

## Convergence of Upwind Schemes for a Stationary Shock

By Jens Lorenz\*

**Abstract.** A nonlinear first-order boundary value problem with discontinuous solutions is considered. It arises in the study of gasflow through a duct and allows, in general, for multiple solutions. New convergence results for three difference schemes are presented and the sharpness of numerical layers is established. For the EO-scheme, stability of a physically correct solution with respect to time evolution is shown.

**1. Introduction.** In this paper we analyze three difference schemes applied to a shock problem

$$(1.1) \quad \frac{d}{dx}f(u(x)) + b(x, u(x)) = 0, \quad 0 \leq x \leq 1, \quad u(0) = \gamma_0, \quad u(1) = \gamma_1.$$

Since the differential equation of first order is supplemented by two boundary conditions, we have to make precise what is meant by a solution of (1.1). In the case

$$(1.2) \quad b_u(x, u) \geq \mu > 0 \quad \text{on } [0, 1] \times \mathbb{R}$$

this is easily done: for all  $\varepsilon > 0$  the second-order problem

$$(1.3) \quad -\varepsilon u'' + f(u)' + b(x, u) = 0, \quad 0 \leq x \leq 1, \quad u(0) = \gamma_0, \quad u(1) = \gamma_1,$$

is uniquely solvable and the solutions  $u_\varepsilon$  tend to a limit function  $U$  of bounded variation.  $U$  is considered as the solution of (1.1). Motivated by the considerations in [3], we are also interested in cases where the condition  $b_u > 0$  is violated. The one-dimensional duct flow equations for an inviscid gas can—for the stationary state—be reduced to a scalar equation for the velocity  $u$ , which has the form (1.1) (see, e.g., [16]). The condition  $b_u > 0$  is violated, e.g., for a converging-diverging duct. We make precise below what we understand by a solution of (1.1) in this case. Since (1.1) describes the stationary states of the hyperbolic problem

$$(1.4) \quad \begin{aligned} u_t + f(u)_x + b(x, u) &= 0, & 0 < x < 1, \quad t > 0, \\ u(0, t) &= \gamma_0, \quad u(1, t) = \gamma_1, & t > 0, \\ u(x, 0) &= \phi(x), & 0 \leq x \leq 1, \end{aligned}$$

the question of stability with respect to time evolution is also of interest. If  $b_u > 0$ , the solution  $U$  is stable; in the converging-diverging duct problem there are often two solutions,  $U^{(1)}$  and  $U^{(2)}$  with shocks;  $U^{(1)}$  is unstable [3]. It seems to be very likely that the solution  $U^{(2)}$ , which has its shock in a region where  $b_u > 0$ , is stable, but this has not yet been shown rigorously.

---

Received June 7, 1984; revised January 31, 1985.

1980 *Mathematics Subject Classification.* Primary 65L10, 34B15.

\*Research supported by NSF Grant #DMS 83-12264.

Possibly the simplest numerical methods for (1.4) are explicit difference schemes

$$(1.5) \quad \begin{aligned} \frac{1}{\tau}(u_i^{n+1} - u_i^n) + \frac{1}{h} \{ g(u_{i+1}^n, u_i^n) - g(u_i^n, u_{i-1}^n) \} \\ + b(ih, u_i^n) = 0, \quad 1 \leq i \leq m, \\ u_0^{n+1} = \gamma_0, \quad u_{m+1}^{n+1} = \gamma_1. \end{aligned}$$

Here  $h = 1/(m + 1)$ ,  $\tau > 0$ ,  $u_i^0 \approx \phi(ih)$ , and  $u_i^n$  approximates  $u(ih, n\tau)$ . The function  $g(\cdot, \cdot)$  is a numerical flux function used to discretize the term  $f(u)_x$ .

We are particularly interested in the flux functions

$$(1.6) \quad g(u, v) = \begin{cases} \max\{f(w): u \leq w \leq v\}, & u \leq v, \\ \min\{f(w): v \leq w \leq u\}, & v \leq u, \end{cases}$$

of Godunov (e.g., [12]) and

$$(1.7) \quad g(u, v) = \int_c^u \min\{f'(s), 0\} ds + \int_c^v \max\{f'(s), 0\} ds$$

of Engquist-Osher [4], [13], [14]. Since these are monotone schemes it is known that they lead to approximations converging in  $L_1$ -norm, at least if  $b \equiv 0$  and if no boundary conditions are present [2]. For the stationary case with Dirichlet boundary conditions (the  $n$ -dependence in (1.5) is dropped) convergence in  $L_1$ -norm has been shown in [9] for all monotone schemes if  $b_u \geq \mu > 0$ . (See [1] for relevant analytical discussions.) These results do not make use of the upwind nature of the EO- and the G-scheme. More specific convergence results for the EO-scheme and an  $O(h^2)$ -modification have recently been shown in [10] under restrictive assumptions. This paper continues [10] and contains a number of new results for the stationary shock problem and discretizations of it.

1. An essential assumption for the convergence result in [10] was that the discrete solution

$$(u_i) = (u_i^h)$$

of the EO-scheme is monotone with respect to  $i$ . In this paper, we can drop this condition completely, allowing for applications to the duct flow problem. Essentially, we can reduce the whole discussion to the consideration of a  $2 \times 2$  nonlinear algebraic system and obtain the same convergence results as in [10]. Especially, it is again established under the milder conditions that a numerical layer for the EO-scheme contains at most two mesh-points. Values at other mesh-points converge uniformly with order  $h$ . For the modification of the EO-scheme introduced in [9], [10] a similar result is shown with  $O(h^2)$ -estimates.

2. The same convergence proof goes through also for the G-scheme. We show even more: Under the conditions made, the values  $u_i^G$  of the G-scheme and the values  $u_i^{\text{EO}}$  of the EO-scheme coincide exactly, with the exception of only one value at an interior layer. The G-scheme marks an interior layer even sharper than the EO-scheme, namely by only one mesh-point. (It should be pointed out, however, that the EO-scheme gives a  $C^1$ -numerical flux function, whereas the flux function for the G-scheme is only Lipschitz continuous. This makes Newton's method difficult for the G-scheme, since one has to distinguish between right and left derivatives.)

3. In a situation as in the converging-diverging duct problem we show for the three schemes existence and local uniqueness of a discrete solution  $(u_i^h)$  converging to the probably stable solution  $U^{(2)}$ . We can show rigorously that the *discrete* solution  $(u_i^h)$  of the EO-scheme is stable in the sense that all eigenvalues of the linearization of the EO-system at  $(u_i^h)$  are positive. Thus, for the method of lines system

$$(1.8) \quad \frac{d}{dt}u(t) + T^h(u(t)) = 0, \quad t \geq 0,$$

the stationary solution  $u^h$  close to  $U^{(2)}$  is an attractor. Here  $T^h$  denotes the EO-discrete analogue of (1.1).

**2. The Continuous Problem.** In this section we define a solution concept for problem (1.1), give a physical example arising in duct flow, and then describe the detailed assumptions under which the behavior of the difference schemes shall be analyzed. With NBV we denote the space of all functions  $u$  of bounded variation on  $[0, 1]$  which are normalized such that

$$u(x) = u(x +) \quad \forall x \in [0, 1), \quad u(1) = u(1 -).$$

Furthermore, let  $\text{sg } \alpha = -1, 0, +1$  for  $\alpha < 0, = 0, > 0$ . Motivated by Theorem 1 of [10], we adopt the following definition:

*Definition 1.* Any function  $U \in \text{NBV}$  which satisfies the following three conditions is a solution of (1.1):

- (i)  $\int_0^1 \{f(U)\phi' - b(x, U)\phi\} dx = 0 \quad \forall \phi \in C_0^\infty(0, 1)$ .
- (ii) For all discontinuities  $y \in (0, 1)$  of  $U$

$$\text{sg}(U(y +) - U(y -))(f(U(y)) - f(k)) \leq 0$$

holds for all  $k$  between  $U(y +)$  and  $U(y -)$ .

- (iii) At the boundary points  $i = 0$  and  $i = 1$

$$0 \leq (-1)^{i+1} \text{sg}(U(i) - \gamma_i)(f(U(i)) - f(k))$$

holds for all  $k$  between  $U(i)$  and  $\gamma_i$ .

We note that (ii) is a special case of Oleřnik's condition (E), see [12], and (iii) is a special case of conditions at the boundary described in [1]. To our knowledge, existence and uniqueness of  $U$  has only been shown for  $b_u \geq \mu > 0$ , see [10]. We accept the solution concept of Definition 1, however, also in cases where  $b_u > 0$  is violated. This is motivated by the following example.

*A gasflow problem.* With  $A(x)$ ,  $0 \leq x \leq 1$ , we denote the cross-sectional area of a duct. Let  $\rho, u, e$ , and  $p$  denote density, velocity, specific internal energy, and pressure. If heat conduction is neglected, the inviscid equations of conservation of mass, momentum, and energy read in the stationary case (e.g, [7], [15]):

$$(2.1) \quad \left( \begin{array}{c} \rho u A \\ (\rho u^2 + p) A \\ \left( \rho e + \rho \frac{u^2}{2} + p \right) u A \end{array} \right)_x + \begin{pmatrix} 0 \\ -p A_x \\ 0 \end{pmatrix} = 0.$$

For definiteness, let  $p$  be given by the perfect gas law,  $p = (\gamma - 1)\rho e$ , where  $\gamma > 1$  is constant. With the two constants

$$c_0 = \rho u A, \quad H = \gamma e + \frac{u^2}{2}$$

the variables  $\rho, e$ , and thus  $p$  can be eliminated from the momentum equation. This yields

$$(2.2) \quad \left( \frac{\gamma + 1}{\gamma - 1} u + \frac{2H}{u} \right)' + \frac{A'(x)}{A(x)} \left( u - \frac{2H}{u} \right) = 0,$$

where for physical reasons  $0 < u < \sqrt{2H}$ . If we try to prescribe  $u(0) = \gamma_0$ ,  $u(1) = \gamma_1$ , we arrive at a problem (1.1). Here  $f(u) = (\gamma + 1)u/(\gamma - 1) + 2H/u$  is convex and  $b(x, u) = A'(x)(u - 2H/u)/A(x)$  fulfills the crucial condition  $b_u \geq 0$  for uniqueness only if the duct is diverging. If a solution  $u(x)$  of (2.2) has a discontinuity at some  $0 < y < 1$ , then

$$f(u(y-)) = f(u(y+))$$

and  $u(y-) > u(y+)$ . This follows from conditions (i) and (ii) of Definition 1. Any solution  $u$  of the scalar equation (2.2) gives rise to a solution  $(\rho, u, e)$  of the system (2.1). One can show that  $(\rho, u, e)$  satisfies the Rankine-Hugoniot jump condition for the stationary system (2.1) and also satisfies the entropy condition at the jump  $y$ , i.e.,

$$u(y-) - c(y-) > 0 > u(y+) - c(y+).$$

Here  $c = \{(\gamma - 1)(H - u^2/2)\}^{1/2}$  denotes the local speed of sound.

In our opinion, these considerations justify—at least partly—the reduction of the system (2.1) to a scalar equation and the solution concept for the scalar equation. It is not claimed, however, that time stability for the system (2.1) is equivalent to time stability for the scalar problem.

Led by the above example, we now make the conditions precise under which the behavior of the difference schemes shall be analyzed. Let (1.1) be given with smooth functions  $f(u)$ ,  $b(x, u)$  and assume

A1. For some  $u^* \in \mathbb{R}$ ,  $a(u) \equiv f'(u) \geq 0$  holds for  $u \geq u^*$ .

A2.  $\gamma_0 > u^* > \gamma_1$ .

A3. The solutions  $U_l$  and  $U_r$  of

$$\begin{aligned} U_l(0) &= \gamma_0, & f(U_l)' + b(x, U_l) &= 0, \\ U_r(1) &= \gamma_1, & f(U_r)' + b(x, U_r) &= 0 \end{aligned}$$

exist on  $[0, 1]$  with

$$U_l(x) > u^* > U_r(x), \quad 0 \leq x \leq 1.$$

Condition A2 means **supersonic** inflow and **subsonic** outflow in the gasflow example. Condition A3 can be relaxed, but it is important that  $U_l$  and  $U_r$  have a common interval of existence; otherwise phenomena like corner layers can appear, which will not be analyzed in this paper. Under the above conditions a crucial quantity is (see [5])

$$J(x) = f(U_l(x)) - f(U_r(x)), \quad 0 \leq x \leq 1.$$

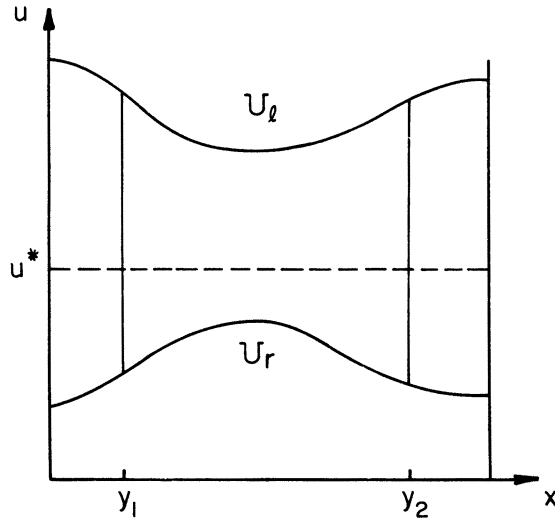


FIGURE 1

If  $J(x) \geq 0$  in  $[0, 1]$ , then  $U = U_l$  solves (1.1) and if  $J(x) \leq 0$  in  $[0, 1]$ , then  $U = U_r$  solves (1.1), as is easily seen by checking condition (iii) of Definition 1. If  $J(y) = 0$  for some  $0 < y < 1$ , then the discontinuous function

$$(2.3) \quad U(x) = \begin{cases} U_l(x), & x < y, \\ U_r(x), & x \geq y, \end{cases}$$

solves (1.1). In the gasflow problem for a converging-diverging duct the function  $J(x)$  often has two zeros  $y_1$  and  $y_2$  with

$$A'(y_1) < 0 < A'(y_2).$$

This leads to *three* solutions of (1.1): one solution equals  $U_l(x)$  for  $x < y_1$  and  $U_r(x)$  for  $x \geq y_1$ ; the second equals  $U_l(x)$  for  $x < y_2$  and  $U_r(x)$  for  $x \geq y_2$ . These are sketched in Figure 1. There is a third solution which equals  $U_r(x)$  for  $0 \leq x \leq 1$ . It corresponds to a solution with a boundary layer at  $x = 0$  for problem (1.3) and is considered irrelevant here, since it contradicts our assumption of supersonic inflow. The observation about two (or more) shock solutions for the duct flow problem is due to [3]. With the interesting idea used in [3] for the model problem

$$u_t + \frac{1}{2}uu_x = a(x)u,$$

one can show that any shock solution of (2.2) with a shock at  $y_1$ ,  $A'(y_1) < 0$ , is unstable as a solution of (1.4).

**3. Analysis of the EO- and the G-scheme.** On a uniform mesh with step-size  $h = 1/(m + 1)$ , the discrete equations for (1.1) read

$$(3.1) \quad \begin{aligned} (T^h u)_i &\equiv \frac{1}{h} g(u_{i+1}, u_i) - g(u_i, u_{i-1}) + b(ih, u_i), & i = 1, \dots, m, \\ (T^h u)_0 &\equiv u_0 = \gamma_0, & (T^h u)_{m+1} &\equiv u_{m+1} = \gamma_1. \end{aligned}$$

For the G-scheme,  $g$  is given in (1.6); for the EO-scheme,  $g$  is given in (1.7). Under the assumption A1 of Section 2 the functions  $g$  simplify. Without loss of generality, we take  $c = u^*$  in (1.7) and assume  $f(u^*) = 0$ . Then

$$g^{\text{EO}}(u, v) = \begin{cases} f(v), & u \geq u^*, v \geq u^*, \\ 0, & u \geq u^*, v < u^*, \\ f(u) + f(v), & u < u^*, v \geq u^*, \\ f(u), & u < u^*, v < u^*, \end{cases}$$

$$g^G(u, v) = \begin{cases} \max\{f(u), f(v)\}, & u < u^*, v \geq u^*, \\ g^{\text{EO}}(u, v), & \text{otherwise.} \end{cases}$$

Our main convergence result for the EO-scheme is:

**THEOREM 1.** *Let conditions A1 to A3 of Section 2 hold and let  $J(x) \equiv f(U_l(x)) - f(U_r(x))$  vanish at some  $y \in (0, 1)$  with  $J'(y) < 0$ . Especially, the function  $U(x)$  given in (2.3) solves (1.1). Then for  $h \leq h_0$  the EO-system has a solution  $u^h = (u_i^h)$  converging to  $U$  with the following estimates: For some  $C$  independent of  $h$  and an index  $j = j(h)$*

$$\begin{aligned} |U_l(ih) - u_i^h| &\leq Ch, & 0 \leq i \leq j - 1, \\ |U_r(ih) - u_i^h| &\leq Ch, & j + 2 \leq i \leq m + 1, u_j^h \geq u^* \geq u_{j+1}^h, \\ |j(h)h - y| &\leq Ch \end{aligned}$$

holds.

In our proof we apply the method of upper and lower solutions to a mapping  $\phi: \mathbb{R}^2 \rightarrow \mathbb{R}^2$ . Since this useful principle is crucial for our arguments we first establish it for completeness. For vectors  $v, w \in \mathbb{R}^k$  define

$$v \leq w \Leftrightarrow v_i \leq w_i \quad \text{for } i = 1, \dots, k.$$

If  $\phi: \mathbb{R}^k \rightarrow \mathbb{R}^k$  is a nonlinear operator, then  $v \in \mathbb{R}^k$  is called a lower solution of the equation  $\phi u = 0$  if  $\phi v \leq 0$ . Similarly,  $\phi w \geq 0$  means that  $w$  is an upper solution. The following lemma is essential.

**LEMMA 1.** *Let  $\phi: \mathbb{R}^k \rightarrow \mathbb{R}^k$  be continuous and outer-diagonally decreasing, i.e., for all  $j \neq i$  the scalar functions*

$$u_i \rightarrow \phi_j(u_1, \dots, u_i, \dots, u_k)$$

*are monotonically decreasing for all  $u_1, \dots, u_{i-1}, u_{i+1}, \dots, u_k$  fixed. If there exist  $v, w \in \mathbb{R}^k$  with  $v \leq w$ ,  $\phi v \leq 0 \leq \phi w$ , then the equation  $\phi u = 0$  has a solution  $\bar{u}$  with  $v \leq \bar{u} \leq w$ .*

*Proof.* Let  $L = \{u \in \mathbb{R}^k: \phi u \leq 0, v \leq u \leq w\}$  denote the set of all lower solutions between  $v$  and  $w$ . For  $i = 1, \dots, k$  define componentwise

$$\bar{u}_i = \sup\{u_i: u \in L\}.$$

It is not difficult to see that  $\phi \bar{u} = 0$ .

*Proof of Theorem 1.* 1. First define discrete values  $u_{il} = u_{il}^h$ ,  $u_{ir} = u_{ir}^h$  in analogy to the continuous functions  $U_l$ ,  $U_r$  by the two-point relations (upwinding):

$$\begin{aligned} u_{0l} = \gamma_0, \quad f(u_{il}) - f(u_{i-1,l}) + hb(ih, u_{il}) &= 0, & i = 1, 2, \dots, m + 1; \\ u_{m+1,r} = \gamma_1, \quad f(u_{i+1,r}) - f(u_{ir}) + hb(ih, u_{ir}) &= 0, & i = m, m - 1, \dots, 0. \end{aligned}$$

If  $h \leq h_0$  it is clear that these values are uniquely defined with  $u_{il} > u^* > u_{ir}$  and

$$\max_i |U_l(ih) - u_{il}| \leq Ch, \quad \max_i |U_r(ih) - u_{ir}| \leq Ch.$$

2. The basic idea of the proof is the following: Fix  $h \leq h_0$  and consider for fixed  $j \in \{1, \dots, m-1\}$  the  $2 \times 2$  system (with  $g$  given in (1.7)):

$$(3.2) \quad \begin{cases} g(u_{j+1}, u_j) - g(u_j, u_{j-1,l}) + hb(jh, u_j) = 0, \\ g(u_{j+2,r}, u_{j+1}) - g(u_{j+1}, u_j) + hb((j+1)h, u_{j+1}) = 0 \end{cases}$$

for  $(u_j, u_{j+1}) \in \mathbb{R}^2$ . Suppose we can find a  $j$  so that this system has a solution  $(\bar{u}_j, \bar{u}_{j+1})$  with

$$(3.3) \quad \bar{u}_j \geq u^* \geq \bar{u}_{j+1}.$$

Then the grid function

$$(3.4) \quad u_i = \begin{cases} u_{il}, & i \leq j-1, \\ \bar{u}_i, & i = j, j+1, \\ u_{ir}, & i \geq j+2, \end{cases}$$

solves the EO-system! Thus the crucial question is: For which  $j$  has the  $2 \times 2$  system (3.2) a solution with (3.3)? We define an  $O(h)$ -approximation to  $J(x) = f(U_l(x)) - f(U_r(x))$  by

$$J^h(i) = f(u_{i-1,l}) - f(u_{i+1,r}) - hb(ih, u^*).$$

Since  $J(y) = 0$  and  $J'(y) < 0$ , there is an index  $j = j(h)$  with  $|j(h)h - y| = O(h)$  and  $J^h(j) \geq 0 \geq J^h(j+1)$ . We claim that for such a  $j$  the  $2 \times 2$  system

$$(3.5) \quad \begin{aligned} \phi_1(p, q) &\equiv g(q, p) - g(p, u_{j-1,l}) + hb(jh, p) = 0, \\ \phi_2(p, q) &\equiv g(u_{j+2,r}, q) - g(q, p) + hb((j+1)h, q) = 0 \end{aligned}$$

has a solution  $(\bar{p}, \bar{q})$  with  $\bar{p} \geq u^* \geq \bar{q}$ .

3. First note that  $\phi$  is outer-diagonally decreasing, i.e.,  $D_2\phi_1 \leq 0$ ,  $D_1\phi_2 \leq 0$ , since  $D_1g \leq 0 \leq D_2g$ . Thus the method of upper and lower solutions can be applied to the system  $\phi(p, q) = 0$ . Now note  $u_{jl} \geq u^* \geq u_{j+1,r}$  and thus

$$\begin{aligned} \phi_1(u^*, u_{j+1,r}) &= g(u_{j+1,r}, u^*) - g(u^*, u_{j-1,l}) + hb(jh, u^*) \\ &= f(u_{j+1,r}) - f(u_{j-1,l}) + hb(jh, u^*) = -J_h(j) \leq 0, \\ \phi_2(u^*, u_{j+1,r}) &= g(u_{j+2,r}, u_{j+1,r}) - g(u_{j+1,r}, u^*) + hb((j+1)h, u_{j+1,r}) \\ &= f(u_{j+2,r}) - f(u_{j+1,r}) + hb((j+1)h, u_{j+1,r}) = 0, \end{aligned}$$

and, similarly,

$$\phi_1(u_{jl}, u^*) = 0, \quad \phi_2(u_{jl}, u^*) = -J^h(j+1) \geq 0.$$

Therefore, there is  $(\bar{p}, \bar{q})$  with  $\phi(\bar{p}, \bar{q}) = 0$  and

$$u^* \leq \bar{p} \leq u_{jl}, \quad u_{j+1,r} \leq \bar{q} \leq u^*.$$

Especially,  $(\bar{u}_j, \bar{u}_{j+1}) = (\bar{p}, \bar{q})$  satisfies (3.3), which completes the proof.  $\square$

What changes if we consider the G-scheme instead of the EO-scheme in a situation as described in Theorem 1? Essentially the same convergence result holds, but the layer becomes even sharper. The EO-values and the G-values coincide exactly except at two mesh-points.

**THEOREM 2.** *Under the conditions of Theorem 1 for  $h \leq h_0$  the G-system has a solution  $u^h = (u_i^h)$  converging to  $U$  with the following estimates: For some  $C$  independent of  $h$  and an index  $k = k(h)$*

$$\begin{aligned} |U_l(ih) - u_i^h| &\leq Ch, & 0 \leq i \leq k, \\ |U_r(ih) - u_i^h| &\leq Ch, & k + 2 \leq i \leq m + 1, \\ |k(h)h - y| &\leq Ch \end{aligned}$$

*holds. Furthermore, if  $(u_i^{\text{EO}})$  and  $(u_i^{\text{G}})$  denote the constructed solutions of the EO-system and the G-system, respectively, then*

$$(3.6) \quad u_i^{\text{EO}} = u_i^{\text{G}} \quad \text{for } j \neq i \neq j + 1,$$

where  $j = j(h)$  is specified in Theorem 1.

*Proof.* 1. Let  $(u_{i,l}), (u_{i,r}), J^h(i)$  and  $j = j(h)$  with  $|j(h)h - y| = O(h)$  and

$$J^h(j) \geq 0 \geq J^h(j + 1)$$

be constructed exactly as in the proof of Theorem 1. Consider again the  $2 \times 2$  system (3.2) where  $g$  is now the G-function (1.6). If (3.2) has a solution  $(\bar{u}_j, \bar{u}_{j+1})$  with (3.3), then again the grid function (3.4) solves the G-system. In (3.5) we rewrote (3.2) as

$$(3.7) \quad \phi(u_j, u_{j+1}) = \phi(p, q) = 0.$$

Again  $\phi: \mathbb{R}^2 \rightarrow \mathbb{R}^2$  is outer-diagonally decreasing, and

$$\phi(u^*, u_{j+1,r}) \leq (0, 0) \leq \phi(u_j, u^*).$$

The method of upper and lower solutions therefore establishes existence of a solution  $(\bar{u}_j, \bar{u}_{j+1})$  of (3.7) with

$$u^* \leq \bar{u}_j \leq u_j, \quad u_{j+1,r} \leq \bar{u}_{j+1} \leq u^*.$$

Since (3.3) is satisfied, we have shown existence of a solution  $(u_i^{\text{G}})$  of the G-system with

$$\begin{aligned} u_i^{\text{G}} &= u_{i,l} = u_i^{\text{EO}}, & i \leq j - 1, \\ u_i^{\text{G}} &= u_{i,r} = u_i^{\text{EO}}, & i \geq j + 2. \end{aligned}$$

2. By (3.7) we have for the G-values:

$$\begin{aligned} 0 &= \phi_1(u_j, u_{j+1}) = \max\{f(u_j), f(u_{j+1})\} - f(u_{j-1,l}) + hb(jh, u_j), \\ 0 &= \phi_2(u_j, u_{j+1}) = f(u_{j+2,r}) - \max\{f(u_j), f(u_{j+1})\} + hb((j+1)h, u_{j+1}). \end{aligned}$$

If  $f(u_{j+1}) \leq f(u_j)$ , then the equation  $0 = \phi_1(u_j, u_{j+1})$  reads

$$f(u_j) - f(u_{j-1,l}) + hb(jh, u_j) = 0,$$

and thus  $u_j = u_{j,l}$  by definition of  $(u_{i,l})$ . Similarly, if  $f(u_{j+1}) \geq f(u_j)$ , then the equation  $0 = \phi_2(u_j, u_{j+1})$  yields  $u_{j+1} = u_{j+1,r}$ . This shows that in both cases the shock is marked by only *one* mesh-point. Therefore, the above convergence result is proved with  $k = j - 1$  or  $k = j$ .  $\square$



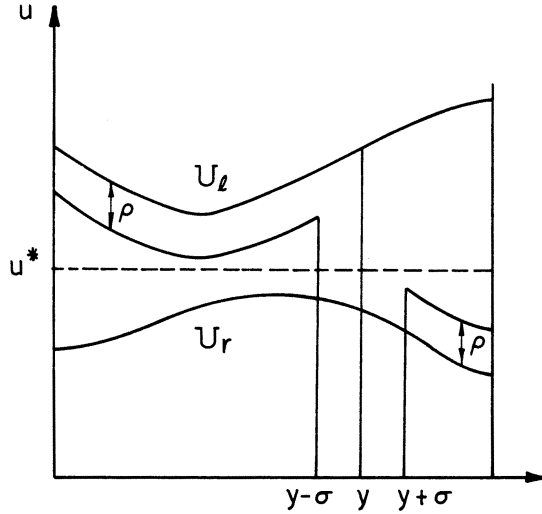


FIGURE 2

Essentially, Theorems 1 and 2 show convergence of order  $h$  in maximum norm and sharpness of the numerical layer for the EO-values and the G-values, if the shock is at a position  $y \in (0, 1)$  with  $J'(y) < 0$ . Since

$$\begin{aligned} J'(y) &= f(U_l(y))' - f(U_r(y))' \\ &= b(y, U_r(y)) - b(y, U_l(y)) \\ &= b_u(y, \xi)(U_r(y) - U_l(y)), \end{aligned}$$

and  $U_l(y) > U_r(y)$ , we find  $J'(y) < 0$  to be satisfied if the usual sign restriction  $b_u(x, u) > 0$  is fulfilled at the position  $x = y$  of the shock. In Theorems 1 and 2 we only stated *existence* of a solution  $u^h$  closed to  $U$ . If the condition  $b_u \geq 0$  does not hold globally, then global uniqueness cannot be expected, in general, as the duct flow problem shows. In order to formulate a local uniqueness result we define for  $\rho > 0, \sigma > 0, y \in (0, 1)$ , a set  $\Omega(h, \rho, \sigma, y) \subset \mathbb{R}^{m+2}$  of mesh functions  $u = (u_i)$  by

$$\begin{aligned} u \in \Omega(h, \rho, \sigma, y) \quad \text{iff} \quad & |U_l(ih) - u_i| \leq \rho \quad \text{for } ih \leq y - \sigma, \\ & |U_r(ih) - u_i| \leq \rho \quad \text{for } ih \geq y + \sigma. \end{aligned}$$

By the convergence result of Theorem 1 we know that for any  $\rho > 0, \sigma > 0$  there is  $h_0 > 0$  such that the EO- and the G-systems both have solutions  $u^h$  in  $\Omega(h, \rho, \sigma, y)$  for  $h \leq h_0$ . We now state a local uniqueness and stability result. For the conditions, see Figure 2.

**THEOREM 3.** *Let the conditions of Theorem 1 hold and assume  $\rho > 0, \sigma > 0$  are fixed such that*

$$\begin{aligned} \min_x U_l(x) - \rho &\geq u^*, & 0 \leq x \leq y - \sigma, \\ \max_x U_r(x) + \rho &\leq u^*, & y + \sigma \leq x \leq 1, \\ (3.8) \quad b_u(x, u) &> 0 & \text{for } y - 2\sigma \leq x \leq y + 2\sigma, u \in \mathbb{R}. \end{aligned}$$



*Extensions.* In [10, Theorem 3.1], we also gave convergence results for the EO-scheme in cases where

$$J(x) = f(U_l(x)) - f(U_r(x))$$

does not vanish in  $(0, 1)$ . Then interior discontinuities for a solution  $U$  of (1.1) are not possible, but  $U$  does, in general, not satisfy both boundary conditions. (For  $\epsilon > 0$  a boundary layer occurs.) The technique used in Theorem 1 of discussing a  $2 \times 2$  nonlinear system to get a convergence result can be used in all cases treated in [10, Theorem 3.1]. Thus one can drop the condition  $b(x, 0) \equiv 0$  which we needed in [10] to guarantee a *monotonic* discrete solution  $(u_i^h)$ . The convergence results are the same as those stated in [10, Theorem 3.1].

**4. Analysis of the  $O(h^2)$ -Modification of the EO-Scheme.** In [9], [10] a three-point modification of the EO-scheme was introduced where the source term  $b(x, u)$  is switched also. With  $g$  given by (1.7) the scheme reads

$$(4.1) \quad \begin{aligned} & \frac{1}{h} \{ g(u_{i+1}, u_i) - g(u_i, u_{i-1}) \} \\ & + \beta_{i-1}^- b(x_{i-1}, u_{i-1}) + (1 - \beta_i^- - \beta_i^+) b(x_i, u_i) \\ & + \beta_{i+1}^+ b(x_{i+1}, u_{i+1}) = 0, \quad i = 1, \dots, m, \\ & u_0 = \gamma_0, \quad u_{m+1} = \gamma_1. \end{aligned}$$

Here  $x_i = ih$  and the coefficients  $\beta_i^-, \beta_i^+$  are defined by

$$\beta_i^- = B(\kappa a(u_i)/\sqrt{h}), \quad \beta_i^+ = B(-\kappa a(u_i)/\sqrt{h}),$$

where  $a(u) \equiv f'(u)$ .

$B(r)$  is the  $C^1$ -function

$$B(r) = \begin{cases} 0 & \text{for } r < 0, \\ r^2 & \text{for } 0 \leq r \leq \frac{1}{2}, \\ \frac{1}{2} - (1 - r)^2 & \text{for } \frac{1}{2} \leq r \leq 1, \\ \frac{1}{2} & \text{for } r > 1 \end{cases}$$

connecting the values 0 and  $\frac{1}{2}$ .  $\kappa > 0$  is a parameter which can be taken independently of  $h$  as follows. Let  $c_0 \leq u \leq c_1$  denote an a priori region containing all  $u$ -values of interest. Furthermore, let

$$|b_u(x, u)| \leq M_1, \quad |a'(u)b(x, u)| \leq M_2 \quad \text{in } [0, 1] \times [c_0, c_1].$$

As shown in [9], [10], any choice of  $\kappa$  and  $h$  with

$$(4.2) \quad 0 < \kappa \leq 1/2\sqrt{M_2}, \quad h \leq 4M_2/M_1^2,$$

leads to an operator (4.1) which is outer-diagonally decreasing in the a priori domain.

In the next theorem, we assume

$$c_0 < U_r(x) < U_l(x) < c_1 \quad \text{for } 0 \leq x \leq 1,$$

and we let  $\kappa$  be fixed with (4.2).

**THEOREM 4.** *Under the conditions of Theorem 1 and for  $h$  sufficiently small the above scheme has a solution  $u^h = (u_i^h)$  converging to  $U$  with the following estimates: for some  $C$  independent of  $h$  and an index  $j = j(h)$*

$$\begin{aligned} |U_l(ih) - u_i^h| &\leq Ch^2, & 0 \leq i \leq j-1, \\ |U_r(ih) - u_i^h| &\leq Ch^2, & j+2 \leq i \leq m+1, \\ |u(h)h - y| &\leq Ch \end{aligned}$$

*holds. Furthermore, a local uniqueness result as in Theorem 3 is valid.*

*Proof.* The proof follows the same steps as the proof of Theorem 1. We only point out some differences. The values  $u_{il}$ ,  $u_{ir}$  are determined according to

$$\begin{aligned} u_{0l} = \gamma_0, \quad f(u_{il}) - f(u_{i-1,l}) + \frac{h}{2} \{b(x_{i-1}, u_{i-1,l}) + b(x_i, u_{il})\} &= 0, \\ & i = 1, 2, \dots, m+1; \\ u_{m+1,r} = \gamma_1, \quad f(u_{i+1,r}) - f(u_{ir}) + \frac{h}{2} \{b(x_i, u_{ir}) + b(x_{i+1}, u_{i+1,r})\} &= 0, \\ & i = m, m-1, \dots, 0. \end{aligned}$$

They are  $O(h^2)$ -close to  $U_l$  and  $U_r$ , respectively.

The  $2 \times 2$  system to be solved for  $(p, q) = (u_j, u_{j+1})$  reads

$$\phi_1(p, q) = \phi_2(p, q) = 0$$

with

$$\begin{aligned} \phi_1(p, q) &= g(q, p) - g(p, u_{j-1,l}) + \frac{h}{2} b(x_{j-1}, u_{j-1,l}) \\ &\quad + h \{1 - B(\kappa a(p)/\sqrt{h})\} b(x_j, p) \\ &\quad + h B(-\kappa a(q)/\sqrt{h}) b(x_{j+1}, q), \\ \phi_2(p, q) &= g(u_{j+2,r}, q) - g(q, p) + h B(\kappa a(p)/\sqrt{h}) b(x_j, p) \\ &\quad + h \{1 - B(-\kappa a(q)/\sqrt{h})\} b(x_{j+1}, q) \\ &\quad + \frac{h}{2} b(x_{j+2}, u_{j+2,r}). \end{aligned}$$

Again, one has to determine an index  $j = j(h)$  for which the system  $\phi(p, q) = 0$  has a solution  $(\bar{p}, \bar{q})$  with  $\bar{p} \geq u^* \geq \bar{q}$ . To find  $j = j(h)$ , we define

$$\begin{aligned} J^h(i) &= f(u_{i-1,l}) - f(u_{i+1,r}) \\ &\quad - \left\{ \frac{h}{2} b(x_{i-1}, u_{i-1,l}) + h b(x_i, u^*) + \frac{h}{2} b(x_{i+1}, u_{i+1,r}) \right\}. \end{aligned}$$

Then there exists  $j = j(h)$  such that

$$J^h(j) \geq 0 \geq J^h(j+1), \quad |j(h)h - y| = O(h).$$

For this  $j$  there are upper and lower solutions for  $\phi(p, q) = 0$ , namely  $\phi(u^*, u_{j+1,r}) \leq (0, 0) \leq \phi(u_j, u^*)$ .  $\phi$  is outer-diagonally decreasing; this follows by the same estimates which show that the operator (4.1) is outer-diagonally decreasing; see [9], [10]. Only here the upper bound for  $\kappa$  is crucial.

**Acknowledgment.** I would like to thank an unknown referee for helpful comments on the first draft of this paper.

Department of Applied Mathematics 217-50  
California Institute of Technology  
Pasadena, California 91125

1. C. BARDOS, A. Y. LEROUX AND J. C. NEDELEC, "First-order quasilinear equations with boundary conditions," *Comm. Partial Differential Equations*, v. 4, 1979, pp. 1017–1034.
2. M. G. CRANDALL & A. MAJDA, "Monotone difference approximations for scalar conservation laws," *Math. Comp.*, v. 34, 1980, pp. 1–21.
3. P. EMBID, J. GOODMAN & A. MAJDA, "Multiple steady states for 1-D transonic flow," *SIAM J. Sci. Statist. Comput.*, v. 5, 1984, pp. 21–41.
4. B. ENGQUIST & S. OSHER, "One-sided difference approximations for nonlinear conservation laws," *Math. Comp.*, v. 36, 1981, pp. 321–351.
5. F. A. HOWES, "Boundary-interior layer interactions in nonlinear singular perturbation theory," *Mem. Amer. Math. Soc.*, No. 203, 1978.
6. L. K. JACKSON, "Subfunctions and second-order ordinary differential inequalities," *Adv. in Math.*, v. 2, 1968, pp. 307–363.
7. H. W. LIEPMANN & A. ROSHKO, *Elements of Gas Dynamics*, Wiley, New York, 1957.
8. J. LORENZ, *Nonlinear Boundary Value Problems with Turning Points and Properties of Difference Schemes* (W. Eckhaus and E. M. de Jager, eds.), Springer Lecture Notes in Math., vol. 942, Springer-Verlag, Berlin and New York, 1982.
9. J. LORENZ, "Numerical solution of a singular perturbation problem with turning points," in *Equadiff 82* (H. W. Knobloch and K. Schmitt, eds.), Lecture Notes in Math., vol. 1017, Springer-Verlag, Berlin and New York, 1983.
10. J. LORENZ, "Analysis of difference schemes for a stationary shock problem," *SIAM J. Numer. Anal.*, v. 21, 1984, pp. 1038–1052.
11. M. NAGUMO, "Über die Differentialgleichung  $y'' = f(x, y, y')$ ," *Proc. Math. Soc. Japan*, v. 19, 1937, pp. 861–866.
12. O. A. OLEĬNIK, "Discontinuous solutions of non-linear differential equations," *Amer. Math. Soc. Transl. (2)*, v. 26, 1963, pp. 95–172.
13. S. OSHER, "Numerical solution of singular perturbation problems and hyperbolic systems of conservation laws," in *Analytical and Numerical Approaches to Asymptotic Problems in Analysis* (O. Axelsson, L. S. Frank and A. von der Sluis, eds.), North-Holland, Amsterdam, 1981.
14. S. OSHER, "Nonlinear singular perturbation problems and one sided difference schemes," *SIAM J. Numer. Anal.*, v. 18, 1981, pp. 129–144.
15. G. R. SHUBIN, A. B. STEPHENS & H. M. GLAZ, "Steady shock tracking and Newton's method applied to one-dimensional duct flow," *J. Comput. Phys.*, v. 39, 1981, pp. 364–374.
16. A. B. STEPHENS & G. R. SHUBIN, "Existence and uniqueness for an exponentially derived switching scheme," *SIAM J. Numer. Anal.*, v. 20, 1983, pp. 885–889.
17. A. B. STEPHENS & G. R. SHUBIN, *Exponentially Derived Switching Schemes for Inviscid Flow*, Applied Math. Branch, Naval Surface Weapons Center, Silver Spring, Maryland, Report, 1981.