

# Convergence of higher order upwind finite volume schemes on unstructured grids for scalar conservation laws in several space dimensions\*

Dietmar Kröner<sup>1</sup>, Sebastian Noelle<sup>2</sup>, Mirko Rokyta<sup>3\*\*</sup>

<sup>1</sup> Freiburg University, Germany; e-mail: dietmar@titan.mathematik.uni-freiburg.de

<sup>2</sup> Bonn University, Germany; e-mail: noelle@sfb256.iam.uni-bonn.de

<sup>3</sup> Charles University, Praha, Czech Republic; e-mail: rokyta@karlin.mff.cuni.cz

Received April 15, 1993 / Revised version received March 13, 1995

**Summary.** We prove convergence of a class of higher order upwind finite volume schemes on unstructured grids for scalar conservation laws in several space dimensions. The result is applied to the discontinuous Galerkin method due to Cockburn, Hou and Shu.

*Mathematics Subject Classification (1991):* 35L60, 35L65, 35L67, 65M12, 76N15

## 1. Introduction

In this paper we shall consider higher order finite volume discretizations for hyperbolic conservation laws in several space dimensions

$$(1.1) \quad \frac{\partial}{\partial t} u(x, t) + \nabla \cdot f(u(x, t)) = 0,$$

with initial values

$$(1.2) \quad u(x, 0) = u_0(x),$$

where  $x \in \mathbb{R}^d$ ,  $t \in \mathbb{R}_+$ ,  $u \in \mathbb{R}$  and  $f \in C^1(\mathbb{R}, \mathbb{R}^d)$  (in fact, in this paper we only need that  $f$  is Lipschitz-continuous). We consider initial data  $u_0 \in L^1 \cap L^\infty(\mathbb{R}^d)$ .

There are essentially three different numerical methods for discretizing (1.1), (1.2). Finite difference methods are defined on cartesian, structured grids. Via dimensional splitting one-dimensional schemes are applied to multidimensional problems [8, 10, 12]. Furthermore there are the finite volume [9, 14, 16, 19, 22, 29] and the finite element methods (streamline diffusion, streamline diffusion shock capturing, transport diffusion) [20, 21, 32, 35] on unstructured grids.

---

\* This work was supported by Deutsche Forschungsgemeinschaft, SFB 256

\*\* Partially supported by the grant No. 201/93/2177 of the Czech Grant Agency

Correspondence to: D. Kröner

The solution of (1.1), (1.2) will in general develop discontinuities which will move through the domain. In order to avoid a very fine global grid for getting higher resolution it is necessary to use *local* mesh refinement near these shocks. A major disadvantage of cartesian grids is that local mesh refinement will either have global effects or produce hanging nodes. On the other hand, it is very simple to refine an unstructured grid locally by dividing a single triangle into two new or four new triangles. Moreover, such grids are very flexible if one wants to discretize bounded domains with more general geometries, which frequently occur in applications. Finally, for improving the resolution of discontinuities it is necessary to use higher order schemes together with flux limiters. Now for finite volume and for finite element schemes mesh refinement and higher order accuracy can be combined much more easily than for dimensional splitting schemes.

Convergence of higher order accurate schemes in several space dimensions was first proved by Johnson and Szepessy [20, 35] for the streamline diffusion shock capturing method on unstructured grids and later by Coquel and LeFloch [9, 10] for a class of dimensional splitting schemes on cartesian grids. For finite volume schemes, there are so far only results for monotone, first order schemes [11, 3, 24]. In this paper, we generalize these results and show convergence of a class of higher order accurate, upwind finite volume schemes on unstructured triangular grids in several space dimensions.

The basic idea for getting higher order is similar as in 1-D. Let us briefly repeat the procedure in 1-D. First we consider a scheme in conservation form of first order.

$$(1.3) \quad u_i^{n+1} := u_i^n - \frac{\Delta t}{\Delta x} (g(u_i^n, u_{i+1}^n) - g(u_{i-1}^n, u_i^n))$$

for the initial value problem

$$(1.4) \quad \partial_t u(x, t) + \partial_x f(u(x, t)) = 0 \quad \text{in } \mathbb{R} \times \mathbb{R}^+,$$

with initial values

$$(1.5) \quad u(x, 0) = u_0(x) \text{ in } \mathbb{R}.$$

The value  $u_i^n$  is assumed to become an approximation of  $u(i \Delta x, n \Delta t)$ , where  $u$  is the exact solution of (1.4), (1.5) and  $\Delta t$ ,  $\Delta x$  refer to a uniform grid. The vector  $(u_i^n)_i$  can be used to define a piecewise constant function  $v_h : \mathbb{R} \rightarrow \mathbb{R}$  as

$$v_h(x, t) := u_i^n \quad \text{if } x_{i-\frac{1}{2}} \leq x < x_{i+\frac{1}{2}} \quad \text{and } t_{n-1} \leq t < t_n.$$

In order to get higher order schemes in 1-D it is usual to replace the piecewise constant function  $v_h$  by a piecewise linear one  $u_h$ , e.g.

$$u_h(x, t) := u_i^n + s_j(x - x_i) \quad \text{for } x_{i-\frac{1}{2}} \leq x < x_{i+\frac{1}{2}} \quad \text{and } t_{n-1} \leq t < t_n.$$

Here  $s_j$  is a suitable slope which has to be chosen very carefully in order to avoid oscillations. One way to do this is given by the following definition. Let

$$\sigma_j = \text{sign}(u_j^n - u_{j-1}^n)$$

and  $r$  be defined by

$$|u_j^n - u_r^n| := \min\{|u_j^n - u_{j-1}^n|, |u_{j+1}^n - u_j^n|\}.$$

Then the slope  $s_j$  is defined by

$$s_j = \frac{1}{\Delta x}(u_j^n - u_r^n) \quad \text{if } \sigma_j \sigma_{j+1} > 0$$

$$s_j = 0 \quad \text{otherwise.}$$

This means that  $s_j$  corresponds to that slope which modulus is minimal, except for local extrema, where  $s_j = 0$ . In this paper we shall consider a generalization of this scheme to 2-D. The convergence of a similar scheme of higher order for 1-D scalar conservation laws with a strictly convex flux function has been proved in [33].

The outline of this paper is as follows: In Sect. 2, we define the class of finite volume schemes which we shall consider in this paper, and state the main convergence result. On the triangles we use piecewise constant values for the approximating solutions. For instance the discontinuous Galerkin method as defined in [6] defines a piecewise smooth function on the grid. The jumps of this function along the edges of the triangles are used to define the numerical fluxes through these edges. It turns out, that for defining the scheme we only need the values of this piecewise smooth function in some integration nodes along the edges. Of course these values should satisfy some properties, which will also be given in this section. As in [11, 3, 24, 35], the main tool for proving the convergence is the concept of measure valued solutions, which was introduced by DiPerna [13]. The details of this theory will be given in Sect. 3. Then in Sections 4–6 we shall show that the approximating sequence satisfies the assumptions of Theorem 3.8 of DiPerna. In particular in Sect. 4 we shall prove the  $L^\infty$ -estimate using similar ideas as in [6]. This estimate implies the weak-star convergence of the approximating sequence to a Young measure. Then in Sect. 5 the convexity of the numerical entropy is used to prove similarly as in [24], that the Young measure also satisfies a set of entropy conditions and therefore is an admissible measure valued solution. The derivation of these entropy inequalities is the central piece of this work. The  $L^1$ -estimate and the consistency with the initial conditions will be given in Sect. 6. DiPerna's theory now implies that the Young measure reduces to the admissible weak solution in the sense of Kruzkov [25]. In Sect. 7 we show that for higher order finite volume schemes, the restrictions imposed on the values at the edges do not affect the order of accuracy. We also remark that the discontinuous Galerkin method as defined in [6] satisfies the conditions of Sect. 2. In Sect. 8, we characterize the numerical fluxes to which our convergence proof applies.

## 2. Statement of the main result

In this section we fix the notation concerning the triangulation, define the class of numerical schemes for which we can show convergence and we summarize the main result.

**Definition 2.1.** The set

$$\mathcal{T} := \{T_i | T_i \text{ is a triangle for } i \in I \subseteq \mathbb{N}\}$$

where  $I \subseteq \mathbb{N}$  is an index set, is called an unstructured grid of  $\Omega \subset \mathbb{R}^2$  if the following two properties are satisfied:

$$1) \quad \Omega = \bigcup_{i \in I} T_i,$$

2) For two different  $T_i, T_j$  we have  $T_j \cap T_i = \emptyset$  or

$$T_j \cap T_i = \text{a common vertex of } T_i, T_j \text{ or}$$

$$T_j \cap T_i = \text{a common edge of } T_i, T_j.$$

In this exposition, we restrict ourselves to two space dimensions. All results can be readily generalized to any number of space dimensions.

**Notation 2.2.** Let  $I$  be an index set and for  $h > 0$  let  $\mathcal{T}_h := \{T_i | i \in I\}$  denote an unstructured triangular grid of  $\mathbb{R}^2$ . We will use the following notation:

$T_i$ : the  $i^{\text{th}}$  triangle.

$|T_i|$ : area of  $T_i$ .

$h := \sup_i \text{diam}(T_i)$ .

$t^n := n \Delta t, n = 0, \dots, N$ : the time after  $n$  time steps.  $N \Delta t = T$ .

$x_i$ : center of gravity of  $T_i$ .

$N_i$ : set of the indices  $j$  of the triangles  $T_j$  neighboring  $T_i$ .

$S_{ij}, j \in N_i$ : joint edge of  $T_i, T_j$ .

$|S_{ij}|$ : length of  $S_{ij}$ .

$n_{ij}$ : outward unit normal to  $T_i$  in the direction of  $j \in N_i$ .

**Assumption 2.3.** We assume that there are constants  $c_1$  and  $c_V$  such that

$$(2.1) \quad 0 < c_1 \leq \frac{\Delta t}{h} \text{ and } \sup_{i \in I} \frac{h^2}{|T_i|} \leq c_V$$

if  $\Delta t, h \rightarrow 0$ .

**Assumption 2.4.** Let  $g_{ij}(u, v)$  be a numerical flux consistent with  $f(u) \cdot n_{ij}$ , i.e.

$$(2.2) \quad g_{ij}(u, u) = f(u) \cdot n_{ij}.$$

We assume that  $g$  is Lipschitz-continuous. In particular, suppose that for all  $M > 0$  there is a constant  $C_g = C_g(M)$  such that for all  $u, u', v, v' \in [-M, M]$

$$(2.3) \quad |g_{ij}(u, v) - g_{ij}(u', v')| \leq C_g(M)(|u - u'| + |v - v'|)$$

and that  $g$  is conservative, i.e.

$$(2.4) \quad g_{ij}(\mathbf{u}, \mathbf{v}) = -g_{ji}(\mathbf{v}, \mathbf{u}).$$

Moreover, assume that  $g_{ij}$  is monotone:

$$(2.5) \quad \frac{\partial}{\partial \mathbf{u}} g_{ij}(\mathbf{u}, \mathbf{v}) \geq 0 \geq \frac{\partial}{\partial \mathbf{v}} g_{ij}(\mathbf{u}, \mathbf{v}).$$

*Example 2.5.* (Engquist-Osher and Lax-Friedrichs numerical fluxes) Let

$$c_{ij}(\mathbf{u}) := f(\mathbf{u}) \cdot \mathbf{n}_{ij},$$

and define

$$c_{ij}^+(\mathbf{u}) := \int_0^{\mathbf{u}} \max\{c'_{ij}(s), 0\} ds, \quad c_{ij}^-(\mathbf{u}) := \int_0^{\mathbf{u}} \min\{c'_{ij}(s), 0\} ds.$$

The Engquist-Osher numerical flux is then given by (see [15]):

$$g_{ij}^{\text{EO}}(\mathbf{u}, \mathbf{v}) := \left( c_{ij}^+(\mathbf{u}) + c_{ij}^-(\mathbf{v}) \right).$$

The Lax-Friedrichs numerical flux is given explicitly by (see [27]):

$$g_{ij}^{\text{LF}}(\mathbf{u}, \mathbf{v}) := \frac{f(\mathbf{u}) + f(\mathbf{v})}{2} \cdot \mathbf{n}_{ij} + \frac{\mathbf{u} - \mathbf{v}}{2\lambda_{ij}},$$

where  $\lambda_{ij}$  are arbitrarily chosen constants satisfying

$$\lambda_{ij} = \lambda_{ji} > \tilde{c} > 0,$$

and

$$\lambda_{ij} \sup_{|u| \leq M} |f'(u) \cdot \mathbf{n}_{ij}| \leq 1.$$

Then  $g_{ij}^{\text{EO}}$ ,  $g_{ij}^{\text{LF}}$  (and all their convex combinations) satisfy the conditions (2.2)–(2.5).

**Assumption 2.6.** Let  $U : \mathbb{R} \rightarrow \mathbb{R}$  be Lipschitz-continuous and convex, and let  $(U, F)$  be an entropy pair for (1.1). We assume that there exists a numerical entropy flux  $G_{ij}(\mathbf{u}, \mathbf{v})$  which is

(i) consistent with  $F \cdot \mathbf{n}_{ij}$ , i.e.

$$(2.6) \quad G_{ij}(\mathbf{u}, \mathbf{u}) = F(\mathbf{u}) \cdot \mathbf{n}_{ij},$$

(ii) Lipschitz-continuous, such that for all  $M > 0$  there is a constant  $C_G = C_G(M)$  such that for all  $\mathbf{u}, \mathbf{u}', \mathbf{v}, \mathbf{v}' \in [-M, M] \subset \mathbb{R}$

$$(2.7) \quad |G_{ij}(\mathbf{u}, \mathbf{v}) - G_{ij}(\mathbf{u}', \mathbf{v}')| \leq C_G(|\mathbf{u} - \mathbf{u}'| + |\mathbf{v} - \mathbf{v}'|),$$

(iii) conservative, i.e. for all  $j \in N_i$

$$(2.8) \quad G_{ij}(u, v) = -G_{ji}(v, u),$$

and

(iv) satisfies the compatibility conditions

$$(2.9) \quad \frac{\partial G_{ij}}{\partial u}(u, v) = U'(u) \frac{\partial g_{ij}}{\partial u}(u, v), \quad \frac{\partial G_{ij}}{\partial v}(u, v) = U'(v) \frac{\partial g_{ij}}{\partial v}(u, v)$$

almost everywhere.

We will require that the numerical flux  $g$  admits an numerical entropy flux satisfying Assumption 2.6. In Sect. 8, we will discuss this assumption and show that it is satisfied for the Engquist-Osher and the Lax-Friedrichs fluxes – in fact, we will characterize all such fluxes.

**Notation 2.7.** We will approximate the cell average of the solution at time  $t^n$  over triangle  $T_i$  by  $u_i^n$ . Below we will often drop the superscript  $n$  and write  $u_i = u_i^n$ . Given  $i, j$ , let  $x_{ijl} \in S_{ij}$  and  $\omega_l \geq 0, l = 1, \dots, m$  be such that the quadrature formula

$$(2.10) \quad \int_{S_{ij}} \varphi(x) dx = |S_{ij}| \sum_{l=1}^m \omega_l \varphi(x_{ijl}) + O(h^3)$$

holds for all smooth  $\varphi$ . Furthermore (see Fig. 1), let

$$(2.11) \quad d_{ij} := x_j - x_i \text{ and } d_{ijl} := x_{ijl} - x_i$$

and choose coefficients  $\alpha_{ijlp}, \beta_{ijlp}, p \in N_i$  such that

$$(2.12) \quad -d_{ijl} = \sum_{p \in N_i} \alpha_{ijlp} d_{ip}, \text{ and } d_{jil} = \sum_{p \in N_i} \beta_{ijlp} d_{ip}.$$

With the  $d_{ij}$  associate

$$\Delta_{ij} := u_j - u_i$$

and define

$$(2.13) \quad a_{ijl} := \sum_{p \in N_i} \alpha_{ijlp} \Delta_{ip} \text{ and } b_{ijl} := \sum_{p \in N_i} \beta_{ijlp} \Delta_{ip}.$$

We will also use the notation

$$I(c_1, \dots, c_n) := [\min\{c_1, \dots, c_n\}, \max\{c_1, \dots, c_n\}].$$

This notation, as well as the following definition, is adapted from Cockburn, Hou and Shu [6].

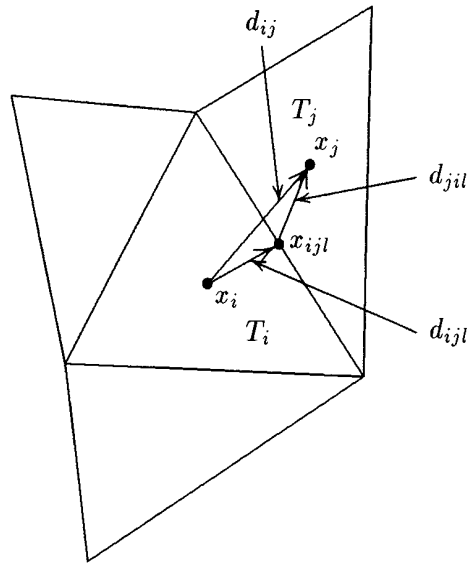


Fig. 1. Triangulation and notation

Remark 2.8. In order to illustrate the meaning of the  $a_{ijl}$  let us consider the case that given a point  $x_{ijl}$ , only two of the values  $\alpha_{ijlp}$  are non-zero, say

$$d_{ijl} = \alpha_{ijlp}d_{ip} + \alpha_{ijlq}d_{iq}.$$

Let  $L : \mathbb{R}^2 \rightarrow \mathbb{R}$  be the linear function defined by

$$L(x_i) = u_i, L(x_p) = u_p, L(x_q) = u_q.$$

Then

$$u_i + a_{ijl} = L(x_{ijl}).$$

Definition 2.9. a) An unstructured triangular grid  $\mathcal{T}_h$  as defined in Definition 2.1 is called a  $B$ -triangulation if the constants  $\alpha_{ijlp}$  and  $\beta_{ijlp}$  defined in (2.12) are nonnegative.

b) A family  $(\mathcal{T}_h)_{h>0}$  of  $B$ -triangulations is called  $B$ -uniform if there exists a constant  $\mu > 0$  such that  $0 \leq \alpha_{ijlp}, \beta_{ijlp} \leq \mu$  as  $h \rightarrow 0$ .

Example 2.10. In ([6], Prop. 2.9), it is shown that if the triangles are acute (no angle is greater than  $\pi/2$ ) and satisfy (2.1), then the family of triangulations is  $B$ -uniform with  $\mu = 2c_V^3(1 + c_V^2)^{2/3}$ .

Now we are ready to define the scheme.

Definition 2.11. (the numerical scheme) Let  $u_0 \in L^\infty(\mathbb{R}^2)$  with compact support, let  $u_i^0$  be defined by

$$(2.14) \quad u_i^0 := \frac{1}{|T_i|} \int_{T_i} u_0(x) dx$$

and assume that  $u_i^k$  are already defined for  $k \leq n$  and  $i \in I$ . Let  $\alpha \in ]\frac{1}{2}, 1]$ . Furthermore assume that there are values  $u_{ijl}^n \in \mathbb{R}$  for  $i \in I, j \in N_i$  and  $l = 1, \dots, m$  such that the following conditions are satisfied:

a) There is a constant  $C_1 > 0$  such that for all  $i, j, l$

$$(2.15) \quad |u_{ijl}^n| \leq C_1 h^\alpha.$$

b) There is a constant  $C_2 > 0$  such that for all  $i, j, l$

$$(2.16) \quad u_{ijl}^n \in I(\pm C_2 h^{2\alpha}, -a_{ijl} \pm C_2 h^{2\alpha})$$

$$(2.17) \quad u_{jil}^n \in I(\pm C_2 h^{2\alpha}, b_{jil} \pm C_2 h^{2\alpha}).$$

c) For all  $i, j, l$ ,

$$(2.18) \quad (u_i - u_j)u_{ijl}^n \leq C_2 h^{2\alpha}.$$

Then we define

$$(2.19) \quad u_i^{n+1} := u_i^n - \frac{\Delta t}{|T_i|} \sum_{j \in N_i} \sum_{l=1}^m \omega_l |S_{ij}| g_{ijl},$$

where

$$(2.20) \quad g_{ijl} := g_{ij}(u_i + u_{ijl}^n, u_j + u_{jil}^n).$$

The following is a natural property of explicit schemes:

**Definition 2.12.** We say that the scheme defined in Definition 2.11 has a finite stencil if there is a constant  $K \in \mathbb{N}$  such that the update  $u_i^{n+1}$  is only affected by the values  $\{u_j \mid |x_j - x_i| < Kh\}$ .

**Theorem 2.13.** Let  $(\mathcal{T}_h := \{T_i \mid i \in I\})_{h>0}$  be a family of  $B$ -triangulations, where  $h := \sup_{i \in I} \text{diam}(T_i)$ . Suppose that Assumptions 2.3, 2.4 and 2.6 and the quadrature rule (2.10) hold. Let  $u_0 \in L^\infty(\mathbb{R}^2)$  with compact support and define  $u_i^0$  by (2.14). Assume that  $u_i^k$  are already defined for  $k \leq n \in \mathbb{N}$  and  $i \in I$ . Moreover assume that for  $i \in I, j \in N_i, l = 1, \dots, m$  the values  $u_{ijl}^n \in \mathbb{R}$  are given and satisfy (2.15)–(2.18) for a given  $\alpha \in ]0, 1]$ . Define  $u_i^{n+1}$  by (2.19)–(2.20) and suppose that the thus defined scheme has a finite stencil in the sense of Definition 2.12. Suppose furthermore that the CFL-condition

$$(2.21) \quad \frac{\Delta t}{h} \leq \frac{1}{c_V C_g(M) \sum_{j \in N_i} \sum_{l=1}^m \omega_l [1 + \sum_{p \in N_i} (\alpha_{ijlp} + \beta_{ijlp})]}$$

is satisfied. Let

$$(2.22) \quad \bar{u}_h(x, t) := u_i^n \quad \text{for } x \in T_i \text{ and } t^n \leq t < t^{n+1}.$$

Then for any given  $T > 0$ ,  $\bar{u}_h \in L^1 \cap L^\infty(\mathbb{R}^2 \times [0, T])$ , and as  $h \rightarrow 0$  there is a subsequence of  $\bar{u}_h$  which converges to the Kruzkov solution strongly in  $L^1_{\text{loc}}(\mathbb{R}^2 \times [0, T])$ .



**Note 2.14.** (i) For general initial data  $u_0 \in L^1 \cap L^\infty(\mathbb{R}^2)$  the statement of Theorem 2.13 remains true (see [23]).

(ii) In [24], the compatibility condition (2.9) was only required in the second argument of  $G$  and  $g$ , because for the first order scheme the numerical flux  $g_{ij}(u_i, u_j)$  depends in the first argument only on  $u_i$ . Since higher order schemes are defined using  $g_{ij}(u_i + u_{ij}^n, u_i + u_{ji}^n)$ , one also needs a compatibility condition in the first argument.

(iii) Assumptions (2.15)–(2.17) are modified from assumption (2.15b) in [6]. They are analogous to a one-dimensional flux-limiter for TVBM schemes (compare for example [4]). Condition (2.18) is an additional strengthening of (2.16), (2.17) which we will need in order to prove that the Young measure established in Sect. 4 is an admissible measure valued solution (compare Theorem 5.1).

(iv) In practice,  $C_1 = C_1(M_1)$  (resp.  $C_2 = C_2(M_2)$ ) where  $M_1$  (resp.  $M_2$ ) is an  $L^\infty$ -bound for the first (resp. second) derivative of  $u_0$  in the regions where  $u_0$  is smooth (compare Sect. 7).

(v) In a recent paper Geiben [17] has removed the requirement of a  $B$ -triangulation and she has developed new high order upwind finite volume schemes following the lines of Theorem 2.13. In the case of smooth solutions of scalar conservation laws the experimental order of convergence was approximately  $h^2$ . These schemes have also been applied to systems, and were combined with local mesh refinement and coarsening in order to resolve discontinuities sharply.

(vi) In [30], Theorem 2.13 could be generalized to irregular grids made of arbitrary polygons in  $\mathbb{R}^d$ ,  $d \geq 1$ , where the condition (2.1) is replaced by

$$\lim_{h \rightarrow 0} h^{2\alpha} / \rho = 0$$

where  $\rho$  is the minimum of all inner diameters of all polygons. Therefore thin polygons, which may become flat in the limit  $h \rightarrow 0$ , are allowed. In [1] numerical experiments indicate that thin triangles which are aligned with the discontinuities of the solution (shocks etc.), improve the resolution.

In the next section, we review the Kruzkov admissible weak solution and DiPerna’s theory of measure-valued solution. In Sect. 4, we prove an  $L^\infty$  bound for  $\bar{u}_h$ , following [6]. Section 5 contains the key entropy inequality, and the convergence proof is finished in Sect. 6. In Sect. 7, we apply the result to the discontinuous Galerkin finite element method.

### 3. Entropy inequalities, the Kruzkov solution, and admissible measure valued solutions

In this section, we briefly review some definitions and results on which our convergence proof is built.

**Definition 3.1.** (i) An entropy is a function  $U : \mathbb{R} \rightarrow \mathbb{R}$  which is Lipschitz and convex. Let  $F : \mathbb{R} \rightarrow \mathbb{R}^2$  be defined by

$$F(s) := \int^s U' f'.$$

The pair  $(U, F)$  is called an entropy-entropy-flux pair or briefly an entropy pair for the conservation law (1.1).

(ii) For any  $k \in \mathbb{R}$ , let  $U(s, k) := |s - k|$ .  $U(\cdot, k)$  is called the Kruzkov entropy, with entropy flux  $F(\cdot, k)$ .

(iii) Let  $u_0 \in L^1 \cap L^\infty(\mathbb{R}^2)$  and  $u \in L^1 \cap L^\infty(\mathbb{R}^2 \times [0, T])$ . If for all  $k \in \mathbb{R}$

$$\partial_t U(u, k) + \nabla \cdot F(u, k) \leq 0$$

in the sense of distributions, and if for all  $R > 0$ ,

$$\lim_{t \downarrow 0} \int_{|x| < R} |u(x, t) - u_0(x)| dx = 0,$$

then  $u$  is called the admissible weak solution of (1.1), (1.2).

*Remark 3.2.* In [25], Kruzkov shows that given  $u_0 \in L^1 \cap L^\infty(\mathbb{R}^2)$ , there exists exactly one admissible weak solution  $u$  of (1.1), (1.2). We will refer to it as the Kruzkov solution.

**Definition 3.3.** A Young measure  $\nu$  is a measurable map

$$\begin{aligned} \nu : \mathbb{R}^m &\rightarrow \text{Prob}(\mathbb{R}^n), \\ y &\mapsto \nu_y. \end{aligned}$$

The starting point for DiPerna's results (Theorems 3.8 and 3.10) is the following theorem of Tartar [37, 38]:

**Theorem 3.4.** Let  $(u_h)_{h>0}$  denote a family of functions mapping  $\mathbb{R}^m \rightarrow \mathbb{R}^n$  that is bounded in  $L^\infty(\mathbb{R}^m)$  with  $\|u_h\|_{L^\infty} \leq M$ . Then there is a subsequence  $u_h$  and a Young measure  $\nu$  supported in a ball of radius  $M$  such that for all continuous  $g$ , the weak-star limit of  $g(u_h)$  exists and

$$w^* - \lim_{h \rightarrow 0} g(u_h) = \langle \nu, g \rangle := \int_{\mathbb{R}^n} g(\lambda) d\nu(\lambda).$$

**Corollary 3.5.** (strong convergence) The sequence  $u_h$  converges to  $u$  strongly in  $L^1_{\text{loc}}$  if and only if the Young measure  $\nu$  reduces at almost all points  $y$  to the Dirac measure concentrated at  $u(y)$ .

The proof of this corollary can be found in [13].

**Definition 3.6.** A Young measure  $\nu : \mathbb{R}^2 \times [0, T] \rightarrow \mathbb{R}$  is called an admissible measure valued solution of (1.1) if there exists a constant  $M > 0$  such that for all  $(x, t)$  the support of  $\nu_{x,t}$  is contained in  $\{\lambda : |\lambda| \leq M\}$ , and if for all Kruzkov entropy pairs,

$$(3.1) \quad \partial_t \langle \nu, U(id, k) \rangle + \nabla \cdot \langle \nu, F(id, k) \rangle \leq 0$$

in the sense of distributions.

*Remark 3.7.* (i) Choosing  $k = \pm M$  in (3.1) one derives

$$\partial_t \langle \nu, id \rangle + \nabla \cdot \langle \nu, f \rangle = 0$$

(ii) Requiring (3.1) only for the Kruzkov entropies is equivalent to requiring it for all entropies (compare [13], p. 239).

The most important tool for proving the convergence will be the following result of DiPerna ([13] Theorems 4.1 and 4.2):

**Theorem 3.8.** *Let  $u_0 \in L^1 \cap L^\infty(\mathbb{R}^2)$  and let  $\nu$  be an admissible measure valued solution of (1.1). Assume that*

(i) *there exists a  $C > 0$  such that for almost every  $t \in [0, T]$ ,*

$$\int_{\mathbb{R}^2} \langle \nu_{x,t}, |id| \rangle dx \leq C$$

(ii)

$$\lim_{t \downarrow 0} \frac{1}{t} \int_0^t \int_{\mathbb{R}^2} \langle \nu_{x,s}, |id - u_0(x)| \rangle dx ds = 0.$$

*Then the Young measure  $\nu$  reduces to a Dirac measure centered at the Kruzkov solution  $u$  of (1.1), (1.2), i.e.*

$$\nu_{x,t} = \delta_{u(x,t)}, \quad a.e. (x, t) \in \mathbb{R}^2 \times [0, T].$$

*Remark 3.9.* This result assumes the existence of Kruzkov’s solution and establishes sufficient conditions which guarantee that an approximating sequence contains a subsequence which converges to the Kruzkov solution.

The following theorem gives a useful sufficient condition for property (ii) of Theorem 3.8:

**Theorem 3.10.** *Assume that  $u_0 \in L^1 \cap L^\infty(\mathbb{R}^2)$  and that  $\nu$  is an admissible measure valued solution of (1.1). Furthermore suppose that condition (i) of Theorem 3.8 is valid and that*

$$(3.2) \quad \lim_{t \downarrow 0} \frac{1}{t} \int_0^t \int_{\mathbb{R}^2} \langle \nu_{x,s}, id \rangle \phi(x) dx ds = \int_{\mathbb{R}^2} u_0(x) \phi(x) dx$$

*for all  $\phi \in C_0^1(\mathbb{R}^2)$ . If in addition*

$$(3.3) \quad \lim_{t \downarrow 0} \frac{1}{t} \int_0^t \int_{\mathbb{R}^2} \langle \nu_{x,s}, U \rangle dx ds \leq \int_{\mathbb{R}^2} U(u_0(x)) dx$$

*holds for one strictly convex continuous function  $U : \mathbb{R} \rightarrow \mathbb{R}$  with  $U(0) = 0$ , then  $\nu$  satisfies condition (ii) of Theorem 3.8.*

The main ideas of the proof of this theorem can be found in Sect. 6 of [13].

#### 4. $L^\infty$ bound

In this section we give an  $L^\infty$ -bound for the piecewise constant function  $\bar{u}_h$  defined in (2.22).

**Theorem 4.1.** *Suppose that the conditions of Theorem 2.13 and  $\|u_0\|_{L^\infty(\mathbb{R}^2)} \leq \frac{1}{2}M$  hold. Let  $T > 0$  be given. Then the following holds:*

*a) There are constants  $\hat{C} = \hat{C}(M) > 0$  and  $h_0 = h_0(T) > 0$  such that if the CFL-condition (2.21) holds, then for all  $h \leq h_0$  and for all  $n$  with  $n\Delta t \leq T$ , we have*

$$(4.1) \quad \sup_{i \in \mathcal{I}} |u_i^n| \leq \frac{1}{2}M + \hat{C}n\Delta th^{2\alpha-1} \leq M.$$

*b) If the family of triangulations is  $B$ -uniform as  $h \rightarrow 0$ , then (4.1) holds under the CFL-condition*

$$(4.2) \quad \frac{\Delta t}{h} \leq \frac{1}{c_V C_g(M) 3(1+4\mu)}.$$

Here  $c_V$  is defined in (2.1),  $C_g(M)$  in (2.3) and  $\mu$  in Definition 2.9.b.

Since the proof is essentially due to [6], we will skip some elementary calculations.

*Proof.* The conditions (2.16) and (2.17) imply that there are constants

$$(4.3) \quad 0 \leq \hat{\alpha}_{ijlp} \leq \alpha_{ijlp}$$

and

$$(4.4) \quad 0 \leq \hat{\beta}_{ijlp} \leq \beta_{ijlp}$$

such that

$$(4.5) \quad |u_{ijl}^n + \hat{a}_{ijl}| \leq C_2 h^{2\alpha}$$

$$(4.6) \quad |u_{jil}^n - \hat{b}_{ijl}| \leq C_2 h^{2\alpha}$$

where

$$(4.7) \quad \hat{a}_{ijl} := \sum_{p \in N_i} \hat{\alpha}_{ijlp} \Delta_{ip}$$

$$(4.8) \quad \hat{b}_{ijl} := \sum_{p \in N_i} \hat{\beta}_{ijlp} \Delta_{ip}$$

From (4.5)–(4.8) and (2.15)–(2.17) we know that

$$(4.9) \quad \hat{a}_{ijl} = O(h^\alpha), \hat{b}_{ijl} = O(h^\alpha).$$

Let

$$(4.10) \quad \hat{g}_{ijl} := g_{ij}(u_i - \hat{a}_{ijl}, u_j + \hat{b}_{ijl}).$$

Then

$$(4.11) \quad |\hat{g}_{ijl} - g_{ijl}| \leq C_g(M)(|\hat{a}_{ijl} + u_{ijl}^n| + |\hat{b}_{ijl} - u_{jil}^n|) = O(h^{2\alpha}).$$

Let  $\hat{\alpha}_{ijl}, \hat{\beta}_{ijl}$  be as in (4.3)–(4.8). Define

$$\hat{u}_i^{n+1} := H(u_i, u_j, j \in N_i) := u_i^n - \frac{\Delta t}{|T_i|} \sum_{j \in N_i} \sum_{l=1}^m \omega_l |S_{ij}| \hat{g}_{ijl}.$$

Now (4.3)–(4.11) and (2.3) imply that there is a constant  $C = C(M)$  independent of  $n$  such that

$$|u_i^{n+1} - \hat{u}_i^{n+1}| \leq Ch^{2\alpha}.$$

Now let  $\hat{C} \geq C/c_1 \geq Ch/\Delta t$ , where  $c_1$  is given in (2.1). Then

$$(4.12) \quad \sup_{i \in I} |u_i^{n+1}| \leq \sup_{i \in I} |\hat{u}_i^{n+1}| + Ch^{2\alpha} \leq \sup_{i \in I} |\hat{u}_i^{n+1}| + \hat{C} \Delta t h^{2\alpha-1}.$$

Choose  $h_0$  so small that  $T\hat{C}h_0^{2\alpha-1} \leq \frac{1}{2}M$ . Inequality (4.12) implies that in order to prove a), it is sufficient to show that under the CFL-condition (2.21),  $\hat{u}_i^{n+1}$  is a convex combination of  $u_i, u_j, j \in N_i$ .

Let us fix  $i, j$  and  $l$ . If  $g_{ij}$  is  $C^1$ , then there are constants  $\xi \in I(u_i - \hat{\alpha}_{ijl}, u_i)$ ,  $\eta \in I(u_j + \hat{\beta}_{ijl}, u_j)$  and  $\zeta \in I(u_i, u_j)$  such that

$$\begin{aligned} \hat{g}_{ijl} - f(u_i) \cdot n_{ij} &= \left[ \partial_u g(\xi, \mu_j + \hat{b}_{ijl}) \sum_{p \in N_i} \hat{\alpha}_{ijlp} - \partial_v g(u_i, \eta) \sum_{p \in N_i} \hat{\beta}_{ijlp} - \partial_v g(u_i, \zeta) \right] u_i \\ &+ \sum_{p \in N_i} \left[ -\partial_u g(\xi, \mu_j + \hat{b}_{ijl}) \hat{\alpha}_{ijlp} + \partial_v g(u_i, \eta) \hat{\beta}_{ijlp} \right] u_p + \partial_v g(u_i, \zeta) u_j. \end{aligned}$$

Let  $\delta_{jp}$  be the Kronecker symbol. Then

$$\begin{aligned} (4.13) \quad \hat{u}_i^{n+1} &= \left\{ 1 - \frac{\Delta t}{|T_i|} \sum_{j \in N_i} \sum_{l=1}^m \omega_l |S_{ij}| \right. \\ &\left. \left[ \partial_u g(\xi, \mu_j + \hat{b}_{ijl}) \sum_{p \in N_i} \hat{\alpha}_{ijlp} - \partial_v g(u_i, \eta) \sum_{p \in N_i} \hat{\beta}_{ijlp} - \partial_v g(u_i, \zeta) \right] \right\} u_i \\ &+ \frac{\Delta t}{|T_i|} \sum_{p \in N_i} \sum_{j \in N_i} \sum_{l=1}^m \omega_l |S_{ij}| \left[ \partial_u g(\xi, \mu_j + \hat{b}_{ijl}) \hat{\alpha}_{ijlp} \right. \\ &\quad \left. - \partial_v g(u_i, \eta) \hat{\beta}_{ijlp} - \partial_v g(u_i, \zeta) \delta_{jp} \right] u_p \\ &=: \gamma_0 u_i + \sum_{p \in N_i} \gamma_p u_p. \end{aligned}$$

Now,  $\gamma_0 + \sum_{p \in N_i} \gamma_p = 1$ , and for  $p \in N_i$ ,  $\gamma_p \geq 0$ . Finally,  $\gamma_0 \geq 0$  follows from (2.21), (4.3) and (4.4), the assumption (2.1) on the triangulation and the monotonicity of  $g$  (see (2.5)). If  $g_{ij}$  is not  $C^1$  replace the derivatives by difference quotients. This proves a). Part b) of the theorem follows directly from a).  $\square$

Using the  $L^\infty$  estimate (4.1) together with Tartar’s theorem, we know that a subsequence of  $\bar{u}_h$  converges weak-star to a Young measure  $\nu$ . In the next section, we show that  $\nu$  is an admissible measure valued solution of (1.1).

## 5. Existence of an admissible measure valued solution

In this section we prove existence of an admissible measure valued solution. First, we prove a cell entropy inequality. This is the main part in our convergence proof. The key tools for obtaining this entropy inequality are (2.9) and (2.18).

**Theorem 5.1.** *Let  $(U, F)$  be a convex entropy pair with  $U \in C^2$  and assume that  $G_{ij}$  satisfies Assumption 2.6. Suppose that the CFL-condition (2.21) holds. Then there are constants  $C = C(C_1, C_2, M, C_g(M), c_V, \frac{\Delta t}{h})$  and  $h_0 > 0$  such that for  $h \leq h_0$*

$$(5.1) \quad U(u_i^{n+1}) - U(u_i) + \frac{\Delta t}{|T_i|} \sum_{j \in N_i} \sum_{l=1}^m \omega_l |S_{ij}| G_{ijl} \leq C \|U''\|_{L^\infty(B_M)} h^{2\alpha},$$

where

$$G_{ijl} := G_{ij}(u_i + u_{ijl}^n, u_j + u_{jil}^n).$$

*Remark 5.2.* a) The proof of this theorem is the central piece of this work. It relies on a careful analysis of the entropy-dissipation and on the properties of the flux-limiters (2.15)–(2.18).

b) A cell entropy inequality similar to (5.1) was derived as early as 1971 by Lax [28] for the first-order Lax-Friedrichs finite difference scheme. Since then, many authors have applied and refined these ideas (see, e.g. [36, 10, 35, 34, 24] and the references therein).

c) Further comments on Theorem 5.1 may be found at the end of the paper in “Added in proof”.

*Proof of Theorem 5.1.* Note that

$$u_i^{n+1} = \frac{1}{3} \sum_{j \in N_i} \sum_{l=1}^m \omega_l [u_i - 3\lambda_{ij}(g_{ijl} - f(u_i) \cdot n_{ij})],$$

where  $\lambda_{ij} := \frac{\Delta t |S_{ij}|}{|T_i|}$ . Since  $U$  is convex,  $\omega_l \geq 0$  and  $\sum_{l=1}^m \omega_l = 1$ ,

$$U(u_i^{n+1}) \leq \frac{1}{3} \sum_{j \in N_i} \sum_{l=1}^m \omega_l U(u_i - 3\lambda_{ij}(g_{ijl} - f(u_i) \cdot n_{ij})).$$

It is thus sufficient to show that for each  $(i, j, l)$ ,

$$\begin{aligned} E &:= U(u_i - 3\lambda_{ij}(g_{ijl} - f(u_i) \cdot n_{ij})) - U(u_i) + 3\lambda_{ij}(G_{ijl} - F(u_i) \cdot n_{ij}) \\ &\leq C \|U''\|_{L^\infty(B_M)} h^{2\alpha}. \end{aligned}$$

From now on let  $(i, j, l)$  be fixed and  $\lambda := \lambda_{ij}$ ,  $u := u_i$ ,  $v := u_j$ ,  $\tilde{u} := u_{ijl}^n$ ,  $\tilde{v} := u_{jil}^n$ ,  $\gamma_0 := f(u) \cdot n_{ij}$  and  $\Gamma_0 := F(u) \cdot n_{ij}$ . For  $0 \leq \tau \leq 1$  let  $\gamma(\tau)$  (respectively  $\gamma_1(\tau), \gamma_2(\tau), \Gamma(\tau), \Gamma_1(\tau), \Gamma_2(\tau)$ ) be the function  $g_{ij}$  (respectively  $\partial_u g_{ij}, \partial_v g_{ij}, G_{ij}, \partial_u G_{ij}, \partial_v G_{ij}$ ) evaluated at  $(u + \tau \tilde{u}, u + \tau(v - u + \tilde{v}))$ . Note that  $\gamma_0 = \gamma(0)$ ,  $\Gamma_0 = \Gamma(0)$ . Let

$$b(\tau) := u - 3\lambda(\gamma(\tau) - \gamma_0).$$

Now

$$E = U(b(1)) - U(u) + 3\lambda(\Gamma(1) - \Gamma_0) = \int_0^1 [U'(b(\tau))b'(\tau) + 3\lambda\Gamma'(\tau)]d\tau$$

From the compatibility conditions (2.9), we have

$$\begin{aligned} \Gamma'(\tau) &= \Gamma_1(\tau)\tilde{u} + \Gamma_2(\tau)(v - u + \tilde{v}) \\ &= U'(u + \tau\tilde{u})\gamma_1(\tau)\tilde{u} + U'(u + \tau(v - u + \tilde{v}))\gamma_2(\tau)(v - u + \tilde{v}). \end{aligned}$$

Therefore,  $E = E_1 + E_2$ , where

$$\begin{aligned} E_1 &:= 3\lambda\tilde{u} \int_0^1 [U'(u + \tau\tilde{u}) - U'(b(\tau))]\gamma_1(\tau)d\tau \\ E_2 &:= 3\lambda(v - u + \tilde{v}) \int_0^1 [U'(u + \tau(v - u + \tilde{v})) - U'(b(\tau))]\gamma_2(\tau)d\tau \end{aligned}$$

We first treat  $E_1$ . Since  $U \in C^2$ , it is possible to choose  $\xi(\tau) \in I(b(\tau), u + \tau\tilde{u})$  such that

$$U'(u + \tau\tilde{u}) - U'(b(\tau)) = U''(\xi(\tau))(u + \tau\tilde{u} - b(\tau)).$$

Let

$$\begin{aligned} \gamma_1^*(\tau) &:= \begin{cases} \frac{g_{ij}(u+\tau\tilde{u},u)-g_{ij}(u,u)}{\tau\tilde{u}} & \text{if } \tau\tilde{u} \neq 0 \\ 0 & \text{otherwise} \end{cases}, \\ \gamma_2^*(\tau) &:= \begin{cases} \frac{g_{ij}(u+\tau\tilde{u},u+\tau(v-u+\tilde{v}))-g_{ij}(u+\tau\tilde{u},u)}{\tau(v-u+\tilde{v})} & \text{if } \tau(v-u+\tilde{v}) \neq 0 \\ 0 & \text{otherwise} \end{cases}. \end{aligned}$$

Then

$$\gamma(\tau) - \gamma_0 = \tau[\gamma_1^*(\tau)\tilde{u} + \gamma_2^*(\tau)(v - u + \tilde{v})]$$

and

$$E_1 = 3\lambda\tilde{u} \int_0^1 \tau U''(\xi(\tau))[(1 + 3\lambda\gamma_1^*(\tau))\tilde{u} + 3\lambda\gamma_2^*(\tau)(v - u + \tilde{v})]\gamma_1(\tau)d\tau.$$

From (2.3),  $|\gamma_1(\tau)|, |\gamma_2(\tau)|, |\gamma_1^*(\tau)|, |\gamma_2^*(\tau)| \leq C_g(M)$ . From (2.1) and (2.21),  $3\lambda C_g(M) \leq 1$ . From (2.15),  $|\tilde{u}|, |\tilde{v}| \leq C_1 h^\alpha$ . From (2.18),  $(v - u)\tilde{u} \geq -C_2 h^{2\alpha}$ , and from the monotonicity of  $g_{ij}$  (2.5),  $\gamma_1(\tau)\gamma_2^*(\tau) \leq 0$ . Therefore,

$$\lambda^2 \gamma_1(\tau)\gamma_2^*(\tau)(v - u)\tilde{u} \leq C_2 h^{2\alpha},$$

which yields

$$E_1 \leq C \|U''\|_{L^\infty(B_M)} h^{2\alpha}.$$

The term  $E_2$  can be treated similarly.  $\square$

**Proposition 5.3.** *Let  $\nu$  be the Young measure established in Sect. 4. Then for any convex entropy pair  $(U, F)$  with  $U \in C^2$ ,*

$$(5.2) \quad \partial_t \langle \nu, U \rangle + \nabla \cdot \langle \nu, F \rangle \leq 0$$

*in the sense of distributions. Moreover, for all  $\varphi \in C_0^\infty(\mathbb{R}^2 \times [0, T])$  with  $\varphi \geq 0$ ,*

$$(5.3) \quad \int_0^T \int_{\mathbb{R}^2} [\langle \nu_{x,t}, U \rangle \partial_t \varphi(x, t) + \langle \nu_{x,t}, F \rangle \cdot \nabla \varphi(x, t)] dx dt + \int_{\mathbb{R}^2} U(u_0(x)) \varphi(x, 0) dx \geq 0.$$

In order to prove this proposition, we need the following

**Lemma 5.4.** *Let  $\varphi \in C_0^\infty(\mathbb{R}^2 \times [0, T])$ ,  $\varphi_i^n := \varphi(x_i, t^n)$ ,  $\varphi_{ijl}^n := \varphi(x_{ijl}, t^n)$ . Then there are constants  $C = C(\varphi, T)$  and  $h_0 > 0$  such that for  $h \leq h_0$*

$$(5.4) \quad \Delta t \sum_{n=0}^N \sum_{i \in I} \sum_{j \in N_i} \sum_{l=1}^m \omega_l |S_{ijl}| |G_{ijl} - F(u_i) \cdot n_{ij}| |\varphi_i^n - \varphi_{ijl}^n| \leq Ch^{\frac{\alpha}{2}}.$$

This lemma will be proved at the end of this section.

*Proof of Proposition 5.3.* Let  $B_R$  be the ball of radius  $R$  in  $\mathbb{R}^2$ , and let  $\chi_R$  be the indicator function of  $B_R$ . Let  $\varphi \in C_0^\infty(B_R \times [0, T])$ ,  $\varphi \geq 0$  and  $\varphi_i^n := \varphi(x_i, t^n)$ . Multiply (5.1) by  $|T_i| \varphi_i^n$  and sum over  $i$  and  $n$ ,  $n \leq N$ , where  $N \Delta t = T$ .

$$\begin{aligned} & \sum_{n=0}^N \sum_{i \in I} |T_i| [U(u_i^{n+1}) - U(u_i^n)] \varphi_i^n + \Delta t \sum_{n=0}^N \sum_{i \in I} \sum_{j \in N_i} \sum_{l=1}^m \omega_l |S_{ijl}| |G_{ijl}| \varphi_i^n \\ & \leq C \sum_{n=0}^N \sum_{i \in I} |T_i| \varphi_i^n h^{2\alpha}. \end{aligned}$$

We have

$$\begin{aligned} & \sum_{n=0}^N \sum_{i \in I} |T_i| [U(u_i^{n+1}) - U(u_i^n)] \varphi_i^n \\ & = -\Delta t \sum_{n=1}^N \sum_{i \in I} |T_i| U(u_i^n) \frac{\varphi_i^n - \varphi_i^{n-1}}{\Delta t} - \sum_{i \in I} |T_i| U(u_i^0) \varphi_i^0 \\ & \geq -\Delta t \sum_{n=1}^N \sum_{i \in I} \int_{T_i} U(\bar{u}_h(x, t^n)) [\partial_t \varphi(x, t^n) + O(\Delta t) \chi_R(x)] dx \\ & \quad - \sum_{i \in I} \int_{T_i} U(u_0(x)) [\varphi(x, 0) + O(\Delta t) \chi_R(x)] dx \\ & = -\Delta t \sum_{n=1}^N \int_{\mathbb{R}^2} U(\bar{u}_h(x, t^n)) \partial_t \varphi(x, t) dx - \int_{\mathbb{R}^2} U(\bar{u}_h(x, 0)) \varphi(x, 0) dx + O(\Delta t) \\ & = -\int_{\Delta t}^T \int_{\mathbb{R}^2} U(\bar{u}_h(x, t)) \partial_t \varphi(x, t) dx dt - \int_{\mathbb{R}^2} U(\bar{u}_h(x, 0)) \varphi(x, 0) dx + O(\Delta t). \end{aligned}$$



Since  $\varphi$  is supposed to have compact support and since  $G_{ijl}(u, v) = -G_{jil}(v, u)$ , we have

$$\sum_{n=0}^N \sum_{i \in I} \sum_{j \in N_i} \sum_{l=1}^m \omega_l |S_{ij}| G_{ijl} \varphi_{ijl}^n = 0$$

and therefore

$$\begin{aligned} \Delta t \sum_{n=0}^N \sum_{i \in I} \sum_{j \in N_i} \sum_{l=1}^m \omega_l |S_{ij}| G_{ijl} \varphi_i^n &= \Delta t \sum_{n=0}^N \sum_{i \in I} \sum_{j \in N_i} \sum_{l=1}^m \omega_l |S_{ij}| G_{ijl} (\varphi_i^n - \varphi_{ijl}^n) \\ &= \Delta t \sum_{n=0}^N \sum_{i \in I} \sum_{j \in N_i} \sum_{l=1}^m \omega_l |S_{ij}| [G_{ijl} - F(u_i^n) \cdot n_{ij}] (\varphi_i^n - \varphi_{ijl}^n) \\ &\quad - \Delta t \sum_{n=0}^N \sum_{i \in I} \sum_{j \in N_i} \left( |S_{ij}| F(u_i^n) \cdot n_{ij} \sum_{l=1}^m \omega_l \varphi_{ijl}^n \right). \end{aligned}$$

From (5.4), the first summand on the RHS is  $O(h^{\frac{\alpha}{2}})$ . The second summand is

$$\begin{aligned} & - \Delta t \sum_{n=0}^N \sum_{i \in I} \sum_{j \in N_i} \left( \int_{S_{ij}} F(\bar{u}_h) \cdot n \varphi + O(h^3) \chi_R(x_i) \right) \\ &= - \Delta t \sum_{n=0}^N \sum_{i \in I} \int_{T_i} F(\bar{u}_h) \cdot \nabla \varphi + TR^2 O(h) = - \int_0^T \int_{\mathbb{R}^2} F(\bar{u}_h) \cdot \nabla \varphi + O(h) \end{aligned}$$

Finally,

$$h^{2\alpha} \sum_{n=0}^N \sum_{i \in I} |T_i| \varphi_i^n \leq Ch^{2\alpha-1} TR^2.$$

Therefore,

$$\begin{aligned} & - \int_{\Delta t}^T \int_{\mathbb{R}^2} U(\bar{u}_h) \partial_t \varphi - \int_0^T \int_{\mathbb{R}^2} F(\bar{u}_h) \cdot \nabla \varphi - \int_{\mathbb{R}^2} U(\bar{u}_h(0)) \varphi(\cdot, 0) \\ & \leq O(h^{\alpha/2}) + O(h^{2\alpha-1}) + O(h) = O(h^\varepsilon) \text{ for some } \varepsilon > 0 \end{aligned}$$

if  $\alpha > \frac{1}{2}$ . Taking the limit as  $h \rightarrow 0$  yields (5.3). For  $\varphi \in C_0^\infty(\mathbb{R}^2 \times (0, T))$ , this implies

$$\partial_t \langle \nu, U \rangle + \nabla \cdot \langle \nu, F \rangle \leq 0$$

in the sense of distributions.  $\square$

**Corollary 5.5.** *The Young measure  $\nu$  established in Sect. 4 is an admissible measure valued solution of (1.1). Moreover,*

$$\begin{aligned} & \int_0^T \int_{\mathbb{R}^2} [\langle \nu_{x,t}, id \rangle \partial_t \varphi(x, t) + \langle \nu_{x,t}, f \rangle \cdot \nabla \varphi(x, t)] dx dt \\ (5.5) \quad & + \int_{\mathbb{R}^2} \langle \nu_{x,0}, id \rangle \varphi(x, 0) dx = 0. \end{aligned}$$

*Proof.* We already know from Tartar’s theorem that  $\text{supp } \nu_{x,t} \subset B_M$  for all  $(x,t) \in \mathbb{R}^2 \times [0,T]$ . It remains to show (3.1) for all Kruzkov entropies  $U(\cdot, k) = |\cdot - k|$ . Given  $k \in \mathbb{R}$  and  $\varepsilon > 0$ , let

$$U_\varepsilon(s) := \begin{cases} -\frac{\varepsilon}{8} [((s - k)/\varepsilon)^4 - 6((s - k)/\varepsilon)^2 - 3] & \text{for } |s - k| < \varepsilon \\ |s - k| & \text{otherwise} \end{cases}$$

$U_\varepsilon \in C^2$ , and as  $\varepsilon \rightarrow 0$ ,  $U_\varepsilon \rightarrow U(\cdot, k)$  pointwise uniformly. Using the dominated convergence theorem one can pass to the limit  $\varepsilon \rightarrow 0$  in (5.3) to obtain

$$\int_0^T \int_{\mathbb{R}^2} \langle \nu_{x,t}, U(id, k) \rangle \partial_t \varphi(x, t) + \langle \nu_{x,t}, F(id, k) \rangle \cdot \nabla \varphi(x, t) + \int_{\mathbb{R}^2} \langle \nu_{x,0}, U(id, k) \rangle \varphi(x, 0) \geq 0.$$

From here, one immediately obtains (3.1). Moreover, setting  $k = \pm M$ , one obtains (5.5).  $\square$

It remains to prove Lemma 5.4.

*Proof of Lemma 5.4.* As before, let  $\lambda_{ij} := \frac{\Delta t |S_{ij}|}{|T_i|}$ . In ([24], Lemma 4.7), it is shown for the first order scheme

$$u_i^{n+1} = u_i^n - \sum_{j \in N_i} \lambda_{ij} g_{ij}(u_i, u_j)$$

that for  $G_{ij}$  corresponding to the entropy  $U(s) = \frac{1}{2}s^2$ ,

$$(5.6) \quad \frac{1}{2} [(u_i^{n+1})^2 - (u_i^n)^2] + \sum_{j \in N_i} \lambda_{ij} G_{ij}(u_i, u_j) + \sum_{j \in N_i} \lambda_{ij}^2 [g_{ij}(u_i, u_j) - f(u_i) \cdot n_{ij}]^2 \leq 0.$$

From (2.20), (2.15) and (2.3),

$$\sum_{j \in N_i} \sum_{l=1}^m \lambda_{ij} \omega_l [g_{ijl} - g_{ij}(u_i, u_j)] = O(h^\alpha).$$

It is now easy to modify the proof of ([24], Lemma 4.7) to show that for the scheme (2.19) one has the estimate

$$(5.7) \quad \frac{1}{2} [(u_i^{n+1})^2 - (u_i^n)^2] + \sum_{j \in N_i} \lambda_{ij} G_{ij}(u_i, u_j) + \sum_{j \in N_i} (\lambda_{ij})^2 [g_{ij}(u_i, u_j) - f(u_i) \cdot n_{ij}]^2 \leq O(h^\alpha).$$

Since the initial data  $u_0$  have compact support, and since the scheme has a finite stencil (see Definition 2.12), there is a constant  $R_0$  such that the support of  $\bar{u}_h(\cdot, t^n)$  is contained in a ball of radius  $R_n := R_0 + nhK$ . Because of Assumption

2.3 and since  $t^n \leq T$ ,  $R_n \leq R_0 + KT/c_1$ . Let  $I_n := \{i \in I : |x_i| \leq R_n\}$ . Multiplying (5.7) by  $|T_i|$  and summing over  $i \in I_n$  we obtain

$$\begin{aligned} & \frac{1}{2} \|\bar{u}_h(t^{n+1})\|_{L^2(\mathbb{R}^2)}^2 - \frac{1}{2} \|\bar{u}_h(t^n)\|_{L^2(\mathbb{R}^2)}^2 \\ & + \sum_{i \in I_n} \frac{\Delta t^2}{|T_i|} \sum_{j \in N_i} |S_{ij}|^2 (g_{ij}(u_i, u_j) - f(u_i) \cdot n_{ij})^2 \leq Ch^\alpha. \end{aligned}$$

Using that  $\frac{(\Delta t)^2}{|T_i|} \geq (c_1)^2 > 0$ , multiplying by  $\Delta t$  and summing over  $n = 0, \dots, N$  we obtain

$$(5.8) \quad \Delta t \sum_{n=0}^N \sum_{i \in I_n} \sum_{j \in N_i} |S_{ij}|^2 [g_{ij}(u_i, u_j) - f(u_i) \cdot n_{ij}]^2 \leq CTh^\alpha.$$

Now let  $U \in C^2$  be any convex entropy with corresponding numerical entropy flux  $G_{ij}$  satisfying Assumption 2.6. As in ([24], Prop.4.3), we use the compatibility condition (2.9) to show that

$$|G_{ij}(u_i, u_j) - F(u_i) \cdot n_{ij}| \leq \|U'\|_{L^\infty(B_M)} |g_{ij}(u_i, u_j) - f(u_i) \cdot n_{ij}|.$$

Applying this to (5.8), we obtain

$$(5.9) \quad \Delta t \sum_{n=0}^N \sum_{i \in I_n} \sum_{j \in N_i} |S_{ij}|^2 [G_{ij}(u_i, u_j) - F(u_i) \cdot n_{ij}]^2 \leq C(\|U'\|_{L^\infty(B_M)})Th^\alpha.$$

Since

$$[G_{ijl} - G_{ij}(u_i, u_j)] = O(h^\alpha),$$

we have that

$$\begin{aligned} & \Delta t \sum_{n=0}^N \sum_{i \in I_n} \sum_{j \in N_i} \sum_{l=1}^m \omega_l |S_{ij}|^2 [G_{ijl} - F(u_i) \cdot n_{ij}]^2 \\ & \leq 2\Delta t \sum_{n=0}^N \sum_{i \in I_n} \sum_{j \in N_i} |S_{ij}|^2 [G_{ij}(u_i, u_j) - F(u_i) \cdot n_{ij}]^2 + C \Delta t \sum_{n=0}^N \sum_{i \in I_n} h^{2+\alpha} \\ (5.10) \quad & \leq CTh^\alpha. \end{aligned}$$

Using (5.10) and Hölder's inequality we derive

$$\begin{aligned} & \Delta t \sum_{n=0}^N \sum_{i \in I} \sum_{j \in N_i} \sum_{l=1}^m \omega_l |S_{ij}| |G_{ijl} - F(u_i) \cdot n_{ij}| |\varphi_i^n - \varphi_{ijl}^n| \\ & \leq \left( \Delta t \sum_{n=0}^N \sum_{i \in I} \sum_{j \in N_i} \sum_{l=1}^m \omega_l |S_{ij}|^2 |G_{ijl} - F(u_i) \cdot n_{ij}|^2 \right)^{\frac{1}{2}} \\ & \quad \left( \Delta t \sum_{n=0}^N \sum_{i \in I} \sum_{j \in N_i} \sum_{l=1}^m \omega_l |\varphi_i^n - \varphi_{ijl}^n|^2 \right)^{\frac{1}{2}} \\ & \leq (CTh^\alpha)^{\frac{1}{2}} (C(\varphi)T)^{\frac{1}{2}} = C(\varphi, T)h^{\frac{\alpha}{2}} \quad \square \end{aligned}$$

### 6. Consistency with the initial conditions

In this section we show that the admissible measure valued solution  $\nu$  of (1.1) equals the Kruzkov solution of (1.1) with initial data (1.2). According to the theory outlined in Sect. 3, it remains to show that assumptions (i) and (ii) of DiPerna’s Theorem 3.8 are satisfied.

**Theorem 6.1.** *For all  $t \in [0, T]$ ,*

$$\int_{\mathbb{R}^2} \langle \nu_{x,t}, |id| \rangle dx \leq \|u_0\|_{L^1(\mathbb{R}^2)},$$

*i.e. condition (i) of Theorem 3.8 holds.*

*Proof.* We know from Tartar’s theorem, that

$$\bar{u}_h \xrightarrow{*} \langle \nu, id \rangle =: u$$

in  $L^\infty(\mathbb{R}^2 \times \mathbb{R}_+)$ . For almost every  $t$ , there is a compactly supported probability measure  $\mu_{\cdot,t}$  such that

$$\bar{u}_h(t) \xrightarrow{*} \langle \mu_{\cdot,t}, id \rangle =: v(\cdot, t).$$

Let  $\sigma \in C_0^\infty([0, T])$ ,  $\varphi \in C_0^\infty(\mathbb{R}^2)$  and  $\psi(x, t) := \sigma(t)\varphi(x)$ . Then

$$\lim_{h \rightarrow 0} \int_{\mathbb{R}^2} \bar{u}_h(x, t)\varphi(x)dx = \int_{\mathbb{R}^2} v(x, t)\varphi(x)dx$$

and therefore

$$\lim_{h \rightarrow 0} \int_0^T \int_{\mathbb{R}^2} \bar{u}_h(x, t)\psi(x, t)dxdt = \int_0^T \int_{\mathbb{R}^2} v(x, t)\psi(x, t)dxdt.$$

On the other hand,

$$\lim_{h \rightarrow 0} \int_0^T \int_{\mathbb{R}^2} \bar{u}_h(x, t)\psi(x, t)dxdt = \int_0^T \int_{\mathbb{R}^2} u(x, t)\psi(x, t)dxdt.$$

and therefore we obtain that  $v(x, t) = u(x, t)$  at all Lebesgue-points  $(x, t)$  of  $u$  and  $v$ . Therefore, we can identify  $v$  with  $u$  and  $\mu_{\cdot,t}$  with  $\nu_{\cdot,t}$ .

Next, for any convex  $U \in C^2$ , let  $(U, F)$  be an entropy-entropy-flux pair and  $G_{ij}$  a numerical entropy-flux consistent with  $F \cdot n_{ij}$  and satisfying Assumption 2.6. From (5.1),

$$U(u_i^{n+1}) - U(u_i^n) \leq -\frac{\Delta t}{|T_i|} \sum_{j \in N_i} \sum_{l=1}^m \omega_l |S_{ij}| G_{ijl} + \|U''\|_{L^\infty([-M, M])} O(h^{2\alpha}).$$

Multiply this by  $|T_i|$  and sum over  $n = 0, \dots, N - 1$  ( $N \Delta t = t$ ) and  $i \in I_N = \{i \in I : |x_i| \leq R_N\}$ . Since for all  $n \leq N$  the support of  $\bar{u}_h(\cdot, t^n)$  is contained in a ball of radius  $R_N$ ,

$$-\Delta t \sum_{n=0}^{N-1} \sum_{i \in I_N} \sum_{j \in N_i} \sum_{l=1}^m \omega_l |S_{ij}| G_{ijl} = 0,$$

and we obtain

$$(6.1) \quad \int_{\mathbb{H}^2} U(\bar{u}_h(x, t)) dx \leq \int_{\mathbb{H}^2} U(\bar{u}_h(x, 0)) dx + C \|U''\|_{L^\infty([-M, M])} h^{2\alpha-1}.$$

Now define

$$U_\varepsilon(s) := \begin{cases} -\frac{\varepsilon}{8}[(s/\varepsilon)^4 - 6(s/\varepsilon)^2 - 3] & \text{for } |s| < \varepsilon \\ |s| & \text{otherwise} \end{cases}$$

Let  $\varepsilon = \varepsilon(h) := h^{2\beta}$ , with  $\beta \in ]0, \alpha - \frac{1}{2}[$ . Note that  $\|U''_{\varepsilon(h)}\|_{L^\infty(B_M)} \leq \frac{C}{\varepsilon} = Ch^{-2\beta}$ . Therefore, (6.1) gives

$$(6.2) \quad \int_{\mathbb{H}^2} U_{\varepsilon(h)}(\bar{u}_h(x, t)) dx \leq \int_{\mathbb{H}^2} U_{\varepsilon(h)}(\bar{u}_h(x, 0)) dx + Ch^{2(\alpha-\beta-\frac{1}{2})}.$$

Using Tartar's theorem and (6.2) we derive

$$\begin{aligned} \int_{\mathbb{H}^2} \langle \nu_{x,t}, |id| \rangle dx &= \lim_{R \rightarrow \infty} \int_{B_R} \langle \nu_{x,t}, |id| \rangle dx \\ &= \lim_{R \rightarrow \infty} \lim_{h \rightarrow 0} \int_{B_R} |\bar{u}_h(x, t)| dx \leq \lim_{R \rightarrow \infty} \lim_{h \rightarrow 0} \int_{B_R} U_{\varepsilon(h)}(\bar{u}_h(x, t)) dx \\ &\leq \lim_{h \rightarrow 0} \int_{\mathbb{H}^2} U_{\varepsilon(h)}(\bar{u}_h(x, t)) dx \leq \lim_{h \rightarrow 0} \left[ \int_{\mathbb{H}^2} U_{\varepsilon(h)}(\bar{u}_h(x, 0)) dx + Ch^{2(\alpha-\beta-\frac{1}{2})} \right] \\ &= \lim_{h \rightarrow 0} \int_{\mathbb{H}^2} U_{\varepsilon(h)}(\bar{u}_h(x, 0)) dx \end{aligned}$$

For every  $h > 0$ ,

$$\int_{\mathbb{H}^2} U_{\varepsilon(h)}(\bar{u}_h(x, 0)) dx \leq \int_{\text{supp}(\bar{u}_h)} \frac{3\varepsilon}{8} dx + \int_{\mathbb{H}^2} |\bar{u}_h(x, 0)| dx \leq Ch^{2\beta} + \int_{\mathbb{H}^2} |u_0(x)| dx$$

so

$$\int_{\mathbb{H}^2} \langle \nu_{x,t}, |id| \rangle dx \leq \int_{\mathbb{H}^2} |u_0(x)| dx,$$

which is the  $L^1$ -bound we were looking for.  $\square$

Next we prove that assumption (ii) of Theorem 3.8 holds. For this purpose, it remains to show that assumptions (3.2) and (3.3) of Theorem 3.10 hold.

**Proposition 6.2.** *For all  $\phi \in C_0^\infty(\mathbb{H}^2)$*

$$(6.3) \quad \lim_{t \downarrow 0} \frac{1}{t} \int_0^t \int_{\mathbb{H}^2} \langle \nu_{x,s}, id \rangle \phi(x) dx ds = \int_{\mathbb{H}^2} u_0(x) \phi(x) dx$$

*i.e. condition (3.2) of Theorem 3.10 holds.*

*Proof.* The proof is taken from ([24], Prop.3.9.). Based on the equality (5.5), one can show that for all  $\phi \in C_0^\infty(\mathbb{R}^2)$  the map

$$t \rightarrow A(t) := \int_{\mathbb{R}^2} \langle \nu_{x,t}, id \rangle \phi(x) dx$$

is absolutely continuous over  $[0, T]$  and that

$$A(0) = \int_{\mathbb{R}^2} u_0(x) \phi(x) dx. \quad \square$$

**Proposition 6.3.** *Let  $U(u) = \frac{1}{2}u^2$ . Then*

$$(6.4) \quad \lim_{t \downarrow 0} \frac{1}{t} \int_0^t \int_{\mathbb{R}^2} \langle \nu_{x,s}, U \rangle dx ds \leq \int_{\mathbb{R}^2} U(u_0(x)) dx ,$$

*i.e. (3.3) holds.*

*Proof.* The proof is a generalization of ([24], Prop.3.13). As in the proof of Theorem 6.1, multiply (5.1) by  $|T_i|$  and sum over  $i \in I_N$ . Then

$$(6.5) \quad \sum_{i \in I_N} |T_i| [U(u_i^{n+1}) - U(u_i^n)] \leq Ch^{2\alpha}$$

Let  $0 \leq t_1 := n_1 \Delta t \leq t_2 := n_2 \Delta t$ . Sum (6.5) over  $0 \leq n' \leq n$ , and let  $t := n \Delta t \leq t_2$ . Then

$$(6.6) \quad \sum_{i \in I_N} |T_i| [U(u_i^n) - U(u_i^0)] \leq nCh^{2\alpha} \leq t_2 Ch^{2\alpha-1}.$$

Sum (6.6) over  $n = n_1, \dots, n_2 - 1$ :

$$\sum_{n=n_1}^{n_2-1} \sum_{i \in I_N} |T_i| U(u_i^n) \leq (n_2 - n_1) \sum_{i \in I_N} |T_i| U(u_i^0) + (n_2 - n_1) t_2 Ch^{2\alpha-1}.$$

Multiplying this by  $\frac{\Delta t}{t_2 - t_1}$  gives

$$\frac{1}{t_2 - t_1} \int_{t_1}^{t_2} \int_{\mathbb{R}^2} U(\bar{u}_h(x, t)) dx dt \leq \int_{\mathbb{R}^2} U(\bar{u}_h(x, 0)) dx + Ct_2 h^{2\alpha-1}$$

Let  $h \rightarrow 0$ . Tartar's theorem 3.4 implies that

$$\frac{1}{t_2 - t_1} \int_{t_1}^{t_2} \int_{\mathbb{R}^2} \langle \nu_{x,t}, U \rangle dx dt \leq \int_{\mathbb{R}^2} U(u_0(x)) dx.$$

The rest of the proof can be found in [24].  $\square$

*Proof of Theorem 2.13.* From Theorem 6.1, condition (i) of Theorem 3.8 holds. From Propositions 6.2 and 6.3, conditions (3.2) and (3.3) of Theorem 3.10 are satisfied. Therefore, we can apply the latter theorem and conclude that condition (ii) of Theorem 3.8 holds. Using this theorem, we obtain that

$$\nu_{x,t} = \delta_{u(x,t)} \text{ a.e. } (x, t) \in \mathbb{R}^2 \times [0, T],$$

where  $u$  is the (unique) Kruzkov solution of (1.1), (1.2). Corollary 3.5 implies that the sequence  $u_h$  converges strongly in  $L^1_{loc}(\mathbb{R}^2 \times [0, T])$  to  $u$ . This finishes the convergence proof for the higher order finite volume schemes.  $\square$

### 7. Applications

In this section we apply our convergence result to the discontinuous Galerkin finite element method. This method was introduced in [4, 5, 6]. Numerical experiments concerning this method for scalar conservation laws are presented in [6], and for the compressible two-dimensional Euler equations in [7]. Very extensive experiments for MUSCL-type reconstruction methods for equations in two and three space dimensions, in particular on unstructured meshes as they appear under local refinement have been conducted by Geiben [18]. Geiben develops a new reconstruction method with flux-limiters along the lines of the present paper. The resulting schemes satisfy the conditions for the convergence theorem. The experiments she has done refer to nonlinear conservation laws in 2-D on unstructured grids. For smooth solutions of a nonlinear scalar equation (where the exact solution is known) she got an experimental order of convergence (EOC) with respect to the  $L_1$ - norm close to 2. If the grid is strongly non-uniform as after a local refinement step, then the EOC is approximately  $\frac{3}{2}$ . For discontinuous solutions the EOC goes down to 1.25 but is still better than for the first order Engquist-Osher scheme (EOC=1.06) or for the Durlovsky-Engquist-Osher scheme(EOC=0.98) [14]. For the shock tube problem she got EOC=1.01 for the higher order and EOC=0.62 for the first order Steger and Warming scheme. Also in this case the  $L_1$ - error is much less than for the Durlovsky-Engquist-Osher scheme [14]. Similar results are obtained on a strongly nonuniform mesh.

In the following,  $k \geq 0$  is fixed. Given a  $B$ -triangulation  $\mathcal{T}_h := \{T_i : i \in I\}$  let

$$V_h := \{v \in L^1 \cap L^\infty(\mathbb{R}^2) : v|_{T_i} \in P_k(T_i)\},$$

where  $P_k$  is the space of polynomials of degree  $\leq k$ . Let  $\bar{V}_h \subset V_h$  denote the subspace of elements which are piecewise constant over the triangles, and decompose  $V_h$  orthogonally with respect to the  $L^2$  scalar product:

$$(7.1) \quad V_h = \bar{V}_h \oplus \tilde{V}_h.$$

Let  $\Pi_k : L^1 \cap L^\infty(\mathbb{R}^2) \rightarrow V_h$  be the  $L^2$ -projection (so  $\Pi_0$  is the projection onto the piecewise constants  $\bar{V}_h$ ). Given  $u \in L^1 \cap L^\infty(\mathbb{R}^2)$  and  $k \geq 0$ , let

$$(7.2) \quad u_h := \Pi_k u \in V_h$$

$$(7.3) \quad \bar{u}_h := \Pi_0 u \in \bar{V}_h$$

$$(7.4) \quad \tilde{u}_h := u_h - \bar{u}_h \in \tilde{V}_h$$

We will approximate the solution  $u(\cdot, t)$  of (1.1) - (1.2) by  $u_h(\cdot, t) \in V_h$ , and we will show that as  $h \rightarrow 0$ , the sequence  $\bar{u}_h$  contains a subsequence which converges to the unique entropy solution of (1.1), (1.2).

*Heuristical derivation of a semidiscrete equation*

In this paragraph, we assume that  $u_h \in C^1([0, T], \mathcal{T}_h)$ , where

$$\mathcal{T}_h := \{v \in L^1 \cap L^\infty(\mathbb{R}^2) : v|_{T_i} \text{ is continuous for all } i \in I\}.$$

If  $u_h \in L^1(\mathbb{R}_+, \mathcal{T}_h)$  is a weak solution of (1.1), it satisfies

$$\frac{d}{dt} \int_{T_i} u_h(\cdot, t) v_h + \int_{T_i} \nabla \cdot f(u_h(\cdot, t)) v_h = 0$$

for all  $i \in I$  and all  $v_h \in V_h$ . Integrating by parts formally yields

$$\frac{d}{dt} \int_{T_i} u_h(\cdot, t) v_h + \sum_{j \in N_i} \int_{S_{ij}} f(u_h(\cdot, t)) \cdot n_{ij} v_h - \int_{T_i} f(u_h(\cdot, t)) \cdot \nabla v_h = 0.$$

Since  $u_h$  is discontinuous across the boundary of  $T_i$ , we replace  $f \cdot n_{ij}|_{S_{ij}}$  in the first integral by a numerical flux  $g_{ij}$  as defined in Definition 2.4. Moreover, we replace the integrals by quadrature formulae

$$\int_{S_{ij}} \varphi \approx \sum_{l=1}^m \omega_l \varphi(x_{ijl}) |S_{ij}|$$

and

$$\int_{T_i} \varphi \approx \sum_{p=1}^r \omega_p \varphi(\underline{x}_{ip}) |T_i|$$

where the  $x_{ijl}$  resp.  $\underline{x}_{ip}$  are integration nodes on  $S_{ij}$  resp. in  $T_i$ , the quadrature formula over the edges is supposed to be exact for polynomials of degree  $\leq 2k+1$  and the one over the triangles for polynomials of degree  $\leq 2k$ . Then as in ([6], (2.4)) we obtain the weak formulation

$$\begin{aligned} \frac{d}{dt} \int_{T_i} u_h(\cdot, t) v_h &= - \sum_{j \in N_i} \sum_{l=1}^m \omega_l g_{ijl}(u_h(\cdot, t)) v_h^{\text{int}}(x_{ijl}) \\ &+ |T_i| \sum_{p=1}^r \omega_p f(u_h(\underline{x}_{ip}, t)) \cdot \nabla v_h(\underline{x}_{ip}), \end{aligned} \quad (7.5)$$

where

$$g_{ijl}(u_h(\cdot, t)) := g_{ij}(u_h^{\text{int}}(x_{ijl}, t), u_h^{\text{int}}(x_{jil}, t)) \quad (7.6)$$

and

$$u_h^{\text{int}}(x_{ijl}, t) := \lim_{x \rightarrow x_{ijl}, x \in T_i} u_h(x, t) \quad (7.7)$$

Given  $u_h(\cdot, t) \in \mathcal{T}_h$  and  $i \in I$  define a linear mapping  $\ell_i : P_k(T_i) \rightarrow \mathbb{R}$  by

$$\ell_i(v_h|_{T_i}) = - \sum_{j \in N_i} \sum_{l=1}^m \omega_l |S_{ij}| g_{ijl}(u_h(\cdot, t)) v_h^{\text{int}}(x_{ijl}) + |T_i| \sum_{p=1}^r \omega_p f(u_h(\underline{x}_{ip}, t)) \cdot \nabla v_h(\underline{x}_{ip}). \quad (7.8)$$



Then there is a  $w_i \in P_k(T_i)$  such that for all  $v_h \in V_h$ ,

$$(7.9) \quad \ell_i(v_h|_{T_i}) = (w_i, v_h)_{L^2(T_i)}.$$

We can now define an operator  $L_h : \mathcal{T}_h \rightarrow V_h$  by

$$(7.10) \quad \begin{aligned} u_h(\cdot, t) &\mapsto L_h u_h(\cdot, t) \\ L_h u_h(x, t) &:= w_i(x) \quad \text{for } x \in T_i. \end{aligned}$$

Combining (7.5) – (7.10) we obtain the semidiscrete equation

$$(7.11) \quad \frac{d}{dt} u_h(t) = L_h u_h(t).$$

*Definition of the discontinuous Galerkin method*

Now we would like to give the definition of the discontinuous Galerkin method. Formally, (7.11) can be discretized with respect to time as

$$u_h(t^{n+1}) = u_h(t^n) + \Delta t L_h u_h(t^n).$$

But in general  $u_h(t^{n+1})$  will not satisfy the constraints given in (2.15)–(2.18). Therefore we have to project  $u_h(t^{n+1})$  onto a space  $\widehat{V}_h$  of functions satisfying all these constraints (see (7.14) below). This projection is given by  $\Lambda_h$  (see (7.18)), and it takes over the role of the slope limiters which are used in the context of upwind finite difference methods.

**Definition 7.1.** (Projection  $\Lambda_h$ ) For  $u(\cdot) \in \mathcal{T}_h$  define

$$(7.12) \quad u_i := \frac{1}{|T_i|} \int_{T_i} u(x) dx$$

$$(7.13) \quad u_{ijl} := \lim_{x \rightarrow x_{ijl}, x \in T_i} u(x) - u_i^n$$

For given  $u(\cdot) \in \mathcal{T}_h$  define  $a_{ijl}, b_{ijl}$  as in (2.13), using  $u_i$  as defined in (7.12). Then for given  $\alpha, C_1, C_2$  define

$$(7.14) \quad \widehat{V}_h(u) := \{v \in \mathcal{T}_h : \text{for all } i \in I \text{ and all } j \in N_i \text{ } v \text{ satisfies (7.15)–(7.17)}\},$$

where

$$(7.15) \quad |\tilde{v}_{ijl}| \leq C_1 h^\alpha$$

$$(7.16) \quad \tilde{v}_{ijl} \in I(\pm C_2 h^{2\alpha}, -a_{ijl} \pm C_2 h^{2\alpha})$$

$$(7.17) \quad (u_i - u_j) \tilde{v}_{ijl} \leq C_2 h^{2\alpha}.$$

Now  $\Lambda_h(u)$  is defined as the  $L^2$ -projection onto  $\widehat{V}_h(u)$ :

$$(7.18) \quad \Lambda_h(u) : L^1 \cap L^\infty(\mathbb{R}^2) \rightarrow \widehat{V}_h(u).$$

**Note 7.2.** a)  $\widehat{V}_h(u)$  is not a subspace of  $V_h$ , but of  $\mathcal{T}_h$ . It is necessary to use  $\mathcal{T}_h$  in the definition of  $\widehat{V}_h(u)$  instead of  $V_h$ , since otherwise the result in Lemma 7.7 will not hold. Compare also Remark 7.9.

b)  $\widehat{V}_h(u)$  is convex, contains  $v \equiv 0$ , and if  $w \in \widehat{V}_h(u)$ , then  $w + \overline{V}_h \subset \widehat{V}_h(u)$ .

**Definition 7.3.** (discontinuous Galerkin finite element method) Let  $\mathcal{T}_h$  be a  $B$ -triangulation, and let  $0 = t^0 < t^1 < \dots < t^N = T$ . For given initial data  $u_0 \in L^\infty(\mathbb{R}^2)$  with compact support define

$$(7.19) \quad \bar{u}_h(x, 0) := \frac{1}{|T_i|} \int_{T_i} u_0(x) dx \quad \text{for } x \in T_i$$

and

$$(7.20) \quad u_h(\cdot, 0) := A_h(\bar{u}_h(\cdot, 0)) \Pi_k u_0.$$

For  $n = 0, \dots, N - 1$  let  $u_h(\cdot, t^n) = \bar{u}_h(\cdot, t^n) + \tilde{u}_h(\cdot, t^n) \in \widehat{V}_h(\bar{u}_h(\cdot, t^n))$  be given and define

$$(7.21) \quad \bar{u}_h(\cdot, t^{n+1}) := \bar{u}_h(\cdot, t^n) + \Delta t \overline{L_h u_h}(\cdot, t^n)$$

$$(7.22) \quad \tilde{u}_h(\cdot, t^{n+1}) := A_h(\bar{u}_h(\cdot, t^{n+1})) [\tilde{u}_h(\cdot, t^n) + \Delta t \widetilde{L_h u_h}(\cdot, t^n)],$$

where we have used the orthogonal decomposition

$$(7.23) \quad L_h u_h(\cdot, t^n) = \overline{L_h u_h}(\cdot, t^n) + \widetilde{L_h u_h}(\cdot, t^n)$$

analogous to (7.4). Now

$$(7.24) \quad u_h(\cdot, t^{n+1}) = \bar{u}_h(\cdot, t^{n+1}) + \tilde{u}_h(\cdot, t^{n+1}) \in \widehat{V}_h(\bar{u}_h(\cdot, t^{n+1})).$$

*Remark 7.4.* 1) Note that for all  $i \in I$ ,

$$\int_{T_i} \tilde{u}_h(x, t^{n+1}) dx = 0.$$

2) While we use an explicit Euler timestep to discretize (7.11), Cockburn, Hou and Shu derive a higher order accurate TVD Runge-Kutta timediscretization.

**Corollary 7.5.** *The definition of the discontinuous Galerkin method implies*

$$(7.25) \quad u_i^{n+1} = u_i^n - \frac{\Delta t}{|T_i|} \sum_{j \in N_i} \sum_{l=1}^m \omega_l |S_{ij}| g_{ijl}(u_h(\cdot, t)),$$

where as before  $u_i^p = \bar{u}_h(x_i, t^p)$  for  $p = 0, \dots, N$  and  $g_{ijl}$  is defined as in (7.6).

*Proof.* Let  $v_h \in V_h, \text{supp } v_h \subset T_i, v_h|_{T_i} \equiv \frac{1}{|T_i|}$ . Using this in (7.8) gives

$$\overline{L_h u_h}(x_i, t^n) = -\frac{1}{|T_i|} \sum_{j \in N_i} \sum_{l=1}^m \omega_l |S_{ij}| g_{ijl}(u_h(\cdot, t)).$$

From here and (7.21), (7.25) follows immediately.  $\square$

*Convergence and spatial accuracy*

We are now ready to prove the following

**Theorem 7.6.** *Suppose that the CFL-condition (2.21) (resp. (4.2) for a B-uniform family of triangulations) is satisfied, and that the regularity assumption (2.1) holds for some  $c_V > 0$ . Then, as  $h \rightarrow 0$ , the sequence  $\bar{u}_h$  produced by the discontinuous Galerkin method contains a subsequence which converges strongly in  $L^1_{loc}(\mathbb{R}^2 \times [0, T])$  to the unique entropy solution  $u$  of (1.1)–(1.2).*

*Proof.* The proof follows immediately from the preceding corollary and Theorem 2.13. It remains only to remark that the method has a finite stencil.  $\square$

Finally, we study the spatial order of consistency of the discontinuous Galerkin method.

**Lemma 7.7.** *Suppose that  $u : [0, T] \times \mathbb{R}^2 \rightarrow \mathbb{R}$  and  $f : \mathbb{R} \rightarrow \mathbb{R}^2$  are smooth, and that  $(\mathcal{T}_h)_{h>0}$  is a B-uniform family of triangulations with respect to a constant  $\mu$  which is regular in the sense of (2.1). For  $t^n \in [0, T]$  let*

$$u_i^n := \frac{1}{|T_i|} \int_{T_i} u(\cdot, t^n).$$

Let  $C_1, C_2$  be sufficiently large and let  $\Lambda_h(\bar{u}_h(\cdot, t^n))$  be the projection defined by (7.18). Then

a)

$$\Lambda_h(\bar{u}_h(\cdot, t^n)) u(\cdot, t^n) = u(\cdot, t^n).$$

b) There is a constant  $C_3 > 0$  depending only on  $\mu, u$  and  $f$  such that

$$(7.26) \quad \sup_{x \in \mathbb{R}^2} |L_h \Lambda_h(\bar{u}_h(\cdot, t^n)) u(x, t^n) + \nabla \cdot f(u(x, t^n))| \leq C_3 h^{k+1}$$

c)

$$(7.27) \quad \sup_{i \in I} \left| \frac{1}{|T_i|} \sum_{j \in N_i} \sum_{l=1}^m \omega_l g_{ijl} (\Lambda_h(\bar{u}_h(\cdot, t^n)) u(\cdot, t^n)) - \frac{1}{|T_i|} \int_{T_i} \nabla \cdot f(u(\cdot, t^n)) \right| \leq C_3 h^{k+1},$$

where  $g_{ijl}(\Lambda_h(\bar{u}_h(\cdot, t^n)) u(\cdot, t^n))$  is defined analogously to (7.6) – (7.7).

*Proof.* Let  $M_l := \|\nabla^l u\|_{L^\infty(\mathbb{R}^2 \times [0, T])}$  and

$$(7.28) \quad u_{ijl}^n := u(x_{ijl}, t^n) - u_i^n$$

Therefore

$$\begin{aligned} u_{ijl}^n &= u(x_{ijl}, t^n) - u(x_i, t^n) + O(h^2) = \nabla u(x_i, t^n)(x_{ijl} - x_i) + O(h^2) \\ &= -\nabla u(x_i, t^n) \sum_{p \in N_i} \alpha_{ijlp} d_{ip} + O(h^2) = -\sum_{p \in N_i} \alpha_{ijlp} [\Delta_{ip} + O(h^2)] + O(h^2) \\ &= -a_{ijl} + O(h^2)(1 + \sum_{p \in N_i} \alpha_{ijlp}). \end{aligned}$$

Since  $\sum_{p \in N_i} \alpha_{ijlp}$  is uniformly bounded and  $\alpha \leq 1$ , this implies

$$|u_{ijl}^n + a_{ijl}| \leq C_2 h^{2\alpha}$$

if  $C_2$  is large enough, i.e. (2.16) holds. Analogously, (2.17) holds with the same choice of  $C_2$ . Conditions (2.15) and (2.18) are satisfied since  $u$  is sufficiently smooth. Therefore,  $C_1, C_2$  can be chosen such that  $u(\cdot, t^n) \in \widehat{V}_h^n$ , so that  $\Lambda_h(\bar{u}_h(\cdot, t^n))u(\cdot, t^n) = u(\cdot, t^n)$ . This proves *a*). Note that in particular,

$$g_{ijl}(\Lambda_h(\bar{u}_h(\cdot, t^n))u(\cdot, t^n)) = f(u(x_{ijl}, t^n)) \cdot n_{ij}.$$

Now we can apply Lemma 2.1 of [6] to obtain *b*). It remains to prove *c*). We have by definition

$$\overline{\nabla \cdot f(u)} = \frac{1}{|T_i|} \int_{T_i} \nabla \cdot f(u(\cdot, t^n)).$$

Decompose  $\nabla \cdot f(u) = \overline{\nabla \cdot f(u)} + \widetilde{\nabla \cdot f(u)}$ . Since this decomposition is orthogonal, we obtain

$$\begin{aligned} &\left[ \frac{1}{|T_i|} \sum_{j \in N_i} \sum_{l=1}^m \omega_l |S_{ij}| g_{ijl}(\Lambda_h(\bar{u}_h(\cdot, t^n))u(\cdot, t^n)) - \overline{\nabla \cdot f(u)} \right]^2 \\ &= \frac{1}{|T_i|} \int_{T_i} [\overline{L_h u} + \overline{\nabla \cdot f(u)}]^2 \\ &\leq \frac{1}{|T_i|} \left\{ \int_{T_i} [\overline{L_h u} + \overline{\nabla \cdot f(u)}]^2 + \int_{T_i} [\widetilde{L_h u} + \widetilde{\nabla \cdot f(u)}]^2 \right\} \\ &= \frac{1}{|T_i|} \int_{T_i} [L_h u + \nabla \cdot f(u)]^2 \leq \|L_h u + \nabla \cdot f(u)\|_{L^\infty(\mathbb{R}^2)}^2 \leq [C_3 h^{k+1}]^2, \end{aligned}$$

where the last inequality follows from (7.26).  $\square$

The following corollary follows immediately from the proof of the preceding Lemma:

**Corollary 7.8.** *If we assume that there is a constant  $C > 0$  such that as  $h \rightarrow 0$ ,*

$$h^{2(1-\alpha)}(1 + \sum_{p \in N_i} \alpha_{ijlp}) \leq C$$

*instead of " $(\mathcal{T}_h)_{h>0}$  is a  $B$ -uniform family of triangulations", then (7.26)–(7.27) hold as well.*

*Remark 7.9.* The point of Lemma 7.7 is that conditions (2.15)–(2.18) are automatically satisfied for smooth functions and that  $A_h(\bar{u}_h(\cdot, t^n))$  reduces to the identity in that case. Inequalities (7.26)–(7.27) do *not* imply that the discontinuous Galerkin method (7.19) – (7.24) is spatially consistent of order  $k + 1$ . The reason is the innocent-looking projection  $\Pi_k$  in (7.20). We recall a counterexample given in [19]: For  $k = 0$ ,

$$L_h(A_h(\bar{u}_h(\cdot, t^n)) \Pi_0 u(\cdot, t^n)) = -\frac{1}{|T_i|} \sum_{j \in N_i} g_{ij}(u_i^n, u_{ij}^n)$$

and in general

$$\left| \frac{1}{|T_i|} \sum_{j \in N_i} g_{ij}(u_i^n, u_{ij}^n) - \frac{1}{|T_i|} \int_{T_i} \nabla \cdot f(u(\cdot, t^n)) \right| = O(1).$$

The same remark applies to Lemma 2.1 of Cockburn, Hou and Shu [6].

### 8. Numerical entropy fluxes

In this section we completely characterize the numerical fluxes  $g_{ij}$  which satisfy Assumption 2.4 and admit a numerical entropy flux  $G_{ij}$  satisfying Assumption 2.6.

**Lemma 8.1.** *Suppose that  $g_{ij}$  is a numerical flux satisfying Assumption 2.4. Then the following are equivalent:*

- (i) *For every Lipschitz-continuous convex entropy  $U$  there exists a  $G_{ij}$  satisfying the Assumptions 2.6.*
- (ii) *For all Kruřkov-entropies  $U(\cdot, k)$  there exists a  $G_{ij}$  satisfying the Assumptions 2.6.*
- (iii) *There are Lipschitz-continuous functions  $\varphi_{ij}, \psi_{ij} : \mathbb{R} \rightarrow \mathbb{R}$  with  $\varphi_{ij}(0) = 0 = \psi_{ij}(0)$ ,*

$$\varphi_{ij} \equiv \psi_{ji}$$

*and constants  $C_\varphi(M), C_\psi(M) > 0$  such that for almost all  $s \in [-M, M]$ ,*

$$C_\varphi \geq \varphi'_{ij}(s) \geq \frac{1}{2} |f'(s) \cdot n_{ij}|$$

$$C_\psi \geq \psi'_{ij}(s) \geq \frac{1}{2} |f'(s) \cdot n_{ij}|$$

*and*

$$(8.1) \quad g_{ij}(u, v) = \frac{f(u) + f(v)}{2} \cdot n_{ij} + \begin{cases} \varphi_{ij}(u) - \varphi_{ij}(v) & \text{for } v \geq u \\ \psi_{ij}(u) - \psi_{ij}(v) & \text{for } v < u \end{cases}$$

*Proof.* That (i) implies (ii) is trivial.

We now show that (iii) implies (i). It is not hard to check that  $g_{ij}$  given by (8.1) satisfies Assumption 2.4. Given an entropy pair  $(U, F)$ , let

$$\begin{aligned} G_{ij}(u, v) &:= F(0) \cdot n_{ij} + \int_0^u U'(s) \partial_u g_{ij}(s, v) ds + \int_0^v U'(t) \partial_v g_{ij}(u, t) dt \\ &= \frac{F(u) + F(v)}{2} \cdot n_{ij} + \begin{cases} \int_v^u U'(s) \varphi'_{ij}(s) ds & \text{for } v \geq u \\ \int_v^u U'(s) \psi'_{ij}(s) ds & \text{for } v < u \end{cases} \end{aligned}$$

Then  $G_{ij}$  satisfies Assumption 2.6.

Finally, we show that (ii) implies (iii). Let  $\phi \in C_0^\infty(\mathbb{R}^2)$ ,  $\text{supp } \phi \subset (-1, 1)^2$ ,  $\phi \geq 0$ ,  $\int \phi = 1$ , and for  $\varepsilon > 0$  define  $\phi_\varepsilon(s, t) := \frac{1}{\varepsilon^2} \phi(\frac{s}{\varepsilon}, \frac{t}{\varepsilon})$ . For  $v \geq u + 4\varepsilon$ , let  $k := \frac{u+v}{2}$ . Then  $u + 2\varepsilon \leq k \leq v - 2\varepsilon$ . Let  $i, j$  be fixed, and drop the subscripts from  $n, g$  and  $G$ . Let  $U(s) := |s - k|$ . Then for  $(s, t) \in [-\varepsilon, \varepsilon]^2$ ,  $U'(u - s) = -1 = -U'(v - t)$ . Since  $g, G$  and  $U$  are Lipschitz, (2.9) gives for almost all  $(s, t) \in [-\varepsilon, \varepsilon]^2$

$$\begin{aligned} G_u(u - s, v - t) &= -g_u(u - s, v - t) \\ G_v(u - s, v - t) &= g_v(u - s, v - t). \end{aligned}$$

Multiplying by  $\phi_\varepsilon(s, t)$  and integrating gives

$$\begin{aligned} \phi_\varepsilon * G_u(u, v) &= -\phi_\varepsilon * g_u(u, v) \\ \phi_\varepsilon * G_v(u, v) &= \phi_\varepsilon * g_v(u, v). \end{aligned}$$

Since  $G$  and  $g$  are Lipschitz, one can apply the dominated convergence theorem to derive

$$\begin{aligned} G_u^\varepsilon(u, v) &= -g_u^\varepsilon(u, v) \\ G_v^\varepsilon(u, v) &= g_v^\varepsilon(u, v), \end{aligned}$$

where  $G^\varepsilon = \phi_\varepsilon * G$ ,  $g^\varepsilon = \phi_\varepsilon * g$ . This gives immediately

$$g_{uv}^\varepsilon(u, v) = 0 \text{ for } v \geq u + 4\varepsilon,$$

and analogously one derives

$$g_{uv}^\varepsilon(u, v) = 0 \text{ for } v \leq u - 4\varepsilon.$$

Let  $v > u + 4\varepsilon > 2\varepsilon$ . Then

$$\begin{aligned} g^\varepsilon(u, v) &= g^\varepsilon(u, u + 4\varepsilon) + \int_{u+4\varepsilon}^v g_v^\varepsilon(u, t) dt \\ &= g^\varepsilon(u, u + 4\varepsilon) + \int_{u+4\varepsilon}^v g_v^\varepsilon(-2\varepsilon, t) dt \\ &= g^\varepsilon(u, u + 4\varepsilon) + g^\varepsilon(-2\varepsilon, v) - g^\varepsilon(-2\varepsilon, u + 4\varepsilon). \end{aligned}$$

As  $\varepsilon \rightarrow 0$ , we obtain for  $v > u > 0$

$$g^\varepsilon(u, v) = f(u) \cdot n - g(0, u) + g(0, v).$$

In the same fashion, we obtain

$$g(u, v) = \begin{cases} f(u) \cdot n - g(0, u) + g(0, v) & \text{for } v > u > 0 \text{ or } v < u < 0 \\ g(u, 0) + g(0, v) - f(0) \cdot n & \text{for } v > 0 > u \text{ or } v < 0 < u \\ g(u, 0) - g(v, 0) + f(v) \cdot n & \text{for } 0 > v > u \text{ or } 0 < v < u \end{cases}$$

From here, one can show that

$$g(u, v) = \frac{f(u) + f(v)}{2} \cdot n + \begin{cases} \varphi(u) - \varphi(v) & \text{for } v \geq u \\ \psi(u) - \psi(v) & \text{for } v < u \end{cases}$$

with

$$(8.2) \quad \varphi(s) = \begin{cases} \frac{f(s) + f(0)}{2} \cdot n - g(0, s) & \text{for } s \geq 0 \\ g(s, 0) - \frac{f(s) + f(0)}{2} \cdot n & \text{for } s < 0 \end{cases}$$

$$(8.3) \quad \psi(s) = \begin{cases} g(s, 0) - \frac{f(s) + f(0)}{2} \cdot n & \text{for } s \geq 0 \\ \frac{f(s) + f(0)}{2} \cdot n - g(0, s) & \text{for } s < 0 \end{cases}$$

It is elementary to check that  $\varphi, \psi$  satisfy the properties required in (iii).  $\square$

*Example 8.2.* The Engquist-Osher flux can be rewritten in the form

$$g_{ij}^{EO}(u, v) = \frac{f(u) + f(v)}{2} \cdot n_{ij} + \frac{1}{2} \int_v^u |f'(s) \cdot n_{ij}| ds.$$

Therefore,

$$\varphi_{ij}^{EO}(s) = \frac{1}{2} \int_0^s |f' \cdot n_{ij}|.$$

For the Lax-Friedrichs flux,

$$\varphi_{ij}^{LF}(s) = \frac{s}{2\lambda_{ij}}.$$

Therefore, if  $\varphi_{ij}$  corresponds to any numerical flux satisfying our assumptions, then

$$\frac{d}{ds} \varphi_{ij}(s) \geq \frac{d}{ds} \varphi_{ij}^{EO}(s),$$

which can be interpreted by saying that the Engquist-Osher flux is the least dissipative one admitted in our proof.

*Added in proof:* We are grateful to the referees for a hint to a recent paper of Cockburn, Coquel and LeFloch [2] which has been published as a preprint of the Ecole Polytechnique (October 1993) while our paper was already being refereed. We would also like to thank Bernado Cockburn for sending us a more recent version of that preprint (June 1994). In this version, Cockburn, Coquel and LeFloch establish an  $L^\infty(L^1)$  error estimate for a general class of finite volume methods under some additional assumptions. The error is proven to be of order  $h^{1/4}$ . They use the Kuznetsov [26] theory in order to get the necessary estimates from the Kruzkov entropy inequalities. Nearly the same result has been obtained

with the same method by Vila [39] (January 1993). The paper [2] refers to higher order and [39] to first order finite volume schemes.

In our paper we show convergence for higher order schemes with a completely different method. Our assumptions are weaker than those made in [2] in the following two ways: First, we can prove convergence for initial data in  $L^1 \cap L^\infty$ , while the Kuznetsov theory (and hence [2]) requires the data to be of bounded variation. Secondly, the result in [2] assumes that the flux function  $f$  is smooth, that  $\partial^2/\partial u^2 f(u)$  is bounded and that the so called *sharp entropy inequalities* hold. In [3, 10], these entropy inequalities could only be derived under the additional assumption that either

(i) there is a  $\delta > 0$  such that for all  $i \in I, j \in N_i$  and all  $u$ ,

$$(8.4) \quad \left| \frac{\partial^2 f(u)}{\partial u^2} \cdot n_{ij} \right| \geq \delta$$

or

(ii) the numerical viscosity coefficient is bounded below by a constant independent of  $h$ ,

$$(8.5) \quad Q_{ij}(u, v) := \lambda_{ij} \frac{f(u) \cdot n_{ij} + f(v) \cdot n_{ij} - 2g_{ij}(u, v)}{v - u} \geq Q_* > 0.$$

The genuine-nonlinearity assumption (8.4) is a restriction both to the flux function  $f$  and the triangulation, and will usually not hold, especially when a grid is automatically refined. The assumption (8.5) is not satisfied by the Engquist-Osher and the Godunov flux. Our entropy inequality (see Theorem 5.1) is of interest in its own right. It does not rely on (8.4) and (8.5). The flux function  $f$  only needs to be Lipschitz continuous, and we can include the Engquist-Osher flux. Furthermore, a recent result of Noelle [30] also applies to Godunov's numerical flux and nonconvex flux functions  $f$ .

We would also like to mention that the technique developed in this paper was already applied to situations with more general triangulations in two recent papers: Geiben [17] removes the requirement of a B-triangulation and enforces the  $L^\infty$  bound by a new flux-limiter, and Noelle [30] generalizes the result to irregular polygonal grids, where cells may become flat in the limit. In a forthcoming paper [31], Noelle also obtains error estimates for such grids, for nonconvex fluxfunctions and general E-fluxes.

## References

1. Beinert, R., Kröner, D. (1994): Finite volume methods with local mesh refinement in 2-D. Notes Numer. Fluid Mech. **46**, 38–53
2. Cockburn, B., Coquel, F., LeFloch, P. (1994): An error estimate for finite volume methods for multidimensional conservation laws. Math. of Comp. **63**, 77–103
3. Cockburn, B., Coquel, F., LeFloch, P., Shu, C.W. (1991): Convergence of finite volume methods. IMA Preprint no. 771
4. Cockburn, B., Shu, C.W. (1989): TVB Runge-Kutta local projection discontinuous Galerkin finite element method for conservation laws II: general framework. Math. of Comp. **52**, 411–435



5. Cockburn, B., Lin, S.Y., Shu, C.W. (1989): TVB Runge-Kutta local projection discontinuous Galerkin finite element method for conservation laws III: one-dimensional systems. *J. of Comp. Physics* **84**, 90–113
6. Cockburn, B., Hou, S., Shu, C.W. (1990): The Runge-Kutta local projection discontinuous Galerkin finite element method for conservation laws IV: the multidimensional case. *Math. of Comp.* **54**, 545–581
7. Cockburn, B., Shu, C.W. (1991): The  $P^1$ -RKDG method for two-dimensional Euler equations of gas-dynamics. ICASE-Report No. 91-32
8. Conway, E., Smoller, J. (1966): Global solutions of the Cauchy problem for quasilinear first-order equations in several space variables. *Comm. Pure Appl. Math.* **19**, 95–105
9. Coquel, F., LeFloch, P. (1993): Convergence of finite difference schemes for conservation laws in several space dimensions: general theory. *SIAM J. Numer. Anal.* **30**, 675–700
10. Coquel, F., LeFloch, P. (1991): Convergence of finite difference schemes for conservation laws in several space dimensions: the corrected antidiffusive flux approach. *Math. of Comp.* **57**, 169–210
11. Coquel, F., LeFloch, P. (1992): The finite volume method on general triangulations converges for general conservation laws. Preprint
12. Crandall, M.C., Majda, A. (1980): Monotone difference approximations for scalar conservation laws. *Math. of Comp.* **34**, 1–21
13. DiPerna, R.J. (1985): Measure valued solutions to conservation laws. *Arch. Rational Mech. Anal.* **88**, 223–270
14. Durlflosky, L.J., Engquist, B., Osher, S. (1992): Triangle based adaptive stencils for the solution of hyperbolic conservation laws. *J. of Comp. Physics* **98**, 64–73
15. Engquist, B., Osher, S. (1981): One sided difference approximations for nonlinear conservation laws. *Math. of Comp.* **36**, 321–351
16. Fezoui, L., Stoufflet, B. (1989): A class of implicit upwind schemes for Euler simulations with unstructured meshes. *J. of Comp. Physics* **84**, 174–206
17. Geiben, M. (1993): Convergence of MUSCL-type upwind finite volume schemes on unstructured triangular grids. SFB256 Preprint no. 318, Bonn, Germany
18. Geiben, M. (1994): Solving the compressible Euler equations in time-dependent geometries. Preprint (to appear in the Proceedings of the Fifth International Conference on Hyperbolic Problems, Stony Brook)
19. Geiben, M., Kröner, D., Rokyta, M. (1991): A Lax-Wendroff type theorem for cell-centered, finite volume schemes in 2-D. Preprint
20. Johnson, C., Szepessy, A. (1987): Convergence of a finite element method for a nonlinear hyperbolic conservation law. *Math. of Comp.* **49**, 427–444
21. Johnson, C., Szepessy, A., Hansbo, P. (1990): On the convergence of shock-capturing streamline diffusion finite element methods for hyperbolic conservation laws. *Math. of Comp.* **54**, 107–129
22. Kröner, D.: Directionally adapted upwind schemes in 2-D for the Euler equations. Priority Research Program, Results 1986-1988. Notes on numerical fluid mechanics II., Vol. 25
23. Kröner, D., Noelle, S., Rokyta, M. (1993): Convergence of higher order upwind finite volume schemes on unstructured grids for scalar conservation laws in several space dimensions. SFB256 preprint no. 268, Bonn, Germany
24. Kröner, D., Rokyta, M. (1994): Convergence of upwind finite volume schemes for scalar conservation laws in 2-D. SFB256 preprint, Bonn, Germany (1992); *SIAM J. Num. Anal.* **31**, 324–343
25. Kruzkov, S.N. (1970): First order quasilinear equations in several independent variables. *Mat.Sbornik Tom 81 (123)*, 1970 and *Math. USSR-Sbornik* **10**, 217–243
26. Kuznetsov, 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.* **16,6** 105–119
27. Lax, P. (1954): Weak solutions of nonlinear hyperbolic equations and their numerical computation. *Comm. Pure Appl. Math.* **VII**, 159–193
28. Lax, P. (1971): Shock waves and entropy. Proceedings of a symposium at the University of Wisconsin, (E.H.Zarantonello, ed.) 603–634
29. Morton, K.W., Paisley, M.F. (1989): A finite volume scheme with shock fitting for the Euler equation. *J. of Comp. Physics* **80**

30. Noelle, S. (1995): Convergence of higher order finite volume schemes on irregular grids. SFB256 Preprint no. 295 (1993); Adv. Comp. Math. **3**, 197–218
31. Noelle, S. (1995): A note on entropy inequalities and error estimates for higher order accurate finite volume schemes on irregular families of grids. SFB256 Preprint (to appear in Math. of Comp., July 1996)
32. Pironneau, O. (1982): On the transport-diffusion algorithms and its applications to the Navier-Stokes equations. Numer. Math. **38**, 309–332
33. Osher, S., Tadmor, E. (1988): On the convergence of difference approximations to scalar conservation laws. Math. of Comput. **50**, 19–51
34. Sonar, T. (1992): Entropy production in second-order three-point schemes. Numer. Math. **62**, 371–390
35. Szepessy, A. (1989): Convergence of a shock capturing streamline diffusion finite element method for a scalar conservation law in two space dimensions. Math. of Comp. **53**, 527–545
36. Tadmor, E. (1984): Numerical viscosity and the entropy condition for conservative difference schemes. Math. of Comp. **43**, 369–381
37. Tartar, L. (1979): Compensated compactness and applications to partial differential equations. In: R.J. Knops, ed., Research Notes in Mathematics, Nonlinear Analysis and Mechanics: Herriot-Watts Symposium, vol. 4. London, Pitman Press, London
38. Tartar, L. (1983): The compensated compactness method applied to systems of conservation laws. J.M. Ball (ed.), Systems of nonlinear PDE, NATO ASI Series, 263–285
39. Vila, J. (1993): Convergence and error estimates in finite volume schemes for general multidimensional scalar conservation laws: I Explicit monotone schemes. Preprint, Université de Nice