Rankine-Hugoniot condition (14), this leads to

$$\mathsf{f}(u_{i+1})-\mathsf{f}(u_i)-\mathsf{u}'_i(f(u_{i+1})-f(u_i)\leq 0.$$

Thus we derive that the dissipation coefficient $Q^L_{i+1/2}$, (12) is nonpositive, since the denominator (13) was nonnegative.

If we consider the expression (11), the convexity requirement for the entropy function $\mathsf{u}(u)$ is

$$(\mathsf{u}_{i+1}-\mathsf{u}_i)-\mathsf{u}'_{i+1}(u_{i+1}-u_i)\leq 0.$$

To assure that

$$-(\mathsf{f}(u_r) - \mathsf{f}(u_l)) + \mathsf{u}'_{i+1}(f(u_r) - f(u_l))$$
$$\geq -(\mathsf{f}(u_r) - \mathsf{f}(u_l)) + \left(\frac{\mathsf{u}(u_r) - \mathsf{u}(u_l)}{u_r - u_l}\right)(f(u_r) - f(u_l)) \geq 0$$

holds it is enough to guarantee

$$\mathsf{u}'_{i+1}(f_{i+1} - f_i) \geq \frac{\mathsf{u}_{i+1} - \mathsf{u}_i}{u_{i+1} - u_i}(f_{i+1} - f_i)$$

which is again true for $f_{i+1} - f_i$ and $u_{i+1} - u_i$ having the same sign and we have

$$-(\mathsf{f}(u_{i+1}) - \mathsf{f}(u_i)) + \mathsf{u}'_i(f(u_{i+1}) - f(u_i) \ \geq \ 0.$$

This leads to the fact, that (11) may be nonpositive. This considerations show that it may be possible that even the maximum of (11),(12) might be negative.

This fact can also be illustrated by an instructive example. First we rewrite the dissipation coefficients as

$$Q^R_{i+1/2} = \frac{\frac{1}{2}[(\mathsf{f}_{i+1} - \mathsf{f}_i) - \mathsf{u}'_{i+1}(f_{i+1} - f_i)]}{[\mathsf{u}'_i(u_{i+1} - u_i) - (\mathsf{u}_{i+1} - \mathsf{u}_i)]},$$

$$Q^L_{i+1/2} = \frac{\frac{1}{2}[(\mathsf{f}_{i+1} - \mathsf{f}_i) - \mathsf{u}'_i(f_{i+1} - f_i)]}{[(\mathsf{u}_{i+1} - \mathsf{u}_i) - \mathsf{u}'_i(u_{i+1} - u_i)]}.$$

By the convexity argument both denominators are positive and we focus on the nominator. If the entropy flux $\mathsf{f}(u)$ has a minimum at $u_{i+1/2}$ i.e. $\mathsf{f}'_{i+1/2} = \mathsf{u}'_{i+1/2} f'_{i+1/2} = 0$ we might have

$$\mathsf{u}'_{i+1}(f_{i+1} - f_i) > (\mathsf{f}_{i+1} - \mathsf{f}_i) = 0,$$
$$\mathsf{u}'_i(f_{i+1} - f_i) > (\mathsf{f}_{i+1} - \mathsf{f}_i) = 0,$$

and both numerators are negative. Since and so the dissipation coefficients.

This considerations show that it may be possible that even the maximum of (11),(12) might be negative. To remedy this we add a positive constant δ which avoids that the dissipation coefficient is negative or vanishes:

$$Q_{i+1/2} = \max(\delta, Q^L, Q^R).$$

We have thus proved the following

Theorem 5 *The dissipation coefficient of the form*

$$Q_{i+1/2} = \max(\delta, Q^L, Q^R), \qquad 0 < \delta << 1, \tag{15}$$

derived from the cell entropy inequalities (11), (12) is consistent with the Roe scheme with entropy fix.

Remark 6 This correction is by no means surprising. Since we have seen that the limit of the cell entropy coefficients (11),(12) gives the dissipation coefficient of the Roe-scheme. This scheme is known to produce entropy-violating solution and needs an entropy fix given by Harten and Hyman [6]. The fix is of the same kind like in this case. If the numerical dissipation coefficient gets to small, an ε takes the role of the dissipation coefficient, i.e.

$$Q^{Roe}_{i+1/2} = \begin{cases} \left|\frac{\Delta_{i+1/2}f}{\Delta_{i+1/2}u}\right| ; & \text{if } \left|\frac{\Delta_{i+1/2}f}{\Delta_{i+1/2}u}\right| > \varepsilon \\ \quad\varepsilon \quad ; & \text{else,} \end{cases} \quad , \quad \varepsilon << 1.$$

Lemma 4 *The dissipation coefficient (15) can be written as*

$$Q_{i+1/2} = \max_{u^\star \in [u_i, u_{i+1}]} \left[\frac{1}{2}\left(\frac{f_{i+1} - f(u^\star)}{u_{i+1} - u^\star}\right), -\frac{1}{2}\left(\frac{f(u^\star) - f_i}{u^\star - u_i}\right), \delta\right].$$

Proof As we know from the result of Krutzkov [8] it is enough to consider the family of entropy pairs

$$\mathsf{u}_i = |u_i - k|, \quad \mathsf{f}_i = \mathsf{sgn}(u_i - k)(f_i - f(k)), \quad \forall k \in \mathbb{R}.$$

If we insert this into the expressions (12) we derive

$$\begin{aligned} Q^L_{i+1/2} &= \frac{1}{2}\left[\frac{\mathsf{f}_{i+1} - \mathsf{f}_i - \mathsf{sgn}(u_i - k)(f_{i+1} - f_i)}{|u_{i+1} - k| - |u_i - k| - \mathsf{sgn}(u_i - k)(u_{i+1} - u_i)}\right] \\ &= \frac{1}{2}\left[\frac{\mathsf{f}_{i+1} - \mathsf{f}_i - \mathsf{sgn}(u_i - k)(f_{i+1} - f(k) + f(k) - f_i)}{|u_{i+1} - k| - |u_i - k| - \mathsf{sgn}(u_i - k)(u_{i+1} - k + k - u_i)}\right] \\ &= \frac{1}{2}\left[\frac{\mathsf{f}_{i+1} - \mathsf{f}_i + \mathsf{f}_i - \mathsf{sgn}(u_i - k)[f_{i+1} - f(k)]}{|u_{i+1} - k| - |u_i - k| + |u_i - k| - \mathsf{sgn}(u_i - k)(u_{i+1} - k)}\right] \\ &= \frac{1}{2}\left[\frac{[\mathsf{sgn}(u_{i+1} - k) - \mathsf{sgn}(u_i - k)](f_{i+1} - f(k))}{[\mathsf{sgn}(u_{i+1} - k) - \mathsf{sgn}(u_i - k)](u_{i+1} - k)}\right] \end{aligned}$$

As we have seen in the proof of Theorem 4 the interesting case is $k \in (u_i, u_{i+1})$ for $u_i < u_{i+1}$ resp. $k \in (u_{i+1}, u_i)$ for $u_i > u_{i+1}$. This gives the desired form

$$\frac{1}{2}\left[\frac{f_{i+1} - f(u^\star)}{u_{i+1} - u^\star}\right], \quad u^\star \in (u_i, u_{i+1}).$$

The same computation for $Q^R_{i+1/2}$ gives the analogous result and finally the form given above.

Corollary 2 *A three-point scheme with numerical dissipation coefficient (15), i.e.*

$$Q^{CE}_{i+1/2} = \max(\delta, Q^L, Q^R), \qquad 0 < \delta << 1, \tag{16}$$

is an E-scheme.

Proof We start from the definition for the flux of an E-scheme we have

$$\mathsf{sgn}(u_{i+1} - u_i)[F_{i+1/2} - f(u)] \leq 0, \quad \forall u \in [u_i, u_{i+1}].$$

If we insert the numerical flux function

$$F^{CE}_{i+1/2} = \tfrac{1}{2}(f_{i+1} + f_i) - Q^{CE}_{i+1/2}(u_{i+1} - u_i)$$

we arrive at

$$\begin{aligned}
&\mathsf{sgn}(u_{i+1} - u_i)[F^{CE}_{i+1/2} - f(u)] \\
&= \mathsf{sgn}(u_{i+1} - u_i)[\tfrac{1}{2}(f_{i+1} + f_i) - Q^{CE}_{i+1/2}(u_{i+1} - u_i) - f(u)] \\
&= \mathsf{sgn}(u_{i+1} - u_i)[\tfrac{1}{2}(f_{i+1} - f(u)) + \tfrac{1}{2}(f_i - f(u)) - Q^{CE}_{i+1/2}(u_{i+1} - u_i) \\
&= \mathsf{sgn}(u_{i+1} - u_i)[\tfrac{1}{2}(f_{i+1} - f(u)) - \tfrac{1}{2}(f(u) - f_i) - Q^{CE}_{i+1/2}(u_{i+1} - u_i) \\
&= \mathsf{sgn}(u_{i+1} - u_i)\left[(u_{i+1} - u)\left(\frac{1}{2}\frac{f_{i+1} - f(u)}{u_{i+1} - u} - Q^{CE}_{i+1/2}\right)\right. \\
&\qquad \left. -(u - u_i)\left(\frac{1}{2}\frac{f(u) - f_i}{u - u_i} - Q^{CE}_{i+1/2}\right)\right] \\
&= \mathsf{sgn}(u_{i+1} - u_i)\left[-(u_{i+1} - u)\underbrace{\left(Q^{CE}_{i+1/2} - \frac{1}{2}\frac{f_{i+1} - f(u)}{u_{i+1} - u}\right)}_{\alpha}\right. \\
&\qquad \left. -(u - u_i)\underbrace{\left(Q^{CE}_{i+1/2} - \frac{1}{2}\frac{f(u) - f_i}{u - u_i}\right)}_{\beta}\right].
\end{aligned}$$

By Lemma 4, the dissipation coefficient $Q^{CE}_{i+1/2}$ majorizes the terms on the right-hand side, i.e. $\alpha, \beta \geq 0$. Thus the terms in the square brackets each are nonnegative and we get

$$\begin{aligned}
&\mathsf{sgn}(u_{i+1} - u_i)[F^{ce}_{i+1/2} - f(u)] \\
&= \mathsf{sgn}(u_{i+1} - u_i)[-(u_{i+1} - u_i)\alpha - (u - u_i)\beta] \\
&\leq -\mathsf{sgn}(u_{i+1} - u_i)[(u_{i+1} - u_i)\max(\alpha, \beta)] \\
&\leq -|u_{i+1} - u_i| \max(\alpha, \beta) \leq 0.
\end{aligned}$$

4 Conclusions

We have presented an alternative derivation of the dissipation model of the Roe scheme for scalar conservation laws by discrete cell entropy inequalities. To this use we have proposed the discretization of the numerical entropy flux similar to the disretization of the corresponding numerical flux function motivated by Lax. We have shown that this choice is consistent with the numerical entropy flux proposed by Crandall and Majda.

As mentioned early all the derivation of the dissipation coefficients from the cell entropy inequalities is based on the restricitve demand that both cell entropy inequalities might be nonnegative. This demand is maybe to restrictive and let some degrees of freedom to derive entropy inequalities which admits numerical schemes with less dissipation. This approach will be examined in detail [5].

References

1. R. Courant, E. Isaacson, and M. Rees. On the solution of nonlinear hyperbolic differential equations by finite differences. *Comm. Pure Appl. Math.*, 5:243–255, 1952.
2. M.G. Crandall and A. Majda. Monotone difference approximations for scalar conservation laws. *Math. Comp.*, 34:1–21, 1980.
3. S.K. Godunov. A finite difference method for the numerical computation of discontinous solutions of the equations of fluid dynamics. *Math. Sb.*, 47:271–306, 1959.
4. Th. Grahs and Th. Sonar. Entropy controlled artificial anisotropic diffusion for the numerical solution of conservation laws based on algorithms from image processing. *J. Vis. Com*, to appear, 2002.
5. Th. Grahs and Th. Sonar. Less restrictive discrete cell entropy inequalities. *Technical report*, TU Braunschweig:to appear, 2002.
6. A. Harten and J. M. Hyman. Self-adjusting grid methods for one-dimensional hyperbolic conservation laws. *J.Comp.Phys.*, 50:235–269, 1983.
7. A. Harten, J.M. Hyman, and P.D. Lax. On finite-difference approximations and entropy conditions for shocks. *Comm. Pure. Appl. Math.*, 29:297–322, 1976.
8. S.N. Kruzkov. First-order quasilinear equations in several independent variables. *Math. USSR Sbornik*, 10:217–243, 1970.

9. P.D. Lax. Hyperbolic systems of conservation laws II. *Comm. Pure Appl. Math.*, 10:537–566, 1957.
10. P.D. Lax. Schock waves and entropy. In E.A. Zarantello, editor, *Contribution to Nonlinear Functional Analysis*, pages 603–634. Academic Press New York San Franzisco London, 1971.
11. P.D. Lax and B. Wendroff. Systems of conservation laws. *Comm. Pure Appl. Math.*, 13:217–237, 1960.
12. M.L. Merriam. Smoothing and the second law. *Comp. Meth. App. Mech. Eng.*, 64(1):177–193, 1987.
13. M.L. Merriam. An entropy-based approach to nonlinear stability. *NASA Technical Memorandum 101086*, 64(1):177–193, 1989.
14. E.M. Murman. Analysis of embedded shock waves calculated by relaxation. *AIAA J.*, 12:626–633, 1974.
15. S. Osher. Riemann solvers, the entropy condition, and difference approximations. *SIAM J. Numer. Anal.*, 21:217–235, 1984.
16. S. Osher and S. Chakravarthy. High resolution schemes and the entropy condition. *SIAM, J. Numer. Anal.*, 21:955–984, 1984.
17. S. Osher and E. Tadmor. On the convergence of difference approximations to scalar conservation laws. *Math. Comp.*, 50:19–51, 1988.
18. P.L. Roe. Approximate Riemann solvers, parameter vectors, and difference schemes. *J.Comp.Phys.*, 43:357–372, 1981.
19. M. Schonbek. Second order conservative schemes and the entropy condition. *Math. Comp.*, 44:423–468, 1985.
20. Th. Sonar. Entropy production in second-order three-point schemes. *Numer. Math*, 62:371–390, 1992.
21. E. Tadmor. The large-time behaviour of the scalar, genuinely nonlinear lax-friedrichs scheme. *Math. Comp.*, 43:353–368, 1984.
22. E. Tadmor. Numerical viscosity and the entropy condition for conservative difference schemes. *Math. Comp.*, 43:369–381, 1984.
23. E. Tadmor. The numerical viscosity of entropy stable schemes for systems of conservation laws i. *Math. Comp.*, 49:91–103, 1987.