Abstract

We present a novel semi-implicit scheme for numerical solutions of time-dependent conservation laws. The core idea of the presented method consists of exploiting and approximating mixed partial derivatives of the solution that occur naturally when deriving higher-order accurate schemes. Such an approach is introduced in the context of the Lax-Wendroff (or Cauchy-Kowalevski) procedure when the time derivatives are not completely replaced by space derivatives using the PDE, but some mixed derivatives are allowed. If approximated in a suitable way, one obtains algebraic systems that have a more convenient structure than the systems derived by standard fully implicit schemes. We derive high-resolution TVD form of the semi-implicit scheme for some representative hyperbolic equations in one-dimensional case including related numerical experiments.

Keywords: conservation laws, finite difference method, semi-implicit method

1. Introduction

One of the most straightforward ways to solve numerically the time dependent hyperbolic partial differential equations (PDEs) is to use a Method Of Lines (MOL) when the spatial and the temporal parts of PDE are discretized separately. Typically, a discretization of the space is performed in the first step that approximates the PDE by a system of ordinary differential equations (ODEs), for which, afterwards, a chosen numerical integration is used [32, 23].

Quite naturally, the first candidates to obtain numerical solutions of ODEs are explicit methods that deliver numerical solutions without the need to solve any algebraic system of equations. Such approximations require a careful choice of discretization steps not only due to accuracy requirements but also for stability reasons. The second requirement is specific to numerical methods and, if not followed, unstable behaviors of numerical solutions can occur even for well-posed problems. For many types of PDEs and the methods for their numerical
solutions, such stability requirements are well understood, typically formulated in the form of a stability condition on the choice of time steps. If the known stability restriction does not limit the choice of discretization steps more than the accuracy requirement, the usage of explicit methods is well justified.

Nevertheless, in several cases such restrictions can be too demanding or simply hard to follow, and consequently implicit time discretization methods for MOL are also considered, see, e.g., [1, 28]. Such methods are receiving increasing attention, especially for models that describe several dynamic processes with different characteristic speeds, of which only those with slow or moderate speed are of practical interest. In such cases, the terms related to the processes with the fastest speed are treated implicitly. As an example we mention here only so called “all Mach number” solvers of Euler equations that are treated, e.g., with fully implicit methods [4] or IMEX methods [2, 7, 37] using also relaxation methods [5, 33].

The implicit methods do not explicitly define the values of numerical solution, but instead formulate algebraic systems to be fulfilled by the discrete numerical values. Therefore, to find the numerical solution, a system of algebraic equations must be solved. Clearly, this is the main price to be paid by the usage of implicit methods that must be well justified, especially if the algebraic systems are nonlinear. Consequently, methods are demanded that simplify the task of (nonlinear) algebraic solvers, and that is the main motivation of our study.

Implicit numerical integrators can be used analogously to the explicit methods applied to the system of ODEs obtained with MOL for PDEs, see, e.g., Diagonally Implicit Runge-Kutta (DIRK) methods [28], adaptive Runge-Kutta methods [1], implicit multi-rate methods [8] and so on. Other types of temporal discretization methods offer a coupled treatment of both discretization methods. We mention explicit methods based on Taylor series expansions like the Lax-Wendroff procedure [29, 39, 9, 10], a local time-space DG discretization [12, 20], and two-, or even multi-derivative implicit schemes [18], see also a review paper [30]. Concerning the methods based on Taylor series, several authors have recognized that it can be advantageous to discretize each replaced term in the Taylor expansion with different spatial discretizations [29, 30, 36, 24, 9, 38]. A natural question arises if such an approach can be used to simplify the algebraic equations resulting from implicit or implicit-explicit time discretization of hyperbolic problems.

In the context of nonconservative advection equation, it has been recognized that a special coupling of temporal and spatial discretization can result in second order accurate numerical schemes that are unconditionally stable and that result in simpler algebraic systems than the ones obtained with fully implicit schemes [17, 15]. Such schemes were introduced under the abbreviation IIODE (“Inflow-Implicit/Outflow-Explicit”) finite volume method for the scalar advection equation in [27] and later successfully applied in, e.g., [16, 19, 21]. The IIODE scheme in [27] resembles the so-called “angle derivative” type methods for constant velocity advection; see the discussion and references in [26].

As the notation “Inflow-Implicit and Outflow-Explicit” suggests, spatial and
temporal discretizations are aware of each other and proposed in a coupled way. In [15, 14] it is shown that the schemes can be derived by considering and approximating mixed spatial-temporal derivatives in the Taylor expansion of the solution when using the Lax-Wendroff (LW) procedure only partially. The idea of not using LW to completely replace the time derivatives with the space derivatives using PDE is used also in different contexts in [13, 9, 10].

Opposite to level set advection equations where the solution is supposed to be continuous, the hyperbolic problems allow for discontinuous solutions. For this type of problems, the numerical scheme must be conservative to approximate correctly the movement of shock waves and the scheme must be nonlinear even for linear problems if higher resolution than the first order scheme with non-oscillatory numerical solutions is required. In this context, we introduce a novel semi-implicit method for some representative models of hyperbolic systems extending the approach for linear problems in [15, 14] with some ideas from [13, 25] for nonlinear problems.

For scalar PDE the proposed method involves a free parameter, for which the method is always second order accurate in time and space for smooth solutions. In the case of discontinuous solutions we propose a predictor-corrector procedure to find solution dependent values of the parameter for which the scheme is Total Variation Diminishing (TVD). For the case of Courant number larger than one, in general, additional limiting must be used, which we propose in a form of flux-corrected transport scheme [11, 22]. The system of algebraic equations is solved efficiently with only one forward and one backward sweep using the fast sweeping method [25]. In this method, a scalar algebraic equation is solved for each grid point that is nonlinear only due to the nonlinearity of the model (if nonlinear). In the case of hyperbolic systems we express the second order correction part of the scheme in characteristic variables as suggested, e.g., in [23, 13].

The paper is organized as follows. In Section 2 we present the method for the scalar case with details for the linear advection equation given in Section 3. In Section 4 we present algorithmic details of the high-resolution method for the scalar nonlinear case, and in Section 5 we give additional details for hyperbolic systems. The Section 6 on numerical experiments presents several test examples that illustrate the properties of the method. Finally, in Section 7 we make some concluding remarks and in Appendix 8 we give some technical details.

2. Scalar conservation laws

In this section, we aim to solve numerically the scalar nonlinear hyperbolic equation written in the form

\[ u_t + f(u)_x = 0, \quad u(x, 0) = u^0(x), \quad x \in \mathbb{R}, \quad t > 0, \]

where \( u = u(x, t) \) is the unknown function with initial values prescribed by a given \( u^0 \) and \( f \) is a given flux function. To discretize [1] we follow the approach of conservative finite difference methods as described, e.g., in [32, 29, 25]. For
that purpose we use the standard notation for grid nodes $x_i$, $i = 0, 1, 2, \ldots, I$ with a uniform step $\Delta x \equiv x_i - x_{i-1}$ and discrete times $0 = t^0 < t^1 < \cdots$ with $\Delta t \equiv t^{n+1} - t^n$, $n = 0, 1, \ldots, N$ where the integers $I$ and $N$ are given. Furthermore, $x_{i+1/2} = x_i + \Delta x/2$, $f^n_i := f(u^n_i)$, etc. Our aim is to find approximations $u^{n+1}_i \approx u(x_i, t^{n+1})$. To do so, we follow the standard form of conservative schemes,

$$u^n_i + \frac{\tau}{\Delta x} \left( F_{i+1/2}^{n+1} - F_{i-1/2}^{n+1} \right) = u^n_i,$$  \hspace{1cm} (2)

where the numerical fluxes will be defined by a numerical flux function.

To propose the numerical flux function, we use the approach of the fractional step scheme presented in [25]. There, the flux function $f$ is split into the sum of two functions having nonnegative and nonpositive derivatives,

$$f = f^+ + f^-, \quad \frac{df^+}{du} \geq 0, \quad \frac{df^-}{du} \leq 0, \quad u \in \mathbb{R}. \hspace{1cm} (3)$$

One choice for (3) is analogous to the Lax-Friedrichs vector splitting,

$$f^+(u) := \frac{1}{2} (f(u) + \alpha u), \quad f^-(u) := \frac{1}{2} (f(u) - \alpha u), \hspace{1cm} (4)$$

where the parameter $\alpha$ is fixed at the maximum value of $|f'(u)|$ over the relevant values of $u$.

Having the splitting, the simplest variant of the fractional step method consists of two partial steps with the first step given by solving the algebraic equations,

$$u^{n+1}_i + \frac{\Delta t}{\Delta x} F^+,_{i+1/2} = u^n_i + \frac{\Delta t}{\Delta x} F^+,_{i-1/2}, \quad i = 1, 2, \ldots, I, \hspace{1cm} (5)$$

where the numerical flux $F^+,_{i+1/2}$ shall be determined from the boundary conditions. The second step is given by the solution of

$$u^{n+1}_{i+1} - \frac{\Delta t}{\Delta x} F^-,_{i-1/2} = u^n_i - \frac{\Delta t}{\Delta x} F^-,_{i+1/2}, \quad i = I - 1, I - 2, \ldots, 0, \hspace{1cm} (6)$$

where the values $u^n_i$ in (6) are equal to the values $u^{n+1}_i$ of the first fractional step (5) that we do not distinguish in the notation. Again, the value $F^-,_{i-1/2}$ in (6) is determined from the boundary conditions. The numerical fluxes in (5) and (6) are given in [25] by the first order accurate upwind approximation,

$$F^+,_{i+1/2} = f^+_i, \quad F^-,_{i-1/2} = f^-_i. \hspace{1cm} (7)$$

In what follows, we propose a second order and a high-resolution extension of the numerical fluxes in (7).

The most important advantage of the proposed method (5) - (7) is that each algebraic equation contains only the single unknown $u^{n+1}_i$. The main disadvantage of (7) is the low order accuracy that we aim to improve here. Note that in our numerical experiments we use the fractional step method in
the first order accurate form (5) - (6), for higher order extensions see a discussion in [25].

The numerical flux functions in our semi-implicit method take the following parametric form,

\[ F^+_{i+1/2} = f^+_{i+1} - \frac{l_i}{2} \left( (1 - \omega_i)(f^+_{i+1} - f^+_{i+1}) + \omega_i(f^+_{i+1} - f^+_{i+1}) \right), \]

\[ F^-_{i-1/2} = f^-_{i-1} - \frac{l_i}{2} \left( (1 - \omega_i)(f^-_{i-1} - f^-_{i-1}) + \omega_i(f^-_{i+1} - f^-_{i+1}) \right), \]

where the parameters \( \omega_i \in [0, 1] \) and \( l_i \in [0, 1] \) shall be chosen. The parameters are different in (8) and (9) (in fact, also in each time step), which we do not emphasize in the notation. For a fixed value of \( \omega_i \equiv \bar{\omega} \) and \( l_i \equiv 1 \) the method is second order accurate for smooth solutions if either \( f \equiv f^+ \) or \( f \equiv f^- \), see the Appendix. In the case of linear advection equation, the scheme is unconditionally stable for \( \omega_i \geq 0 \) having no restriction on the choice of \( \tau \) due to the stability, see the proof in [14].

Note that replacement of (7) by the definitions (8) and (9) again results in a fully upwinded form in the implicit part of the schemes (5) and (6) for any particular choice of parameters. Consequently, the left hand sides of (8) and (9) contain again only single unknown values \( u^{n+1}_i \) if computed in the order defined in (5) and (6).

Any constant choice of parameters \( \omega_i \) gives a scheme with a fixed stencil that can result for non-smooth problems in numerical solutions with unphysical oscillations that are not diminishing with a grid refinement. To suppress such behavior, we define variable values of \( \omega_i \) depending on the numerical solution that results in a nonlinear numerical scheme even for the linear advection equation. In what follows, we propose such a dependency of \( \omega \) in (5) - (6) on \( u^{n+1}_i \) to obtain a Total Variation Diminishing (TVD) scheme [23, 35].

3. Linear advection equation

For clarity of presentation, we derive the nonlinear numerical scheme for the simplest case of the linear advection equation with positive constant velocity \( \bar{v} \), when \( f(u) \equiv f^+(u) = \bar{v}u \). The Courant number is denoted by

\[ C = \frac{\bar{v} \Delta t}{\Delta x}. \]

Our aim is to propose a function \( \omega = \omega(r) \) such that \( \omega_i = \omega(r_i) \) and

\[ r_i = \frac{u^{n+1}_i - u^n_i}{u^{n+1}_i - u^n_{i+1}}, \]

if \( u^{n+1}_i \neq u^n_{i+1} \). As \( r_i \) in (10) depends on the unknown value \( u^{n+1}_i \), the resulting scheme will be nonlinear even for the linear advection equation.
Using (10) and \( l_i \equiv 1 \) we can express the numerical fluxes as follows,

\[
F_{i-1/2}^{+,n+1} = \bar{v} \left( u_{i-1}^{n+1} - \frac{1}{2} (1 - \omega_{i-1} + \omega_{i-1} r_{i-1}) (u_{i-1}^{n+1} - u_i^n) \right),
\]

(11)

\[
F_{i+1/2}^{+,n+1} = \bar{v} \left( u_i^{n+1} - \frac{1}{2} r_i (1 - \omega_i + \omega_i r_i) (u_{i-1}^{n+1} - u_i^n) \right),
\]

(12)

if \( r_i \neq 0 \) (that we suppose for the rest of this derivation and comment later).

Denoting

\[ \Psi_i = 1 - \omega_i + \omega_i r_i, \]

one can express the numerical fluxes in (11) and (12) using the form

\[
F_{i-1/2}^{+,n+1} = \bar{v} \left( u_{i-1}^{n+1} - \frac{1}{2} \Psi_i (u_{i-1}^{n+1} - u_i^n) \right),
\]

\[
F_{i+1/2}^{+,n+1} = \bar{v} \left( u_i^{n+1} - \frac{1}{2} \Psi_i (u_i^{n+1} - u_{i+1}^n) \right).
\]

The values \( \Psi_i \) can be viewed as the so-called flux limiters [23, 25, 13, 28], when the scheme (5) can be written formally in the form,

\[
u_{i}^{n+1} - u_{i}^{n} + C \left( u_{i}^{n+1} - u_{i-1}^{n+1} - \frac{1}{2} \left( \Psi_i - \Psi_{i-1} \right) \right) (u_{i-1}^{n+1} - u_i^n) = 0. \]

(13)

Now, using

\[
u_{i}^{n+1} - u_{i}^{n} = (u_{i}^{n+1} - u_{i}^{n}) - (u_{i-1}^{n+1} - u_{i-1}^{n}), \]

the scheme (13) can be written in the form

\[
\left( 1 - \frac{C}{2} \left( \frac{\Psi_i}{r_i} - \Psi_{i-1} \right) \right) (u_{i}^{n+1} - u_{i}^{n}) + \]

\[
C \left( 1 + \frac{1}{2} \left( \frac{\Psi_i}{r_i} - \Psi_{i-1} \right) \right) (u_{i}^{n+1} - u_{i-1}^{n}) = 0. \]

(14)

If the coefficients before \((u_{i}^{n+1} - u_{i}^{n})\) and \((u_{i}^{n+1} - u_{i-1}^{n})\) in (14) are fixed and nonnegative, then the scheme is TVD [13, 28]. In what follows, we propose \( \omega = \omega(r) \) so that this property is fulfilled for \( C \leq 1 \). For larger Courant numbers, to preserve the TVD property, we have to consider \( l_i \in [0,1] \), see later their definition inspired by flux-corrected type methods [22, 11].

**Remark 1.** To derive (14) we have supposed, among others, that \( u_{i-1}^{n+1} \neq u_{i}^{n} \). As we show later, the case \( u_{i-1}^{n+1} = u_{i}^{n} \) can happen only if \( \omega_{i-1} = 0 \), when the scheme (5) takes the simpler form,

\[
u_{i}^{n+1} - u_{i}^{n} + C \left( u_{i}^{n+1} - u_{i-1}^{n+1} - \frac{1}{2} (1 - \omega_i) (u_{i}^{n+1} - u_{i+1}^{n}) \right) = 0. \]

(15)

To fulfill the TVD property, we choose \( \omega_i = 1 \) in (15), which results in the first order scheme. If by chance the result of (15) is \( u_{i}^{n+1} = u_{i+1}^{n} \), then \( \omega_i \) in (15) can be arbitrarily chosen, so we set \( \omega_i = 0 \).
To have positive coefficients in (14) we require the following,

\[-1 \leq \Psi_{i-1} \leq 2,\]  
\[
\Psi_{i-1} - 2 \leq \frac{\Psi_i}{r} \leq \Psi_{i-1} + 2,
\]

where the inequalities in (17) must be satisfied for an arbitrary nonzero \( r \in R \).

Note that the inequalities in (16) are, in fact, required to fulfill (17) for two special cases that can occur: \( \Psi_i = 0 \) and \( \Psi_i = r \). Note that for accuracy reasons, we require \( \Psi(1) = 1 \) [23, 13, 28].

One of the simplest choices is used in [13], where

\[
\Psi(r) = \begin{cases}  
  r & |r| \leq 1 \\
  1 & |r| > 1
\end{cases}
\]

or, equivalently,

\[
\omega(r) = \begin{cases} 
  1 & |r| \leq 1 \\
  0 & |r| > 1
\end{cases}
\]

In the case of fully explicit or fully implicit schemes, the choice (18) can be viewed as the second order ENO reconstruction [32, 13].

We propose the function \( \omega = \omega(r) \) following a strategy of modified ENO schemes [31]. That is, we assume that a preferable constant value \( \bar{\omega} \in (0, 1] \) of \( \omega \) is chosen to be used in (8) that should be considered for each \( \omega_i \) if the TVD property is not destroyed. In what follows, we choose \( \bar{\omega} = 1 \) which gives the upwinded form of fluxes in (8) - (9) that we prefer for larger Courant numbers.

In particular, we define

\[
\omega(r) = \begin{cases} 
  \frac{1}{r-1} & 2 \leq r \\
  \frac{2}{r} & -1 \leq r \leq 2
\end{cases}
\]

or, equivalently,

\[
\Psi(r) = \begin{cases} 
  -1 & r \leq -1 \\
  \frac{1}{r-1} & 2 \leq r \\
  \frac{2}{r} & -1 \leq r \leq 2
\end{cases}
\]

Clearly, if \( \Psi_{i-1} \) and \( \Psi_i \) are defined by (20), then the inequalities (16) - (17) are fulfilled and the scheme is TVD.

Next, we comment on how to solve the nonlinear algebraic equations (5) with (8) and (19). We propose it in the form of an iterative predictor and corrector procedure. Firstly, we predict the value of \( u_{i}^{n+1} \) by \( u_{i}^{n+1,0} \) that is obtained, e.g., from the first order scheme or with the second order scheme fixing \( \omega_i \) at some chosen value, see discussions in the section on numerical experiments.
Suppose that some predicted value \( u_i^{n+1,k} \approx u_i^{n+1} \) for \( k \geq 0 \) is available, then we define the value of \( r_i = r^k \) by replacing \( u_i^{n+1} \) with \( u_i^{n+1,k} \) in (10). Similarly, the value \( \omega_i = \omega_i^k \) from (19) or \( \Psi_i = \Psi_i^k \) from (20) is obtained. Now, solving the linear algebraic equation for the unknown \( u_i^{n+1} \):

\[
u_i^{n+1} + Cu_i^{n+1} = u_i^n + C \left( u_{i-1}^{n+1} + \frac{1}{2} \left( \frac{\Psi_i}{r_i} - \Psi_{i-1} \right) (u_{i-1}^{n+1} - u_i^n) \right), \tag{21}\]

one obtains the corrected value \( u_i^{n+1,k+1} \). If the difference \( |u_i^{n+1,k+1} - u_i^{n+1,k}| \) is acceptable, set \( u_i^{n+1} := u_i^{n+1,k+1} \). If \( u_i^{n+1} \neq u_i^{n+1,k+1} \), one has to choose between the flux \( F_{i+1/2}^{n+1} \) that preserves the TVD property, i.e. obtained from (21),

\[
F_{i+1/2}^{n+1} = C \left( u_i^{n+1} - \frac{1}{2} \left( (1 - \omega_i^k)(u_i^{n+1,k} - u_{i+1}^n) + \omega_i^k(u_{i-1}^{n+1} - u_i^n) \right) \right), \tag{22}
\]

or the locally conservative flux,

\[
F_{i+1/2}^{n+1} = C \left( u_i^{n+1} - \frac{1}{2} \left( (1 - \omega_i^k)(u_i^{n+1} - u_{i+1}^n) + \omega_i^k(u_{i-1}^{n+1} - u_i^n) \right) \right). \tag{23}
\]

In our numerical experiments, we prefer the conservative one, especially if only one corrector step is used, therefore a small violation of TVD property for the numerical solution can occur in general.

Note that to find the numerical solution in each time step we have to visit each point in the grid only once, and we compute the value \( u_i^{n+1} \) explicitly for the linear advection equation in each corrector step (very often only a single one).

Finally, we have to solve the case where \( C > 1 \). For that purpose we introduced in (8) and (9) the factors \( l_i \in [0,1] \) in the spirit of flux corrected transport schemes. In particular, instead of (14) we obtain

\[
\left( 1 - \frac{C}{2} \frac{l_i \Psi_i}{r_i} - l_{i-1} \Psi_{i-1} \right) (u_i^{n+1} - u_i^n) + C \left( 1 + \frac{1}{2} \frac{l_i \Psi_i}{r_i} - l_{i-1} \Psi_{i-1} \right) (u_i^{n+1} - u_{i-1}^{n+1}) = 0. \tag{24}
\]

Clearly, if \( l_i = l_{i-1} = 1 \), we obtain the origin (uncorrected) scheme. To have positive coefficients in (24) if \( C > 1 \), we have to require more restrictive inequalities than (16) and (17), namely,

\[
- \frac{1}{C} < l_{i-1} \Psi_{i-1} \leq 2, \quad -2 + l_{i-1} \Psi_{i-1} \leq \frac{l_i \Psi_i}{r} \leq \frac{2}{C} + l_{i-1} \Psi_{i-1}. \tag{25}
\]

Therefore, we define

\[
\omega_i = \begin{cases} 
\frac{1}{r-1} & \text{if } 2 \leq r \\
\frac{1+C}{C(1-r)} & \text{if } -\frac{1}{C} \leq r \leq \frac{2}{C} + l_{i-1} \Psi_{i-1} \\
1 & \text{otherwise}
\end{cases} \tag{26}
\]
or equivalently
\[ \Psi_i = \begin{cases} 2 & 2 \leq r \\ -1/C & r \leq \frac{-1}{C} \\ r & \text{otherwise} \end{cases} \] .

(27)

Finally,
\[ l_i = \min \left\{ 1, \max \left\{ 0, \frac{r_i}{\Psi_i} \left( \frac{2}{C} + l_{i-1}\Psi_{i-1} \right) \right\} \right\} .
\]

(28)

Using (26) - (28), one obtains the inequalities in (25) for arbitrary \( C \geq 1 \).

Before formulating the method with the semi-implicit scheme for scalar case in the next section, we comment briefly the case of nonlinear flux function \( f \). If, for example, \( f'(u) \geq 0 \) for \( u \in \mathbb{R} \), then we generalize (24) to the form
\[ (1 - \frac{1}{2} \frac{f_{i+1}^n - f_i^n}{u_{i+1}^n - u_i^n} \left( \frac{l_i \Psi_i}{r_i} - l_{i-1}\Psi_{i-1} \right)) (u_{i+1}^n - u_i^n) + \frac{f_{i+1}^n - f_{i-1}^n}{u_{i+1}^n - u_i^n} \left( 1 + \frac{1}{2} \frac{l_i \Psi_i}{r_i} - l_{i-1}\Psi_{i-1} \right) (u_{i+1}^n - u_{i-1}^n) = 0 ,
\]
when the indicators \( r_i \) are now determined by
\[ r_i = \frac{f_{i+1}^n - f_i^n}{f_{i+1}^n - f_{i-1}^n} .
\]

(30)

The important difference from (14) is that the constant Courant number \( C \) is replaced in (30) by nonlinear terms. However, if an estimate of the maximal Courant number \( C \) is available, the TVD property is preserved.

In the next section, we describe the details of the semi-implicit scheme for the scalar (nonlinear) hyperbolic equation including some algorithmic aspects.

4. The semi-implicit high resolution scheme

For simplicity, we suppose that the solution \( u \) of (1) has a compact support, so we can set
\[ u_0^{n+1} = u_0^n , \ u_i^{n+1} = u_i^n , \]
and
\[ u_I^{n+1} = u_I^n , \ u_{I-1}^{n+1} = u_{I-1}^n . \]

Furthermore, we suppose that Courant numbers \( C^+ \geq 0 \) and \( C^- \geq 0 \) are available such that
\[ \frac{\tau}{h} \max_u \frac{d}{du} f^+(u) \leq C^+ \quad \text{and} \quad \frac{\tau}{h} \max_u \frac{d}{du} f^-(u) \geq -C^- . \]

Moreover, \( \epsilon > 0 \) denotes a small enough constant. The default values for each time step, if not recomputed, are \( \omega_i = 0 , \ l_i = 1 , \) and \( \Psi_i = 1 . \) The scheme (5) using (8) for \( i = 2, 3, \ldots, I-2 \) or (6) using (9) for \( i = I-2, I-3, \ldots, 2 \) is then iteratively solved at the \( n \)-th time step as follows.
1. Compute
\[ \Delta^{up} = f^{+,n+1}_{i-1} - f^{+,n}_i \] or \[ \Delta^{up} = f^{-,n+1}_{i+1} - f^{-,n}_i. \]
If \( |\Delta^{up}| \leq \epsilon \) then set \( \omega_i = 1 \) and solve the algebraic equation (5) or (6) for the unknown \( u^{n+1}_i \). Continue with the step 1 for \( i + 1 \) or \( i - 1 \).

2. If \( |\Delta^{up}| > \epsilon \) then set an initial guess \( u^0 \approx u^{n+1}_i \), e.g., by solving for \( u 
\frac{\tau}{h} f^+(u) = u^n_i + \frac{\tau}{h} F^{+,n+1}_{i-1/2} 
\) or \n\frac{\tau}{h} f^-(u) = u^n_i - \frac{\tau}{h} F^{-,n+1}_{i+1/2}. 
\)
3. For the value \( u^k \approx u^{n+1}_i \) for some \( k \geq 0 \) compute
\[ \Delta^{dw,k} = f^{+,n+1}_i - f_i^{+,n} \] or \[ \Delta^{dw,k} = f^{-,n+1}_i - f_i^{-,n}. \]
If \( |\Delta^{dw,k}| \leq \epsilon \) then proceed with the step 5.
4. If \( |\Delta^{dw,k}| > \epsilon \) then compute
\[ r^k = \frac{\Delta^{up}}{\Delta^{dw,k}}, \]
and
\[ \omega^k_i = \begin{cases} 
\frac{1}{r^k - 1} & \text{if } 2 \leq r^k \\
\frac{1 + C}{\omega^k_i (r^k - 1)} & \text{if } r^k \leq -\frac{1}{C} \\
1 & \text{otherwise,}
\end{cases} \tag{31} \]
with \( C = \max\{1, C^+\} \) or \( C = \max\{1, C^-\} \), respectively. Furthermore, \n\psi^k_i = 1 - \omega^k_i + \omega^k_i r^k \)
and if \( \psi^k_i \neq 0 \) then \n\ell^k_i = \min\{1, \max\{0, \frac{r^k}{\psi^k_i} \left( \frac{2}{C} + \ell_i-1 - 1 \right) \} \}, \)
or \n\ell^k_i = \min\{1, \max\{0, \frac{r^k}{\psi^k_i} \left( \frac{2}{C} + \ell_i+1 - 1 \right) \} \}.
\)
5. Having set all the \( k \)-th estimates of the input parameters, we solve the algebraic equation (5) or (6) and denote its solution by \( u^{n+1,k+1}_i \). If a chosen stopping criterion is fulfilled, we set \( u^{n+1}_i = u^{n+1,k+1}_i \) and proceed with the step 1 for \( i + 1 \) or \( i - 1 \). If not, we proceed with the step 3.

We note that to improve accuracy, one can replace \( C^+ \) in (31) by \( C^+_i \) defined by \n\frac{f^+(u^{n+1+1,k}) - f^+(u^n_i)}{u^{n+1+1,k}_i - u^n_i} \quad u^{n+1,k}_i \neq u^n_i 
\) and analogously for \( C^- \).
5. Hyperbolic systems

Concerning systems of hyperbolic equations, one has to take the steps defined for the scalar case in previous sections for each component of the system. Similarly to experiences in [23, 13], we prefer to express the second order update of numerical fluxes with the help of characteristic variables and characteristic speeds (the eigenvalues). In what follows, we explain some details.

We solve the system for $f : \mathbb{R}^m \rightarrow \mathbb{R}^m$

$$\partial_t u + \partial_x f(u) = 0,$$

where we suppose that the Jacobian $f'(u)$ has only nonnegative real eigenvalues $\lambda^p, p = 1, 2, \ldots, m$. The systems with nonpositive eigenvalues are treated analogously, the general case can be solved using the fractional step method as explained before in (5) - (6) [25].

Let the columns of the matrix $R = R(u)$ be given by the eigenvectors $r^p, p = 1, 2, \ldots, m$. Due to hyperbolicity, the matrix $R$ is regular for each considered value of $u$. Let $u$ be the last estimate of $u$ and let $R^{-1}$ be the inverse matrix of $R(u)$. We express the terms in the second order update of the semi-implicit scheme (8) as a linear combination of the eigenvectors, namely

$$\alpha_i = R^{-1} \cdot (f^{k,n+1}_i - f^n_i) , \quad \beta_i = R^{-1} \cdot (f^{n+1}_i - f^n_i).$$

The idea is that the weights in $w_i = (w^1_i, w^2_i, \ldots, w^m_i)$ and $l_i = (l^1_i, l^2_i, \ldots, l^p_i)$ are now associated with the coefficients $\alpha_i$ and $\beta_i$, so the fluxes in (8) take the form,

$$F^{n+1}_{i+1/2} = f^n_i + \frac{1}{2} \sum_p l^p_i \left((1 - w^p_i) \alpha_i^p + w^p_i \beta_i^p \right) r^p_i. \quad (32)$$

Having the form (32), the high-resolution form of the scalar case is used for each component of the system with the indicators defined by

$$r^p_i = \frac{\beta^p_i}{\alpha^p_i}, \quad p = 1, 2, \ldots, m. \quad (33)$$

Furthermore, the Courant numbers $C^+$ in (31) are replaced by the corresponding values of the eigenvalues for each component.

6. Numerical experiments

In what follows, we illustrate numerical resolutions of the proposed semi-implicit high-resolution scheme for several standard test problems.

When computing examples for Burgers equation with $f(u) = u^2/2$, we use the approach of [25] when the splitting (3) is obtained by

$$f^+(u) := \frac{1}{2} \left(f(u) + |u| \frac{u}{2}\right), \quad f^-(u) := \frac{1}{2} \left(f(u) - |u| \frac{u}{2}\right).$$
6.1. Linear advection

To illustrate the TVD property, we solve the test example \cite{3, 6} with non-smooth solutions for the advection with constant unity speed. The initial condition consists of four different segments - a Gaussian, a triangle, a square-wave and a semi-ellipse, see \cite{3, 6} for a complete definition. The problem is solved with an integer Courant number, and numerical solutions are shifted backward after each time step to return to the initial position. To compute the predicted value \( u_{n+1,i} \), the scheme with \( w_{i}^{0} = 0 \) is used and only one correction step is computed.

For a visual comparison see Figure 1 where a clear improvement with respect to the first order scheme can be seen. Moreover, no over- or under-shootings with a magnitude larger than rounding errors are observed.

6.2. The smooth solution of Burgers equation

In this example, we test the method for the fixed choice of \( \omega \equiv 1 \) in the case of a smooth solution of Burgers equation. Namely, we define

\[
 f(u) = \frac{u^2}{2}, \quad u(x, 0) = 1 + \frac{1}{8} \sin(2\pi x), \quad x \in [0, 1],
\]
and we solve the equation for $t \in [0, 1]$. The exact solution is computed numerically using the method of characteristics by solving the algebraic equations for $u = u(x_i, t^n)$:

$$u = 1 + \frac{1}{8} \sin(2\pi(x_i - ut^n)).$$

In Figure 3, a comparison of the exact and numerical solutions obtained with the first and the second order method are compared at the final time for two grids. The global $l_1$ discrete error in time and space

$$E^N_I := h\tau \sum_{i=0}^{I} \sum_{n=1}^{N} |u^n_i - u(x_i, t^n)|$$

is equal for the coarse grid $E^{20}_{80} = 9.09 \cdot 10^{-4}$ and the EOC (the Experimental Order of Convergence) equals for $I = 160$ and $I = 320$ to 2.08 and 2.17, respectively.

6.3. Slowly moving shock of Burgers equation

Inspired by [25], we present numerical solutions of Riemann problem with a slowly moving shock. The initial discontinuity of the piecewise constant function is placed at $x = -0.5$ with the left value $u_L = 20$ and the right value $u_R = -18$. Consequently, the shock speed is equal to 1. We present the comparison of the first order and the high-resolution scheme at $t = 1$ in Figure 3 for two rather coarse meshes with $I = 20$ and $I = 40$ with the time step $\tau = h/20$ corresponding to the maximal Courant number equal to 10. One can observe a significantly improved approximation of the shock speed for numerical solutions obtained with the high-resolution scheme when compared with the first order scheme.

6.4. Burgers equation with interacting shock and rarefaction

The last example of Burgers equation is taken from [25] and contains typical features of solutions for Riemann problems. The initial condition is given by

$$u(x, 0) = \begin{cases} 
1 & 0.3 < x < 0.6 \\
-0.2 & \text{otherwise}
\end{cases}$$
Figure 4: The comparison of exact solutions (orange) with numerical solutions obtained with the first order method (green) and the high-resolution method (blue) for the example in Section 6.4. The first line presents the results at $t = 0.5$ (left) and $t = 1$ (right) and the second line at $t = 1.5$ and $t = 2$. The number of mesh points is $I = 640$, the maximal Courant number is 4 and the numerical solutions are obtained with only one corrector step with the predictor using the first order scheme.
and the exact solution can be found in \cite{25}. At time $t = 1$ the end points of a rarefaction and a shock wave merge and the solution evolves further with a triangular profile. In Figure 4 one can see that the first order scheme approximates the exact solution at $t = 1$ with a visibly larger error than the high-resolution scheme. Probably, this is a reason why the position of the shock for $t > 1$ is significantly better approximated with the high-resolution method. Both methods converge to the exact solution with respect to the error defined in (34), see Table 1.

6.5. Linear hyperbolic system

To test the method for systems of conservation laws, we begin with a simple linear system having a constant matrix,

$$f = f(q) = f(q_1, q_2) = A \cdot q,$$

where $A = \begin{pmatrix} 1.1 & -0.9 \\ -0.9 & 1.1 \end{pmatrix}$. The matrix has positive eigenvalues 1 and 0.1. The initial condition consists of rectangular profiles, see the first row in Figure 5.

The example is computed with Courant number 10, so only the slowly moving waves shall be well resolved by numerical solutions. The predicted values are computed with the second order scheme using $\omega = 0$, afterwards only one corrector step is used. One can clearly see that numerical solutions do not contain visible oscillations and that the contact discontinuities are well resolved for slowly moving waves and smeared for fast moving discontinuities.

6.6. Shallow water equation

Finally, we test the method for the simple example \cite{23} of nonlinear shallow water system using two equations

$$\partial_t h + \partial_x (hu) = 0, \quad h(x, 0) = 1 + 0.4 \exp -5(x - 5)^2,$$

$$\partial_t (hu) + \partial_x (hu^2 + 0.5h^2) = 0, \quad u(x, 0) = 0,$$

for $x \in [0, 10]$ and $t \in [0, 2]$. The system is discretized with conservative variables $(h, hu)$ using the Lax-Friedrichs splitting (4) and $\alpha = 1.3$. A comparison of
Figure 5: The comparison of exact solutions $q_1$ (green) and $q_2$ (red) with numerical solutions (blue and orange, respectively) obtained with the high-resolution method for the example in Section 6.5. The first line is the initial condition, the second row is for $I = 400$ and $t = 0.15$ (the first column) and $t = 0.4$ (the second column), and the third row is for $I = 800$ and the analogous times. The constant Courant number is 10 and the numerical solutions are obtained with only one corrector step.
Figure 6: The comparison of numerical solutions obtained with the first order method (orange) and the high-resolution method (blue) for the example in Section 6.6. The first column is for $t = 1$, the second one for $t = 2$. The first row compares $h$ for $I = 400$, the second one $h$ for $I = 800$, the third one $hu$ for $I = 400$ and the fourth one $hu$ for $I = 800$. The maximal Courant number is always 6.21.
Figure 7: The comparison of numerical solutions obtained with the first order method (orange) and the high-resolution method (blue) at $t = 1$ (left) and $t = 2$ (right). The first order method was computed with $I = 800$ and the high-resolution one with $I = 200$. The maximal Courant number is always $6.21$. 
results at $t = 1$ and $t = 2$ for the first order \[25\] method and the high-resolution method is given in Figure 6 for two fine grids. The maximal Courant number is around 6.21, the predicted values are computed with the second order semi-implicit method using $\omega = 0$, and only one corrector step is used. In Figure 6 one can see a significantly improved resolution of shock and rarefaction waves when comparing the high-resolution method with the first order accurate one. The results resemble well those presented in \[23\].

To make the difference in the resolution even clearer, we compare in Figure 7 the results obtained on a coarse grid with the high-resolution method and the results obtained by the first order accurate method on two times uniformly refined grid that still do not have the quality of the high-resolution method.

7. Conclusion

We have presented the semi-implicit conservative finite difference method for hyperbolic problems in one-dimensional case. The method shares the advantageous properties of the first order accurate implicit method. Namely, the method is unconditionally stable for the linear advection equation and non-oscillatory numerical solutions are obtained explicitly after two sweeps of the fast sweeping method. In the case of nonlinear scalar hyperbolic PDEs one has to solve for each grid point a single nonlinear algebraic equation with nonlinearity only due to the nonlinear flux function. All these properties are preserved in the proposed high-resolution (TVD) method, which is second order accurate if the solution is smooth. Although the TVD limiters depend on the single unknown per each grid point, the nonlinearity can be typically resolved with one predictor and one corrector step for each algebraic equation. The method is applied successfully for the linear system of hyperbolic PDE and for the shallow water equations by expressing second order correction terms in the scheme using characteristic variables and speeds.

The proposed semi-implicit methods can be used and extended for problems where up to now fully implicit or explicit-implicit schemes appeared useful. In addition to accuracy requirements, the method does not require formal restrictions on the choice of time steps $\Delta t$ for stability reasons. A possible restriction on $\Delta t$ due to slow or no convergence of the nonlinear algebraic solver is shared with the first order accurate implicit method. We plan to extend the method analogously to \[34\] with high-order WENO type spatial reconstruction and Lax-Wendroff type of time discretization.

8. Appendix

In what follows, we motive the form of fluxes \[8\] - \[9\] for the semi-implicit method. Let $u$ be sufficiently smooth solution of \[1\] with smooth flux function $f \equiv f^+$. The first order accurate scheme takes the form

$$u_i^{n+1} - u_i^n + \frac{\Delta t}{\Delta x} (f_i^{n+1} - f_i^n) = 0.$$  \hspace{1cm} (35)
Using finite Taylor series we can express the dominant error term of the scheme \((35)\) by
\[
\Delta t \Delta x \frac{\partial^2 f(u(x_i, t^{n+1}))}{\partial x^2} - \frac{\Delta t^2}{2} \frac{\partial f(u(x_i, t^{n+1}))}{\partial x}.
\]
(36)

Note that we keep the mixed derivative in \((36)\), so we follow the Lax-Wendroff procedure of replacing every time derivative using the PDE \(\partial_t u = -\partial_x f\) only partially \([13, 9, 15, 14]\).

Now applying the following approximations in \((36)\)
\[
\Delta x \frac{\partial f(u(x_i, t^{n+1}))}{\partial x} - \Delta t \frac{\partial f(u(x_i, t^{n+1}))}{\partial t} \approx f^{n+1}_i - f^{n+1}_{i-1} - (f^{n+1}_i - f^n_i) \quad (37)
\]
and the parametric approximation (analogously for \((x_{i-1}, t^{n+1})\))
\[
\Delta x \frac{\partial f(u(x_i, t^n))}{\partial x} \approx (1 - \omega)(f^{n+1}_{i+1} - f^n_i) + \omega(f^n_i - f^n_{i-1}), \quad (38)
\]
the numerical fluxes \((8)\) are recovered for \(l_i \equiv 1\).

References

[1] T. Arbogast, C.-S. Huang, X. Zhao, and D. N. King, A third order, implicit, finite volume, adaptive Runge–Kutta WENO scheme for advection–diffusion equations, Comput. Meth. Appl. Mech. Eng., 368 (2020), pp. 113–155.

[2] S. Avgerinos, F. Bernard, A. Iollo, and G. Russo, Linearly implicit all Mach number shock capturing schemes for the Euler equations, J. Comp. Phys., 393 (2019), pp. 278–312.

[3] D. S. Balsara and C.-W. Shu, Monotonicity preserving weighted essentially non-oscillatory schemes with increasingly high order of accuracy, J. Comp. Phys., 160 (2000), pp. 405–452.

[4] W. Barsukow, P. V. Edelmann, C. Klingenberg, F. Miczek, and F. K. Röpke, A numerical scheme for the compressible low-Mach number regime of ideal fluid dynamics, Journal of Scientific Computing, 72 (2017), pp. 623–646.

[5] C. Berthon, C. Klingenberg, and M. Zenk, An all Mach number relaxation upwind scheme, The SMAI J. Comp. Math., 6 (2020), pp. 1–31.

[6] R. Borges, M. Carmona, B. Costa, and W. S. Don, An improved weighted essentially non-oscillatory scheme for hyperbolic conservation laws, J. Comp. Phys., 227 (2008), pp. 3191–3211.

[7] W. Boscheri, G. Dimarco, R. Loubère, M. Tavelli, and M.-H. Vignal, A second order all Mach number IMEX finite volume solver for the three dimensional Euler equations, J. Comp. Phys., 415 (2020), p. 109486.
[8] L. D. Carciopolo, L. Bonaventura, A. Scotti, and L. Formaggia, *A conservative implicit multirate method for hyperbolic problems*, Comput. Geosci., 23 (2019), pp. 647–664.

[9] H. Carrillo and C. Parés, *Compact approximate Taylor methods for systems of conservation laws*, Journal of Scientific Computing, 80 (2019), pp. 1832–1866.

[10] H. Carrillo, C. Parés, and D. Zorío, *Lax–Wendroff approximate Taylor methods with fast and optimized weighted essentially non-oscillatory reconstructions*, Journal of Scientific Computing, 86 (2021), pp. 1–41.

[11] M. Q. de Luna and D. I. Ketcheson, *Maximum principle preserving space and time flux limiting for Diagonally Implicit Runge-Kutta discretizations of scalar convection-diffusion equations*, arXiv preprint arXiv:2109.08272, (2021).

[12] M. Dumbser, C. Enaux, and E. F. Toro, *Finite volume schemes of very high order of accuracy for stiff hyperbolic balance laws*, J. Comp. Phys., 227 (2008), pp. 3971–4001.

[13] K. Duraisamy and J. D. Baeder, *Implicit Scheme for Hyperbolic Conservation Laws Using Nonoscillatory Reconstruction in Space and Time*, SIAM J. Sci. Comput., 29 (2007), pp. 2607–2620.

[14] P. Frolkovič, S. Kríšková, M. Rohová, and M. Žeravý, *Semi-implicit methods for advection equations with explicit forms of numerical solution*, arXiv:2106.15474, (2021). accepted to JJIAM.

[15] P. Frolkovič and K. Mikula, *Semi-implicit second order schemes for numerical solution of level set advection equation on Cartesian grids*, Appl. Num. Math., 329 (2018), pp. 129–142.

[16] P. Frolkovič, K. Mikula, and J. Urbán, *Semi-implicit finite volume level set method for advective motion of interfaces in normal direction*, Appl. Num. Math., 95 (2015), pp. 214–228.

[17] P. Frolkovič, *Semi-implicit methods based on inflow implicit and outflow explicit time discretization of advection*, in Proc. ALGORITMY, Spektrum STU Bratislava, 2016, pp. 165–174.

[18] S. Gottlieb, Z. J. Grant, J. Hu, and R. Shu, *High Order Strong Stability Preserving MultiDerivative Implicit and IMEX Runge-Kutta Methods with Asymptotic Preserving Properties*, SIAM J. Numer. Anal., 60 (2022), pp. 423–449.

[19] J. Hahn, K. Mikula, P. Frolkovič, M. Medl’a, and B. Basara, *Iterative inflow-implicit outflow-explicit finite volume scheme for level-set equations on polyhedron meshes*, Comput. Math. with Appl., 77 (2019), pp. 1639–1654.
[20] M. Han Veiga, P. Öffner, and D. Torlo, *Dec and Ader: similarities, differences and a unified framework*, Journal of Scientific Computing, 87 (2021), pp. 1–35.

[21] G. Ibolya and K. Mikula, *Numerical Solution of the 1d Viscous Burgers’ and Traffic Flow Equations by the Inflow-Implicit/Outflow-Explicit Finite Volume Method*, in Proc. ALGORITMY, Spektrum STU Bratislava, 2020, pp. 191–200.

[22] D. Kuzmin, M. Quezada de Luna, D. I. Ketcheson, and J. Gröll, *Bound-preserving flux limiting for high-order explicit Runge–Kutta time discretizations of hyperbolic conservation laws*, Journal of Scientific Computing, 91 (2022), pp. 1–34.

[23] R. J. Leveque, *Finite Volume Methods for Hyperbolic Problems*, Cambridge UP, 2nd ed., 2004.

[24] J. Li, *Two-stage fourth order: temporal-spatial coupling in computational fluid dynamics (CFD)*, Adv. Aerodyn., 1 (2019), p. 3.

[25] E. Lozano and T. D. Aslam, *Implicit fast sweeping method for hyperbolic systems of conservation laws*, J. Comp. Phys., 430 (2021), p. 110039.

[26] B. J. McCartin, *The method of angled derivatives*, Appl. Num. Math., 170 (2005), pp. 440–461.

[27] K. Mikula, M. Ohlberger, and J. Urbán, *Inflow-implicit/outflow-explicit finite volume methods for solving advection equations*, Appl. Numer. Math., 85 (2014), pp. 16–37.

[28] G. Puppo, M. Semplice, and G. Visconti, *Qinpi: Integrating Conservation Laws with CWENO Implicit Methods*, Commun. Appl. Math. Comput., (2022).

[29] J. Qiu and C.-W. Shu, *Finite Difference WENO Schemes with Lax–Wendroff-Type Time Discretizations*, SIAM J. Sci. Comp., 24 (2003).

[30] D. C. Seal, Y. Güçlü, and A. J. Christlieb, *High-Order Multiderivative Time Integrators for Hyperbolic Conservation Laws*, J. Sci. Comput., 60 (2014), pp. 101–140.

[31] C.-W. Shu, *Numerical experiments on the accuracy of ENO and modified ENO schemes*, J. Sci. Comput., 5 (1990), pp. 127–149.

[32] C.-W. Shu, *Essentially non-oscillatory and weighted essentially non-oscillatory schemes for hyperbolic conservation laws*, in Advanced Numerical Approximation of Nonlinear Hyperbolic Equations, Lecture Notes in Mathematics, Springer, Berlin, Heidelberg, 1998, pp. 325–432.
[33] A. Thomann, A. Iollo, and G. Puppo, Implicit relaxed all Mach number schemes for gases and compressible materials, arXiv preprint arXiv:2112.14126, (2021).

[34] V. Titarev and E. Toro, Weno schemes based on upwind and centred TVD fluxes, Computers & Fluids, 34 (2005), pp. 705–720.

[35] E. F. Toro, Riemann solvers and numerical methods for fluid dynamics: a practical introduction, Springer, Dordrecht; New York, 3rd ed., 2009.

[36] A. Y. J. Tsai, R. P. K. Chan, and S. Wang, Two-derivative Runge–Kutta methods for PDEs using a novel discretization approach, Numer. Alg., 65 (2014), pp. 687–703.

[37] J. Zeifang, J. Schuetz, K. Kaiser, A. Beck, M. Lukáčová-Medviďová, and S. Noelle, A novel full-euler low Mach number IMEX splitting, Communications in Computational Physics, 27 (2020), pp. 292–320.

[38] J. Zeifang and J. Schütz, Implicit two-derivative deferred correction time discretization for the discontinuous Galerkin method, Journal of Computational Physics, (2022), p. 111353.

[39] D. Zorio, A. Baeza, and P. Mulet, An Approximate Lax–Wendroff-Type Procedure for High Order Accurate Schemes for Hyperbolic Conservation Laws, J. Sci. Comput., 71 (2017), pp. 246–273.