# CONVERGENCE OF IMPLICIT FINITE DIFFERENCE METHODS APPLIED TO NONLINEAR MIXED SYSTEMS

H. J. SCHROLL *

— *SIAM J. Numer. Anal. 33(3), 1996, pp 997-1013.* —

**Abstract.** A new technique to prove convergence of finite difference methods applied to nonlinear PDEs arising in computational fluid dynamics is presented. The underlying systems may be hyperbolic, parabolic or of mixed type like the Navier–Stokes equations. Implicit finite difference methods are analyzed. The essential idea leading to success is the introduction of a pilot function, that is highly attractive to the numerical approximation and converges itself to the solution of the underlying system.

**Key words.** finite difference method, mixed system, stability and convergence

**AMS(MOS) subject classifications.** 65M10

**1. Introduction.** The purpose of this paper is to introduce the *pilot function* as a tool in nonlinear convergence analysis. The technique presented here is a synthesis of two earlier contributions by Strang [12] and López–Marcos and Sanz–Serna [7]. Strang analyzed explicit methods applied to nonlinear hyperbolic systems. He constructed a high order approximate solution to the difference scheme by an ansatz

$$u^{pi} = \widetilde{u} + \Delta x f + \Delta x^2 g + \dots, \tag{1}$$

where $f$ and $g$ are smooth functions of space and time. $\Delta x$ is the spatial mesh size. $\widetilde{u}$ denotes the solution of the underlying Cauchy problem which is assumed to exist and to be smooth. Applying stability of the scheme linearized at $\widetilde{u}$ Strang proved the exact solution of the method $U$ to converge to $u^{pi}$.

We define the pilot function by the same ansatz (1) but, in contrast to Strang, we linearize the scheme at $u^{pi}$. By means of energy estimates we show $L^2$–stability of this linearization. Applying ideas of López–Marcos and Sanz–Serna we establish a family of step size dependent stability regions around the pilot function. This combination of Strang's ansatz and the stability concept of López–Marcos and Sanz–Serna is the novelty of the present paper. As a result of the theory, the computed approximation $U$ stays inside the shrinking stability region as the step size tends to zero. Consequently, it converges to $u^{pi}$ as well as to the desired solution $\widetilde{u}$.

The union of all stability regions defines a "stability corridor" that guarantees a "safe approach" to $\widetilde{u}$. With this interpretation $u^{pi}$ acts like a pilot boat leading a sea ship to the port. We will demonstrate that this argument leads to

* RWTH-Aachen, Numerische Mathematik, Templergraben 55, D-52056 Aachen, Germany. E-mail: schroll@igpm.rwth-aachen.de

convergence results not only for nonlinear hyperbolic, but also for parabolic and mixed systems.

The basic concept is similar to the modified equation approach [1] but, to the knowledge of the author, the modified equation approach has not yet been successfully applied to nonlinear mixed systems. The point is, that the convergence of the solution of the parabolized equation, which is the obvious modified equation, to the solution of the original system may be a problem in general.

Here, we should mention that Sanz–Serna [9] refers to the paper by Strang [12] in cases where the theory developed in [7] and [9] does not apply due to a lack of consistency. In fact, Strang's idea was used earlier by Sanz–Serna [8] but in combination with Stetter's local stability theory [11]. In a later paper [6] López–Marcos and Sanz–Serna discussed why their stability concept should be favored in general.

This paper is organized as follows: In the next section the underlying systems and the finite difference methods are defined. Furthermore, the main result is formulated. Since the analysis relies heavily on the step size dependent stability theory by López–Marcos and Sanz–Serna, we will review their result in section 3. In section 4 we define the pilot function. Energy estimates are enclosed in section 5. The more technical Lipschitz estimates are delayed to section 6. Finally, in section 7 we prove the main result.

**2. Mixed systems and difference approximations.** The book by Kreiss and Lorenz [5] provides a very good survey on the theory of hyperbolic, parabolic and mixed systems. We will apply some of the basic concepts and the existence result introduced in that book.

Mixed systems are PDEs consisting of two kinds of subsystems. If we ignore the nonlinear coupling, the first part is a parabolic and the second a hyperbolic system. To be more concrete, we consider first order systems

$$u_t + A(u,t,x)\,u_x = C(u,t,x), \tag{2}$$

second order systems

$$u_t + A(u,t,x)\,u_x = B(u,t,x)\,u_{xx} + C(u,t,x) \tag{3}$$

and finally systems of mixed type

$$u_t + \begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix}(u,t,x)\,u_x = \begin{pmatrix} B_{11} & 0 \\ 0 & 0 \end{pmatrix}(u,t,x)\,u_{xx} + C(u,t,x). \tag{4}$$

In order to have unique solutions the equations have to be supplemented by initial and boundary conditions. Mathematically, the easiest case is the spatial periodic Cauchy problem with smooth initial data

$$u(0,x) = z(x), \quad x \in I\!R, \quad z \in C_p^\infty(I\!R, I\!R^m). \tag{5}$$

By $C_p^\infty$ we denote $C^\infty$–functions which are 1–periodic with respect to $x$. The type of the above systems depends on the coefficients $A, B$ and therefore on $u$ itself. Concerning Cauchy problems the type is defined in a neighborhood of the initial data

$$U_\eta := \{u \in I\!R^m : |u - u(0,x)| < \eta \text{ for some } x \in I\!R\}, \quad \eta > 0.$$

The coefficients are assumed to be smooth and 1–periodic

$$A, B \in C_p^\infty(U_\eta \times I\!R^+ \times I\!R, I\!R^{m\times m}), \quad C \in C_p^\infty(U_\eta \times I\!R^+ \times I\!R, I\!R^m). \tag{6}$$

In the following, the term periodic or 1–periodic instead of 1–periodic with respect to $x$ will be frequently used. For symmetric matrices we write $A \le B$ if and only if $u^T A u \le u^T B u$ for all $u \in I\!R^n$. If $\alpha \in I\!R$, $A \le \alpha$ is an abbreviation for $A \le \alpha I$.

DEFINITION 2.1. *The system (2) is strongly hyperbolic, if there exists a symmetric matrix function $H \in C_p^\infty(U_\eta \times I\!R^+ \times I\!R, I\!R^{m\times m})$ and a positive constant $\alpha$ such that $\alpha \le H$ and $HA$ is symmetric. $A$ is called symmetrizable and $H$ is the symmetrizer.*

DEFINITION 2.2. *The system (3) is strongly parabolic, if there exists a constant $\beta$ such that $B + B^T \ge 2\beta > 0$.*

DEFINITION 2.3. *The system (4) is a mixed system, if there exists a symmetrizer to $A_{22}$ and a constant $\beta$ such that $B_{11} + B_{11}^T \ge 2\beta > 0$.*

Mixed type systems arise in fluid dynamics. A well known example are the compressible Navier–Stokes equations. If thermodynamic processes can be neglected the flow is called isentropic. Such a flow can be described by the conservation of momentum and mass. If $\rho$ is the density, $v$ the velocity and $p = p(\rho)$ the pressure of the flowing medium, the equations in one space dimension read

$$\begin{array}{ccccl} (\rho v)_t & + & (\rho v^2 + p)_x & = & \epsilon v_{xx}, \quad 0 < \epsilon \ll 1 \\ \rho_t & + & (\rho v)_x & = & 0. \end{array}$$

The viscosity coefficient $\epsilon$ is a small positive constant. Using $(\rho v)_t = \rho_t v + \rho v_t$ and inserting $\rho_t$ from the continuity equation this system can be written in non–conservative form

$$\begin{pmatrix} v \\ \rho \end{pmatrix}_t + \begin{pmatrix} v & p\prime(\rho)/\rho \\ \rho & v \end{pmatrix} \begin{pmatrix} v \\ \rho \end{pmatrix}_x = \begin{pmatrix} \epsilon/\rho & 0 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} v \\ \rho \end{pmatrix}_{xx}.$$

Since $\epsilon/\rho$ is positive, this is a mixed system in the sense of Definition 2.3. In the case of a polytropic medium, one has to add a additional equation for the temperature $T$

$$\begin{pmatrix} T \\ v \\ \rho \end{pmatrix}_t + \begin{pmatrix} v & (\gamma-1)T & 0 \\ p_T & v & p_\rho \\ 0 & \rho & v \end{pmatrix} \begin{pmatrix} T \\ v \\ \rho \end{pmatrix}_x = \frac{1}{\rho Re} \begin{pmatrix} \gamma k/Pr & 0 & 0 \\ 0 & 2\mu + \lambda & 0 \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} T \\ v \\ \rho \end{pmatrix}_{xx}.$$

Here the pressure $p = \rho RT$ is a function of $\rho$ and $T$. $R$ is the gas constant, $\gamma$ is the ratio of specific heats, $k$ is the heat conduction constant, $\mu$ and $\lambda$ are the Lamé constants, $Re$ is the Reynolds number, and $Pr$ the Prandtl number. Again, the system is of mixed type. An other example are the viscous shallow water equations; cf. [2].

Now, we are in the position to state the existence result. It can be found in [5] (theorems 5.2.3, 3.4.3, 3.3.3 and 3.7.2).

THEOREM 2.1. *In case of a strongly hyperbolic, parabolic or mixed system the periodic Cauchy problem (5) and (6) has locally a unique solution: There exists a time $T > 0$ and a unique solution $\tilde{u} \in C_p^\infty([0,T] \times I\!R, U_\eta)$.*
*In case of linear systems, this solution exists for all time $t > 0$ and the Cauchy problem is well posed in $L^2$.*

Recently Hagstrom and Lorenz [3] proved all–time existence of smooth solutions for a class of nonlinear mixed systems including the compressible Navier–Stokes equations. They assume initial data close to a constant state.

Unfortunately, only a few convergence, existence and uniqueness results concerning more general weak solutions of nonlinear systems are known. Therefore, it is reasonable to study whether a special method can at least approximate the solution as long as it stays smooth. This follows exactly the spirit of Strang and Michelson.

To specify the finite difference methods, we introduce some notation. First the domain $\Omega := [0,T] \times (0,1]$ is covered by a regular grid

$$\begin{array}{lcl} \Omega_{\Delta t} & := & \{t_n = n\Delta t,\ n = 0,1,\dots N,\ \Delta t = T/N\} \subset [0,T] \\ \Omega_{\Delta x} & := & \{x_i = i\Delta x,\ i = 1,2,\dots M,\ \Delta x = 1/M\} \subset (0,1] \\ \Omega_h & := & \Omega_{\Delta t} \times \Omega_{\Delta x} \subset \Omega \\ h & := & (\Delta t, \Delta x)^T. \end{array} \tag{7}$$

Let $\mathcal{F}_p(\Omega_*, Y)$ denote the finite dimensional vector space of periodic grid functions $u : \Omega_* \to Y$. $\Omega_*$ may either be $\Omega_h$ or $\Omega_{\Delta x}$ and $Y$ stands for $U_\eta$ or $I\!R^m$. For $u \in \mathcal{F}_p(\Omega_h, Y)$ we denote by $u_n$ the periodic function $u(t_n, \cdot) \in \mathcal{F}_p(\Omega_{\Delta x}, Y)$. If $u \in \mathcal{F}_p(\Omega_{\Delta x}, Y)$ then $u_i$ is the value $u(x_i) \in I\!R^m$. Furthermore, $u|_h$ or $u|_{\Delta x}$ denotes the restriction of a continuous function to $\Omega_h$ or $\Omega_{\Delta x}$ respectively. Using the shift operator

$$(Eu)_i := u_{i+1}, \quad (E^j u)_i := u_{i+j}, \quad u \in \mathcal{F}_p(\Omega_{\Delta x}, I\!R^m),$$

we define forward, backward

$$D_+ := \frac{(E - I)}{\Delta x}, \quad D_- := \frac{(I - E^{-1})}{\Delta x}$$

and central finite difference operators

$$D_0 := \frac{(D_+ + D_-)}{2}, \quad D_+ D_- := \frac{(E - 2I + E^{-1})}{\Delta x^2}.$$

The difference methods that will be analyzed, are constructed by replacing the partial derivatives by finite difference operators. Replacing only space derivatives –in a first step– results in the method of lines system which is a high dimensional ODE system given by

$$\frac{\partial}{\partial t}\, u(t,\cdot) + P_{\Delta x}(u(t,\cdot)) = 0. \tag{8}$$

Here $P_{\Delta x}$ represents a nonlinear, finite difference operator discretizing the space variable. This operator acts on $\mathcal{F}_p(\Omega_{\Delta x}, U_\eta)$. Three typical examples will be given below. In order to discretize the time variable, we concentrate on a implicit Euler method. The fully discrete method can then be written as a root equation

$$\Phi_h(u) = 0 \tag{9}$$

for the operator

$$\begin{aligned} \Phi_h \;:\; D_h := \mathcal{F}_p(\Omega_h, U_\eta) \quad &\rightarrow \quad \mathcal{F}_p(\Omega_h, I\!R^m) \\ u \quad &\mapsto \quad \Phi_h(u) := \begin{pmatrix} u_0 - z|_{\Delta x} \\ \dfrac{u_n - u_{n-1}}{\Delta t} + P_{\Delta x}(u_n) \\ n = 1, 2, \ldots N \end{pmatrix}. \end{aligned} \tag{10}$$

One can also apply algebraically stable implicit Runge–Kutta methods, but our intention here is just to describe the main principles. Further details can be found in [10].

We will discuss three important types of space discretizations applied to a hyperbolic, parabolic or mixed system of the general form

$$u_t + A(u)\, u_x = B(u)\, u_{xx}. \tag{11}$$

The easiest case, and therefore the first one, is a central difference operator

$$P_{\Delta x}(u)_i := A(u_i)(D_0 u)_i - (B(u_i) + \delta \Delta x I)(D_+ D_- u)_i,\ \delta \geq 0, u \in \mathcal{F}_p(\Omega_{\Delta x}, U_\eta) \tag{12}$$

Here, $\delta \Delta x I D_+ D_- u$ is a artificial damping term.

With regard to the hyperbolic character of the equation so–called upwind methods are widely used. The idea is to decompose the hyperbolic part into two matrices such that the corresponding spectra have a unique sign. In the more general non symmetric case we assume

$$\begin{aligned} &A_{22} = A_{22}^+ + A_{22}^-, \quad A_{22}^\pm \text{ simultanously symmetrizable,} \\ &\text{and } Re(F(\pm A_{22}^\pm)) \subset [\gamma, \infty), \quad \gamma > 0, u \in U_\eta \end{aligned} \tag{13}$$

where $F(A) := \{x^T A x, x \in I\!R^m, |x| = 1\}$ is the field of values of a square matrix $A$. Then each part of the convection term is differenced in the appropriate direction

$$\begin{aligned} P_{\Delta x}(u)_i \;&=\; A_1(u_i)(D_* u)_i + A_2^+(u_i)(D_- u)_i + A_2^-(u_i)(D_+ u)_i, \\ &-\; B(u_i)(D_+ D_- u)_i, \quad u \in \mathcal{F}_p(\Omega_{\Delta x}, I\!R^m). \end{aligned} \tag{14}$$

$D_*$ may be one of the first order operators $D_\pm, D_0$ and we have used the abbreviation

$$A_1 := \begin{pmatrix} A_{11} & A_{12} \\ A_{21} & 0 \end{pmatrix}, \quad A_2^\pm := \begin{pmatrix} 0 & 0 \\ 0 & A_{22}^\pm \end{pmatrix}. \tag{15}$$

Many mixed systems have the property, that not only the block $A_{22}$, but even the whole matrix $A$ is symmetrizable. That is the case when the system is hyperbolic, in the limit without viscosity, eg. the Euler and Navier–Stokes equations. In this case, one would like to approximate $A\ u_x$ by an upwind method. Our third and last model is an example of this type

$$\begin{aligned} P_{\Delta x}(u)_i &= A^+(u_i)(D_-u)_i + A^-(u_i)(D_+u)_i, \\ &- B(u_i)(D_+D_-u)_i, \quad u \in \mathcal{F}_p(\Omega_{\Delta x}, I\!R^m). \end{aligned} \tag{16}$$

Here, we assume $A$ and the decomposition to be symmetric

$$A = A^+ + A^-, (A^\pm)^T = A^\pm \quad \text{and} \quad \sigma(\pm A^\pm) \subset [\gamma, \infty), \quad \gamma > 0, u \in U_\eta. \tag{17}$$

Our main results are:

THEOREM 2.2. *Each of the methods (9) & (10) combined with one of the space operators (12), (13) & (14) & (15) or (16) & (17) has for a sufficiently small step size a unique solution $U$; ($\Delta t \le (2C)^{-1}$, where $C$ is the energy constant of the linearized scheme cf. Section 5 and $\Delta t = \mu \Delta x$, $\mu$ arbitrary). This solution converges third order to a properly defined pilot function $u^{pi}$, i.e.*

$$\|U - u^{pi}|_h\|^h_{\infty,2} = O(\Delta x^3), \quad \Delta x \to 0.$$

*By construction, the pilot function and $U$ converge first order to the desired solution $\tilde{u}$ of the periodic Cauchy problem (5) and (6) for a strongly hyperbolic, strongly parabolic or a mixed system*

$$\|U - \tilde{u}|_h\|^h_{\infty,2} = O(\Delta x), \quad \Delta x \to 0.$$

The pilot function will be constructed in section 4. Throughout this paper the following scalar products, norms and their operator norms will be used:
In $I\!R^m$:

$$< u, v > := u^T v, \quad |u| := (< u, u >)^{1/2}.$$

In $\mathcal{F}_p(\Omega_{\Delta x}, I\!R^m)$:

$$(u, v)_2^{\Delta x} := \Delta x \textstyle\sum_{x \in \Omega_{\Delta x}} < u(x), v(x) >, \quad \|u\|_2^{\Delta x} := ((u, u)_2^{\Delta x})^{1/2},$$

$$\|u\|_1^{\Delta x} := \Delta x \textstyle\sum_{x \in \Omega_{\Delta x}} |u(x)|, \quad \|u\|_\infty^{\Delta x} := \max_{x \in \Omega_{\Delta x}} |u(x)|.$$

In $\mathcal{F}_p(\Omega_h, I\!R^m)$:

$$\|u\|^h_{\infty,2} := \max_{t \in \Omega_{\Delta t}} \|u(t, \cdot)\|_2^{\Delta x}, \quad \|u\|^h_{\infty,\infty} := \max_{t \in \Omega_{\Delta t}} \|u(t, \cdot)\|_\infty^{\Delta x}.$$

The corresponding continuous norms do not carry the step size as superscript.

**3. Local stability and convergence.** An important contribution towards the desired convergence result is the observation by López–Marcos and Sanz–Serna that the discretization process itself can introduce some amount of stability that need not necessarily to be present in the continuous problem. In [7] they consider a discrete method given by a nonlinear operator

$$\Phi_h \ : \ D_h := \mathcal{F}_p(\Omega_h, U_\eta) \to \mathcal{F}_p(\Omega_h, I\!R^m).$$

If the method is successful, a solution $U$ of $\Phi_h(u) = 0$, close to $\tilde{u}$ exists and they are led to consider the linearized scheme

$$\Phi_h(\tilde{u}|_h) + D\Phi_h(\tilde{u}|_h)(u - \tilde{u}|_h) = 0. \tag{18}$$

Even though $\tilde{u}$ is the natural choice in this context, there is no need to insert the exact solution of the differential problem here. In fact, it turns out that we avoid some difficulty that occurs at a later stage, if we replace $\tilde{u}$ by a proper smooth function of $t$ and $x$ that assumes values in $U_\eta$, say $u^{pi}$. As pointed out below, this function has to be highly consistent with the scheme, i.e.

$$\|\Phi_h(u^{pi}|_h)\| \leq C_1 \Delta t^r + C_2 \Delta x^s, \tag{19}$$

where $r$ and $s$ are big enough and $C_{1,2}$ are independent of small step sizes $h$. Within this section $\|\cdot\|$ is just any norm in $\mathcal{F}_p(\Omega_h, I\!R^m)$. In the following, we denote by $\mathcal{B}(u, R)$ the open ball $\{v \in \mathcal{F}_p(\Omega_h, I\!R^m) : \|u - v\| < R\}$.

López–Marcos and Sanz–Serna introduced the concept of stability with $h$–dependent threshold. The advantage of this kind of stability is, that it can be inferred from the stability of the linearized scheme (18).

THEOREM 3.1. *Assume:*

- $\Phi_h$ *is continuously differentiable in* $\mathcal{B}(u^{pi}|_h, C_{\Delta x}) \subset D_h$ *and the derivative is Lipschitz continuous*

$$\|D\Phi_h(u^{pi}|_h) - D\Phi_h(u)\| \leq Lip_{\Delta x} \|u^{pi}|_h - u\|, \quad u \in \mathcal{B}(u^{pi}|_h, C_{\Delta x}). \tag{20}$$

- *Furthermore, the scheme linearized at* $u^{pi}$ *is stable, that means: For small step sizes* $D\Phi_h(u^{pi}|_h)$ *is invertible and* $h$*–independent bounded:*

$$\|D\Phi_h(u^{pi}|_h)^{-1}\| \leq L.$$

*Then* $\Phi_h$ *is locally stable near* $u^{pi}$, *i.e. there exists a constant* $S > L$ *and the so–called stability radius* $R_{\Delta x} := \min(C_{\Delta x}, (L^{-1} - S^{-1})/Lip_{\Delta x})$, *such that*

$$\|u - v\| \leq S \|\Phi_h(u) - \Phi_h(v)\|, \quad u, v \in \mathcal{B}(u^{pi}|_h, R_{\Delta x}). \tag{21}$$

Notice, that if one can apply the stability inequality (21) to $U$, convergence follows easily

$$\|u^{pi}|_h - U\| \leq S \|\Phi_h(u^{pi}|_h)\| \leq C_1 \Delta t^r + C_2 \Delta x^s.$$

Consequently, the key problem is whether the implicit equation $\Phi_h(u) = 0$ has a solution inside the stability range $\mathcal{B}(u^{pi}|_h, R_{\Delta x})$. Therefore, one has to estimate the stability radius $R_{\Delta x}$.

$\Phi_h$ contains negative powers of time and space step sizes. Taking the derivative, linear parts become constant and disappear in the difference $D\Phi_h(u^{pi}|_h) - D\Phi_h(u)$. Since the time derivative in the original PDE appears linearly with respect to $u$ the time step size cancels out, in contrast to $\Delta x$. The typical situation is

$$Lip_{\Delta x} \leq \frac{C}{\Delta x^q}, \quad q > q_P,$$

where $C$ is independent of $h$ and $q_P$ means the highest order of any finite difference expression occurring nonlinear in $P_{\Delta x}$ or equivalently the highest order of derivative that occurs nonlinearly in (11). The actual estimate of the Lipschitz constant is technical, we delay it to section 6. Since $Lip_{\Delta x}$ is unbounded the stability radius decays as $h \to 0$ and the convergence analysis seems to be useless, but appearances are deceptive. In fact, it is possible to balance this lack of stability by consistency. López–Marcos and Sanz–Serna apply a lemma due to Stetter [11] to show their convergence result:

THEOREM 3.2. *Let $\Phi_h$ be continuous, locally stable and*

$$\|\Phi_h(u^{pi}|_h)\| = o(R_{\Delta x}), \quad h \to 0, \tag{22}$$

*then the discrete equations $\Phi_h(u) = 0$ possess (for $h$ small enough) a unique solution $U$ inside the stability range. This solution converges to $u^{pi}$.*

Condition (22) implies two consequences. The first one is, that the step sizes can no longer be independent. If for example $\Delta t$ is fixed and the problem is time dependent, that means the consistency error due to the time derivative is not zero, (22) must fail if $\Delta x$ is sufficiently small. In other words, time and space step must be in relation to each other. If the stability radius is bounded from below

$$R_{\Delta x} \geq C\ \Delta x^q, \quad q > 0,$$

then

$$\Delta t^r = o(\Delta x^q) \quad \text{or} \quad \Delta t = o(\Delta x^{q/r}), \quad \Delta x \to 0.$$

would be sufficient. We will assume a linear relation between the two step sizes

$$\Delta t = \mu \Delta x = O(\Delta x), \quad \Delta x \to 0. \tag{23}$$

Consequently, the order of consistency will be $p := \min(r, s)$

$$\|\Phi_h(u^{pi}|_h)\| = O(\Delta x^p), \quad \Delta x \to 0$$

and (22) is fulfilled if $p > q$. (23) is not a severe condition, since there will be no restriction on $\mu$.

The second more severe implication is that low order methods are excluded from the analysis if one chooses $u^{pi}$ to be the solution of the continuous Cauchy problem. Unfortunately, most of the numerical methods designed to compute not only smooth but also discontinuous shock solutions are not of high order. Nevertheless, concerning this analysis we assume the solutions to be smooth.

As indicated earlier, the idea to avoid this second consequence is to insert Strang's ansatz (1) into the local stability theory by López–Marcos and Sanz–Serna and to show convergence to this pilot function. This concept will be carried out in the following sections. First the pilot functions are defined and then the linear stability will be discussed.

**4. Construction of the pilot function.** The general procedure to define a high order approximative solution to the discrete equation is the following:

- Substitute the ansatz $u^{pi} = \tilde{u} + \Delta x f + \Delta x^2 g$ into $\Phi_h(u)$,
- expand the finite differences with respect to the step size,
- set to zero the coefficients of $\Delta x$ and $\Delta x^2$.

As a result one obtains defining Cauchy problems for $f$ and $g$. The initial condition turns out to be zero, because of $u_0^{pi} = \tilde{u}_0 + \Delta x f_0 + \Delta x^2 g_0 - z$ has to be zero and $\tilde{u}_0 = z$.

We illustrate the above concept in case of the central space operator (12). The difference expression is

$$\frac{u_{n,i}^{pi} - u_{n-1,i}^{pi}}{\Delta t} + A(u_{n,i}^{pi})(D_0 u_n^{pi})_i - (B(u_{n,i}^{pi} + \alpha \Delta x I)(D_+ D_- u_n^{pi})_i.$$

Replacing the difference quotients by Taylor series and inserting (23) the result is

$$\begin{aligned} \| u_t^{pi} \; & - \; \mu \Delta x/2 \; u_{tt}^{pi} + \mu^2 \Delta x^2/3! \; u_{ttt}^{pi} \\ & + \; A(u^{pi})(u_x^{pi} + \Delta x^2/3! \; u_{xxx}^{pi}) \\ & - \; B(u^{pi})(u_{xx}^{pi} + 2\Delta x^2/4! \; u_{xxxx}^{pi}) - \alpha \Delta x \; u_{xx}^{pi} \|_{\infty,\infty} \\ = O(\Delta x^3), & \quad \Delta x \to 0. \end{aligned}$$

Expanding $A$ and $B$ and inserting the ansatz (1) results in

$$\begin{aligned} & \| \; \tilde{u}_t + A(\tilde{u}) \; \tilde{u}_x - B(\tilde{u}) \; \tilde{u}_{xx} \\ & \; +\Delta x \; \{ f_t \; + \; A(\tilde{u}) \; f_x + DA(\tilde{u}) \; f \; \tilde{u}_x - B(\tilde{u}) \; f_{xx} - DB(\tilde{u}) \; f \; \tilde{u}_{xx} \\ & \qquad\qquad - \; \mu/2 \; \tilde{u}_{tt} - \alpha \; \tilde{u}_{xx} \} \\ & \; +\Delta x^2 \{ g_t \; + \; A(\tilde{u})(g_x + 1/6 \; \tilde{u}_{xxx}) + DA(\tilde{u}) \; f \; f_x + DA(\tilde{u}) \; g \; \tilde{u}_x \\ & \qquad\qquad + \; 1/2 D^2 A(\tilde{u}) \; f \; f \; \tilde{u}_x \\ & \qquad\qquad - \; B(\tilde{u})(g_{xx} + 1/12 \; \tilde{u}_{xxxx}) - DB(\tilde{u}) \; f \; f_{xx} - DB(\tilde{u}) \; g \; \tilde{u}_{xx} \\ & \qquad\qquad - \; 1/2 D^2 B(\tilde{u}) \; f \; f \; \tilde{u}_{xx} - \alpha \; f_{xx} - \mu/2 \; f_{tt} + \mu^2/3! \; \tilde{u}_{ttt} \} \; \|_{\infty,\infty} \\ & = O(\Delta x^3), \quad \Delta x \to 0. \end{aligned}$$

If $f$ is defined as the solution of the linear system

$$
\begin{aligned}
f(0,x) &= 0, \\
f_t &+ A(\tilde{u})\ f_x + DA(\tilde{u})\ f\ \tilde{u}_x \\
&= B(\tilde{u})\ f_{xx} + DB(\tilde{u})\ f\ \tilde{u}_{xx} + \mu/2\ \tilde{u}_{tt} + \alpha\ \tilde{u}_{xx}, \\
& (t,x) \in [0,T] \times I\!R
\end{aligned} \tag{24}
$$

and $g$ as the solution of

$$
\begin{aligned}
g(0,x) &= 0, \\
g_t &+ A(\tilde{u})(g_x + 1/6\ \tilde{u}_{xxx}) + DA(\tilde{u})\ f\ f_x + DA(\tilde{u})\ g\ \tilde{u}_x \\
&+ 1/2 D^2 A(\tilde{u})\ f\ f\ \tilde{u}_x \\
&= B(\tilde{u})(g_{xx} + 1/12\ \tilde{u}_{xxxx}) + DB(\tilde{u})\ f\ f_{xx} + DB(\tilde{u})\ g\ \tilde{u}_{xx} \\
&+ 1/2 D^2 B(\tilde{u})\ f\ f\ \tilde{u}_{xx} + \alpha\ f_{xx} + \mu/2\ f_{tt} - \mu^2/3!\ \tilde{u}_{ttt}, \\
& (t,x) \in [0,T] \times I\!R,
\end{aligned} \tag{25}
$$

then $u^{pi}$ is third order consistent with $\Phi_h$. (24) and (25) are systems of exactly the same type as the original PDE, but linear. According to the existence theorem 2.1 $f$ and $g$ are well defined on arbitrarily large time intervals. We also observe, that $f$ and $g$ are non–zero because of forcing terms which are caused by consistency errors of the method.

**5. Energy estimates and linear stability.** In order to analyze the stability of the scheme linearized at the pilot function, we will demonstrate that the linearized method of lines systems are energy bounded:

DEFINITION 5.1. *A method of lines system $u_t + L_{\Delta x}(t)u = 0$ and the operator $L_{\Delta x}$ are called energy bounded, if there exists an energy norm and a constant $C$ such that*

$$-(u, L_{\Delta x}(t)u)_H^{\Delta x} \le C(\|u\|_H^{\Delta x})^2, \quad u \in \mathcal{F}_p(\Omega_{\Delta x}, I\!R^m), t \in [0,T].$$

The scalar product and norm are defined as follows:

DEFINITION 5.2. *Every symmetric, bounded matrix function*

$$H : [0,T] \times [0,1]^2 \to I\!R^{m \times m}, \quad 0 < \alpha \le H(t,x,\Delta x) \le \|H\|_{\infty,\infty,\infty}$$

*defines a scalar product*

$$(u,v)_H^{\Delta x} := (u, H(t,\cdot,\Delta x)v)_2^{\Delta x}, \quad u,v \in \mathcal{F}_p(\Omega_{\Delta x}, I\!R^m).$$

*The induced norm $\|u\|_H^{\Delta x} := ((u,u)_H^{\Delta x})^{1/2}$ is called the energy norm.*

*Remark.* The energy norm is equivalent to the $L^2$–norm:

$$(\|u\|_H^{\Delta x})^2 = (u, Hu)_2^{\Delta x} \le \|H\|_{\infty,\infty,\infty} (\|u\|_2^{\Delta x})^2$$

$$\alpha(\|u\|_2^{\Delta x})^2 = (u, \alpha u)_2^{\Delta x} \le (u, Hu)_2^{\Delta x} = (\|u\|_H^{\Delta x})^2.$$

Proving stability of an implicit Euler method applied to an energy bounded system, is an easy estimate.

**5.1. Linear stability.** Consider an energy bounded method of lines system

$$u_t(t,\cdot) + L_{\Delta x}(t)\, u(t,\cdot) = 0, \quad t \in [0,T] \tag{26}$$

with smooth, periodic initial condition

$$u(0,x) = z(x) = z(x+1), \quad x \in \Omega_{\Delta x}.$$

An implicit Euler step applied to this ODE system reads

$$\begin{aligned} u_0 &= z|_{\Delta x}, \\ \frac{u_n - u_{n-1}}{\Delta t} + L_{\Delta x}(t_n)\, u_n &= 0, \\ n &= 1, 2, \ldots N. \end{aligned} \tag{27}$$

Let $u \in \mathcal{F}_p(\Omega_h, I\!R^m)$ be a solution of the implicit system (27). Using the energy bound of $L_{\Delta x}$ it follows

$$\left(u_n, \frac{u_n - u_{n-1}}{\Delta t}\right)_H^{\Delta x} = -(u_n, L_{\Delta x}(t_n)\, u_n)_H^{\Delta x} \le C(\|u_n\|_H^{\Delta x})^2.$$

This implies

$$(\|u_n\|_H^{\Delta x})^2 - (u_n, u_{n-1})_H^{\Delta x} \le \Delta t C(\|u_n\|_H^{\Delta x})^2,$$

$$(1 - \Delta t C)(\|u_n\|_H^{\Delta x})^2 \le \|u_n\|_H^{\Delta x}\ \|u_{n-1}\|_H^{\Delta x}$$

and finally

$$\|u_n\|_H^{\Delta x} \le \frac{\|u_{n-1}\|_H^{\Delta x}}{1 - \Delta t C} \le (1 + 2\Delta t C)\|u_{n-1}\|_H^{\Delta x}, \quad \Delta t \le (2C)^{-1}.$$

It follows, that every single Euler step has a unique solution.

By formal manipulation one can verify, that the complete method written as $\Psi_h\, u = 0$ is stable in the sense $\|\Psi_h^{-1}\|_{\infty,2} \le (1+T)C_1 \exp(C_2 T)$, where $C_{1,2}$ are $h$–independent constants. Thus, we have the result:

LEMMA 5.1. *A backward Euler step applied to an energy bounded method of lines system (26) is stable. If $\Delta t \le (2C)^{-1}$ the method has a unique solution.*

**5.2. Energy estimates.** The technique for deriving energy estimates is closely related to the continuous case where one uses "integration by parts": If $u, v \in C_p^1$, then $(u_x, v) = -(u, v_x)$. The reader should compare sections 3.1.1, 3.2.2, 3.3.1 and 3.4 in [5]. The important point here is, that despite of the fact that the coefficients of the linearized operators depend on $\Delta x$, we obtain step size independent estimates.

First of all we consider a central differencing operator

$$L_{\Delta x}(t)\, u := A D_0 u$$

where $A$ is a smooth, periodic function of $(t,x) \in \Omega$. Since we want to apply energy estimates to schemes which are linearized at the pilot function, the coefficients will depend on $\Delta x$ also

$$A \in C_p^\infty(\Omega \times [0,1], I\!R^{m\times m}).$$

We assume $A$ to be uniformly symmetrizable with symmetrizer $H$, that means the symmetrizer depends smoothly on $\Delta x$ and is uniformly bounded

$$(28) H \in C_p^\infty(\Omega \times [0,1], I\!R^{m\times m}), \quad 0 < \alpha \leq H(t,x,\Delta x), \quad (t,x,\Delta x) \in \Omega \times [0,1].$$

The aim is to estimate $-(u, HAD_0 u)_2^{\Delta x}$ independent of the step size. Since $D_0$ is antisymmetric $D_0^T = -D_0$ like the partial derivative $\frac{\partial}{\partial x}$, there is a discrete analogue of the "integration by parts" formula: $(D_0 u, v)^{\Delta x} = -(u, D_0 v)^{\Delta x}$ for $u, v \in \mathcal{F}_p(\Omega_{\Delta x}, I\!R^m)$. Therefore this estimate is a discrete version of the corresponding estimate for continuous solutions (cf. [5] section 3.3.1). We apply the discrete Leibniz' product rule

$$D_0(HAu) = HAD_0 u + \frac{1}{2}[D_+(HA)Eu + D_-(HA)E^{-1}u],$$

the symmetry of $HA$ and the antisymmetry of $D_0$ to obtain

$$\begin{aligned} -2(u, HAD_0 u)^{\Delta x} &= \frac{1}{2}\left[(u, D_+(HA)Eu)^{\Delta x} + (u, D_-(HA)E^{-1}u)^{\Delta x}\right] \\ &\leq \frac{1}{2}\left[\|D_+(HA)\|_\infty^{\Delta x} + \|D_-(HA)\|_\infty^{\Delta x}\right](\|u\|_2^{\Delta x})^2. \end{aligned}$$

As $H$ and $A$ are smooth functions, the finite differences $D_\pm(HA)$ are bounded by the derivative $(HA)_x$ and there exists a $\Delta x$–independent constant $C$, such that the energy estimate

$$-(u, AD_0 u)_H^{\Delta x} \leq C(\|u\|_H^{\Delta x})^2 \tag{29}$$

holds.

In a next step we consider a second order operator

$$L_{\Delta x}(t)\, u := -(B + \delta \Delta x I) D_+ D_- u, \quad B \in C_p^\infty(\Omega \times [0,1], I\!R^{m\times m}), \delta \geq 0$$

where $B$ is uniformly strongly parabolic in the sense

$$B(t,x,\Delta x) + B^T(t,x,\Delta x) \geq 2\beta, \quad (t,x,\Delta x) \in \Omega \times [0,1]. \tag{30}$$

Here, we estimate directly the discrete $L^2$–scalar product

$$(u, (B + \delta \Delta x I) D_+ D_- u)_2^{\Delta x} = (u, BD_+ D_- u)_2^{\Delta x} + \Delta x \delta (u, D_+ D_- u)_2^{\Delta x}.$$

Since $D_+^T = -D_-$, the second term is obviously negative

$$\Delta x \delta (u, D_+ D_- u)_2^{\Delta x} = -\Delta x \delta (D_- u, D_- u)_2^{\Delta x} = -\Delta x \delta (\|D_- u\|_2^{\Delta x})^2.$$

To estimate the first term we apply the Leibniz' formula for $D_+$

$$D_+(BD_- u) = BD_+ D_- u + (D_+ B) E D_- u$$

to obtain

$$\begin{aligned}(u, BD_+ D_- u)_2^{\Delta x} &= -(D_- u, BD_- u)_2^{\Delta x} - (u, (D_+ B) E D_- u)_2^{\Delta x} \\ &= -\frac{1}{2}(D_- u, (B + B^T) D_- u)_2^{\Delta x} - (u, (D_+ B) E D_- u)_2^{\Delta x}.\end{aligned}$$

Using (30) and the fact that the norm of a periodic function is invariant with respect to the shift operator we get

$$(u, BD_+ D_- u)_2^{\Delta x} \leq -\beta (\|D_- u\|_2^{\Delta x})^2 + C \|u\|_2^{\Delta x} \, \|D_- u\|_2^{\Delta x}. \tag{31}$$

Now, a simple version of Youngs inequality

$$ab \leq \frac{\sigma^2 a^2}{2} + \frac{b^2}{2\sigma^2}, \quad \sigma > 0, a, b \in I\!R \tag{32}$$

can be used to eliminate $D_- u$. The resulting energy estimate is

$$(u, (B + \delta \Delta x I) D_+ D_- u)_2^{\Delta x} \leq \frac{C^2}{4\beta} \, (\|u\|_2^{\Delta x})^2.$$

Combining the previous two estimates leads to the energy estimate for a mixed type operator

$$\begin{aligned}&L_{\Delta x}(t) \, u = A_1 D_* u + A_2 D_0 u - (B + \delta \Delta x I) D_+ D_- u, \quad \delta \geq 0 \\ &A_1 := \begin{pmatrix} A_{11} & A_{12} \\ A_{21} & 0 \end{pmatrix}, \quad A_2 := \begin{pmatrix} 0 & 0 \\ 0 & A_{22} \end{pmatrix}, \quad B := \begin{pmatrix} B_{11} & 0 \\ 0 & 0 \end{pmatrix}. \\ &A := A_1 + A_2, B \in C_p^\infty(\Omega \times [0,1], I\!R^{m\times m}).\end{aligned} \tag{33}$$

Lemma 5.2. *The mixed operator (33), where $A_{22}$ is uniformly symmetrizable (28) and $B_{11}$ is uniformly strongly parabolic (30), is energy bounded.*

*Proof.* Let $H_{22}$ be the symmetrizer of $A_{22}$,

$$H := \begin{pmatrix} I & 0 \\ 0 & H_{22} \end{pmatrix}.$$

and consider

$$\begin{aligned}-(u, L_{\Delta x}(t) \, u)_H^{\Delta x} = \; & (u, HBD_+ D_- u)_2^{\Delta x} + \delta \Delta x (u, HD_+ D_- u)_2^{\Delta x} \\ & -(u, HA_1 D_* u)_2^{\Delta x} - (u, HA_2 D_0 u)_2^{\Delta x}.\end{aligned}$$

Now, decompose $u^T = (v, w)$ into parabolic and hyperbolic components and collect the terms in a proper order

$$\begin{aligned}
-(u, L_{\Delta x}(t)u)_H^{\Delta x} &= \quad (v, B_{11}D_+D_-v)_2^{\Delta x} + \delta\Delta x(u, HD_+D_-u)_2^{\Delta x} \\
&\quad -(v, A_{11}D_*v)_2^{\Delta x} - (v, A_{12}D_*w)_2^{\Delta x} \\
&\quad -(w, H_{22}A_{21}D_*v)_2^{\Delta x} \\
&\quad -(w, H_{22}A_{22}D_0w)_2^{\Delta x} \\
&= \quad (I) + (II) + (III) + (IV) + (V).
\end{aligned}$$

$$\begin{aligned}
(I) &:= \quad (v, B_{11}D_+D_-v)_2^{\Delta x} - (v, A_{11}D_*v)_2^{\Delta x} \\
(II) &:= -(v, A_{12}D_*w)_2^{\Delta x} \\
(III) &:= -(w, H_{22}A_{21}D_*v)_2^{\Delta x} \\
(IV) &:= -(w, H_{22}A_{22}D_0w)_2^{\Delta x} \\
(V) &:= \quad \delta\Delta x(u, HD_+D_-u)_2^{\Delta x}.
\end{aligned}$$

Since $H$ is symmetric and uniformly bounded, $(V)$ is a parabolic term. For bounded $\delta$ it can be estimated by

$$(V) \le C_1(\|u\|_2^{\Delta x})^2$$

where $C_1$ depends on $H$ and its bounds, but is independent of $\delta \in [0, \delta_{max}]$. Expression $(IV)$ is of hyperbolic type and has already been estimated (cf. (29)):

$$(IV) \le C_2(\|u\|_2^{\Delta x})^2.$$

In case of $(III)$ we apply the Cauchy–Schwarz inequality and

$$\|D_0u\|_2^{\Delta x} \le \|D_-u\|_2^{\Delta x} = \|ED_-u\|_2^{\Delta x} = \|D_+u\|_2^{\Delta x} \tag{34}$$

to get

$$(III) \le C_3(\|w\|_2^{\Delta x})^2(\|D_-v\|_2^{\Delta x})^2.$$

Since there is no $\Delta x$ independent diffusion that could compensate $D_*w$ in $(II)$, we apply partial summation to move the finite difference operator to the $v$ component. In case of $D_* = D_\pm$ we have

$$A_{12}D_\pm w = D_\pm(A_{12}w) - (D_\pm A_{12})E^{\pm 1}w,$$

and

$$-(v, A_{12}D_\pm w)_2^{\Delta x} = (D_\mp v, A_{12}w)_2^{\Delta x} + (v, (D_\pm A_{12})E^{\pm 1}w)_2^{\Delta x}.$$

If $D_*$ is the central difference $D_0$ we apply

$$A_{12}D_0w = D_0(A_{12}w) - \frac{1}{2}[(D_+A_{12})Ew + (D_-A_{12})E^{-1}w]$$

and obtain

$$\begin{aligned} -(v, A_{12}D_0 w)_2^{\Delta x} &= (D_0 v, A_{12} w)_2^{\Delta x} \\ &\quad + \frac{1}{2}(v, (D_+A_{12})Ew)_2^{\Delta x} + \frac{1}{2}(v, (D_-A_{12})E^{-1}w)_2^{\Delta x}. \end{aligned}$$

In any case it follows

$$(II) \le C_4 \|w\|_2^{\Delta x} \; \|D_- v\|_2^{\Delta x} + C_5 \|v\|_2^{\Delta x} \; \|w\|_2^{\Delta x}.$$

It remains to estimate the parabolic part $(I)$. Using (31), Cauchy–Schwarz and (34) we have

$$(I) \le -\beta(\|D_- v\|_2^{\Delta x})^2 + C_6 \|v\|_2^{\Delta x} \; \|D_- v\|_2^{\Delta x}.$$

Summing all up and applying once again Youngs inequality (32) yields

$$\begin{aligned} -(u, L_{\Delta x}(t)u)_H^{\Delta x} &\le -\beta(\|D_- v\|_2^{\Delta x})^2 \\ &\quad + C_7 \|u\|_2^{\Delta x} \; \|D_- v\|_2^{\Delta x} + C_8(\|u\|_2^{\Delta x})^2 \\ &\le C_9(\beta)(\|u\|_2^{\Delta x})^2. \end{aligned}$$

Since $C_9$ (and all other $C's$) do not depend on step sizes, time or $u$ Lemma 5.2 is proved. □

To show the convergence result for the upwind type operators (14) and (16) we estimate next

$$L_{\Delta x}(t)\; u = A^+ D_- u + A^- D_+ u, \quad A^\pm \in C_p^\infty(\Omega \times [0,1], I\!R^{m \times m}).$$

The assumptions are

$$\text{(35)} \qquad \begin{array}{l} A^\pm \text{ are simultaneously, uniformly symmetrizable} \\ \text{and } Re(F(\pm A^\pm)) \subset [\gamma, \infty), \quad \gamma > 0. \end{array}$$

Application of the discrete product rule leads to

$$(u, HA^\pm D_\mp u)_2^{\Delta x} = -(D_\pm u, HA^\pm u)_2^{\Delta x} - (u, D_\mp(HA^\pm)E^{\mp 1}u)_2^{\Delta x}.$$

Unfortunately, the one sided operators $D_\pm$ are not antisymmetric, therefore the differences of $u$ do not cancel as easy as in the central case. There remains

$$\begin{aligned} &-2(u, HA^\pm D_\mp u)_2^{\Delta x} \\ &= -(u, HA^\pm D_\mp u)_2^{\Delta x} + (D_\pm u, HA^\pm u)_2^{\Delta x} + (u, D_\mp(HA^\pm)E^{\mp 1}u)_2^{\Delta x} \\ &= \quad (u, HA^\pm(D_\pm - D_\mp)u)_2^{\Delta x} + (u, D_\mp(HA^\pm)E^{\mp 1}u)_2^{\Delta x}. \end{aligned}$$

The last term is no problem. If we show that $\pm HA^\pm$ is positive definite, then the first one is strongly parabolic, since $D_+ - D_- = \Delta x D_+ D_-$. As $H$ is positive definite, the inclusion

$$F(\pm HA^\pm) \subset F(H)F(\pm A^\pm)$$

for the field of values holds (cf. [4] Corollary 1.7.7). Especially $H$ is symmetric, therefore $F(H) \subset [\lambda_{min}(H), \lambda_{max}(H)] \subset I\!R^+$. According to our assumption, the spectrum of $\pm HA^\pm$ is positive

$$\sigma(\pm HA^\pm) \subset Re(F(\pm HA^\pm)) \subset [\gamma\lambda_{min}(H), \infty).$$

Of course $\pm HA^\pm$ is symmetric and consequently $\pm HA^\pm + (\pm HA^\pm)^T = \pm 2HA^\pm \geq 2\bar{\gamma}$, where $\bar{\gamma} := \gamma\alpha > 0$. We can proceed like estimating a parabolic term (cf. (31))

$$\begin{aligned} (u, HA^\pm(D_\pm - D_\mp)u)_2^{\Delta x} &= \pm\Delta x(u, HA^\pm D_+ D_- u)_2^\pm \\ &\leq -\bar{\gamma}(\|D_- u\|_2^{\Delta x})^2 + C\|u\|_2^{\Delta x}\|D_- u\|_2^{\Delta x}. \end{aligned}$$

Again Youngs inequality (32) gives the result

$$-(u, A^+ D_- u)_H^{\Delta x} - (u, A^- D_+ u)_H^{\Delta x} \leq C(\bar{\gamma})(\|u\|_2^{\Delta x})^2. \tag{36}$$

Now, we can state the result for a linear version of (14).

$$\begin{aligned} & L_{\Delta x}(t)\, u = A_1 D_* u + A_2^+ D_- u + A_2^- D_+ u - BD_+ D_- u, \\ & A_1 := \begin{pmatrix} A_{11} & A_{12} \\ A_{21} & 0 \end{pmatrix}, \quad A_2^\pm := \begin{pmatrix} 0 & 0 \\ 0 & A_{22}^\pm \end{pmatrix}, \quad B := \begin{pmatrix} B_{11} & 0 \\ 0 & 0 \end{pmatrix}. \\ & A := A_1 + A_2^+ + A_2^-, B \in C_p^\infty(\Omega \times [0,1], I\!R^{m\times m}). \end{aligned} \tag{37}$$

Lemma 5.3. *The mixed operator (37), where $A_{22}^\pm$ satisfy (35) and $B_{11}$ is uniformly strongly parabolic (30), is energy bounded.*

*Proof.* Let $H_{22}$ be the common symmetrizer of $A_{22}^\pm$ and set

$$H := \begin{pmatrix} I & 0 \\ 0 & H_{22} \end{pmatrix}.$$

As in the proof of Lemma 5.2 we have

$$-(u, L_{\Delta x}(t)u)_H^{\Delta x} = (I) + (II) + (III) + (IV').$$

$(I)$, $(II)$ and $(III)$ are exactly the same as above, but $(IV)$ is replaced by

$$(IV') := -(w, H_{22}A_{22}^+ D_- w)_2^{\Delta x} - (w, H_{22}A_{22}^- D_+ w)_2^{\Delta x}.$$

This term was estimated in (36), therefore the proof is complete. $\square$

It is also not difficult to estimate

$$(38)\quad \begin{aligned} &L_{\Delta x}(t)\,u = A^+D_-u + A^-D_+u - BD_+D_-u,\\ &A := A^+ + A^-, B := \begin{pmatrix} B_{11} & 0\\ 0 & 0\end{pmatrix} \in C_p^\infty(\Omega\times[0,1], I\!R^{m\times m}).\end{aligned}$$

LEMMA 5.4. *If $A^\pm$ are symmetric, $\sigma(\pm A^\pm)\subset[\gamma,\infty), \gamma>0$ and $B_{11}$ is uniformly strongly parabolic (30), the operator (38) is energy bounded.*

*Proof.* We have

$$-(u, L_{\Delta x}(t)u)_2^{\Delta x} = (u, BD_+D_-u)_2^{\Delta x} - (u, A^+D_-u)_2^{\Delta x} - (u, A_-D_+u)_2^{\Delta x}.$$

The parabolic part can be estimated by

$$(u, BD_+D_-u)_2^{\Delta x} = (v, B_{11}D_+D_-v)_2^{\Delta x} \le C(\|v\|_2^{\Delta x})^2 \le C(\|u\|_2^{\Delta x})^2.$$

The hyperbolic part has already been estimated in (36). □

**6. The Lipschitz constant of $D\Phi_h$.** The purpose of this section is to study the Lipschitz "constant" in (20) as a "function" of $\Delta x$. The derivative of the operator (10) is given by

$$(39)\quad \begin{aligned} &D\Phi_h(u) = \frac{1}{\Delta t}\begin{pmatrix} \Delta t I & & & \\ -I & Q_1 & & \\ & \ddots & \ddots & \\ & & -I & Q_N \end{pmatrix},\\ &Q_n := I + \Delta t DP_{\Delta x}(u_n), \quad n = 1,2,\dots N, \quad u\in D_h.\end{aligned}$$

Therefore, the difference is

$$D\Phi_h(u^{pi}|_h) - D\Phi_h(u) = \text{ blockdiag }(0, DP_{\Delta x}(u_n^{pi}|_{\Delta x}) - DP_{\Delta x}(u_n), n = 1,2,\dots N).$$

Taking the $\infty$–norm in time one has to estimate the Lipschitz "constant" of $DP_{\Delta x}$. The most general form of space operators we are concerned with is

$$(40)\quad \begin{aligned} P_{\Delta x}(u)_i \;=\;& A_1(u_i)(D_0u)_i + A_2(u_i)(D_-u)_i + A_3(u_i)(D_+u)_i\\ &-(B(u_i) + \delta\Delta x I)(D_+D_-u)_i, \quad \delta\ge 0, u\in\mathcal{F}_p(\Omega_{\Delta x}, U_\eta).\end{aligned}$$

Since linear terms have no effect on the Lipschitz constant, we may set $\delta$ to zero. The basic concept of the estimate is as follows:

- Since $P_{\Delta x}$ has a finite, three point stencil, the Jacobian $DP_{\Delta x}$ is (at inner grid points) block three diagonal.
- Each of these blocks is $h$–independent Lipschitz continuous, therefore the $\infty$–norm of $DP_{\Delta x}$ and $DP_{\Delta x}^T$ can be estimated by

$$(41)\quad \left.\begin{aligned} \|DP_{\Delta x}(u_n^{pi}|_{\Delta x}) - DP_{\Delta x}(u_n)\|_\infty^{\Delta x}\\ \|DP_{\Delta x}(u_n^{pi}|_{\Delta x})^T - DP_{\Delta x}(u_n)^T\|_\infty^{\Delta x}\end{aligned}\right\} \le \Delta x^{-q_P}\|u_n^{pi}|_{\Delta x} - u_n\|_\infty^{\Delta x}.$$

- Application of the estimate $\|A\|_2^{\Delta x} \leq \|A\|_\infty^{\Delta x}\|A^T\|_\infty^{\Delta x}$ on the left hand side and $\|u\|_\infty^{\Delta x} \leq \Delta x^{-1/2}\|u\|_2^{\Delta x}$ on the right gives the desired estimate (20).

Let us estimate the nonlinear diffusion term $(\widetilde{P_{\Delta x}}(u_n))_i := B(u_{n,i})(D_+D_-u_n)_i$ in detail. $B$ is defined on $U_\eta$. First we have to construct a $\|\cdot\|_{\infty,2}$–ball $\mathcal{B}(u^{pi}|_h, C_{\Delta x})$ contained in $D_h = \mathcal{F}_p(\Omega_h, U_\eta)$.

Since $\tilde{u}$ and $u^{pi}$ are smooth functions and $U_\eta$ is an open set, there exist $\eta^* \in (0,\eta)$ and $\Delta x_0 \in (0,1]$ such that $u^{pi}$ takes values in $U_{\eta^*}$ and $u^{pi}|_h \in \mathcal{F}_p(\Omega_h, U_{\eta^*}) \subset D_h$ for all $\Delta x \in [0, \Delta x_0]$. Furthermore, the ball with radius $C_{\Delta x} := \frac{\eta-\eta^*}{2}\sqrt{\Delta x}$ is contained in $D_h$. This is observed by selecting an arbitrary $v \in \mathcal{B}(u^{pi}|_h, C_{\Delta x})$. For this $v$ we have

$$\|u^{pi}|_h - v\|_{\infty,\infty}^h < \frac{\eta-\eta^*}{2}.$$

Since $u^{pi}(t,x,\Delta x) \in U_{\eta^*}$, there is some $x_0 \in (0,1]$ such that $|u^{pi}(t,x,\Delta x) - z(x_0)| < \eta^*$. As a consequence $|v(t,x) - z(x_0)| < \frac{\eta-\eta^*}{2} < \eta$, that means $v \in D_h$.

Now, write down the Jacobian of $\widetilde{P_{\Delta x}}$ at $u_n$ where $u \in \mathcal{B}(u^{pi}|_h, C_{\Delta x})$

$$D\widetilde{P_{\Delta x}}(u_n) =$$

$$\frac{1}{\Delta x^2}\begin{pmatrix} -2B(u_{n,1}) & B(u_{n,1}) & & & B(u_{n,1}) \\ B(u_{n,2}) & -2B(u_{n,2}) & B(u_{n,2}) & & \\ & \ddots & \ddots & \ddots & \\ & & B(u_{n,M-1}) & -2B(u_{n,M-1}) & B(u_{n,M-1}) \\ B(u_{n,M}) & & & B(u_{n,M}) & -2B(u_{n,M}) \end{pmatrix}$$

$$+\frac{1}{\Delta x^2}\,\text{blockdiag}\,(DB(u_{n,i})(u_{n,i+1} - 2u_{n,i} + u_{n,i-1}),\ x_i \in \Omega_{\Delta x}).$$

We have to estimate the difference of this derivative at $u_n^{pi}|_{\Delta x}$ and $v_n$ where $v \in \mathcal{B}(u^{pi}|_h, C_{\Delta x})$ ie. $\|u^{pi}|_h - v\|_{\infty,\infty}^h < \frac{\eta-\eta^*}{2}$. Since $B$ is smooth, it is Lipschitz continuous on the compact set

$$K := \left\{u \in I\!R^m : \exists (t,x) \in \Omega, |u - u^{pi}(t,x,\Delta x)| \leq \frac{\eta-\eta^*}{2}, \forall \Delta x \in [0,\Delta x_0]\right\} \subset U_\eta.$$

This leads to the estimate (41) for the $B$ differences. The terms involving derivatives of $B$ can be estimated by

$$\begin{aligned} &|DB(u_{n,i}^{pi})(u_{n,i+1}^{pi} - 2u_{n,i}^{pi} + u_{n,i-1}^{pi}) - DB(v_{n,i})(v_{n,i+1} - 2v_{n,i} + v_{n,i-1})| \\ &= \ |DB(v_{n,i})(u_{n,i+1}^{pi} - 2u_{n,i}^{pi} + u_{n,i-1}^{pi} - v_{n,i+1} + 2v_{n,i} - v_{n,i-1}) \\ &\quad\ +(DB(u_{n,i}^{pi}) - DB(v_{n,i}))(u_{n,i+1}^{pi} - 2u_{n,i}^{pi} + u_{n,i-1}^{pi})| \\ &\leq \ |DB(v_{n,i})|\{|u_{n,i+1}^{pi} - v_{n,i+1}| + 2|u_{n,i}^{pi} - v_{n,i}| + |u_{n,i-1}^{pi} - v_{n,i-1}|\} \\ &\quad\ +C_1\,|u_{n,i}^{pi} - v_{n,i}| \cdot 4\,\|u_n^{pi}\|_\infty^{\Delta x} \\ &\leq C_2\,\|u_n^{pi}|_{\Delta x} - v_n\|_\infty^{\Delta x}, \end{aligned}$$

where $C_{1,2}$ are independent of $v \in \mathcal{B}(u^{pi}|_h, C_{\Delta x})$ and independent of $h$. There is no essential difficulty estimating the first order terms. The final result is:

LEMMA 6.1. *Consider the nonlinear operator $\Phi_h$ defined in (10) and the space discretization of general form (40). If the coefficients are smooth matrix functions, the Lipschitz inequality (20) holds with*

$$Lip_{\Delta x} \leq \frac{C}{\Delta x^{q_P+0.5}}, \quad C_{\Delta x} = \frac{\eta - \eta^*}{2}\sqrt{\Delta x} \quad \textit{and} \quad \eta^* \in (0, \eta).$$

*$q_P$ is the highest order of any finite difference expression occurring nonlinear in (40).*

**7. Proof of the main result.** With the help of the previous results this proof consists in arranging the arguments in the right order. We want to apply the convergence theorem 3.2, therefore the linearized scheme has to be stable. The original method $\Phi_h(u) = 0$ linearized at the pilot function $u^{pi}$ is stable if and only if

$$D\Phi_h(u^{pi}|_h)\, u = 0 \tag{42}$$

is stable. This formula represents a backward Euler step applied to the linear method of lines system

$$\frac{\partial}{\partial t}\, u(t,\cdot) + DP_{\Delta x}(u^{pi}(t,\cdot,\Delta x))\, u(t,\cdot) = 0, \quad u(0,\cdot) = z|_{\Delta x}.$$

In case of the central operator (12) the linearized space operator reads

$$\begin{aligned}
&(DP_{\Delta x}(u^{pi}|_{\Delta x}(t,\cdot,\Delta x))\, u)_i \\
= \;& A(u^{pi}(t,x_i,\Delta x))(D_0 u)_i - (B(u^{pi}(t,x_i,\Delta x)) + \delta \Delta x I)(D_+D_- u)_i \\
&+ DA(u^{pi}(t,x_i,\Delta x))(D_0 u^{pi}(t,\cdot,\Delta x))_i\, u_i \\
&- DB(u^{pi}(t,x_i,\Delta x))(D_+D_- u^{pi}(t,\cdot,\Delta x))_i\, u_i,
\end{aligned}$$

$$u \in \mathcal{F}_p(\Omega_{\Delta x}, I\!R^m), \quad (t,x,\Delta x) \in [0,T]\times\Omega_{\Delta x}\times(0,1].$$

This operator is of the general form

$$P_{\Delta x}(u^{pi}|_{\Delta x})\, u = L_{\Delta x}(t)\, u - C\, u \tag{43}$$

where $L_{\Delta x}$ is an operator of type (33). The matrix $C$ is smooth and periodic, but actually not defined at $\Delta x = 0$. Nevertheless, the $\infty$–norm with respect to all three variables is bounded

$$\|C\|_{\infty,\infty,\infty} := \sup_{\Delta x \in (0,1]} \|C(\cdot,\cdot,\Delta x)\|_{\infty,\infty} < \infty, \tag{44}$$

since finite differences of a smooth function are bounded by the corresponding derivative. We can estimate

$$\|D_0 u^{pi}(\cdot,\cdot,\Delta x)\|^h_{\infty,\infty} \leq \max_{\Delta x \in [0,1]} \|u^{pi}_x(\cdot,\cdot,\Delta x)\|_{\infty,\infty}$$

and

$$\|D_+D_-u^{pi}(\cdot,\cdot,\Delta x)\|^h_{\infty,\infty} \leq \max_{\Delta x\in[0,1]} \|u^{pi}_{xx}(\cdot,\cdot,\Delta x)\|_{\infty,\infty}$$

independent of $t$ and $\Delta x$. Furthermore, the parabolic part of this mixed operator is uniformly strongly parabolic since $B_{11}(t,x,\Delta x) = B_{11}(u^{pi}(t,x,\Delta x))$ and for $\Delta x$ small enough $u^{pi}(t,x,\Delta x)$ is in $U_\eta$, independent of $\Delta x$. In a similar way the hyperbolic part $A_{22}$ is uniformly symmetrizable. Following Lemma 5.2, $L_{\Delta x}$ is energy bounded. The same is true for $DP_{\Delta x}(u^{pi}|_{\Delta x})$, since $(u, Cu)_H$ can be estimated by the Cauchy Schwarz inequality. Lemma 5.1 states that (42) is stable with $\|\cdot\|_{\infty,2}$.

The linearization of the upwind operator (14) can be written as (43) where $L_{\Delta x}$ is of type (37). Condition (13) ensures (35) for the linear counterpart. Again the parabolic part is uniformly strongly parabolic and Lemma 5.3 with 5.1 gives linear stability.

In case of (16) & (17) linear stability follows from Lemma 5.4 and 5.1.

In all three cases Theorem 3.1 and Lemma 6.1 establish local stability regions around $u^{pi}$ with radius

$$R_{\Delta x} \geq C\ \Delta x^{q_P+0.5}$$

and $C$ independent of $h$. Since $q_P \leq 2$ and the order of consistency of the pilot function is $p = 3$, we have

$$\frac{\|\Phi_h(u^{pi}|_h)\|_{\infty,2}}{R_{\Delta x}} = O(\Delta x^{0.5}), \quad \Delta x \to 0$$

and (22) is fulfilled. Finally application of Theorem 3.2 completes the proof. □

**8. Acknowledgement.** The author wants to thank Rolf Jeltsch and Christian Lubich, who gave advice and support. Last but not least special thanks to Ragnar Winther, who read the manuscript and made a number of helpful suggestions for corrections and improvements.

## REFERENCES

[1] D. F. Griffiths and J. M. Sanz-Serna, *On the scope of the method of modified equations*, SIAM J. Sci. Stat. Comput., 7 (1986), pp. 994–1008.

[2] B. Gustafsson and A. Sundström, *Incompletely parabolic problems in fluid dynamics*, SIAM J. Appl. Math., 35 (1978), pp. 343–357.

[3] T. Hagstrom and J. Lorenz, *All-time existence of smooth solutions to PDEs of mixed type and the invariant subspace of uniform states*, Adv. Appl. Math., 16 (1995), pp. 219–257.

[4] R. A. Horn and C. R. Johnson, *Topics in Matrix Analysis*, Cambridge University Press, Cambridge, 1991.

[5] H.-O. Kreiss and J. Lorenz, *Initial–Boundary Value Problems and the Navier–Stokes Equations*, Academic Press, London, 1989.

[6] J. C. López-Marcos and J. M. Sanz-Serna, *A definition of stability for nonlinear problems*, in Numerical Treatment of Differential Equations, K. Strehmel, ed., Teubner, Leipzig, 1988, pp. 216–226.

[7] J. C. López-Marcos and J. M. Sanz-Serna, *Stability and convergence in numerical analysis III: Linear investigation of nonlinear stability*, IMA J. Numer. Anal., 8 (1988), pp. 71–84.

[8] J. M. Sanz-Serna, *Convergence of the Lambert-McLeod trajectory solver and of the Celf method*, Numer. Math., 45 (1984), pp. 173–182.

[9] ——, *Two topics in nonlinear stability*, in Advances in Numerical Analysis, Vol. 1, W. Light, ed., Clarendon Press, Oxford, 1991, pp. 147–174.

[10] H. J. Schroll, *Konvergenz finiter Differenzenverfahren für nichtlineare hyperbolisch–parabolische Systeme.* Dissertation, ETH Zürich, 1993.

[11] H. J. Stetter, *Analysis of Discretization Methods for Ordinary Differential Equations*, Springer-Verlag, Berlin, 1973.

[12] G. Strang, *Accurate partial difference methods II. Non–linear problems*, Numer. Math., 6 (1964), pp. 37–46.