

ELIMINATION OF FIRST ORDER ERRORS IN TIME DEPENDENT SHOCK CALCULATIONS*

MALIN SIKLOSI[†] AND GUNILLA KREISS[†]

Abstract. First order errors downstream of shocks have been detected in computations with higher order shock-capturing schemes in one and two dimensions. We use matched asymptotic expansions to analyze the phenomenon for one dimensional time dependent hyperbolic systems and show how to design the artificial viscosity term in order to avoid the first order error. Numerical computations verify that second order accurate solutions are obtained.

Key words. hyperbolic conservation laws, shock wave, artificial viscosity, asymptotic analysis

AMS subject classifications. 35L65, 35L67, 65M06, 65M02

DOI. 10.1137/S0036142902400457

1. Introduction. In many cases, solutions of conservation laws obtained by formally higher order methods are only first order accurate downstream of shocks; see, e.g., [2], [5], and [4]. Basically, errors from the shock region follow outgoing characteristics and pollute the solution downstream. Examples in one space dimension in which this effect can be seen are steady-state calculations for systems with a source term and time dependent calculations for systems with nonconstant solution. The effect cannot be seen in one dimensional Riemann problems, because the exact global conservation determines the postshock states.

This degeneration in accuracy is troublesome, even though the first order term for reasonable mesh-sizes seems to be small in many cases. In some applications, e.g., aeroacoustics, where waves with small amplitude need to be computed accurately, it is particularly important to achieve very high accuracy. It is also important to understand the phenomenon more deeply in order to be able to design new methods which do not suffer from this deficiency.

The aim of this paper is to show that the first order error can be understood by matched asymptotic analysis of the modified equation and that the analysis can be used to construct methods that yield second order accurate solutions.

We consider the case of systems with time dependent solutions. We assume that the numerical solution can be modeled by a slightly viscous equation, a so-called modified equation. In the shock layer, the coefficient of the viscous term is $\mathcal{O}(h)$, where h is the grid size. We analyze the solution of the modified equation using matched asymptotic expansions. It is assumed that an inner solution is valid in the shock region, and an outer solution is valid elsewhere. The two solutions are matched in a so-called matching zone. From the analysis, we see that generally the outer solution contains a term of $\mathcal{O}(h)$ downstream of the shock. We also see that if the inner solution satisfied a certain condition, the $\mathcal{O}(h)$ term would be eliminated. Based on this observation, we design a matrix valued viscosity coefficient, which gives the inner solution the right shape to eliminate the $\mathcal{O}(h)$ downstream term. We construct a numerical scheme, using this matrix valued viscosity coefficient, and show in numerical

*Received by the editors January 2, 2002; accepted for publication (in revised form) April 4, 2003; published electronically November 25, 2003.

<http://www.siam.org/journals/sinum/41-6/40045.html>

[†]Department of Numerical Analysis and Computer Science, Royal Institute of Technology, SE-100 44 Stockholm, Sweden (malins@nada.kth.se, gunillak@nada.kth.se).

experiments that the first order downstream error really is eliminated. However, we do not claim to have constructed an efficient and robust numerical method which can be used in realistic computations.

Similar analysis and construction of a matrix valued viscosity coefficient is done in [8] for the case of a steady-state solution of a system with a source term. In [3], matched asymptotic expansions for a problem that is very similar to the problem studied in this paper are analyzed for other purposes. The phenomenon has also been studied by other methods in [5] and [2]. In [5], analytic examples are constructed where the numerical solution is only first order accurate downstream of a shock, although the numerical scheme is formally second order. It is also shown that a converging numerical method will yield solutions having the formal order of accuracy in domains where no characteristics have passed through a shock. In [2], the first order downstream error is numerically detected in solutions of a shock-sound interaction problem solved by a fourth order ENO method. A scalar, linear equation is used to model the problem. It can be seen that the solution of the model problem computed with the fourth order ENO method behaves qualitatively differently depending on whether the discontinuity is located on a cell interface or in the interior of a cell. In the first case, the solution is fourth order in all of the domain, but in the second case the solution is only first order downstream of the discontinuity. Based on this observation, the numerical method is modified such that the shock position will always be on a cell interface, and fourth order accuracy of the solution of the shock-sound interaction problem is obtained both upstream and downstream. Also in [1], shock wave solutions are analyzed, and it is concluded that the structure in the shock region is of crucial importance for the solution outside the shock region. However, the analysis in [1] concerns another numerical phenomenon and considers methods where the shock is so narrow that it is not well modeled by the solution of a slightly viscous equation.

This paper is organized as follows. In section 2, we use asymptotic analysis to explain the first order downstream error and derive a matrix valued viscosity coefficient that eliminates it. In section 3, we implement a numerical method using the matrix valued viscosity coefficient and show in computations that the first order downstream error is eliminated.

2. Analysis.

2.1. The inviscid problem. Consider the inviscid problem

$$\begin{aligned} (1) \quad & \mathbf{u}_t + \mathbf{f}(\mathbf{u})_x = 0, \quad 0 \leq x \leq x_{\text{end}}, \\ (2) \quad & \mathbf{u}(x, 0) = \mathbf{g}(x), \end{aligned}$$

where $\mathbf{u}(x, t), \mathbf{g}(x) \in \mathbf{R}^n$, $\mathbf{f} : \mathbf{R}^n \rightarrow \mathbf{R}^n$, and $\mathbf{g}$ is a piecewise smooth function. We denote the Jacobian of the flux function $\mathbf{f}'(\mathbf{u})$ by $J(\mathbf{u})$. We assume that the eigenvalues of $J(\mathbf{u})$, denoted $\lambda_i(\mathbf{u}), i = 1, 2, \dots, n$, are real and ordered in increasing order and that the eigenvectors span $\mathbf{R}^n$.

The initial and boundary conditions are chosen such that a shock forms at some inner point $s(t)$. At the shock, the solution satisfies the Rankine–Hugoniot condition

$$\dot{s}[\mathbf{u}] = [\mathbf{f}(\mathbf{u})].$$

Here $[\mathbf{u}] = \mathbf{u}^+ - \mathbf{u}^-$, where $\mathbf{u}^\pm = \lim_{\delta \rightarrow 0^+} \mathbf{u}(s(t) \pm \delta, t)$. Corresponding notation for other quantities will be used frequently.

We assume that the shock is a classical Lax 1-shock, i.e.,

$$\begin{aligned}\dot{s} &< \lambda_1^-, \\ \lambda_1^+ &< \dot{s} < \lambda_2^+, \end{aligned}$$

and that the matrix

$$(3) \quad D = \left(\begin{array}{c|c} S_{II}^+ & [\mathbf{u}] \end{array} \right)$$

is nonsingular. Here the columns of S_{II}^+ are the eigenvectors of J^+ corresponding to the eigenvalues $\lambda_2^+, \lambda_3^+, \dots, \lambda_n^+$.

To complete the problem we also need boundary conditions. At each boundary we need as many boundary conditions as there are ingoing characteristics. We consider pointwise boundary conditions, i.e., boundary conditions where the quantities involved are prescribed pointwise at the boundary to some function of time. One example of such boundary conditions is when the ingoing characteristic variables are prescribed as a function of time. We call these boundary conditions mathematical boundary conditions to distinguish them from numerical boundary conditions. For more details concerning mathematical boundary conditions for hyperbolic equations, we refer to [9].

Remark. For 1-shocks and n -shocks there is just one downstream side. Hence, the first order error appears on only one side of the shock. For other Lax shocks, both sides of the shock are downstream sides, and first order errors appear on both sides. The phenomenon can be analyzed by the same method in both cases, but the analysis becomes less involved when only one side must be considered. Hence, here we analyze a 1-shock.

2.2. The slightly viscous model. We intend to study the behavior of numerical solutions of (1); i.e., we want to study the behavior of discrete functions that are the solutions of difference equations. A useful technique for studying the behavior of solutions to difference equations is to model the difference equation by a differential equation. Such a differential equation is often called a modified equation; see, e.g., [11], [6]. Many numerical solutions of (1) can be viewed as higher order accurate solutions of the modified equation

$$\mathbf{u}_t^\varepsilon + \mathbf{f}(\mathbf{u}^\varepsilon)_x = (\Gamma \mathbf{u}_x^\varepsilon)_x, \quad 0 \leq x \leq x_{\text{end}}.$$

In the shock region, the modified equation can be shown to be valid only for weak shocks; see, e.g., [7]. However, our computations indicate that it applies also for strong shocks. In the neighborhood of a shock layer we must have $\Gamma = \mathcal{O}(h)$, where h is the grid size, in order to avoid oscillations in the solution. Outside the shock region, Γ can be smaller. In this paper we consider methods which can be modeled by

$$(4) \quad \mathbf{u}_t^\varepsilon + \mathbf{f}(\mathbf{u}^\varepsilon)_x = \varepsilon(\phi \mathbf{u}_x^\varepsilon)_x + c_2 \varepsilon^2 \mathbf{u}_{xx}^\varepsilon,$$

where $\varepsilon = c_1 h$ and c_1 and c_2 a scalar constants. Here ϕ is a smooth function of $(x - s(t))/\varepsilon$ satisfying

$$\phi\left(\frac{x - s(t)}{\varepsilon}\right) = \begin{cases} 1 & \text{for } \left|\frac{x - s(t)}{\varepsilon}\right| \leq K_0, \\ 0 & \text{for } \left|\frac{x - s(t)}{\varepsilon}\right| \geq K_1, \end{cases}$$

where $K_0 < K_1$ are constants with K_0 sufficiently large.

We must also model the initial data. In computations, the shape of the shock profile will depend on the method. If the initial data does not have exactly the right shape, the profile will after a short time adjust and obtain the right shape. In this process, small diffusion waves appear and flow out of the shock region, following the outgoing characteristics; see [12] and the references therein. We are not interested in studying this initial effect, and consequently we assume that the initial profile is exactly the right profile for the method that is modeled. We specify the initial profile in (12) and (13).

We consider the same mathematical boundary conditions for $\mathbf{u}^\varepsilon$ as for $\mathbf{u}$. When (1) is solved numerically, the mathematical boundary conditions must be augmented by numerical boundary conditions. Correspondingly, additional boundary conditions that model the numerical boundary conditions are needed for $\mathbf{u}^\varepsilon$. Numerical boundary conditions can introduce boundary layers in the solution. We consider numerical boundary conditions where such effects are $\mathcal{O}(h^2)$ or smaller, e.g., extrapolation of outgoing characteristic variables.

We define the position of the viscous shock layer as the smallest x -value such that $\mathbf{u}^{\varepsilon(1)}(x, t) = (\mathbf{u}^{-(1)} + \mathbf{u}^{+(1)})/2$, and denote this point by x_ε ; i.e., the viscous shock position is defined as the point where the first component of the viscous solution $\mathbf{u}^\varepsilon$ is halfway between the right and left states in the corresponding inviscid shock.

2.3. Asymptotic expansions. We assume the following: The solution of (4) can be described by an inner solution, valid in the shock layer, and an outer solution, valid elsewhere. These solutions can be expanded in powers of ε and matched in a region of overlap. Also, the position of the shock layer can be expanded in ε . To leading order, the outer solution is equal to the solution of the corresponding inviscid problem.

We will now show that the outer solution downstream of the shock contains an $\mathcal{O}(h)$ term; i.e., downstream, the solution of (4) is just a first order approximation of the solution of the corresponding inviscid problem (1). There is no $\mathcal{O}(h)$ term upstream.

The inner solution is expressed using the variables $(\tilde{x}, \tilde{t})$, where

$$\begin{aligned}\tilde{x} &= \frac{x - s(t)}{\varepsilon}, \\ \tilde{t} &= t.\end{aligned}$$

Thus we have expansions of the form

$$\begin{aligned}(5) \quad \text{Outer:} \quad \mathbf{u}^\varepsilon &\sim \mathbf{u}(x, t) + \varepsilon \mathbf{u}_1(x, t) + \varepsilon^2 \mathbf{u}_2(x, t) + \cdots, \\ \text{Inner:} \quad \mathbf{u}^\varepsilon &\sim \mathbf{U}_0(\tilde{x}, \tilde{t}) + \varepsilon \mathbf{U}_1(\tilde{x}, \tilde{t}) + \varepsilon^2 \mathbf{U}_2(\tilde{x}, \tilde{t}) + \cdots, \\ (6) \quad \text{Position:} \quad x_\varepsilon &\sim s(t) + \varepsilon x_1(t) + \varepsilon^2 x_2(t) + \cdots.\end{aligned}$$

In [3], analysis of the asymptotic expansions for a very similar problem is presented, and also the existence of an asymptotic expansion is treated. For a detailed presentation of matched asymptotic expansions, we refer to [10].

We will match the inner and the outer solutions at an upstream and a downstream matching point, $x_m^-(t)$ and $x_m^+(t)$. The matching points must satisfy $\lim_{\varepsilon \rightarrow 0} |x_m^\pm - s| = 0$. We will also need $e^{\mp \tilde{x}_m^\pm} = \mathbf{o}(1)$. Choosing $x_m^\pm = s \mp \varepsilon \log(\varepsilon)$, we have $e^{\mp \tilde{x}_m^\pm} = \mathcal{O}(\varepsilon)$, and both requirements are satisfied.

The viscous problem (4) models a method which is a second order accurate approximation of (1) away from the shock region. We claim that the solution will be

second order accurate upstream of the shock, but in general only first order downstream. Hence we must show that $\mathbf{u}_1 = 0$ upstream and $\mathbf{u}_1 \neq 0$ downstream. To do this we need equations, initial data, and boundary conditions for $\mathbf{u}_1$. Via the boundary conditions in the shock region, the outer solution will be coupled to the inner solution. Specifically, to derive boundary conditions for $\mathbf{u}_1$ we need information about $\mathbf{U}_0$. Hence, we derive equations and boundary conditions also for $\mathbf{U}_0$.

To obtain equations for the terms in the outer and inner expansions we substitute the expansions into (4), Taylor expand, and collect terms multiplying the same power of ε . The equation for $\mathbf{U}_0$ is

$$(7) \quad (\phi \mathbf{U}_{0\tilde{x}})_{\tilde{x}} + \dot{s} \mathbf{U}_{0\tilde{t}} - \mathbf{f}(\mathbf{U}_0)_{\tilde{x}} = 0, \quad -\infty < \tilde{x} < \infty,$$

where we have used that the relations between derivatives in x and t and derivatives in $\tilde{x}$ and $\tilde{t}$ are

$$\begin{aligned} \frac{\partial}{\partial x} &= \frac{1}{\varepsilon} \frac{\partial}{\partial \tilde{x}}, \\ \frac{\partial}{\partial t} &= -\frac{\dot{s}}{\varepsilon} \frac{\partial}{\partial \tilde{x}} + \frac{\partial}{\partial \tilde{t}}. \end{aligned}$$

The inner and outer expansions of $\mathbf{u}^\varepsilon$ are assumed to be valid in a region of overlap containing the matching points $x_m^\pm$. Hence, in the matching points, we must have $\lim_{\varepsilon \rightarrow 0} |\mathbf{U}_0 - \mathbf{u}| = 0$, i.e.,

$$\lim_{\varepsilon \rightarrow 0} |\mathbf{U}_0(\mp \log(\varepsilon), \tilde{t}) - \mathbf{u}(s \mp \varepsilon \log(\varepsilon), t)| = 0,$$

where we have used that $x_m^\pm = s \mp \varepsilon \log(\varepsilon)$. Evaluating the limit, we arrive at the matching conditions

$$(8) \quad \mathbf{U}_0(\pm\infty, \tilde{t}) = \mathbf{u}^\pm(t),$$

where $\mathbf{U}_0(\pm\infty, \tilde{t}) = \lim_{\tilde{x} \rightarrow \pm\infty} \mathbf{U}_0(\tilde{x}, \tilde{t})$. Note that (7) and (8) determine the shape of $\mathbf{U}_0$ but not the exact position of the shock layer.

We define $\hat{\mathbf{U}}(\tilde{x}, \tilde{t})$ by

$$(9) \quad \hat{\mathbf{U}}_{\tilde{x}\tilde{x}} + \dot{s} \hat{\mathbf{U}}_{\tilde{t}} - \mathbf{f}(\hat{\mathbf{U}})_{\tilde{x}} = 0, \quad -\infty < \tilde{x} < \infty,$$

$$(10) \quad \hat{\mathbf{U}}(\tilde{x}, \tilde{t}) = \mathbf{u}(s \pm 0, \tilde{t}) \quad \text{as } \tilde{x} \rightarrow \pm\infty,$$

$$(11) \quad \hat{\mathbf{U}}^{(1)}(0, \tilde{t}) = (\mathbf{u}^{-(1)} + \mathbf{u}^{+(1)})/2.$$

We see that $\hat{\mathbf{U}}$ differs from $\mathbf{U}_0$ in two ways. First, $\hat{\mathbf{U}}$ is independent of ϕ , which makes the equation for $\hat{\mathbf{U}}$ much easier to analyze. Second, the position of $\hat{\mathbf{U}}$ is fixed at $\tilde{x} = 0$. We note that both problems are independent of ε .

Let us first compare the shape of solutions of (7) and (8) with the shape of solutions of (9) and (10), disregarding the difference in shock position. It is easy to show that $\hat{\mathbf{U}}$ approaches its limit values exponentially fast as $\tilde{x} \rightarrow \pm\infty$. If one constructs the equations for the difference $\hat{\mathbf{U}} - \mathbf{U}_0$ and uses the exponential behavior of $\hat{\mathbf{U}}$, one can conclude that $\mathbf{U}_0$ also approaches its limit values exponentially fast and that $|\hat{\mathbf{U}} - \mathbf{U}_0|_\infty < e^{-K}$, where K is a large constant.

Since the position of the shock layer has the expansion (6), we have, except for exponentially small terms, to leading order

$$\mathbf{U}_0(\tilde{x}, \tilde{t}) = \hat{\mathbf{U}}(\tilde{x} - x_1(\tilde{t}), \tilde{t}).$$

Below we will derive an ordinary differential equation for $x_1(\tilde{t})$. The initial value of $x_1(\tilde{t})$ is determined by the initial condition $\mathbf{g}^\varepsilon$, which we now specify:

$$(12) \quad \text{Outer region: } \mathbf{g}^\varepsilon(x) = \mathbf{g}(x),$$

$$(13) \quad \text{Inner region: } \mathbf{g}^\varepsilon(\tilde{x}) = \hat{\mathbf{U}}(\tilde{x}, 0).$$

This is sufficient for our purposes. However, if one considers more terms in the inner expansion, one would have to add the corresponding terms to (13). Note that (13) means that $x_1(0) = 0$.

The equation for $\mathbf{u}_1$ is

$$(14) \quad \mathbf{u}_{1t} + (\mathbf{f}'(\mathbf{u})\mathbf{u}_1)_x = 0, \quad x \in \text{outer region},$$

where we have used that $\phi = 0$ in the outer region. We also need initial data and boundary conditions for $\mathbf{u}_1$. The initial conditions for $\mathbf{u}^\varepsilon$, (12), gives $\mathbf{u}_1(x, 0) = 0$. Since $\mathbf{u}^\varepsilon$ and $\mathbf{u}$ satisfy the same mathematical boundary conditions and since boundary layer effects, due to numerical boundary conditions, are assumed to be $\mathcal{O}(h^2)$ or smaller, we conclude that all boundary conditions for $\mathbf{u}_1$ at $x = 0$ and $x = x_{\text{end}}$ are homogeneous. At the upstream side of the shock, no further boundary conditions are needed since all characteristics of (14) are going into the shock. Since $\mathbf{u}_1$ in the upstream region is the solution of a homogeneous equation with homogeneous initial data and homogeneous boundary conditions, we have $\mathbf{u}_1 \equiv 0$ in the upstream region. To determine $\mathbf{u}_1$ in the downstream region, we also need boundary conditions at $x = s^+$. We derive such boundary conditions in the next section.

2.4. Downstream boundary condition for the first order outer term.

Integration of the viscous equation (4) over the shock layer, from matching point x_m^- to matching point x_m^+ , gives

$$(15) \quad \int_{x_m^-}^{x_m^+} \mathbf{u}_t^\varepsilon dx + [\mathbf{f}(\mathbf{u}^\varepsilon)]_{x_m^-}^{x_m^+} = \mathcal{O}(\varepsilon^2),$$

where we have used that ϕ vanishes in the matching regions. Using the outer expansion of $\mathbf{u}^\varepsilon$, we obtain

$$(16) \quad [\mathbf{f}(\mathbf{u}^\varepsilon)]_{x_m^-}^{x_m^+} = [\mathbf{f}(\mathbf{u})]_{x_m^-}^{x_m^+} + \varepsilon[J(\mathbf{u})\mathbf{u}_1]_{x_m^-}^{x_m^+} + \mathcal{O}(\varepsilon^2).$$

By integrating the inviscid (1) over the same interval, we obtain

$$(17) \quad [\mathbf{f}(\mathbf{u})]_{x_m^-}^{x_m^+} = \dot{s}[\mathbf{u}] - \int_{x_m^-}^{s^-} \mathbf{u}_t dx - \int_{s^+}^{x_m^+} \mathbf{u}_t dx.$$

Note that $\mathbf{u}$ is discontinuous at $x = s(t)$ and the Rankine–Hugoniot condition applies across the discontinuity. After taking into account that $\mathbf{u}_1 \equiv 0$ to the left of the shock layer and introducing (16) and (17) into (15), we arrive at

$$(18) \quad \dot{s}[\mathbf{u}] + \varepsilon J(\mathbf{u}(x_m^+, t))\mathbf{u}_1(x_m^+, t) + I_1 = \mathcal{O}(\varepsilon^2),$$

where we have introduced the notation

$$I_1 = \int_{x_m^-}^{s^-} (\mathbf{u}_t^\varepsilon - \mathbf{u}_t) dx + \int_{s^+}^{x_m^+} (\mathbf{u}_t^\varepsilon - \mathbf{u}_t) dx.$$

After Taylor expansion of $\mathbf{u}$ and $\mathbf{u}_1$ around $x = s^+$, (18) can be rewritten as

$$(19) \quad \dot{s}[\mathbf{u}] + \varepsilon J(\mathbf{u}(s^+, t))\mathbf{u}_1(s^+, t) + I_1 = \mathbf{o}(\varepsilon).$$

In the coordinate system $(\tilde{x}, \tilde{t})$ we have

$$I_1 = -\dot{s}A + \varepsilon I_2,$$

where

$$A = \int_{\tilde{x}_m^-}^{0^-} (\mathbf{u}^\varepsilon - \mathbf{u})_{\tilde{x}} d\tilde{x} + \int_{0^+}^{\tilde{x}_m^+} (\mathbf{u}^\varepsilon - \mathbf{u})_{\tilde{x}} d\tilde{x},$$

$$I_2 = \int_{\tilde{x}_m^-}^{0^-} (\mathbf{u}^\varepsilon - \mathbf{u})_{\tilde{t}} d\tilde{x} + \int_{0^+}^{\tilde{x}_m^+} (\mathbf{u}^\varepsilon - \mathbf{u})_{\tilde{t}} d\tilde{x}.$$

Evaluating the integral yields

$$A = [\mathbf{u}] + [\mathbf{u}^\varepsilon - \mathbf{u}]_{\tilde{x}_m^+}^{\tilde{x}_m^-}.$$

By using the outer expansion of $\mathbf{u}^\varepsilon$, taking into account that $\mathbf{u}_1$ is zero upstream and Taylor expanding $\mathbf{u}_1$ around $x = s^+$, we obtain

$$A = [\mathbf{u}] + \varepsilon \mathbf{u}_1^+ + \mathbf{o}(\varepsilon).$$

Next, consider I_2 . Using the inner expansion of $\mathbf{u}^\varepsilon$, the Taylor expansion of $\mathbf{u}$ around $x = s \pm 0$, and $\mathbf{U}_0(\tilde{x}, \tilde{t}) = \hat{\mathbf{U}}(\tilde{x} - x_1, \tilde{t})$, we obtain

$$I_2 = \int_{\tilde{x}_m^-}^0 (\hat{\mathbf{U}}(\tilde{x} - x_1, \tilde{t}) - \mathbf{u}^-)_{\tilde{t}} d\tilde{x} + \int_0^{\tilde{x}_m^+} (\hat{\mathbf{U}}(\tilde{x} - x_1, \tilde{t}) - \mathbf{u}^+)_{\tilde{t}} d\tilde{x} + \mathbf{o}(1).$$

We rewrite I_2 in two steps. First we make the substitution $\hat{x} = \tilde{x} - x_1$. Next, we use the fact that $\hat{\mathbf{U}}$ approaches the limit values exponentially fast, and the matching points are chosen such that $e^{\mp \tilde{x}_m^\pm} = \mathcal{O}(\varepsilon)$. Hence we can extend the integration interval to infinity, still keeping the remainder term $\mathbf{o}(1)$. We obtain

$$I_2 = I_{3\tilde{t}} - (x_1[\mathbf{u}])_{\tilde{t}} + \mathbf{o}(1),$$

where

$$I_3(\tilde{t}) = \int_{-\infty}^0 (\hat{\mathbf{U}}(\tilde{x}, \tilde{t}) - \mathbf{u}^-) d\tilde{x} + \int_0^{\infty} (\hat{\mathbf{U}}(\tilde{x}, \tilde{t}) - \mathbf{u}^+) d\tilde{x}.$$

Since I_2 , I_3 , x_1 , and $[\mathbf{u}]$ are functions of $\tilde{t}$ only, and since $\tilde{t} = t$, this can be written as

$$I_2(t) = \frac{\partial}{\partial t} (I_3(t) - x_1(t)[\mathbf{u}](t)) + \mathbf{o}(1).$$

Hence we have

$$I_1 = -\dot{s}[\mathbf{u}] + \varepsilon(-\dot{s}\mathbf{u}_1^+ + I_{3t} - (x_1[\mathbf{u}])_t) + \mathbf{o}(\varepsilon).$$

Substituting this into (19) and rearranging, we obtain

$$(J^+ - \dot{s}I)\mathbf{u}_1^+ - (x_1[\mathbf{u}])_t + I_{3t} = \mathbf{o}(1).$$

Hence the equations for $\mathbf{u}_1^+$ and $x_1(t)$ are

$$(20) \quad (J^+ - \dot{s}I)\mathbf{u}_1^+(t) - (x_1(t)[\mathbf{u}])_t = -I_{3t},$$

$$(21) \quad x_1(0) = 0.$$

The two equations (20) and (21) constitute the boundary conditions for $\mathbf{u}_1$ at $x = s^+$. To make (20) and (21) easier to understand, we rewrite them using the characteristic variables of $\mathbf{u}_1$. Let $\mathbf{w}_I$ be the characteristic variable of $\mathbf{u}_1$ going into the shock, and $\mathbf{w}_{II}$ be the characteristic variables going out of the shock. We then have

$$\mathbf{u}_1^+ = (S_I^+ S_{II}^+) \begin{pmatrix} \mathbf{w}_I^+ \\ \mathbf{w}_{II}^+ \end{pmatrix},$$

where S_I^+ is the eigenvector of J^+ corresponding to the eigenvalue λ_1^+ and the columns of S_{II}^+ are the eigenvectors of J^+ corresponding to the eigenvalues $\lambda_2^+, \lambda_3^+, \dots, \lambda_n^+$. Expressed in the characteristic variables, the boundary condition is

$$(22) \quad \begin{pmatrix} \mathbf{w}_{II}^+ \\ \dot{x}_1 \end{pmatrix} = \begin{pmatrix} \Lambda_{II}^+ - \dot{s}I & 0 \\ 0 & -1 \end{pmatrix}^{-1} D^{-1} (-I_{3t} + x_1[\mathbf{u}]_t - S_I^+(\lambda_1^+ - \dot{s})\mathbf{w}_I^+),$$

$$(23) \quad x_1(0) = 0,$$

where $\Lambda_{II}^+ = \text{diag}(\lambda_2^+, \lambda_3^+, \dots, \lambda_n^+)$ and D is defined by (3).

By solving (22) and (23) for $x_1(t)$ and then substituting the solution into (22) again, we can express $\mathbf{w}_{II}^+$ in $\mathbf{w}_I^+$ and known functions of time. The energy method (see [9]) shows that the equation, boundary conditions, and initial data for $\mathbf{w}$ constitute a well-posed problem. Well-posedness implies that for any I_{3t} there exists a unique solution. The boundary condition for $\mathbf{w}$ at $x = s^+$ is homogeneous if $I_{3t} \equiv 0$, and nonhomogeneous otherwise. Since $\mathbf{w}$ is a transformation of $\mathbf{u}_1$, the same applies for $\mathbf{u}_1$.

It is now clear that I_{3t} is crucial for the order of accuracy of $\mathbf{u}^\varepsilon$. In the special case $I_{3t} \equiv 0$ we have $\mathbf{u}_1 \equiv 0$ in the downstream region, since $\mathbf{u}_1$ is the solution of a homogeneous equation with homogeneous initial data and boundary conditions. From (5) it then follows that $\mathbf{u}^\varepsilon$ is a second order accurate approximation of $\mathbf{u}$. However, in the general case, we have $\mathbf{u}_1(x, t) \neq 0$ for $x > s$, and $\mathbf{u}^\varepsilon$ will be a first order accurate approximation of $\mathbf{u}$.

2.5. A matrix valued viscosity coefficient eliminating the $\mathcal{O}(h)$ error.

We will now investigate whether it is possible to design the viscosity term such that the first order downstream error is eliminated and second order accurate solutions are obtained. We consider a method which has the modified equation

$$(24) \quad \mathbf{u}_t^\varepsilon + \mathbf{f}(\mathbf{u}^\varepsilon)_x = \varepsilon(\phi(x)E(\mathbf{u}^\varepsilon)\mathbf{u}_x^\varepsilon)_x + c_2\varepsilon^2\mathbf{u}_{xx}^\varepsilon,$$

where $E(\mathbf{u}^\varepsilon)$ is a matrix valued function. The solutions given by such a method can be analyzed in the same way as in the previous sections. The only point which will change in the analysis is the equation for $\hat{\mathbf{U}}$. The new equation for $\hat{\mathbf{U}}$ is

$$(25) \quad (E(\hat{\mathbf{U}})\hat{\mathbf{U}}_{\hat{x}})_{\hat{x}} + \dot{s}\hat{\mathbf{U}}_{\hat{x}} - \mathbf{f}(\hat{\mathbf{U}})_{\hat{x}} = 0,$$

together with the conditions (10) and (11). The boundary condition for $\mathbf{u}_1$ at $x = s^+$ is still given by (22) and (23). If $E(\hat{\mathbf{U}})$ can be chosen such that $I_{3t} \equiv 0$, we will have $\mathbf{u}_1(x, t) \equiv 0$ also in the downstream region.

We note that if $\hat{\mathbf{U}} = \hat{\mathbf{U}}^*$ with

$$(26) \quad \hat{\mathbf{U}}^* = \mathbf{u}^- + \gamma(\tilde{x})[\mathbf{u}],$$

where γ is a scalar smooth function, we obtain

$$I_3 = c_\gamma[\mathbf{u}],$$

where

$$c_\gamma = \int_{-\infty}^0 \gamma(\tilde{x}) d\tilde{x} + \int_0^\infty (\gamma(\tilde{x}) - 1) d\tilde{x}.$$

If γ is antisymmetric around $(0, 0.5)$ with

$$\gamma(-\infty) = 0, \quad \gamma'(-\infty) = 0, \quad \gamma(\infty) = 1, \quad \gamma'(\infty) = 0,$$

then $c_\gamma = 0$ and the boundary conditions (10) and (11) are satisfied.

It now remains to investigate whether it is possible to choose the matrix valued function $E(\hat{\mathbf{U}})$ such that $\hat{\mathbf{U}}^*$ satisfies (25). Integrating (25) from $-\infty$ to $\tilde{x}$ and substituting $\hat{\mathbf{U}}^*$ gives

$$(27) \quad \gamma'(\tilde{x})E(\hat{\mathbf{U}}^*)[\mathbf{u}] = \mathbf{q}(\hat{\mathbf{U}}^*),$$

where

$$\mathbf{q}(\mathbf{U}) = \mathbf{f}(\mathbf{U}) - \mathbf{f}(\mathbf{u}^-) - \dot{s}(\mathbf{U} - \mathbf{u}^-).$$

Note that $E(\hat{\mathbf{U}})$ is a function of $\hat{\mathbf{U}}$ only, with no explicit $\tilde{x}$ dependence. Hence, in order to solve (27) for $E(\hat{\mathbf{U}})$, we must be able to express γ' as a function of $\hat{\mathbf{U}}$. This is the case if we can express γ' in terms of γ , and if γ is monotone. Now solving (27) for $E(\hat{\mathbf{U}}^*)$ gives

$$(28) \quad E(\hat{\mathbf{U}}^*) = \frac{1}{\gamma'} \frac{\mathbf{q}(\hat{\mathbf{U}}^*)\mathbf{q}^T(\hat{\mathbf{U}}^*)}{\mathbf{q}^T(\hat{\mathbf{U}}^*)[\mathbf{u}]}.$$

To ensure that $E(\mathbf{u}^\varepsilon)$ is bounded as $\tilde{x} \rightarrow \pm\infty$ we must also require

$$\lim_{\tilde{x} \rightarrow -\infty} \frac{\gamma}{\gamma'} = M^-, \quad \lim_{\tilde{x} \rightarrow \infty} \frac{\gamma - 1}{\gamma'} = M^+,$$

where $|M^\pm| < \infty$.

Note that in order to evaluate $E(\hat{\mathbf{U}})$ the quantities $\dot{s}$, $\mathbf{u}^-$, and $\mathbf{u}^+$ must be known or estimated.

Remark. Prescribing the viscous profile as above means that the solution follows a straight line in phase space between the upstream and the downstream states. Many other shapes of the solution, and hence, paths in phase space, would also be possible. The properties of $E(\mathbf{u}^\varepsilon)$ will depend on which path is chosen. In order to obtain a stable method, it is necessary that the total viscosity coefficient of the method be positive definite. Since the term $c_2\varepsilon^2\mathbf{u}_{xx}^\varepsilon$ is also present, it is sufficient that $E(\mathbf{u}^\varepsilon)$ be positive semidefinite. We have found that the choice (28) is not positive semidefinite for all problems. In order to design a robust numerical method, we must further investigate what paths in phase space to use. Probably, this will differ depending on the equation that is to be solved. However, we are interested only in showing that it is possible to obtain second order accuracy also downstream, and for this purpose it is good enough to use $E(\mathbf{u}^\varepsilon)$ defined by (28).

3. Numerical experiments. In this section we test how the matrix valued viscosity coefficient derived in the previous section behaves in computations, and compare the results to corresponding computations with a scalar viscosity coefficient.

3.1. The test problems. We consider two test problems. In both problems, the equations, domain, initial data, and boundary condition at $x = x_{\text{end}}$ are the same, while the boundary condition at $x = 0$ differs.

We consider the Euler equations with Riemann initial data connected by a 1-shock. That is, we consider (1) and (2) with

$$(29) \quad \mathbf{u} = \begin{pmatrix} \rho \\ \rho u \\ E \end{pmatrix}, \quad \mathbf{f}(\mathbf{u}) = \begin{pmatrix} \rho u \\ \rho u^2 + p \\ u(E + p) \end{pmatrix}, \quad x_{\text{end}} = 6,$$

$$\mathbf{u}(x, 0) = \begin{cases} \mathbf{u}_L & \text{for } x \leq s_0, \\ \mathbf{u}_R & \text{for } x > s_0, \end{cases}$$

where E and p are connected by the equation of state for a polytropic gas

$$E = \frac{p}{\gamma - 1} + \frac{1}{2}\rho u^2, \quad \gamma = 1.4.$$

Since $\mathbf{u}_L$ and $\mathbf{u}_R$ are connected by a 1-shock, they are fully determined if ρ_L , u_L , and p_L —the initial density, velocity, and pressure at $x \leq s_0$ —and p_R , the initial pressure at $x > s_0$, are specified. We have used $\rho_L = 3$, $u_L = 1.2$, $p_L = 2$, and $p_R = 5$. This gives a 1-shock with speed ≈ -0.26 . We have not specified the initial shock position s_0 explicitly. In the computations, which will be further described below, we started the computation at $t = -1$, with the profile (31) and the shock located at $x = 1.75$. We computed for one time unit using $\mathbf{u}(0, t) = \mathbf{u}_L$. In this way we obtained a good initial profile. We do this to avoid pollution of the numerical solution by disturbances due to nonperfect initial data. We have also used a rather large x_{end} in order to avoid reflection of such disturbances at the boundary $x = x_{\text{end}}$.

At $x = x_{\text{end}}$ we have the boundary condition

$$R_1(x_{\text{end}}, t) = R_1(x_{\text{end}}, 0),$$

where $R_1 = u - 2c/(\gamma - 1)$ is the Riemann invariant connected to λ_1 , and $c = \sqrt{\gamma p/\rho}$ is the local speed of sound.

At $x = 0$ the boundary condition is specified by

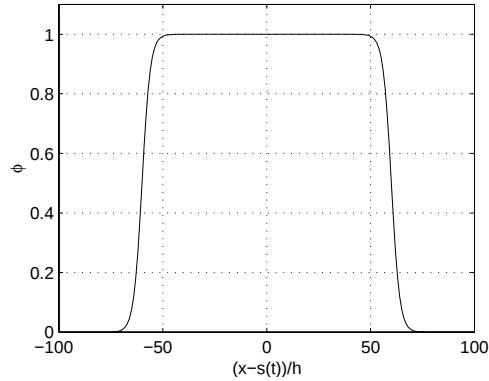
$$\begin{aligned} p(0, t) &= p_L(1 + \alpha d(t)), \\ \rho(0, t) &= \rho_L \left(\frac{p(0, t)}{p_L} \right)^{1/\gamma}, \\ u(0, t) &= u_L + \frac{2}{\gamma - 1}(c(0, t) - c_L) \end{aligned}$$

(see [2]); i.e., a disturbance with amplitude α is introduced into the Riemann invariant R_1 , while the two other Riemann invariants are held constant. If α is small, this models an acoustic disturbance. We have considered the following two test problems:

$$\text{Test problem 1:} \quad \alpha = -0.2, \quad d(t) = (1 - e^{-5t}) \sin 2t;$$

$$\text{Test problem 2:} \quad \alpha = -0.1, \quad d(t) = e^{-10(t-0.7)^2}.$$

In the test problems, α is rather large, in order to make the first order effect more visible.

FIG. 1. The function ϕ , with $s_1 = 60$ and $s_2 = 4$.

3.2. The standard method. A common way to solve (1) is to discretize in space using central differences and add artificial viscosity. To avoid oscillations in the solution, the viscosity must be $\mathcal{O}(h)$ in the shock layer. Outside the shock region, the viscosity can be smaller. We obtain a formally second order method, whose solutions can be modeled by (4), using the semidiscrete scheme

$$(30) \quad (\mathbf{u}_j)_t + D_0 \mathbf{f}(\mathbf{u}_j) = \kappa_1 h D_+ \phi_j D_- \mathbf{u}_j + \zeta h^2 D_+ D_- \mathbf{u}_j.$$

For test problem 1, we used $\kappa_2 = 1$ and $\zeta = 20$, and for test problem 2, we used $\kappa_1 = 0.5$ and $\zeta = 40$. We discretized in space by introducing $x_j = jh$, $h = 1/N$, $j = 0, 1, \dots, N$, where $\mathbf{u}_j(t)$ is a grid function with $\mathbf{u}_j(t) \approx \mathbf{u}^\varepsilon(x_j, t)$. The system of ODEs (30) was solved with the classical fourth order Runge–Kutta method. The time step was chosen as $k = 0.5h$, i.e., CFL-number 0.5.

The switch ϕ was

$$\phi(x) = \begin{cases} 0.5 \tanh((x - s(t) + s_1 h)/s_2 h) + 0.5, & x \leq s(t), \\ 0.5 \tanh((x - s(t) - s_1 h)/s_2 h) + 0.5, & x > s(t), \end{cases}$$

with $s_1 = 60$ and $s_2 = 4$; see Figure 1. Generally, there will be approximately $2s_1$ points where $\phi > 0.5$, and hence we have used a very wide switch. The parameter s_2 determines how steep the gradient of ϕ is in the transition area. The shock position $s(t)$ was numerically determined.

At $x = 6$ we used the mathematical boundary condition

$$R_1(6, t) = R_1(6, 0)$$

and the numerical boundary conditions

$$R_i(6, t) = 2R_i(6 - h, t) - R_i(6 - 2h, t), \quad i = 2, 3,$$

where the Riemann invariants R_2 and R_3 are

$$R_2 = \frac{p}{\rho^\gamma}, \quad R_3 = u + \frac{2c}{\gamma - 1}.$$

The initial data was obtained in the following way. We started the computations at $t = -1$ with the profile (31) and the shock located at $x = 1.75$. We computed for one time unit using $\mathbf{u}(0, t) = \mathbf{u}_L$.

We will refer to this method as the standard method.

3.3. The matrix viscosity method. We will now introduce a method which can be modeled by (24), and we will refer to it as the matrix viscosity method. The matrix viscosity method is the same as the standard method, except that (30) is replaced by

$$(\mathbf{u}_j)_t + D_0 \mathbf{f}(\mathbf{u}_j) = \kappa_2 h D_+ \phi_j E_j D_- \mathbf{u}_j + \zeta h^2 D_+ D_- \mathbf{u}_j.$$

Here $E_j \approx E(\mathbf{u}^\varepsilon(x_j, t))$. Our implementation is described below. When solving test problem 1, we used $\kappa_2 = 15$, $\zeta = 20$, and CFL-number 0.05. For test problem 2, we used $\kappa_2 = 7$, $\zeta = 40$, and CFL-number 0.1.

To implement E_j in a robust and accurate way is difficult. The expression (28) is not suited for computations. The solution changes rapidly in the shock layer from being close to $\mathbf{u}^-$ to being close to $\mathbf{u}^+$. The quantities $\dot{s}$, $\mathbf{u}^-$, and $\mathbf{u}^+$ must be numerically determined; hence it is difficult to compute $\mathbf{q}$ with high accuracy. Also, both $\mathbf{q}$ and γ' tend rapidly to zero as $\tilde{x} \rightarrow \pm\infty$. However, for large $\tilde{x}$ we can linearize the expression for $\mathbf{q}$ and find

$$\mathbf{q} = \begin{cases} \gamma(J^- - \dot{s}I)[\mathbf{u}] & \text{as } \tilde{x} \rightarrow -\infty, \\ (\gamma - 1)(J^+ - \dot{s}I)[\mathbf{u}] & \text{as } \tilde{x} \rightarrow \infty. \end{cases}$$

By the assumptions on γ we find

$$E = \begin{cases} E^- & \text{as } \tilde{x} \rightarrow -\infty, \\ E^+ & \text{as } \tilde{x} \rightarrow \infty, \end{cases}$$

where

$$E^\pm = M^\pm \frac{(J^\pm - \dot{s}I)[\mathbf{u}][\mathbf{u}]^T (J^\pm - \dot{s}I)^T}{[\mathbf{u}]^T (J^\pm - \dot{s}I)^T [\mathbf{u}]}.$$

We have used

$$\gamma(\tilde{x}) = \frac{1}{2}(\tanh(\tilde{x}) + 1),$$

and hence we have $M^- = 1/2$ and $M^+ = -1/2$.

In the computations we have used

$$E_j = (1 - \gamma(\tilde{x}_j))E^- + \gamma(\tilde{x}_j)E^+.$$

The quantities $\dot{s}$, $\mathbf{u}^-$, and $\mathbf{u}^+$ were numerically determined. First, we approximated $\mathbf{u}^\pm$ by simply taking the value of the numerical solution 20 points upstream and downstream, respectively, of the approximated shock position. The shock speed $\dot{s}$ was approximated by

$$\dot{s}_{\text{approx}} = \sum_{k=1}^3 [\mathbf{f}^{(k)}(\mathbf{u})] / \sum_{k=1}^3 [\mathbf{u}^{(k)}].$$

By adding the jumps in the different components of $\mathbf{f}(\mathbf{u})$ and dividing by the sum of the jumps in the different components of $\mathbf{u}$, we avoid introducing large errors in $\dot{s}_{\text{approx}}$ due to rounding effects. This method for approximating $\mathbf{u}^\pm$ and $\dot{s}$ was used

when test problem 1 was solved, and we obtain second order accuracy both upstream and downstream. The results are further presented in section 3.4.

The approximation of $\mathbf{u}^\pm$ mentioned above has an error which is small, but independent of h . Hence, there will be a small $\mathcal{O}(1)$ error in our approximation of $E(\mathbf{u}^\varepsilon)$, which will cause a small $\mathcal{O}(h)$ error in the solution. This first order effect became evident as we tried to solve test problem 2. In order to eliminate it, we needed to use an estimate of $\mathbf{u}^\pm$ with an error which is $\mathcal{O}(h)$. We obtain this if we make use of the fact that the solution in the shock region follows a straight line in phase space between $\mathbf{u}^-$ and $\mathbf{u}^+$, and that the function γ determines how fast the solution is approaching the limit values. We pick the value of the solution $2\kappa_2$ points upstream and downstream of the approximate shock position. Since γ is known, we know how far from the end states these values are and correct for this. The shock speed $\dot{s}$ is still computed as above. Test problem 2 was solved using this improved approximation of $\mathbf{u}^\pm$, and we obtained second order accuracy both upstream and downstream.

The implementation of $E(\mathbf{u}^\varepsilon)$ described above requires a fixed number of computations, independent of h .

As initial profile at $t = -1$, we used

$$(31) \quad \mathbf{u}_j = \mathbf{u}_L + \gamma(\tilde{x}_j)(\mathbf{u}_R - \mathbf{u}_L);$$

i.e., in the shock region the profile satisfies (26). The initial profile (31) is used to avoid diffusion waves in the solution (see section 2.2). Ideally, using this initial profile, no such waves should appear. However, in our numerical computations we see diffusion waves, but they are very small.

In computations not reported here, we have also tried to use

$$(32) \quad \mathbf{u}(x, -1) = \begin{cases} \mathbf{u}_L & \text{for } x \leq 1.75, \\ \mathbf{u}_R & \text{for } x > 1.75. \end{cases}$$

As expected, the profile rapidly adjusts, and diffusion waves appear and move out of the shock region following outgoing characteristics. Also as expected, for the matrix method, the diffusion waves are much larger with (32) as initial data than if (31) is used. However, the order of accuracy of the solution behind the diffusion waves is the same.

3.4. Results. We have numerically investigated the rate of convergence of the standard method and of the matrix viscosity method by solving the two test problems described in section 3.1 with successively refined space step.

First, consider test problem 1. In Figure 2 we see the solution at $t = 1.25$. We have solved test problem 1 with successively halved space step h with both methods. We started with $h = 0.02$. In all, we computed six solutions with the matrix viscosity method and eight solutions with the standard method.

The computational order of accuracy, r_h , was estimated in the standard way,

$$r_h = \log \left(\frac{\|\rho u_{4h} - \rho u_{2h}\|}{\|\rho u_{2h} - \rho u_h\|} \right) / \log 2,$$

where ρu_h denotes the discrete approximation of ρu with space step h , and the norm used was the discrete L_2 -norm on the interval $(0, 0.7)$ in the upstream region and $(1.4, 2.2)$ in the downstream region. In Table 1 we see that the standard method is second order accurate upstream, but only first order accurate downstream of the

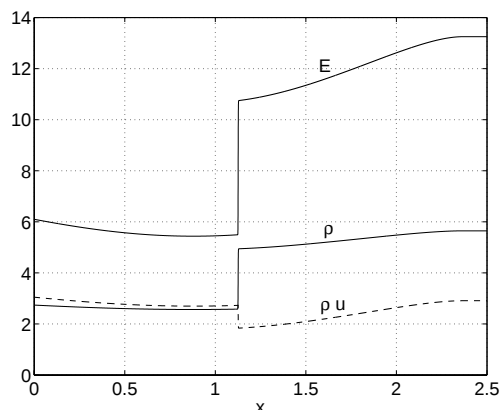


FIG. 2. The solution of test problem 1 at $t = 1.25$. The solution is computed numerically using the standard method with $h = 1.5625 \cdot 10^{-4}$.

TABLE 1

Estimated order of accuracy (r_h) and absolute error ($\|e_h\|$) for the pu -component of the solution of test problem 1, computed by the standard method.

h	Upstream		Downstream	
	r_h	$\ e_h\ $	r_h	$\ e_h\ $
$1 \cdot 10^{-2}$		$1.5 \cdot 10^{-3}$		$3.1 \cdot 10^{-2}$
$5 \cdot 10^{-3}$	3.92	$1.0 \cdot 10^{-4}$	2.17	$6.9 \cdot 10^{-3}$
$2.5 \cdot 10^{-3}$	2.38	$1.9 \cdot 10^{-5}$	2.15	$1.6 \cdot 10^{-3}$
$1.25 \cdot 10^{-3}$	2.00	$4.6 \cdot 10^{-6}$	1.46	$5.7 \cdot 10^{-4}$
$6.25 \cdot 10^{-4}$	2.00	$1.2 \cdot 10^{-6}$	1.20	$2.5 \cdot 10^{-4}$
$3.125 \cdot 10^{-4}$	2.00	$2.9 \cdot 10^{-7}$	1.08	$1.2 \cdot 10^{-4}$
$1.5625 \cdot 10^{-4}$	2.00	$7.2 \cdot 10^{-8}$	1.03	$5.7 \cdot 10^{-5}$

TABLE 2

Estimated order of accuracy (r_h) and absolute error ($\|e_h\|$) for the pu -component of the solution of test problem 1, computed by the matrix viscosity method.

h	Upstream		Downstream	
	r_h	$\ e_h\ $	r_h	$\ e_h\ $
$1 \cdot 10^{-2}$		$6.3 \cdot 10^{-3}$		$1.8 \cdot 10^{-2}$
$5 \cdot 10^{-3}$	4.71	$2.4 \cdot 10^{-4}$	1.93	$4.8 \cdot 10^{-3}$
$2.5 \cdot 10^{-3}$	3.68	$1.9 \cdot 10^{-5}$	1.96	$1.2 \cdot 10^{-3}$
$1.25 \cdot 10^{-3}$	2.01	$4.6 \cdot 10^{-6}$	2.00	$3.1 \cdot 10^{-4}$
$6.25 \cdot 10^{-4}$	2.00	$1.2 \cdot 10^{-6}$	1.99	$7.7 \cdot 10^{-5}$

shock. The matrix viscosity method is second order accurate, both upstream and downstream of the shock; see Table 2.

In Figure 3 we see an overview of how the ρ -component of the solution converges, and in Figure 4 we see a close-up. Note that the aim when designing the matrix viscosity method was to avoid the first order error outside the shock region. Hence, the matrix viscosity method performs better than the standard method for fine grids, where the first order downstream error destroys the convergence rate of the standard method. For coarse grids, however, the matrix viscosity method is not better.

If the order of accuracy is r , then the error in the ρu -component of the solution

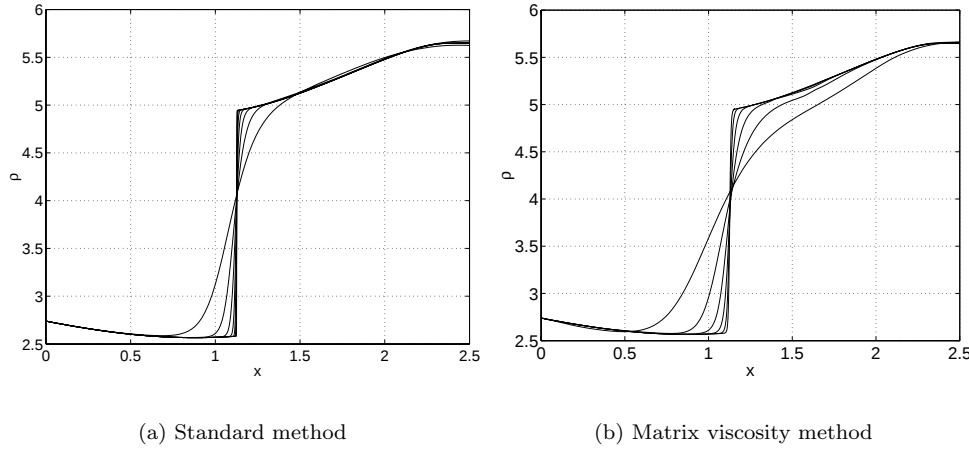


FIG. 3. Overview of the convergence of test problem 1. In the plots we see the ρ -component of the solution. In both cases, the most viscous solution is computed using $h = 0.02$. Additional solutions are computed using successively halved space step. For the standard method we see eight different solutions, and for the matrix viscosity method six solutions.

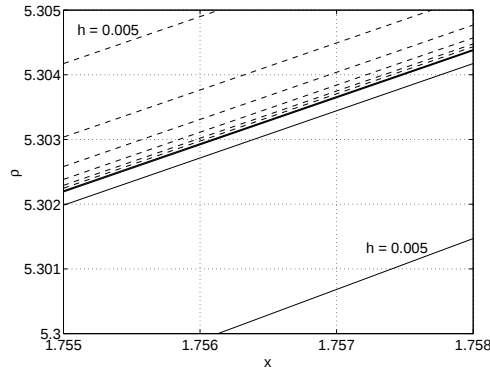


FIG. 4. Close-up of the convergence of the ρ -component for test problem 1. Solid lines: the matrix viscosity method; dashed lines: the standard method. For both methods, the two coarsest solutions are not seen in the close-up. As h is successively halved, the solutions from the matrix viscosity method increase and the solutions from the standard method decrease.

can be estimated:

$$\|e_h\| = \frac{1}{2^r - 1} \|u_{2h} - u_h\|.$$

Hence, $\|e_h\|$ is computed with $r = 2$ in the upstream region for the standard method, and both upstream and downstream for the matrix method. For the standard method downstream of the shock, we have used $r = 1$. Again we use the discrete L_2 -norm, on the same interval as above.

Corresponding computations for the ρ - and E -components of the solution give qualitatively the same result.

By plotting the shock profile in phase space (see Figure 5), we see that the shock

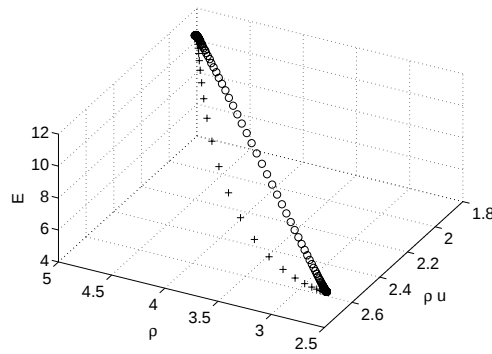


FIG. 5. Numerical phase diagram of the shock profile computed by the matrix viscosity method (o) and the standard method (+). Both solutions are computed using $h = 6.25 \cdot 10^{-4}$. The shock profile computed by the matrix viscosity method follows a straight line in phase space quite closely.

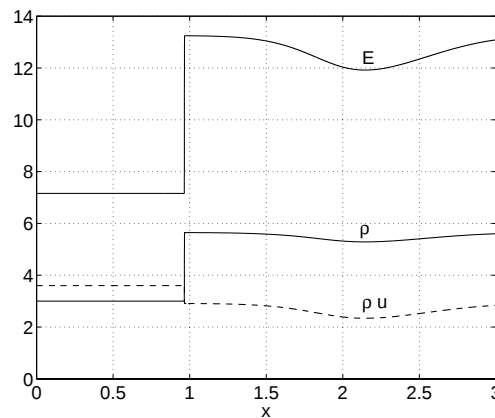


FIG. 6. The solution of test problem 2 at $t = 1.9$. The solution is computed numerically using the standard method with $h = 1.5625 \cdot 10^{-4}$.

profile obtained by the matrix viscosity method approximately follows a straight line between the shock states. The shock profile of the standard method clearly has another shape.

Next, consider test problem 2. In Figure 6, we see the solution of test problem 2 at $t = 1.9$. Test problem 2 was also solved with successively halved space step h with both methods, starting with $h = 0.02$. Again, in all we computed eight solutions with the standard method and six solutions with the matrix viscosity method. In Table 3, we see the estimated order of accuracy for the standard method. Again, we have used the interval $(0, 0.7)$ in the upstream region. In the downstream region the discrete L_2 -norm was computed on the interval $(1.1, 3)$. Upstream, the solution is second order. Downstream the convergence is slower, and the order of accuracy is slowly approaching one. In Table 4 we see that the matrix viscosity method is second order accurate both upstream and downstream. In Figure 7 we see an overview of how the ρ -component converges, and in Figure 8 we see a close-up. In phase space, the shock profiles of the solutions of test problem 2 are qualitatively the same as in Figure 5.

TABLE 3

Estimated order of accuracy (r_h) and absolute error ($\|e_h\|$) for the pu -component of the solution of test problem 2, computed by the standard method.

h	Upstream		Downstream	
	r_h	$\ e_h\ $	r_h	$\ e_h\ $
$1 \cdot 10^{-2}$		$6.2 \cdot 10^{-4}$		$5.0 \cdot 10^{-2}$
$5 \cdot 10^{-3}$	7.93	$2.6 \cdot 10^{-6}$	1.95	$1.3 \cdot 10^{-2}$
$2.5 \cdot 10^{-3}$	2.58	$4.3 \cdot 10^{-7}$	1.96	$3.3 \cdot 10^{-3}$
$1.25 \cdot 10^{-3}$	2.33	$8.5 \cdot 10^{-8}$	1.91	$8.8 \cdot 10^{-4}$
$6.25 \cdot 10^{-4}$	2.01	$2.1 \cdot 10^{-8}$	1.88	$2.4 \cdot 10^{-4}$
$3.125 \cdot 10^{-4}$	2.00	$5.3 \cdot 10^{-9}$	1.69	$7.4 \cdot 10^{-5}$
$1.5625 \cdot 10^{-4}$	2.00	$1.3 \cdot 10^{-9}$	1.39	$2.8 \cdot 10^{-5}$

TABLE 4

Estimated order of accuracy (r_h) and absolute error ($\|e_h\|$) for the pu -component of the solution of test problem 2, computed by the matrix viscosity method.

h	Upstream		Downstream	
	r_h	$\ e_h\ $	r_h	$\ e_h\ $
$1 \cdot 10^{-2}$		$1.5 \cdot 10^{-3}$		$2.1 \cdot 10^{-2}$
$5 \cdot 10^{-3}$	4.24	$7.9 \cdot 10^{-5}$	2.06	$4.9 \cdot 10^{-3}$
$2.5 \cdot 10^{-3}$	8.01	$3.1 \cdot 10^{-7}$	2.33	$9.8 \cdot 10^{-4}$
$1.25 \cdot 10^{-3}$	1.86	$8.5 \cdot 10^{-8}$	2.23	$2.1 \cdot 10^{-4}$
$6.25 \cdot 10^{-4}$	2.01	$2.1 \cdot 10^{-8}$	2.00	$5.2 \cdot 10^{-5}$

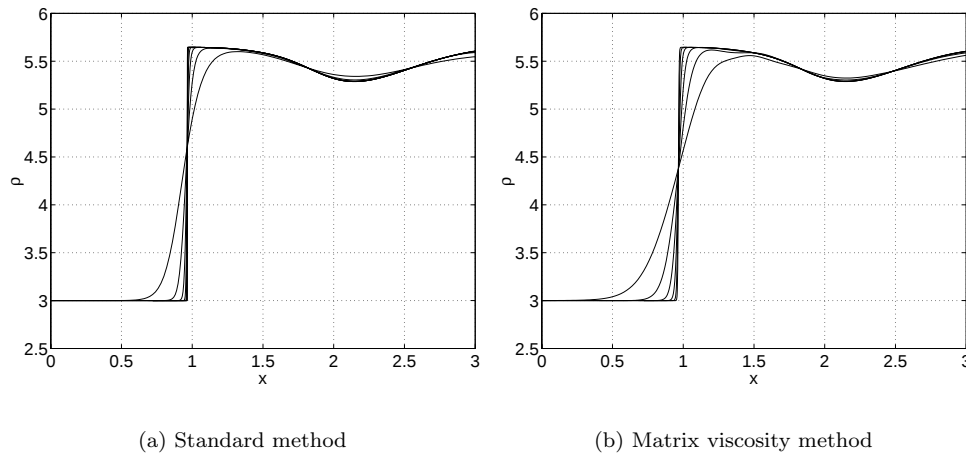


FIG. 7. Overview of the convergence of test problem 2. We see the ρ -component of the solution. In both cases, the most viscous solution is computed using $h = 0.02$. Additional solutions are computed using successively halved space step. For the standard method we see eight different solutions, and for the matrix viscosity method six solutions.

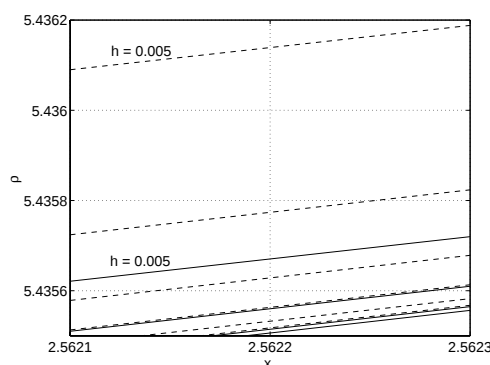


FIG. 8. Close-up of the convergence of the p -component for test problem 2. Solid lines: the matrix viscosity method; dashed lines: the standard method. For both methods, the two coarsest solutions are not seen in the close-up. As h is successively halved, the solutions from both methods decrease.

REFERENCES

- [1] M. ARORA AND P. L. ROE, *On postshock oscillations due to capturing schemes in unsteady flows*, J. Comput. Phys., 130 (1997), pp. 25–40.
- [2] J. CASPER AND M. H. CARPENTER, *Computational considerations for the simulation of shock-induced sound*, SIAM J. Sci. Comput., 19 (1998), pp. 813–828.
- [3] G. EFRAIMSSON AND G. KREISS, *Approximate solutions to slightly viscous conservation laws*, Quart. Appl. Math., to appear.
- [4] G. EFRAIMSSON, J. NORDSTRÖM, AND G. KREISS, *Artificial Dissipation and Accuracy Downstream of Slightly Viscous Shocks*, Paper 2001-2608, American Institute of Aeronautics and Astronautics, Reston, VA, 2001.
- [5] B. ENGQUIST AND B. SJÖGREEN, *The convergence rate of finite difference schemes in the presence of shocks*, SIAM J. Numer. Anal., 35 (1998), pp. 2464–2485.
- [6] J. GOODMAN AND A. MAJDA, *The validity of the modified equation for nonlinear shock waves*, J. Comput. Phys., 58 (1985), pp. 336–348.
- [7] S. KARNI AND S. ČANIĆ, *Computations of slowly moving shocks*, J. Comput. Phys., 136 (1997), pp. 132–139.
- [8] G. KREISS, G. EFRAIMSSON, AND J. NORDSTRÖM, *Elimination of first order errors in shock calculations*, SIAM J. Numer. Anal., 38 (2001), pp. 1986–1998.
- [9] H.-O. KREISS AND J. LORENZ, *Initial-Boundary Value Problems and the Navier–Stokes Equations*, Academic Press, New York, 1989.
- [10] P. A. LAGERSTROM, *Matched Asymptotic Expansions*, Springer-Verlag, New York, 1988.
- [11] R. J. LEVEQUE, *Numerical Methods for Conservation Laws*, Birkhäuser Verlag, Basel, Switzerland, 1992.
- [12] A. SZEPESSY AND Z. XIN, *Nonlinear stability of viscous shock waves*, Arch. Ration. Mech. Anal., 122 (1993), pp. 53–103.