A new finite-volume approach for the solution of convection-diffusion problems

Dietmar Hietel* **and Jürgen Witzel****

Abstract

For the unsteady convection-diffusion equation in two dimensions we derive a new cell-based semi-discretization which is founded on the method of lines and a finite volume approach. Moreover, we present a second semidiscretization technique, a nonconforming Petrov-Galerkin method with exponentially fitted trial and test functions. If we use appropriate quadrature rules both approaches are equivalent. We obtain differential-algebraic equations of index 1 and present consistency, stability and convergence properties for these semidiscrete systems. For the time integration we use implicit methods and state the convergence of the fully discrete systems. Finally, we consider several test problems and present numerical results concerning the quality of our discretization method.

AMS Subject Classification: 65M20, 65M55, 76D30

Key words: convection-diffusion equation, finite volume method, exponential fitting, nonconforming Petrov-Galerkin method.

1 Introduction

Convection-diffusion problems appear in many fields of applications. Especially in computational fluid dynamics they serve as a model problem for the derivation of numerical schemes. The aim of this article is to present a new finite-volume approach for the solution of the unsteady linear convection-diffusion equation of the form

$$u_t + Lu = q \quad \text{with} \quad Lu = \boldsymbol{\nabla} \cdot (\boldsymbol{\beta} u - \epsilon \boldsymbol{\nabla} u) \tag{1}$$

on a bounded domain $\Omega \subseteq \mathbb{R}^2$ with constant diffusion coefficient $\epsilon > 0$ and a given divergence free velocity field $\boldsymbol{\beta}(\boldsymbol{x})$ in $\mathbb{R}^2$. The initial and boundary conditions are given by

$$\begin{array}{rcll} u(\boldsymbol{x},0) & = & u_0(\boldsymbol{x}) & \text{on} \quad \Omega \\ u(\boldsymbol{x},t) & = & u_D(\boldsymbol{x}) & \text{on} \quad \Gamma_D \supseteq \Gamma_{\text{in}} := \{\boldsymbol{x} \in \partial\Omega\,,\ \boldsymbol{\beta}(\boldsymbol{x}) \cdot \boldsymbol{n} \leq 0\} \\ \partial_n u(\boldsymbol{x},t) & = & 0 & \text{on} \quad \Gamma_N \subseteq \partial\Omega \setminus \Gamma_D\,. \end{array} \tag{2}$$

In the convection dominated case ($|\boldsymbol{\beta}| \gg \epsilon$) boundary layers can occur due to Dirichlet conditions on the outflow boundary if $\Gamma_D \setminus \Gamma_{in} \neq \emptyset$ and internal layers have to be considered when using nonsmooth Dirichlet data u_D.

The third section presents the derivation of a related Petrov-Galerkin method. Moreover, it is shown that both methods are equivalent under certain conditions.

*Institut für Techno- und Wirtschaftsmathematik e.V. (ITWM), Erwin-Schrödinger-Straße, 67663 Kaiserslautern, GERMANY, e-mail: hietel@itwm.uni-kl.de

**Technische Universität Darmstadt, Fachbereich Mathematik, AG Numerische Mathematik, Schloßgartenstr. 7, 64289 Darmstadt, GERMANY, e-mail: witzel@mathematik.tu-darmstadt.de

The properties of the semidiscrete systems are stated in the fourth section. We also treat the time integration of these systems and conclude with a consideration of the fully discrete systems.

Finally we describe several test problems and discuss our numerical results.

2 Cell-based semi-discretization

The convection-diffusion equation (1) is a parabolic conservation law in divergence form. Using Gauss' theorem it can be written in integral form

$$\frac{d}{dt}\int_{\tilde\Omega} u\,dV + \oint_{\partial\tilde\Omega} [\boldsymbol{\beta}\cdot\boldsymbol{n}\,u - \epsilon\,\partial_n u]\;ds = \int_{\tilde\Omega} q\,dV \quad \text{for all } \tilde\Omega \subseteq \Omega \tag{3}$$

where $\tilde\Omega$ is a domain with piecewise smooth boundary $\partial\tilde\Omega$. This equation is one possibility of a weak form [2].

The cell-based semidiscretization is based on a finite volume approach for (3) using the method of lines. In the following we will outline the main steps of its derivation under the simplifying assumption of a constant velocity field $\boldsymbol{\beta}$ with nonvanishing components.

For a rectangular cell V with edges S_i (see figure 1) (3) leads to

$$\frac{d}{dt}\int_V u\,dV + \sum_{i=1}^{4}\boldsymbol{\beta}\cdot\boldsymbol{n}_i\int_{S_i} u\,ds - \epsilon\int_{S_i}\partial_n u\,ds = \int_V q\,dV\,. \tag{4}$$

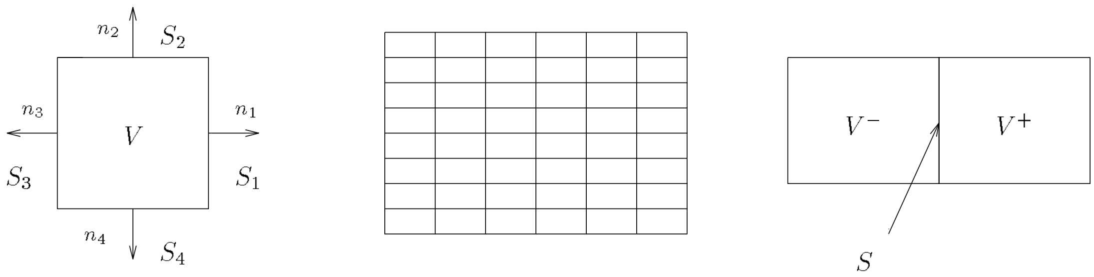


Figure 1: local cell, grid and neighbouring cells

Equation (4) motivates the use of average values of u and their semidiscrete analogues

$$\begin{aligned}
u_V &\approx & u(V) &:= \frac{1}{|V|}\int_V u\,dV & &\text{(cell average)}\,,\\
u_S &\approx & u(S) &:= \frac{1}{|S|}\int_S u\,ds & &\text{(edge average)}\,,\\
\partial_n u_S &\approx & \partial_n u(S) &:= \frac{1}{|S|}\int_S \partial_n u\,ds & &\text{(edge average of normal derivatives)}
\end{aligned} \tag{5}$$

where $|V|$ and $|S|$ indicate the measure of V and S, respectively.

For a rectangular grid with cells $V \in \mathcal{V}$ and edges $S \in \mathcal{S}$ (see figure 1) the finite volume equation (4) has the semidiscrete counterpart

$$|V|\frac{d}{dt}u_V + \sum_{i=1}^{4} |S_i| \left[\boldsymbol{\beta} \cdot \boldsymbol{n}_i \, u_{S_i} - \epsilon \partial_n u_{S_i}\right] \; = |V| q_V \tag{6}$$

for all $V \in \mathcal{V}$.

Here, q_V denotes the exact evaluation of $q(V)$ or – for practical purposes – an approximation by a cubature formula. Rewriting the boundary conditions in the form

$$\begin{aligned} u_S(0) &= u_D(S) \quad \text{for} \quad S \in \partial\mathcal{S}_D := \{S \in \mathcal{S} : S \subseteq \Gamma_D\}, \\ \partial_n u_S(0) &= 0 \qquad\;\; \text{for} \quad S \in \partial\mathcal{S}_N := \{S \in \mathcal{S} : S \subseteq \Gamma_N\} \end{aligned} \tag{7}$$

we get an underdetermined system. Thus, we have to define additional conditions for edge averages u_S, $\partial_n u_S$.

Standard cell-center finite volume techniques approximate the convective flux $\boldsymbol{\beta} \cdot \boldsymbol{n}_S \, u_S$ by a numerical flux function and the diffusive flux $-\varepsilon \partial_n u_S$ by a difference formula using the cell averages u_V of the cells in the neighbourhood of S.

In contrast to this explicit flux representation we use an implicit one given by the representation of the normal derivative $\partial_n u_S$ as a linear combination of the cell and edge averages in a cell V in the form

$$\partial_n u_S = \alpha_V u_V + \sum_{i=1}^{4} \alpha_{S_i} u_{S_i} \, . \tag{8}$$

The coefficients are determined by postulating the validity of (8) for the averages on a local trial space

$$\mathbb{V}_V = \operatorname{span}\left\{1, x_1, x_2, \exp\left(\frac{\beta_1}{\epsilon} x_1\right), \exp\left(\frac{\beta_2}{\epsilon} x_2\right)\right\}. \tag{9}$$

The basis functions of $\mathbb{V}_V$ are inhomogeneous exponentially fitted L-splines [11, 16] varying only in one coordinate direction.

The values for α_V and α_{S_i}, $i = 1, \ldots, 4$ are determined by solving a 5×5 linear system. The result is that the normal derivatives are given by

$$\partial_n u_{S_i} = \frac{1}{2h_i}\left\{\left[u_{S_i} - u_{S_{i+2}}\right] + \theta_i \left(1 + \frac{1}{\coth(\theta_i) - 1/\theta_i}\right)\left[u_{S_i} + u_{S_{i+2}} - 2u_V\right]\right\} \tag{10}$$

$$\partial_n u_{S_{i+2}} = -\frac{1}{2h_i}\left\{\left[u_{S_i} - u_{S_{i+2}}\right] + \theta_i \left(1 - \frac{1}{\coth(\theta_i) - 1/\theta_i}\right)\left[u_{S_i} + u_{S_{i+2}} - 2u_V\right]\right\} \tag{11}$$

for $i = 1, 2$. Here $\theta_i := \dfrac{\beta_i |S_{i+1}|}{2\varepsilon}$ is the local Peclet number.

This representation enables us to eliminate of the normal derivatives in (6) and (7).

For each interior edge S belonging to two cells V^+, V^- (see figure 1) the normal derivatives $\partial_n u_S$ have different signs so that we get additional constraints in terms of the unknowns $\boldsymbol{u}_\mathcal{V} := (u_V)_{V\in\mathcal{V}}$ and $\boldsymbol{u}_\mathcal{S} := (u_S)_{S\in\mathcal{S}}$ only. These equations combined with the finite volume equations (4) and boundary conditions altered by the elimination of the normal derivatives define a differential-algebraic system of the form

$$\boldsymbol{M}\frac{d}{dt}\boldsymbol{u} + \boldsymbol{B}\boldsymbol{u} = \boldsymbol{b} \iff \begin{pmatrix} \boldsymbol{V} & 0 \\ 0 & 0 \end{pmatrix} \frac{d}{dt}\begin{pmatrix} \boldsymbol{u}_\mathcal{V} \\ \boldsymbol{u}_\mathcal{S} \end{pmatrix} + \begin{pmatrix} \boldsymbol{B}_{11} & \boldsymbol{B}_{12} \\ \boldsymbol{B}_{21} & \boldsymbol{B}_{22} \end{pmatrix} \begin{pmatrix} \boldsymbol{u}_\mathcal{V} \\ \boldsymbol{u}_\mathcal{S} \end{pmatrix} = \begin{pmatrix} \boldsymbol{b}_1 \\ \boldsymbol{b}_2 \end{pmatrix} \qquad (12)$$

where $\boldsymbol{V} = \mathrm{diag}(|V|)_{V\in\mathcal{V}}$.

Here the differential part consists of the finite volume equations and the algebraic part is composed of the boundary conditions and the additional constraints for the interior edges.

3 A nonconforming Petrov-Galerkin method

Here, we restrict ourselves to the case of homogeneous Dirichlet boundary conditions. For the well-known extension to the general inhomogeneous case we refer to [3, 5].

With the aim to obtain a weak formulation for the original differential equation we multiply (1) with a function $w = w(\boldsymbol{x}) \in \mathbb{W}$ and integrate over the domain Ω. We choose $H_0^1 \subset \mathbb{V} := \{v \in H^1 | v = 0 \text{ on } \Gamma_D\} \subset H^1(\Omega)$ and $\mathbb{V} = \mathbb{W}$. Using Gauss' Theorem we observe that $u(t) \in \mathbb{V}$ is now the solution of the weak formulation

$$\frac{d}{dt}\int_\Omega u\, w\, d\Omega + \int_\Omega (\boldsymbol{\beta}\cdot\boldsymbol{\nabla} u\, w + \varepsilon \boldsymbol{\nabla} u \cdot \boldsymbol{\nabla} w)\; d\Omega \;\; = \;\; \int_\Omega q w\, d\Omega \quad \forall w \in \mathbb{W}$$

which can be written shortly as

$$\frac{d}{dt}(u,w) + a(u,w) = (q,w) \quad \forall w \in \mathbb{W}\,.$$

The nonsymmetric bilinear form $a(\cdot,\cdot)$ is defined by

$$a(u,w) \;\; := \;\; \int_\Omega \boldsymbol{\beta}\cdot\boldsymbol{\nabla} u\, w + \varepsilon\boldsymbol{\nabla} u\cdot\boldsymbol{\nabla} w\, d\Omega$$

and $(\cdot,\cdot)$ denotes the L_2-scalar product.

To obtain a Petrov-Galerkin method we replace the spaces $\mathbb{V}$ and $\mathbb{W}$ by finite spaces $\mathbb{V}_h$ and $\mathbb{W}_h$. The construction of the trial space $\mathbb{V}_h$ is based on the local spaces $\mathbb{V}_V$ from (9). We require that a trial function $v \in \mathbb{V}_h$ restricted to a cell V belongs to $\mathbb{V}_V$ and that the evaluation of the average for an interior edge S using these restrictions is independent of the cell under consideration. Therefore we can define a canonical basis of $\mathbb{V}$ by using the cell and edge averages as 'nodal values'.

Canonical trial functions ϕ^V, $V \in \mathcal{V}$ and ϕ^S, $S \in \mathcal{S}$ as a basis of $\mathbb{V}_h$ are defined via

$$\begin{aligned} \phi^V(\tilde V) = \frac{1}{|\tilde V|}\int_{\tilde V} \phi^V(\boldsymbol{x})d\tilde V &= \delta_{V\tilde V} \quad \forall \tilde V \in \mathcal{V}\,, \\ \phi^V(\tilde S) = \frac{1}{|\tilde S|}\int_{\tilde S} \phi^V(\boldsymbol{x})d\tilde S &= 0 \qquad \forall \tilde S \in \mathcal{S}\,, \end{aligned}$$

and

$$\begin{aligned} \phi^S(\tilde V) = \frac{1}{|\tilde S|}\int_{\tilde S} \phi^S(\boldsymbol{x})d\tilde S &= \delta_{S\tilde S} \quad \forall \tilde S \in \mathcal{S}\,, \\ \phi^S(\tilde S) = \frac{1}{|\tilde V|}\int_{\tilde V} \phi^S(\boldsymbol{x})d\tilde V &= 0 \qquad \forall \tilde V \in \mathcal{V}\,. \end{aligned}$$

In analogy to the definition of the trial functions we construct a test space $\mathbb{W}_h$ which is for each $V \in \mathcal{V}$ based on the local test space

$$\mathbb{W}_V = \operatorname{span}\Big\{1, x_1, x_2, \exp\Big(-\frac{\beta_1}{\epsilon}x_1\Big), \exp\Big(-\frac{\beta_2}{\epsilon}x_2\Big)\Big\}\,. \tag{13}$$

Here, $\mathbb{W}_V$ consists of inhomogeneous exponentially fitted L^*-splines where L^* denotes the adjoint operator of L.

Canonical test functions ψ^V, $V \in \mathcal{V}$ and ψ^S, $S \in \mathcal{S}$ are defined analogously to the canonical trial functions. The trial and the test functions are plotted in figure 2.

We observe that the trial and test functions are discontinuous at interelement boundaries. Consequently, we have to deal with nonconforming spaces and, moreover, we have to change the bilinear form $a(\cdot,\cdot)$ – by summing up over all grid cells – into

$$a_h(u,w) := \sum_{V\subset\Omega}\int_V \boldsymbol{\beta}\cdot\boldsymbol{\nabla} u\, w + \varepsilon \boldsymbol{\nabla} u\cdot\boldsymbol{\nabla} w\, dV\,. \tag{14}$$

Hence the resulting nonconforming Petrov-Galerkin method reads:

Find $u(t) \in \mathbb{V}_h$ such that

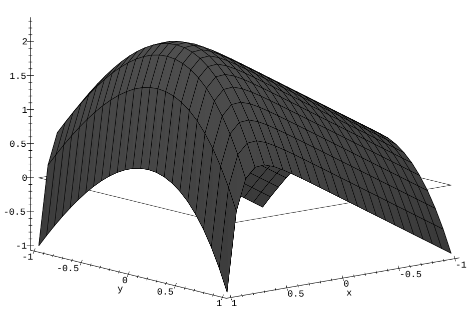
$$\frac{d}{dt}(u,w) + a_h(u,w) = (q,w) \quad \forall w \in \mathbb{W}_h\,. \tag{15}$$

Describing $u(t)$ by its nodal values as

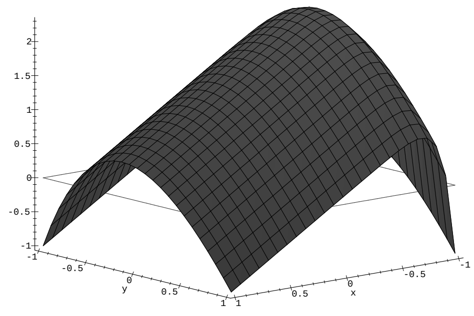
$$u(t) = \sum_{V\in\mathcal{V}} u_V(t)\Phi_V + \sum_{S\in\mathcal{S}} u_S(t)\Phi_S$$

and inserting the canonical test functions into (15) we obtain by setting $\boldsymbol{u}_{\mathcal{V}} := (u_V)_{V\in\mathcal{V}}$ and $\boldsymbol{u}_{\mathcal{S}} := (u_S)_{S\in\mathcal{S}}$

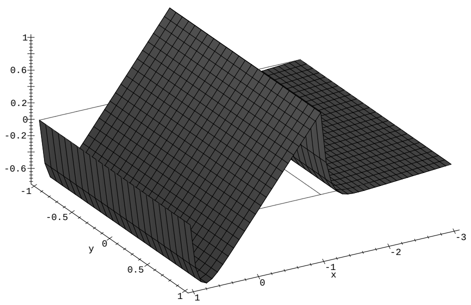
$$\hat{\boldsymbol{M}}\frac{d}{dt}\boldsymbol{u} + \boldsymbol{B}\boldsymbol{u} = \boldsymbol{b} \iff \hat{\boldsymbol{M}}\frac{d}{dt}\boldsymbol{u} + \begin{pmatrix} \boldsymbol{B_{11}} & \boldsymbol{B_{12}} \\ \boldsymbol{B_{21}} & \boldsymbol{B_{22}} \end{pmatrix}\begin{pmatrix} \boldsymbol{u}_{\mathcal{V}} \\ \boldsymbol{u}_{\mathcal{S}} \end{pmatrix} = \begin{pmatrix} \boldsymbol{b}_1 \\ \boldsymbol{b}_2 \end{pmatrix}\,. \tag{16}$$



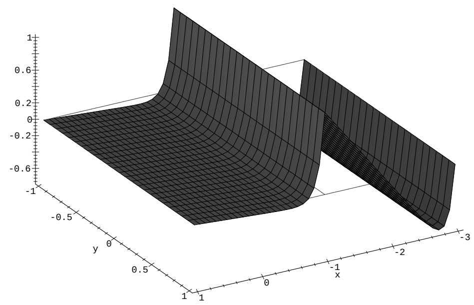
Cell trial function



Cell test function



Edge trial function



Edge test function

Figure 2: Trial and test functions in two dimensions

Here the stiffness matrix $\boldsymbol{B}$ is exactly the same as in the finite volume case of the preceding section (see equation (12)). This is shown in [16]. If we approximate the L_2-scalar product appearing in the mass matrix $\hat{\boldsymbol{M}}$ and the right hand side $\boldsymbol{b}$ of (16) by the quadrature rule

$$(f, g) \approx \sum_{V \in \mathcal{V}} |V| f(V) g(V)$$

we achieve completely the same system as in the finite volume case. An identical mass matrix is also obtained by using a mass lumping technique because of

$$\sum_{V \in \mathcal{V}} \Phi^V + \sum_{S \in \mathcal{S}} \Phi^S \equiv 1$$

in all cells.

For both approaches presented up to now the stiffness matrix $\boldsymbol{B}$ in (12) can be assembled by the cell matrices

$$
\begin{aligned}
\boldsymbol{B}_{V,i} \;&=\; \frac{|S_i|}{2}\left\{ \operatorname{cof}\left(\beta, \frac{2\varepsilon}{|S_{i+1}|}\right)\begin{pmatrix} 4 & -2 & -2 \\ -2 & 1 & 1 \\ -2 & 1 & 1 \end{pmatrix} + \frac{2\varepsilon}{|S_{i+1}|}\begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & -1 \\ 0 & -1 & 1 \end{pmatrix} \right. \\
&\left. \quad + \beta_i \begin{pmatrix} 0 & -2 & 2 \\ 2 & -1 & -1 \\ -2 & 1 & 1 \end{pmatrix} \right\} \text{ for } V \in \mathcal{V},\, i = 1, 2.
\end{aligned}
$$

The cell matrices $\boldsymbol{B}_{V,i}$ are connected to the cell vectors $\boldsymbol{u}_{V,i} := (u_V, u_{S_{i+2}}, u_{S_i})^T$ ignoring rows and columns belonging to Dirichlet data u_S for $S \in \partial\mathcal{S}_D$. The scalar B_S is associated with u_S for Neumann boundaries $S \in \mathcal{S}_N$.

The function cof is defined by

$$
\operatorname{cof}(\xi, \eta) \;=\; \begin{cases} \xi\left(\coth(\dfrac{\xi}{\eta}) - \dfrac{\eta}{\xi}\right)^{-1}, & \xi, \eta \neq 0 \\ |\xi|\,, & \eta = 0 \\ 3\eta\,, & \xi = 0 \end{cases} .
$$

At this point we mention that the finite volume and also the Petrov-Galerkin approach can be extended to the case of an arbitrary, divergence free velocity field $\boldsymbol{\beta}$ [7, 16].

4 Convergence of the semidiscretization

As pointed out at the end of section 2 we obtain a semidiscrete system with a differential and an algebraic part of the form

$$
\begin{aligned}
\boldsymbol{V}\frac{d}{dt}\boldsymbol{u}_{\mathcal{V}} + \boldsymbol{B}_{11}\boldsymbol{u}_{\mathcal{V}} + \boldsymbol{B}_{12}\boldsymbol{u}_{\mathcal{S}} &= \boldsymbol{b}_1(t) \\
\boldsymbol{B}_{21}\boldsymbol{u}_{\mathcal{V}} + \boldsymbol{B}_{22}\boldsymbol{u}_{\mathcal{S}} &= \boldsymbol{b}_2(t)
\end{aligned} \tag{17}
$$

Such systems are called differential-algebraic systems [1]. Due to the regularity of $\boldsymbol{B}_{22}$ it is a semi-implicit system of index 1 [7, 16].

For a DAE not all initial conditions do admit smooth solutions. We have to choose consistent initial conditions. This means that the corresponding initial data $\boldsymbol{u}(0) = \boldsymbol{u}^0$ should be consistent in the sense that the algebraic equations are satisfied, i.e.

$$
\boldsymbol{B}_{21}\boldsymbol{u}_{\mathcal{V}}^0 + \boldsymbol{B}_{22}\boldsymbol{u}_{\mathcal{S}}^0 = \boldsymbol{b}_2(0)\,.
$$

Resolving the algebraic part of (12) for the variables $\boldsymbol{u}_{\mathcal{S}}$ and substituting into the differential part we get the underlying system

$$
\boldsymbol{V}\frac{d}{dt}\boldsymbol{u}_{\mathcal{V}} + [\boldsymbol{B}/\boldsymbol{B}_{22}]\boldsymbol{u}_{\mathcal{V}} = \boldsymbol{b}_1 - \boldsymbol{B}_{12}\boldsymbol{B}_{22}^{-1}\boldsymbol{b}_2 \tag{18}
$$

of ordinary differential equations where $[\boldsymbol{B}/\boldsymbol{B}_{22}] := \boldsymbol{B}_{11} - \boldsymbol{B}_{12}\boldsymbol{B}_{22}^{-1}\boldsymbol{B}_{21}$ denotes the Schur complement of $\boldsymbol{B}_{22}$ for the matrix $\boldsymbol{B}$.

In the case of ODEs as semidiscrete systems sufficient conditions for stability ($T < \infty$) and asymptotic stability ($T = \infty$) in the sense of Liapunov are well known (see [7, 13]). To generalize these criteria to DAEs of index 1 we can use the connection to ODEs described by the decoupling of the underlying ODE (18) and the algebraic equations (17b) described in [7, 8].

The equivalence between (17) and

$$\boldsymbol{V}\frac{d}{dt}\boldsymbol{u}_{\mathcal{V}} + [\boldsymbol{B}/\boldsymbol{B}_{22}]\boldsymbol{u}_{\mathcal{V}} = \boldsymbol{b}_1(t) - \boldsymbol{B}_{12}\boldsymbol{B}_{22}^{-1}\boldsymbol{b}_2(t) \tag{19}$$

$$\boldsymbol{u}_{\mathcal{S}} = \boldsymbol{B}_{22}^{-1}\Big(\boldsymbol{b}_2(t) - \boldsymbol{B}_{21}\boldsymbol{u}_{\mathcal{V}}\Big) \tag{20}$$

leads to an extension of the stability results known for ODE semidiscretizations [13] to DAE semidiscretizations.

In the following we consider so-called lub-norms $\|\boldsymbol{A}\|$ defined by $\|\boldsymbol{A}\| := \max\limits_{\|\boldsymbol{x}\|=1} \|\boldsymbol{A}\boldsymbol{x}\|$ where $\|\cdot\|$ is a given vector norm.

Definition 4.1
The logarithmic norm according to a lub-norm is defined by

$$\mu(\boldsymbol{A}) := \lim_{\varepsilon\to 0^+} \frac{\|I + \varepsilon\boldsymbol{A}\| - 1}{\varepsilon}.$$

□

The proof of the following theorem can be found in [7].

Theorem 4.2
The semidiscretization (17) is asymptotically stable if there exist constants $\ell < 0$ and $L < \infty$ independent of $h > 0$ with

$$\mu(-\boldsymbol{V}^{-1}[\boldsymbol{B}/\boldsymbol{B}_{22}]) \le \ell \quad , \quad \|\boldsymbol{B}_{22}^{-1}\boldsymbol{B}_{21}\| \le L\,.$$

□

The main key for the application of this theorem is to determine $\ell < 0$ which can be interpreted as a generalized one-sided Lipschitz constant. For this purpose, we cite two estimates proven in [7].

Theorem 4.3
Given a norm $\|\boldsymbol{u}\| := \|\big(\|\boldsymbol{u}_{\mathcal{V}}\|_{\mathcal{V}}, \|\boldsymbol{u}_{\mathcal{S}}\|_{\mathcal{S}}\big)\|_{(0)}$ with $\|.\|_{(0)}$ to be monotone, $\mu(\boldsymbol{C}) \le 0$ and $\boldsymbol{C}_{22}$ regular, then

$$\mu_{\mathcal{V}}([\boldsymbol{C}/\boldsymbol{C}_{22}]) \le \mu(\boldsymbol{C})\,.$$

Let $\boldsymbol{K}$ be a symmetric matrix with $\boldsymbol{K}_{22}$ negative definite and $\mu_2(\boldsymbol{C} - \boldsymbol{K}) \le 0$ then

$$\mu_2([\boldsymbol{C}/\boldsymbol{C}_{22}]) \le \mu_2([\boldsymbol{K}/\boldsymbol{K}_{22}])\,.$$

□

By transformation the latter estimate can also be used for norms given by scalar products without coupling between differential and algebraic variables.

Based on the equivalence between the semi-implicit DAE (17) and the underlying ODE (18) including perturbations one can define an appropriate consistency of the DAE to get the expected result

$$\text{DAE consistency } + \text{asymptotic stability } \Longrightarrow \text{ DAE convergence}.$$

A general framework of this kind for the nonlinear case is presented in [7]. Here we cite only the main results in a version for linear systems with constant coefficients and sufficient conditions for DAE convergence which can be checked for the cell-based discretization directly. For a DAE semidiscretization (17) depending on a grid parameter h the truncation error is given by

$$\begin{array}{rcl} \boldsymbol{\tau}^h_{\mathcal{V}}(t) & = & \frac{d}{dt}\boldsymbol{u}^h_{\mathcal{V}}(t) + \boldsymbol{V}^{-1}\left\{\boldsymbol{B}_{11}\boldsymbol{u}^h_{\mathcal{V}}(t) + \boldsymbol{B}_{12}\boldsymbol{u}^h_{\mathcal{S}}(t) - \boldsymbol{b}_1(t)\right\} \\ \boldsymbol{\tau}^h_{\mathcal{S}}(t) & = & \boldsymbol{B}_{21}\boldsymbol{u}^h_{\mathcal{V}}(t) + \boldsymbol{B}_{22}\boldsymbol{u}^h_{\mathcal{S}}(t) - \boldsymbol{b}_2(t) \end{array} \tag{21}$$

where $\boldsymbol{u}^h = (\boldsymbol{u}^h_{\mathcal{V}}, \boldsymbol{v}^h_{\mathcal{S}})^T$ is the projection of the solution of the partial differential equation onto the semidiscrete space.

Theorem 4.4 (DAE Convergence)
Suppose that the family of DAEs (17) which depends on h fulfills

i) $\max\limits_{t\geq 0}\|\boldsymbol{\tau}^h_{\mathcal{V}}\|_{\mathcal{V}} = \mathcal{O}(h^{k_1}), \quad \max\limits_{t\geq 0}\|\boldsymbol{\tau}^h_{\mathcal{S}}\|_{\mathcal{S}} = \mathcal{O}(h^{k_2})$,

ii) $\|\boldsymbol{V}^{-1}\boldsymbol{B}_{12}\|_{\mathcal{V},\mathcal{S}} = \mathcal{O}(h^{k_{12}})$,

iii) $\|\boldsymbol{B}_{21}\|_{\mathcal{S},\mathcal{V}} = \mathcal{O}(1)$,

iv) $\mu_{\mathcal{V}}(-\boldsymbol{V}^{-1}[\boldsymbol{B}/\boldsymbol{B}_{22}]) \leq l_1 < 0$,

v) $\mu_{\mathcal{S}}(-\boldsymbol{B}_{22}) \leq l_2 < 0$

with l_1, l_2 independent of h and the vector norm $\|\cdot\|$ defined as in Theorem 4.3. Then the DAE semidiscretization converges, i.e.

$$\max_{t\geq 0}\|\boldsymbol{u}^h_{\mathcal{V}} - \boldsymbol{u}_{\mathcal{V}}\|_{\mathcal{V}} = \mathcal{O}(h^k), \quad \max_{t\geq 0}\|\boldsymbol{u}^h_{\mathcal{S}} - \boldsymbol{u}_{\mathcal{S}}\|_{\mathcal{S}} = \mathcal{O}(h^k)$$

with $k := \min\{k_1, k_2 - k_{12}\}$. □

The results of the following part are proven in [7]. Here we quote the main theorems, only. For the cell based semidiscretization we choose the following norms

$$\|\boldsymbol{u}_{\mathcal{V}}\|^2_{\mathcal{V}} := \boldsymbol{u}^T_{\mathcal{V}}\boldsymbol{V}\boldsymbol{u}_{\mathcal{V}}$$

and

$$\|\boldsymbol{u}_{\mathcal{S}}\|^2_{\mathcal{S}} := h^2\boldsymbol{u}^T_{\mathcal{S}}\boldsymbol{u}_{\mathcal{S}}$$

with $\|\cdot\|_{(0)} := \|\cdot\|_2$. This gives now

$$\|\boldsymbol{u}\|^2 = \boldsymbol{u}^T\boldsymbol{H}\boldsymbol{u} \quad \text{with} \quad \boldsymbol{H} = \begin{pmatrix} \boldsymbol{V} & \boldsymbol{0} \\ \boldsymbol{0} & h^2\boldsymbol{I} \end{pmatrix}.$$

Lemma 4.5
For the cell-based semidiscretization we have

$$\begin{aligned}\max_{t\geq 0} \|\boldsymbol{\tau}_{\mathcal{V}}^h\|_{\mathcal{V}} &= \mathcal{O}(h^2)\\ \max_{t\geq 0} \|\boldsymbol{\tau}_{\mathcal{S}}^h\|_{\mathcal{S}} &= \mathcal{O}(h^{k_2})\end{aligned}$$

with $k_2 = 3$ on arbitrary grids, and $k_2 = 4$ for problems on quasi-uniform grids with Dirichlet boundary conditions. $\square$

The proof of this statement is based on a generalized Taylor's expansion for the cell and the edge averages. We summarize the estimates for the norms of the submatrices as follows.

Lemma 4.6
For the cell-based semidiscretization we obtain

i) $\|\boldsymbol{V}^{-1}\boldsymbol{B}_{12}\|_{\mathcal{V},\mathcal{S}} = \varepsilon C_1 h^{-2}$,

ii) $\|\boldsymbol{B}_{21}\|_{\mathcal{S},\mathcal{V}} = \varepsilon C_2$,

iii) $\mu_{\mathcal{V}}(-\boldsymbol{V}^{-1}[\boldsymbol{B}/\boldsymbol{B}_{22}]) \leq -\varepsilon C_3$,

iv) $\mu_{\mathcal{S}}(-\boldsymbol{B}_{22}) \leq -\varepsilon C_4$

with constants $C_1, C_2, C_3, C_4 > 0$. $\square$

From Theorem 4.4 and the lemmas from above we deduce

Corollary 4.7
The cell-based semidiscretization converges of order 1 on arbitrary grids and of order 2 for Dirichlet problems on quasi-uniform grids. $\square$

For the convection case ($\varepsilon = 0$) it is possible to show convergence of order 1 on a finite time interval $0 \leq t \leq T$.

5 Time integration and convergence

For systems of ordinary differential equations we distinguish between one-step and multistep methods. These methods have to be transferred to DAEs. For further reading we refer to [6] and [1].

In [7] Runge-Kutta methods as an important subclass of one-step methods and backward differentiation formulas as a subclass of multistep methods are treated.

We restrict ourselves to multistep methods. Concerning Runge-Kutta methods we refer to [6, 7].

For the semi-implicit DAE (17) we define for abbreviation

$$\begin{aligned} \boldsymbol{f_1}(\boldsymbol{u}_{\mathcal{V}}, \boldsymbol{u}_{\mathcal{S}}, t) &:= \boldsymbol{b}_1(t) - \boldsymbol{B}_{11}\boldsymbol{u}_{\mathcal{V}} - \boldsymbol{B}_{12}\boldsymbol{u}_{\mathcal{S}}, \\ \boldsymbol{f_2}(\boldsymbol{u}_{\mathcal{V}}, \boldsymbol{u}_{\mathcal{S}}, t) &:= \boldsymbol{b}_2(t) - \boldsymbol{B}_{21}\boldsymbol{u}_{\mathcal{V}} - \boldsymbol{B}_{22}\boldsymbol{u}_{\mathcal{S}}\,. \end{aligned} \tag{22}$$

A general multistep method for the semi-implicit system (17) reads

$$\begin{aligned} \boldsymbol{V}(\Delta t)^{-1}\sum_{j=0}^{m} a_j \boldsymbol{u}_{\mathcal{V}}^{n+1-j} &= \sum_{j=0}^{m} b_j \boldsymbol{f}_1(\boldsymbol{u}_{\mathcal{V}}^{n+1-j}, \boldsymbol{u}_{\mathcal{S}}^{n+1-j}, t_{n+1-j}), \\ \boldsymbol{0} &= \sum_{j=0}^{m} b_j \boldsymbol{f}_2(\boldsymbol{u}_{\mathcal{V}}^{n+1-j}, \boldsymbol{u}_{\mathcal{S}}^{n+1-j}, t_{n+1-j}) \end{aligned} \tag{23}$$

where Δt denotes a constant step size.

We assume that the algebraic equations are fulfilled at previous time steps, i.e.

$$\boldsymbol{0} = \boldsymbol{f}_2(\boldsymbol{u}_{\mathcal{V}}^{n+1-j}, \boldsymbol{u}_{\mathcal{S}}^{n+1-j}, t_{n+1-j}) \quad \text{for } j = 1, \ldots, m \tag{24}$$

and, hence, the second equation of (23) becomes

$$\boldsymbol{0} = \boldsymbol{f}_2(\boldsymbol{u}_{\mathcal{V}}^{n+1}, \boldsymbol{u}_{\mathcal{S}}^{n+1}, t_{n+1})\,. \tag{25}$$

A special subclass of the multistep methods defined above are the *backward differentiation formulas* (BDF). They are defined via $b_0 = 1$ and $b_j = 0$ for $j = 1, \ldots, m$. The coefficients a_i are deduced by using backward differencing of consistency order m.

From [1] we cite

Theorem 5.1 (BDF Convergence)
A m-step BDF method with $m \leq 6$ and constant step size Δt applied to a semi-implicit DAE system (17) converges if the initial values $\boldsymbol{u}^j, j = 0, \ldots, m-1$ are correct to $\mathcal{O}(\Delta t^m)$ accuracy and if the linear systems are solved to $\mathcal{O}(\Delta t^{m+1})$ accuracy. □

The coefficients of multistep methods are usually written in the form

$$\begin{array}{c|c} a_0 & b_0 \\ \vdots & \vdots \\ a_m & b_m \end{array}$$

and linear one-step methods can now be represented via

$$\begin{array}{r|c} 1 & \varpi \\ -1 & 1-\varpi \end{array}$$

with $\varpi \in \mathbb{R}$.

For $\varpi = 1$ we obtain the first order consistent *backward Euler method* and for $\varpi = \frac{1}{2}$ the second order consistent *Crank-Nicolson method.* A drawback of the latter methods is that it shows an oscillating behaviour as confirmed by our numerical results.

Additionally, we consider the following 2-step BDF method

$$\begin{array}{r|l} \frac{3}{2} & 1 \\ -2 & 0 \\ \frac{1}{2} & 0 \end{array}$$

which is second order consistent, too. This method is also called *method of Gear.*

At each time step it is necessary to solve one or more systems of linear equations having the form

$$\boldsymbol{A}\boldsymbol{u} = \boldsymbol{f} \quad \text{with} \quad \boldsymbol{A} = \frac{\gamma}{\Delta t}\boldsymbol{M} + \boldsymbol{B} \tag{26}$$

for constant time steps Δt.

The parameter $\gamma > 0$ depends on the chosen method, only. The vector $\boldsymbol{u}$ is either an approximation on the next time level or an increment.

Since our semidiscretization is of order 2 we also choose time integration methods of order 2. Studying the quality of linear solvers for (26) it is adequate to restrict ourselves to the method of Gear (2-step backward differentiation formula) and the Crank-Nicolson scheme. Several other schemes like multistep methods and implicit Runge-Kutta methods are discussed in [7] with regard to the cell-based discretization.

With similar techniques as in the semidiscrete context the convergence theorem for the cell-based discretization can be extended to the full discretization by using contractivity and convergence results for time integration methods (see [7]).

Theorem 5.2
The cell-based discretization is convergent in the grid parameter h with the order stated in Corollary 4.7 and of order 2 in Δt. For the method of Gear there is no restriction on the coupling of h and Δt. $\square$

6 Test problems and numerical results

For our numerical computations we consider three test problems [4, 7, 10, 15] for the linear convection-diffusion equation (1, 2).

Problem 6.1 (Berkeley problem)

$$\boldsymbol{\nabla} \cdot (\boldsymbol{\beta} u - \epsilon \boldsymbol{\nabla} u) = 0 \quad \text{in } \Omega = (-1, 1) \times (0, 1)$$

with $\boldsymbol{\beta} = (-\psi_y, \psi_x)^T$ given by the stream function $\psi = (1 - x^2)(1 - y^2)$ and

$$\begin{aligned} u(x,y) &= \begin{cases} 1 + \tanh(10[2x+1]) & \text{on } \Gamma_{\text{inlet}} := \{(x,y) \in \Gamma : -1 \le x \le 0, y = 0)\}, \\ 0 & \text{on } \Gamma_D \setminus \{\Gamma_{\text{inlet}} \cup \Gamma_N\}, \end{cases} \\ \partial u_n(x,y) &= 0 \quad \text{on } \Gamma_N := \{(x,y) \in \Gamma : 0 \le x \le 1, y = 0)\}. \end{aligned}$$

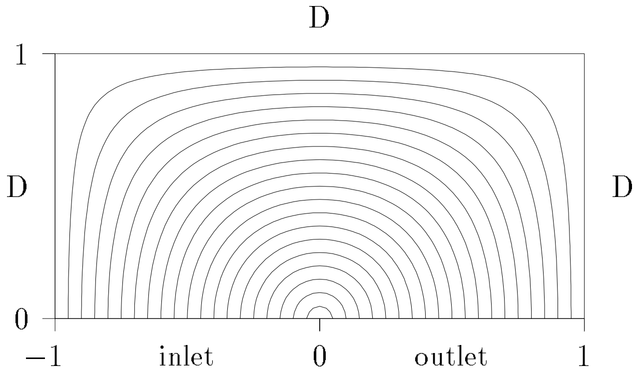


Figure 3: Streamlines and boundary conditions for problem 6.1

Problem 6.2 (Rotating hump problem)

$$u_t + \boldsymbol{\beta} \cdot \boldsymbol{\nabla} u = 0 \quad \text{in } \Omega = (-1,1) \times (-1,1)$$

with $\boldsymbol{\beta} = (-\psi_y, \psi_x)^T$, $\psi = \frac{x^2+y^2}{2}$ and

$$\begin{aligned} u(\boldsymbol{x},0) &= \begin{cases} \frac{1}{2}\left(1+\cos(\frac{5}{2}\pi r)\right) & \text{in } r = \|\boldsymbol{x}-\boldsymbol{x}_0\|_2 \leq \frac{2}{5}\,, \\ 0 & \text{else}\,, \end{cases} \\ u(\boldsymbol{x},t) &= 0 \quad \text{on } \Gamma\,. \end{aligned}$$

Problem 6.3 (Convection-diffusion problem on the unit square)

$$\boldsymbol{\beta} \cdot \boldsymbol{\nabla} u - \epsilon\, \Delta\, u = q \quad \text{in } \Omega = (0,1) \times (0,1)$$

with $\boldsymbol{\beta} = (\cos\frac{\pi}{8}, \sin\frac{\pi}{8})$ and a diffusion coefficient $\varepsilon = 10^{-6}$. We require homogeneous Dirichlet data at the inflow $x = 0$, $y = 0$ and homogeneous Neumann data on the two remaining sides. We adapt the right hand side q by inserting the function $u = \sin\left(\frac{\pi}{2}x\right)\sin\left(\frac{\pi}{2}y\right)$. Thus u is the solution of this problem.

The results presented here were obtained by using a linear multigrid method with problem-adapted components [16]. Especially in combination with Krylov subspace methods [12] this yields a fast and robust solution method. Different Krylov subspace methods and different versions of the multigrid approach are studied in [17].

First of all we have to illustrate the quality of the cell-based discretization. This is influenced by the approximation in space mainly. Therefore it is convenient to consider the steady problem 6.1. Figure 4 shows the results on a 40×40 grid and various values of ϵ. The isolines of the computed solution depicted on the left are indicating that there is no artifical crosswind diffusion. And on the right hand side the good coincidence between the nodal values of the approximation and the exact solution can be observed [14].

In the pure convection problem 6.2 the initial condition is transported a long the streamlines which are concentric circles. Here the initial 'hump' should be rotated with pe-

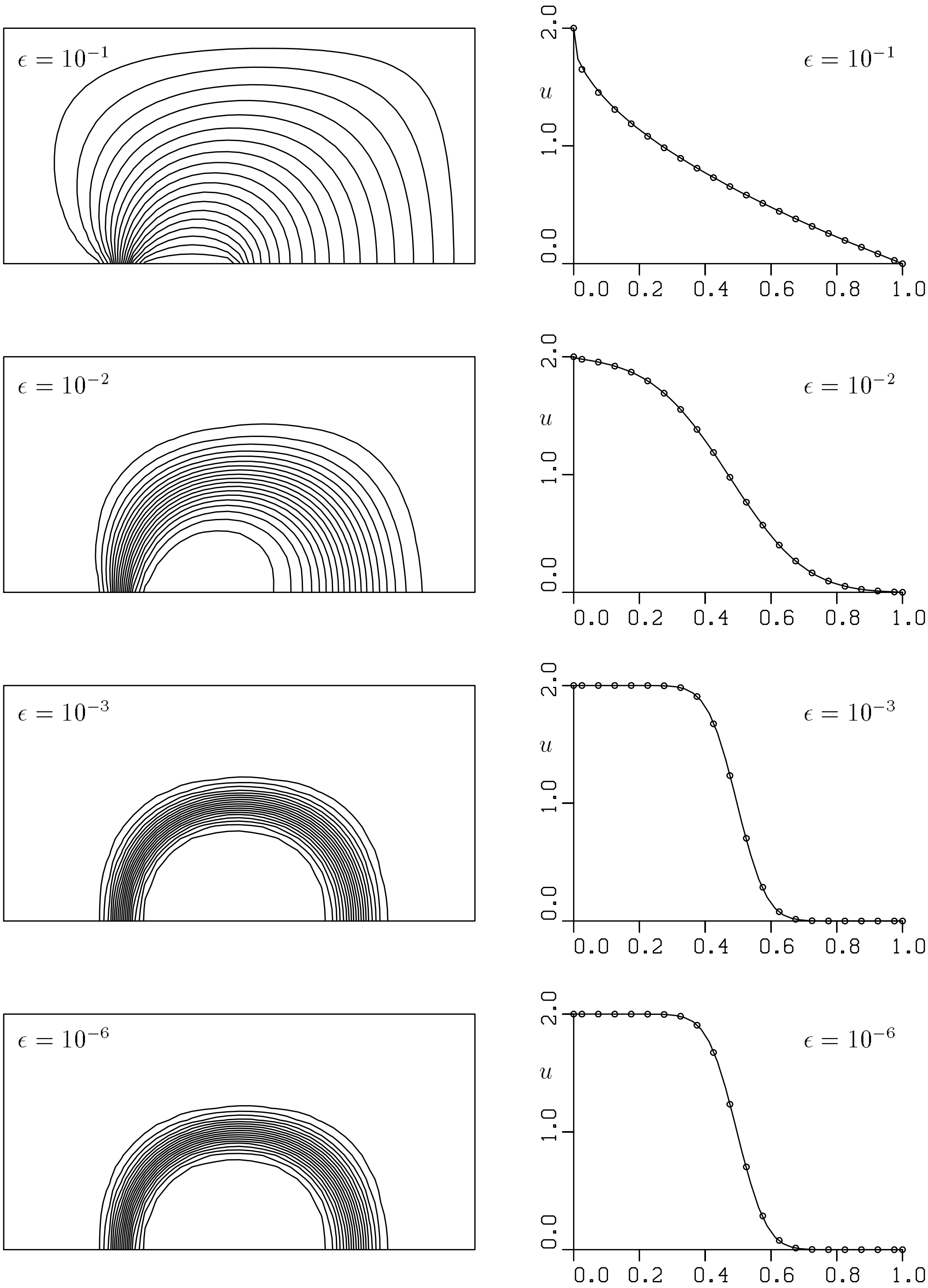


Figure 4: The quality of the cell orientated discretization

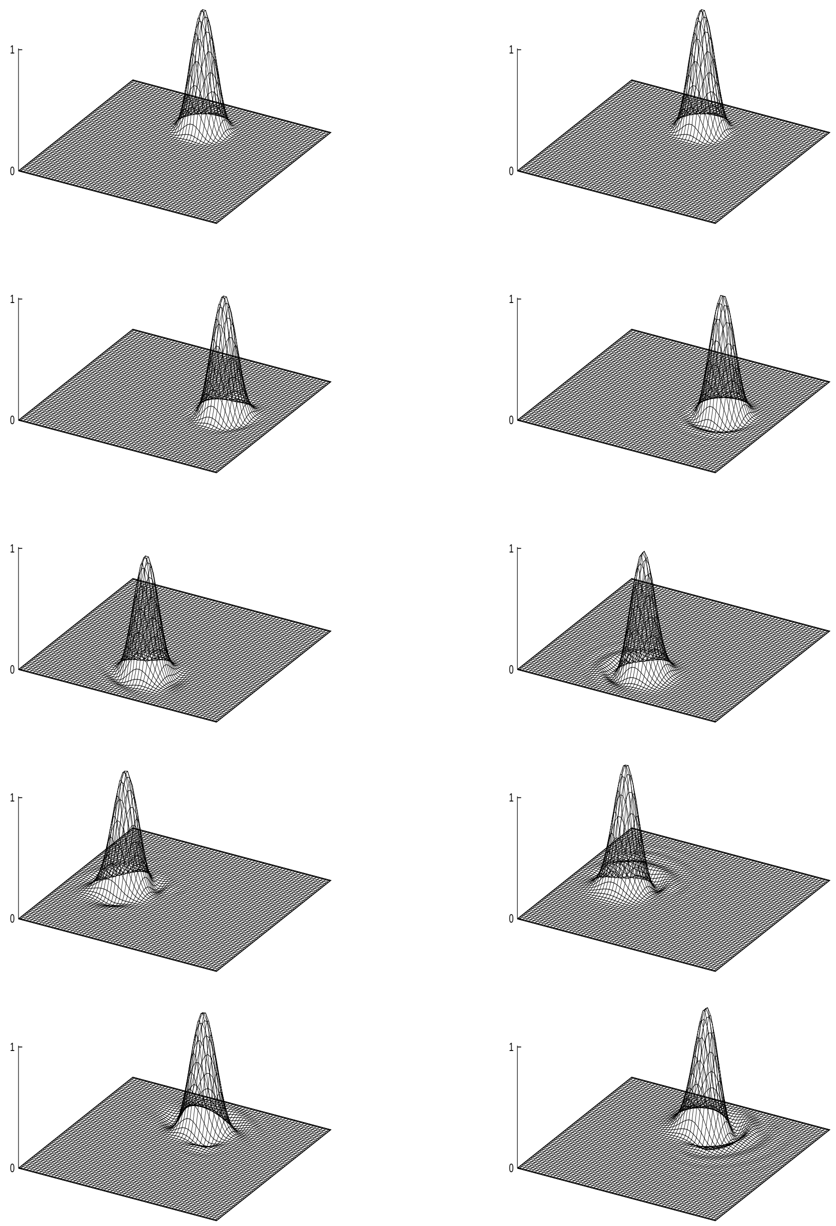


Figure 5: Different time integrators (Gear and Crank-Nicolson method)

riodicity of 2π. Figure 5 shows the results with Gear and Crank-Nicolson method for one rotation at times $t = i\frac{\pi}{2}$, $i = 0, \ldots, 4$ on a 64×64-grid with 200 time steps. The height of the hump decreases about $3\,\%$ for the Gear method which is an effect of its damping behaviour. The shape of the hump is rotated by both methods without significant deformation or translation. The quality is good in comparison with results shown in [9, 10].

	Stabilized Galerkin method		Petrov-Galerkin method		Cell-based semidiscretization		
M	$\|\boldsymbol{\eta}\|_V$	$\|\boldsymbol{\eta}\|_\infty$	$\|\boldsymbol{\eta}\|_V$	$\|\boldsymbol{\eta}\|_\infty$	$\|\boldsymbol{\eta}_1\|_V$	$\|\boldsymbol{\eta}_1\|_\infty$	k_{app}^{SD}
10	$5.07 \cdot 10^{-4}$	$2.45 \cdot 10^{-3}$	$1.92 \cdot 10^{-3}$	$4.26 \cdot 10^{-3}$	$1.03 \cdot 10^{-3}$	$2.04 \cdot 10^{-3}$	—
20	$1.12 \cdot 10^{-4}$	$6.13 \cdot 10^{-4}$	$4.95 \cdot 10^{-4}$	$1.15 \cdot 10^{-3}$	$2.57 \cdot 10^{-4}$	$5.13 \cdot 10^{-4}$	2.00
30	$4.74 \cdot 10^{-5}$	$2.74 \cdot 10^{-4}$	$2.22 \cdot 10^{-4}$	$5.18 \cdot 10^{-4}$	$1.14 \cdot 10^{-4}$	$2.28 \cdot 10^{-4}$	2.00
40	$2.58 \cdot 10^{-5}$	$1.56 \cdot 10^{-4}$	$1.25 \cdot 10^{-4}$	$2.91 \cdot 10^{-4}$	$6.42 \cdot 10^{-5}$	$1.28 \cdot 10^{-4}$	2.00
60	$1.08 \cdot 10^{-5}$	$7.11 \cdot 10^{-5}$	$5.46 \cdot 10^{-5}$	$1.26 \cdot 10^{-4}$	$2.86 \cdot 10^{-5}$	$5.71 \cdot 10^{-5}$	1.99

Table 6: Comparison of the error $\boldsymbol{\eta}$ in the energy and the maximum norm

Table 6 shows results for problem 6.3 on $M \times M$ grids for varying grid parameter $h = M^{-1}$. We conclude that the cell-based semidiscretization yields good results on an equidistant grid in comparison to the results of [4]. Under the assumption $\|\boldsymbol{\eta}_{1,h}\| \approx Ch^{k^{SD}}$ we can estimate the order of convergence by a comparison of two grids with grid parameter h resp. $\tilde{h}$ by

$$k_{app}^{SD} = \log \frac{\|\boldsymbol{\eta}_{1,h}\|}{\|\boldsymbol{\eta}_{1,\tilde{h}}\|} \left[\log \frac{h}{\tilde{h}} \right]^{-1} \tag{27}$$

This example gives a hint that the convergence order of 2 for Dirichlet problems is also realistic for problems with Neumann data.

7 Conclusion

We have been presenting a new finite-volume technique for the discretization of unsteady convection-diffusion problems. This approach is equivalent to a nonconforming Petrov-Galerkin method with exponentially fitted L-splines as trial and L^*-splines test functions which was also described.

Numerical experiments show that the quality of this new discretization approach is good, also in comparison to methods presented in literature. The convergence results for the semidiscrete and the fully discrete systems stated here are given without proofs. These can be found in [7] and will be published in a forthcoming article.

References

[1] K. E. Brenan, S. L. Campbell, L. R. Petzold. *Numerical Solution of Initial-Value Problems in Differential-Algebraic Equations*, Classics In Appl. Math. 14. SIAM, Philadelphia 1996.

[2] G. Bruhn. *Erhaltungssätze und schwache Lösungen in der Gasdynamik.* Math. Methods Appl. Sci. **7** (1985) 740–479.

[3] P. G. Ciarlet. *The Finite Element Method for Elliptic Problems.* North-Holland, Amsterdam 1978.

[4] P. P. N. de Groen, M. van Veldhuizen. *A stabilized Galerkin method for convection-diffusion problems.* SIAM J. Sci. Stat. Comput. **10** (1989) 274–297.

[5] D. F. Griffiths, A. R. Mitchell. *Nonconforming elements.* In D. F. Griffiths, *The mathematical basis of finite element methods*, Oxford Science Publications. Oxford University Press, Oxford 1984.

[6] E. Hairer, C. Lubich, M. Roche. *The Numerical Solution of Differential-Algebraic Systems by Runge-Kutta Methods*, Lect. Notes Math. 1409. Springer, Berlin 1989.

[7] D. Hietel. *Ein neuer Typ zellorientierter Diskretisierungsverfahren zur numerischen Lösung von Konvektions-Diffusions-Problemen.* Ph.D. thesis, Technische Hochschule Darmstadt 1991.

[8] D. Hietel. *Stability and Convergence for Semidiscretization leading to Differential Algebraic Equations.* Z. Angew. Math. Mech. **76** (1996) 421–422.

[9] W. J. Minkowycz, E. M. Sparrow, G. E. Schneider, R. H. Pletcher. *Handbook of Numerical Heat Transfer.* John Wiley & Sons, Chichester 1988.

[10] O. Pironneau. *Finite Element Methods for Fluids.* John Wiley & Sons, Chichester 1989.

[11] H.-G. Roos, M. Stynes, L. Tobiska. *Numerical Methods for Singularly Perturbed Differential Equations*, Springer Ser. Comput. Math. 24. Springer, Berlin 1996.

[12] Y. Saad. *Iterative Methods for Sparse Linear Systems.* PWS Publishing Company, Boston 1996.

[13] J. M. Sanz-Serna, J. G. Verwer. *Stability and convergence at the PDE/stiff ODE interface.* Appl. Numer. Math. **5** (1989) 117–132.

[14] R. M. Smith, A. G. Hutton. *The numerical treatment of advection: A performance comparison of current methods.* Numer. Heat Transfer **5** (1982) 439–461.

[15] P. Sonneveld. *CGS, a fast Lanczos-type solver for nonsymmetric linear systems.* SIAM J. Sci. Stat. Comput. **10** (1989) 36–52.

[16] J. Witzel. *Numerical Solution of Linear Systems Arising from a New Petrov-Galerkin Discretization of Convection-Diffusion Problems* . Shaker Verlag, Aachen 1997. Ph.D. thesis.

[17] J. Witzel. *A numerical study of linear solvers for the solution of convection-diffusion problems.* Preprint, Technische Universität Darmstadt 1997.