Univerzita Karlova v Praze
Matematicko–fyzikální fakulta

DIPLOMOVÁ PRÁCE

Miloslav Vlasák

Numerická analýza nespojité Galerkinovy metody pro řesení konvektivně-difusních rovnic

Katedra numerické matematiky
Vedoucí diplomové práce: Doc. RNDr. Vít Dolejší, Ph.D
Studijní obor: Výpočtová matematika

Charles Univerzity Prague
Faculty of Mathematics and Physics

MASTER THESIS

Miloslav Vlasák
Numerical Analysis of Discontinuous Galerkin Finite Element Method for Convection-Diffusion Equations
Department of Numerical Mathematics
Supervisor: Doc. RNDr. Vít Dolejší, Ph.D
Branch: Computational Mathematics

Děkuji vedoucímu diplomové práce, Doc. RNDr. Vítu Dolejšímu, Ph.D, za zajímavé téma a za trpělivou a účinnou spolupráci při tvorbě diplomové práce. Dále bych chtěl poděkovat svému bratrovi, Vaškovi Vlasákovi a Zuzce Kasarové za korektůru a za pomoc při výpočtech. V neposlední řadě děkuji i svým rodičům za hmotnou i morální podporu po celou dobu studií.

Prohlašuji, že jsem svou diplomovou práci napsal samostatně a výhradně s použitím citovaných pramenů. Souhlasím se zapůjčováním práce.

V Praze dne Miloslav Vlasák

Název práce: Numerická analýza nespojité Galerkinovy metody pro řesení konvektivně-difusních rovnic
Autor: Miloslav Vlasák
Katedra: Katedra numerické matematiky
Vedoucí diplomové práce: Doc. RNDr. Vít Dolejší, Ph.D
e-mail vedoucího: dolejsi@karlin.mff.cuni.cz
Abstrakt: Práce se zabývá numerickým řešením skalární nelineární konvektivně-difusní rovnice. Diskretizace byla provedena pomocí nespojité Galerkin metody (DGFEM) v nesymetrickém tvaru stabilizacních členů a s vnitřní a vnější penalizací. Časová diskretizace je provedena pomocí zpětné diferenční formule (BDF), kde difusní a stabilizační členy jsou vyjádřeny implicitně a nelinearní konvektivní člen explicitně. Odvozujeme apriorní asymptotické odhady chyb v diskrétní $L^\infty(L^2)$–normě a $L^2(H^1)$–seminormě pro BDF řádu $k = 2, 3$ vzhledem ke kroku sítě h a časovému kroku τ. Je též prezentován numerický experiment demonstrující přesnost schématu pro BDF řadu $k = 1, 2, 3$.
Klíčová slova: nelineární konvektivně-difusní rovnice, nespojitá Galerkinova metoda, vnitřní a vnější penalizace, semiimplicitní schéma, apriorní odhady chyby, experimentální řád konvergence, zpětná diferenční formule

Title: Numerical Analysis of Discontinuous Galerkin Finite Element Method for Convection-Diffusion Equations
Author: Miloslav Vlasák
Department: Department of Numerical Mathematics
Supervisor: Doc. RNDr. Vít Dolejší, Ph.D
Supervisor's e-mail address: dolejsi@karlin.mff.cuni.cz
Abstract: The thesis deals with the numerical solution of a scalar nonlinear convection-diffusion equation. The discretization is carried out by a semi-implicit discontinuous Galerkin finite element method (DGFEM) with nonsymmetric treatment of stabilization terms and interior and boundary penalty. The time discretization is carried out by backward differential formulae, diffusive and stabilization terms are treated implicitly whereas the nonlinear convective terms explicitly. We derive a priori asymptotic error estimates in the discrete $L^\infty(L^2)$–norm and $L^2(H^1)$–seminorm for BDF of orders $k = 2, 3$ with respect to the mesh size h and time step τ. A numerical example demonstrating the accuracy of the scheme with BDF of order $k = 1, 2, 3$ is presented.
Keywords: nonlinear convection-diffusion equation, discontinuous Galerkin finite element method, interior and boundary penalty, semi-implicit scheme, a priori error estimates, experimental order of convergence, backward differential formulae

Contents

Chapter 1

Introduction

The investigation of convection-diffusion problems is a very topical subject. These problems play an important role in fluid dynamics, hydrology, heat and mass transfer, environmental protection and other fysical-related problems at the one side and financial mathematics, image processing at the other side.

Our aim is to develop a sufficiently robust and accurate numerical method for the solution of nonlinear convection-diffusion equations. To obtain such a method we use discontinuous Galerkin finite element method (DGFEM) which can be viewed as the mix of ideas of finite volume (FV) and finite element (FE) methods.

Conforming (i.e., continuous) finite element method is suitable for problems with sufficiently regular solutions. However, singularly perturbed problems or nonlinear conservation laws have solutions with steep gradients or discontinuities and their approximations by conforming finite elements may suffer from the Gibbs phenomenon manifested by spurious oscilations.

Discontinuous Galerkin finite element method appears to be very suitable for problems with solutions containing discontinuities or steep gradients. The DGFEM is based on piecewise polynomial discontinuous approximations. It uses adventages of the FV as well as FE methods. Similarly as in the FV method, the DGFEM uses discontinuous approximations and boundary fluxes are evaluated with the aid of a numerical flux, which allow a precise capturing of discontinuities and steep gradients. Similarly as in the FE method, the DGFEM uses higher degree of polynomial approximations of solutions, which produces an accurate resolution in regions, where the solution is smooth. There are a number of works devoted to theory and applications of DGFEM. Let us mention, e.g. [1], [7], [14].

In this thesis we shall be concerned with the analysis of the DGFEM of nonlinear convection-diffusion initial-boundary value problem. There are several variants of the DGFEM. The method can be stabilized with the aid of symmetric or nonsymmetric treatment of diffusion term, combined with an interior and boundary penalty. We consider here the nonsymmetric variant with the interior and boundary penalty (denoted as NIPG method). Nonsymmetric variant was investigated in [7] and [8]. This approach is simpler regarding to choice of penalty coefficient σ, but suffers from the suboptimal estimates of order of convergence for elliptic problems (see [1], [14]).

The time discretization can be carried out by many ways. Runge-Kutta methods, which are very favourite for solving ordinary differential equations, have high order of accuracy and they are simple for implementation. But the resulting scheme is conditionally stable and the time step is drastically limited. In order to avoid this disadventage, it seems suitable to apply an implicit method, which allows to use a much longer time step. However, fully implicit DGFEM leads to a large, strongly nonlinear algebraic system, whose solution is rather complicated. This is the reason that we propose here a semi-implicit scheme, which appears quite efficient and robust. The linear diffusion and stabilization terms are treated implicitly, whereas the nonlinear convective terms explicitly using an extrapolation in such a way that we do not loose the order of accuracy in time.

As a suitable class of implicit methods we have chosen backward differential formulae (BDF), In this thesis we are mainly concerned with theoretical analysis of error estimates of this method. We go out from [6], where the error estimates of this method for the first order in time are proven. We formulate the method generally and prove the error estimates up to the third order in time.

The contents of this paper is the following. In Chapter 2, we formulate the initial-boundary value problem for scalar nonlinear convection-diffusion equation in the classical meaning. Then we reformulate this problem in the weak sence and formulate adequate regularity conditions for proving the error estimates. In Chapter 3, we carry out the discretization of the problem by DGFEM and achieve the semi-discretized problem. Then we derive some properties of BDF methods, which will be needed in the next sections. At the end of this chapter we formulate fully-discrete solution of our problem and prove the existence and uniqueness of this numerical solution. Chapter 4 contains some auxiliary results, namely assumptions on the space discretization (allowing even nonconforming grids with nonconvex star-shaped elements) and some important inequalities and estimates. All these results are used in Chapter 5, where error estimates in the discrete $L^\infty(L^2)$–norm and $L^2(H^1)$–seminorm are proven. In Chapter 6 we present a numerical example to demonstrate and verify the proven errors numericaly. In Chapter 7 we introduce some concluding remarks and formulate open problems.

Chapter 2

Continuous problem

Let $\Omega \subset I\!R^d$ ($d = 2$ or 3) be a bounded polyhedral domain and $T > 0$. (For $d = 2$ under the concept of a polyhedral domain we mean a polygonal domain.) We set $Q_T = \Omega \times (0, T)$. By $\overline{\Omega}$ and $\partial\Omega$ we denote the closure and boundary of Ω, respectively. Let us consider the following *initial-boundary value problem*: Find $u : Q_T \to I\!R$ such that

$$\frac{\partial u}{\partial t} + \sum_{s=1}^{d} \frac{\partial f_s(u)}{\partial x_s} = \varepsilon\, \Delta u + g \quad \text{in } Q_T, \tag{2.0.1}$$

$$u\Big|_{\partial\Omega \times (0,T)} = u_D, \tag{2.0.2}$$

$$u(x, 0) = u^0(x), \quad x \in \Omega. \tag{2.0.3}$$

We assume that the data satisfy the following conditions:

$$\begin{aligned}
&\text{a)} \quad f_s \in C^1(I\!R),\ f_s(0) = 0,\ s = 1, \ldots, d, \\
&\text{b)} \quad \varepsilon > 0, \\
&\text{c)} \quad g \in C([0, T]; L^2(\Omega)), \\
&\text{d)} \quad u_D \text{ is the trace of some } u^* \in C([0, T]; H^1(\Omega)) \cap L^\infty(Q_T) \\
&\qquad \text{on } \partial\Omega \times (0, T), \\
&\text{e)} \quad u^0 \in L^2(\Omega).
\end{aligned} \tag{2.0.4}$$

We use the standard notation for function spaces (see, e. g. [11]): $L^p(\Omega)$, $L^p(Q_T)$ denote the Lebesgue spaces, $W^{k,p}(\Omega)$, $H^k(\Omega) = W^{k,2}(\Omega)$ are the Sobolev spaces, $L^p(0, T; X)$ is the Bochner space of functions p-integrable over the interval $(0, T)$ with values in a Banach space X, $C([0, T]; X)$ ($C^1([0, T]; X)$) is the space of continuous (continuously differentiable) mappings of the interval $[0, T]$ into X. By $H^1_0(\Omega)$ we denote the subspace of all functions from $H^1(\Omega)$ with zero traces on $\partial\Omega$.

The assumption that $f_s(0) = 0$, $s = 1, \ldots, d$, does not cause any loss of generality, as can be seen from equation (2.0.1). The functions f_s, called fluxes, represent convective terms, $\varepsilon > 0$ is the diffusion coefficient. The diffusion term can be, in general, more complicated (in some cases even nonlinear as in [14]). It is also possible to consider mixed Dirichlet–Neumann boundary conditions. For simplicity we prescribe the Dirichlet condition on the whole boundary.

A sufficiently regular function satisfying (2.0.1) – (2.0.3) pointwise is called a *classical solution.* It is suitable to introduce the concept of a weak solution. To this end we use the following *notation*:

$$(u, w) = \int_\Omega uw \,\mathrm{d}x, \quad u,\, w \in L^2(\Omega) \tag{2.0.5}$$

(scalar product in $L^2(\Omega)$),

$$\|u\|_{L^2(\Omega)} = (u, u)^{1/2} \tag{2.0.6}$$

(norm in $L^2(\Omega)$),

$$a(u, w) = \varepsilon \int_\Omega \nabla u \cdot \nabla w \,\mathrm{d}x, \quad u,\, w \in H^1(\Omega), \tag{2.0.7}$$

$$b(u, w) = \int_\Omega \sum_{s=1}^{d} \frac{\partial f_s(u)}{\partial x_s} w \,\mathrm{d}x, \quad u \in H^1(\Omega) \cap L^\infty(\Omega),\ w \in L^2(\Omega), \tag{2.0.8}$$

$$\|u\|_{H^1(\Omega)} = \left(\int_\Omega \left(|u|^2 + |\nabla u|^2 \right) \mathrm{d}x \right)^{1/2}, \quad u \in H^1(\Omega), \tag{2.0.9}$$

(norm in $H^1(\Omega)$),

$$|u|_{H^1(\Omega)} = \left(\int_\Omega |\nabla u|^2 \,\mathrm{d}x \right)^{1/2}, \quad u \in H^1(\Omega), \tag{2.0.10}$$

(seminorm in $H^1(\Omega)$). It is known that $|\cdot|_{H^1(\Omega)}$ is a norm on $H_0^1(\Omega)$ equivalent to $\|\cdot\|_{H^1(\Omega)}$.

Definition 1 *We say that a function u is a* weak solution *of (2.0.1) – (2.0.3), if the following conditions are satisfied*

$$\begin{aligned} &a) \quad u - u^* \in L^2(0, T; H_0^1(\Omega)), \quad u \in L^\infty(Q_T), \\ &b) \quad \frac{\mathrm{d}}{\mathrm{d}t}(u(t), w) + b(u(t), w) + a(u(t), w) = (g(t), w) \\ &\qquad \textit{for all } w \in H_0^1(\Omega) \textit{ in the sense of distributions on } (0, T), \\ &c) \quad u(0) = u_0 \quad \textit{in } \Omega. \end{aligned} \tag{2.0.11}$$

By $u(t)$ we denote the function on Ω such that $u(t)(x) = u(x, t)$, $x \in \Omega$.

With the aid of techniques from [12] and [13], it is possible to prove that there exists a unique weak solution. Moreover, it satisfies the condition $\partial u/\partial t \in L^2(Q_T)$. Then (2.0.11), b) can be rewritten as

$$\left(\frac{\partial u(t)}{\partial t}, w \right) + b(u(t), w) + a(u(t), w) = (g(t), w) \tag{2.0.12}$$

for all $w \in H_0^1(\Omega)$ and almost every $t \in (0, T)$.

We shall assume that the weak solution u is sufficiently regular, namely,

$$\begin{aligned} &u \in W^{k,\infty}(0, T; H^{p+1}(\Omega)) \\ &u^{(k+1)} \in L^\infty(0, T; L^2(\Omega)). \end{aligned} \tag{2.0.13}$$

where $u^{(s)} = \frac{\partial^s u}{\partial t^s}$, an integer $p \geq 1$ will denote a given degree of polynomial approximations and an integer k is a desired order of time aproximation. Such a solution satisfies problem (2.0.1) – (2.0.3) pointwise.

Chapter 3

Discretization of the problem

3.1 Triangulations

Let $\mathcal{T}_h$ $(h > 0)$ be a partition of the closure $\overline{\Omega}$ of the domain Ω into a finite number of closed d-dimensional (convex or nonconvex) polyhedra K with mutually disjoint interiors. We call $\mathcal{T}_h$ a triangulation of $\overline{\Omega}$ and *do not* require the conforming properties from the finite element method. In 2D problems we choose usually $K \in \mathcal{T}_h$ as triangles or quadrilaterals. In 3D, $K \in \mathcal{T}_h$ can be, e. g., tetrahedra, pyramids or hexahedra, but we can construct even more general elements K, as dual finite volumes from [10]. In the analysis carried out in Chapter 4 it is important to assume that all elements $K \in \mathcal{T}_h$ are star-shaped.

In our further considerations we shall use the following *notation*. We set $h_K = \mathrm{diam}(K)$, $h = \max_{K\in\mathcal{T}_h} h_K$. By ρ_K we denote the radius of the largest d-dimensional ball inscribed into K and by $|K|$ we denote the d-dimensional Lebesgue measure of K. All elements of $\mathcal{T}_h$ will be numbered so that $\mathcal{T}_h = \{K_i\}_{i\in I}$, where $I \subset Z^+ = \{0, 1, 2, \ldots\}$ is a suitable index set. If two elements K_i, $K_j \in \mathcal{T}_h$, $i \neq j$, contain a nonempty common open face, we call them *neighbours*. We set in this case

$$\Gamma_{ij} = \partial K_i \cap \partial K_j. \tag{3.1.1}$$

(See Figure 3.1, showing a possible 2D situation.) For $i \in I$ we set

$$s(i) = \{j \in I; K_j \text{ is a neighbour of } K_i\}. \tag{3.1.2}$$

The boundary $\partial\Omega$ is formed by a finite number of faces of elements K_i adjacent to $\partial\Omega$. We denote all these boundary faces by S_j, where $j \in I_b \subset Z^- = \{-1, -2, \ldots\}$ and set

$$\begin{aligned} \gamma(i) &= \{j \in I_b; S_j \text{ is a face of } K_i\}, \\ \Gamma_{ij} &= S_j \text{ for } K_i \in \mathcal{T}_h \text{ such that } S_j \subset \partial K_i,\ j \in I_b. \end{aligned} \tag{3.1.3}$$

For K_i not containing any boundary face S_j we set $\gamma(i) = \emptyset$. Obviously, $s(i) \cap \gamma(i) = \emptyset$ for all $i \in I$. Now, if we write $S(i) = s(i) \cup \gamma(i)$, we have

$$\partial K_i = \bigcup_{j\in S(i)} \Gamma_{ij}, \quad \partial K_i \cap \partial\Omega = \bigcup_{j\in\gamma(i)} \Gamma_{ij}. \tag{3.1.4}$$

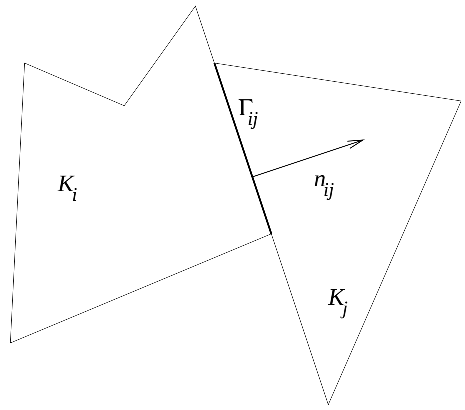

Figure 3.1: Neighbouring elements $K_i,\ K_j$

Furthermore, we use the following notation: $\boldsymbol{n}_{ij} = ((n_{ij})_1, \dots, (n_{ij})_d)$ – unit outer normal to ∂K_i on the face Γ_{ij} (see Figure 3.1), $|\Gamma_{ij}|$ – $(d-1)$-dimensional measure of Γ_{ij} (i. e., the length of Γ_{ij}, if $d = 2$, and the area of Γ_{ij}, if $d = 3$), $d(\Gamma_{ij}) = \operatorname{diam}(\Gamma_{ij})$. Obviously, $\boldsymbol{n}_{ij} = -\boldsymbol{n}_{ji}$ and, from the point of view of point sets, we have $\Gamma_{ij} = \Gamma_{ji}$, if $i \in I,\ j \in s(i)$ (and, hence, $j \in I,\ i \in s(j)$).

It is obvious that for $i \in I$ and $j \in S(i)$ we have

$$|K_i| \le h_{K_i}^d \le h^d, \tag{3.1.5}$$

$$d(\Gamma_{ij}) \le h_{K_i} \le h. \tag{3.1.6}$$

3.2 Broken Sobolev spaces

Over the triangulation $\mathcal{T}_h$ we define the so-called *broken Sobolev space*

$$H^z(\Omega, \mathcal{T}_h) = \{w; w|_K \in H^z(K)\ \forall\, K \in \mathcal{T}_h\} \tag{3.2.7}$$

and define there the norm

$$\|w\|_{H^z(\Omega,\mathcal{T}_h)} = \left(\sum_{K\in\mathcal{T}_h} \|w\|^2_{H^z(K)} \right)^{1/2} \tag{3.2.8}$$

and the seminorm

$$|w|_{H^z(\Omega,\mathcal{T}_h)} = \left(\sum_{K\in\mathcal{T}_h} |w|^2_{H^z(K)} \right)^{1/2}. \tag{3.2.9}$$

For $w \in H^1(\Omega, \mathcal{T}_h)$, $i \in I$ and $j \in s(i)$ we introduce the following notation:

$$w|_{\Gamma_{ij}} \;=\; \text{the trace of } w|_{K_i} \text{ on } \Gamma_{ij}, \tag{3.2.10}$$

$$\begin{aligned} w|_{\Gamma_{ji}} &= \text{the trace of } w|_{K_j} \text{ on } \Gamma_{ji}, \\ \langle w\rangle_{\Gamma_{ij}} &= \frac{1}{2}\left(w|_{\Gamma_{ij}} + w|_{\Gamma_{ji}}\right), \\ [w]_{\Gamma_{ij}} &= w|_{\Gamma_{ij}} - w|_{\Gamma_{ji}}. \end{aligned}$$

Obviously, $\langle w\rangle_{\Gamma_{ij}} = \langle w\rangle_{\Gamma_{ji}}$, but $[w]_{\Gamma_{ij}} = -[w]_{\Gamma_{ji}}$. On the other hand, $[w]_{\Gamma_{ij}}\boldsymbol{n}_{ij} = [w]_{\Gamma_{ji}}\boldsymbol{n}_{ji}$.

3.3 Space discretization

In order to carry out the space discretization, we start from the weak solution u satisfying (2.0.13), multiply equation (2.0.1) by an arbitrary $w \in H^2(\Omega, \mathcal{T}_h)$, integrate over each $K_i \in \mathcal{T}_h$ and apply Green's theorem. Summing over all $K_i \in \mathcal{T}_h$, using (3.1.4), (2.0.2),

$$[u]_{\Gamma_{ij}} = 0, \quad \langle \nabla u\rangle_{\Gamma_{ij}} = \nabla u|_{\Gamma_{ij}} = \nabla u|_{\Gamma_{ji}} \tag{3.3.11}$$

and

$$\int_{\Gamma_{ij}} \langle \nabla w\rangle \cdot \boldsymbol{n}_{ij}\, [u]\, \mathrm{d}S = 0, \ \int_{\Gamma_{ij}} [u][w]\, \mathrm{d}S = 0 \quad \forall j \in s(i)\ \forall i \in I, \tag{3.3.12}$$

valid for the solution u satisfying (2.0.13), we obtain the identity

$$\begin{aligned} \left(\frac{\partial u}{\partial t}(t), w\right) \ + \ & \sum_{i\in I} \Bigg\{ \sum_{j\in S(i)} \int_{\Gamma_{ij}} \sum_{s=1}^{d} f_s(u(t))\, (n_{ij})_s\, w|_{\Gamma_{ij}}\, \mathrm{d}S \\ & - \int_{K_i} \sum_{s=1}^{d} f_s(u(t))\, \frac{\partial w}{\partial x_s}\, \mathrm{d}x + \varepsilon \int_{K_i} \nabla u(t) \cdot \nabla w\, \mathrm{d}x \\ & - \varepsilon \sum_{\substack{j\in s(i) \\ j<i}} \int_{\Gamma_{ij}} \Big(\langle \nabla u(t)\rangle \cdot \boldsymbol{n}_{ij}\, [w] - \langle \nabla w\rangle \cdot \boldsymbol{n}_{ij}\, [u(t)] \Big)\, \mathrm{d}S \\ & - \varepsilon \sum_{j\in\gamma(i)} \int_{\Gamma_{ij}} \Big(\nabla u(t) \cdot \boldsymbol{n}_{ij}\, w - \nabla w \cdot \boldsymbol{n}_{ij}\, u \Big)\, \mathrm{d}S \\ & + \varepsilon \sum_{j\in s(i)} \int_{\Gamma_{ij}} \sigma [u]\, [w]\, \mathrm{d}S + \sum_{j\in\gamma(i)} \int_{\Gamma_{ij}} \sigma\, u\, w\, \mathrm{d}S \Bigg\} \\ = \ & \int_{\Omega} g(t)\, w\, \mathrm{d}x + \varepsilon \sum_{i\in I} \sum_{j\in\gamma(i)} \int_{\Gamma_{ij}} \Big(\nabla w \cdot \boldsymbol{n}_{ij}\, u_D + \sigma\, u_D\, w \Big)\, \mathrm{d}S, \end{aligned} \tag{3.3.13}$$

where the weight function $\sigma : \bigcup_{i\in I} \bigcup_{j\in S(i)} \Gamma_{ij} \to I\!R$ is defined by

$$\sigma|_{\Gamma_{ij}} = \frac{1}{d(\Gamma_{ij})}. \tag{3.3.14}$$

We see that the terms

$$\varepsilon \int_{\Gamma_{ij}} \left(\nabla w \cdot \boldsymbol{n}_{ij}\, u(t) + \sigma\, u(t)\, w \right)\, \mathrm{d}S$$

and

$$\varepsilon \int_{\Gamma_{ij}} \left(\nabla w \cdot \boldsymbol{n}_{ij}\, u_D(t) + \sigma\, u_D(t)\, w\right) \,\mathrm{d}S,$$

$i \in I,\ j \in \gamma(i)$, on the left-hand side and the right-hand side of (3.3.13), respectively, are equal. Due to the boundary condition (2.0.2), these terms are equal. Above and in the sequel, in the integrals $\int_{\Gamma_{ij}}$, under $\langle\cdot\rangle$ and $[\cdot]$ we understand $\langle\cdot\rangle_{\Gamma_{ij}}$ and $[\cdot]_{\Gamma_{ij}}$, respectively.

Let us note that the form (3.3.13) represents a *nonsymmetric variant* of DG approximation of the diffusion terms.

The above considerations lead us to the definition of the following forms. For $u, w \in H^2(\Omega, \mathcal{T}_h)$ we set

$$\begin{aligned} a_h(u,w) &= \varepsilon \sum_{i\in I} \Bigg\{ \int_{K_i} \nabla u \cdot \nabla w \,\mathrm{d}x \\ &\quad - \sum_{\substack{j\in s(i)\\ j<i}} \int_{\Gamma_{ij}} \Big(\langle \nabla u\rangle \cdot \boldsymbol{n}_{ij} [w] - \langle \nabla w \rangle \cdot \boldsymbol{n}_{ij} [u] \Big) \,\mathrm{d}S \\ &\quad - \sum_{j\in\gamma(i)} \int_{\Gamma_{ij}} \Big((\nabla u \cdot \boldsymbol{n}_{ij})\, w - (\nabla w \cdot \boldsymbol{n}_{ij})\, u \Big) \,\mathrm{d}S \Bigg\}, \end{aligned} \tag{3.3.15}$$

$$\tilde{b}_h(u,w) = \sum_{i\in I} \left\{ \sum_{j\in S(i)} \int_{\Gamma_{ij}} \sum_{s=1}^{d} f_s(u)\, (n_{ij})_s\, w|_{\Gamma_{ij}} \,\mathrm{d}S - \int_{K_i} \sum_{s=1}^{d} f_s(u)\, \frac{\partial w}{\partial x_s} \,\mathrm{d}x \right\}, \tag{3.3.16}$$

$$\ell_h(w)\,(t) = (g(t), w) + \varepsilon \sum_{i\in I} \sum_{j\in\gamma(i)} \int_{\Gamma_{ij}} \left(\nabla w \cdot \boldsymbol{n}_{ij}\, u_D(t) + \sigma u_D(t)\, w\right) \,\mathrm{d}S, \tag{3.3.17}$$

$$J_h^\sigma(u,w) = \sum_{i\in I} \left\{ \sum_{j\in s(i)} \int_{\Gamma_{ij}} \sigma [u]\, [w] \,\mathrm{d}S + \sum_{j\in\gamma(i)} \int_{\Gamma_{ij}} \sigma u w \,\mathrm{d}S \right\}. \tag{3.3.18}$$

In view of (3.3.13), the exact solution with properties (2.0.13) satisfies the identity

$$\left(\frac{\partial u}{\partial t}(t), w\right) + a_h(u(t),w) + \tilde{b}_h(u(t),w) + \varepsilon J_h^\sigma(u(t),w) = \ell_h(w)\,(t) \quad \forall w \in H^2(\Omega, \mathcal{T}_h)\ \forall t \in (0,T). \tag{3.3.19}$$

Now we shall approximate the form $\tilde{b}_h$ in such a way that similarly as in the finite volume method, for $i \in I$, $j \in S(i)$, the fluxes $\int_{\Gamma_{ij}} \sum_{s=1}^{d} f_s(u)\, n_s\, w \,\mathrm{d}S$ are approximated with the aid of the so-called *numerical flux* $H(u, u', \boldsymbol{n})$ by the expression $\int_{\Gamma_{ij}} H(u|_{\Gamma_{ij}}, u|_{\Gamma_{ji}}, \boldsymbol{n}_{ij})\, w|_{\Gamma_{ij}} \,\mathrm{d}S$. Of course, if $j \in \gamma(i)$, then $\Gamma_{ij} \subset \partial\Omega$ and it is necessary to specify the meaning of $u|_{\Gamma_{ji}}$. Here we use the extrapolation, i.e. we set $u|_{\Gamma_{ji}} := u|_{\Gamma_{ij}}$. In this way we obtain the approximation $b_h(u,w)$ of the convection form $\tilde{b}_h(u,w)$:

$$b_h(u,w) = \sum_{i\in I} \left(\sum_{j\in s(i)} \int_{\Gamma_{ij}} H(u|_{\Gamma_{ij}}, u|_{\Gamma_{ji}}, \boldsymbol{n}_{ij})\, w|_{\Gamma_{ij}} \,\mathrm{d}S \right. \tag{3.3.20}$$

$$+\sum_{j\in\gamma(i)}\int_{\Gamma_{ij}} H(u|_{\Gamma_{ij}}, u|_{\Gamma_{ij}}, \boldsymbol{n}_{ij})\, w|_{\Gamma_{ij}}\, \mathrm{d}S\Bigg)$$

$$-\sum_{i\in I}\int_{K_i}\sum_{s=1}^{d} f_s(u)\,\frac{\partial w}{\partial x_s}\,\mathrm{d}x, \quad u,\, w \in H^1(\Omega, \mathcal{T}_h), \quad u \in L^\infty(\Omega).$$

We shall assume that the numerical flux has the following properties:

Assumptions (H):

1. $H(u, w, \boldsymbol{n})$ is defined in $I\!R^d \times \boldsymbol{S}_1$, where $\boldsymbol{S}_1 = \{\boldsymbol{n} \in I\!R^d; |\boldsymbol{n}| = 1\}$, and *Lipschitz-continuous* with respect to u, w: there exists a constant $C_1 > 0$ such that

$$|H(u, w, \boldsymbol{n}) - H(u^*, w^*, \boldsymbol{n})| \le C_1(|u - u^*| + |w - w^*|), \qquad (3.3.21)$$
$$u,\, w,\, u^*,\, w^* \in I\!R,\ \boldsymbol{n} \in \boldsymbol{S}_1.$$

2. $H(u, w, \boldsymbol{n})$ is *consistent*:

$$H(u, u, \boldsymbol{n}) = \sum_{s=1}^{d} f_s(u)\, n_s, \quad u \in I\!R,\ \boldsymbol{n} = (n_1, \ldots, n_d) \in \boldsymbol{S}_1. \qquad (3.3.22)$$

3. $H(u, w, \boldsymbol{n})$ is *conservative*:

$$H(u, w, \boldsymbol{n}) = -H(w, u, -\boldsymbol{n}), \quad u,\, w \in I\!R,\ \boldsymbol{n} \in \boldsymbol{S}_1. \qquad (3.3.23)$$

In virtue of (3.3.21) and (3.3.22), the functions f_s, $s = 1, \ldots, d$, are Lipschitz-continuous with constant $L_f = 2C_1$.

From (2.0.4), a) and (3.3.22) we see that

$$H(0, 0, \boldsymbol{n}) = 0 \quad \forall\, \boldsymbol{n} \in \boldsymbol{S}_1. \qquad (3.3.24)$$

By (3.3.16), (3.3.20) and (3.3.22),

$$b_h(u, w) = \tilde{b}_h(u, w) \quad \forall u \in H^1(\Omega) \cap L^\infty(\Omega),\ \forall w \in H^1(\Omega, \mathcal{T}_h). \qquad (3.3.25)$$

Now we define the space of discontinuous piecewise polynomial functions

$$S_h = S^{p,-1}(\Omega, \mathcal{T}_h) = \{w; w|_K \in P_p(K)\ \forall\, K \in \mathcal{T}_h\}, \qquad (3.3.26)$$

where $P_p(K)$ denotes the space of all polynomials on K of degree $\le p$, where the integer $p \ge 1$ is a given degree of approximation.

Using (3.3.19) and (3.3.25), we find that the exact solution u with property (2.0.13) satisfies the identity

$$\left(\frac{\partial u}{\partial t}(t), w_h\right) + a_h(u(t), w_h) + b_h(u(t), w_h) + \varepsilon J_h^\sigma(u(t), w_h) = \ell_h(w_h)\,(t) \qquad (3.3.27)$$

for all $w_h \in S_h$ and all $t \in (0, T)$.

3.4 Linear multistep methods

Let us assume an ordinary differential equation

$$\begin{aligned} y'(t) &= F(t, y(t)) \\ y(a) &= \mu, \end{aligned} \tag{3.4.28}$$

where $F : [a, b] \times R \to R$ is a continuous function. Let us also assume that this equation has an unique solution. Let $t_n = a + n\tau$ be a partition of $[a, b]$ with the step $\tau > 0$. Then we define general linear multistep method by formula

$$\sum_{v=0}^{k} \alpha_v y_{n+v} = \sum_{v=0}^{k} \beta_v F_{n+v}, \tag{3.4.29}$$

where y_m is the approximate value of the solution $y(t_m)$ and $F_m = F(t_m, y_m)$, k is integer and α_v and β_v are real constants. We suppose that

$$\begin{aligned} \alpha_k &\neq 0 \\ |\alpha_0| + |\beta_0| &> 0. \end{aligned} \tag{3.4.30}$$

Then we call method (3.4.29) linear k step method.
If $\beta_k = 1$ and $\beta_v = 0$ for $v = 0, \ldots, k-1$ we speak about backward differential formula.
Using (3.4.29) to solve (3.4.28) we compute y_{n+k} from the known values $y_n, \ldots, y_{n+k-1}$. We can see that at the beginnig we need to know $y_0, \ldots, y_{k-1}$. This values are called initial conditions. Usually we set $y_0 = \mu$ and the other values $y_1, \ldots, y_{k-1}$ are given by some suitable one step method which is started with $y_0 = \mu$.
We say that a method (3.4.29) is an explicit method when $\beta_k = 0$ and we say that a method (3.4.29) is an implicit method when $\beta_k \neq 0$. When we use some implicit method we cannot express y_{n+k} by simple arithmetic means, because the value y_{n+k} takes part in term $\beta_k F_{n+k}$.
For every multistep method (3.4.29) we consider characteristic polynomial

$$\rho(t) = \sum_{v=0}^{k} \alpha_v t^v. \tag{3.4.31}$$

Definition 2 *The linear multistep method (3.4.29) is called D-stable if no root of the characteristic polynomial $\rho(t)$ has modulus greater than one and every root of modulus one has multiplicity one.*

Definition 3 *We say that a method (3.4.29) has order $p \geq 0$, if*

$$\sum_{v=0}^{k} \alpha_v = 0 \tag{3.4.32}$$

$$\sum_{v=0}^{k} v^s \alpha_v = s \sum_{v=0}^{k} v^{s-1} \beta_v \tag{3.4.33}$$

holds for $s = 1, \ldots, p$.
We call a method (3.4.29) consistent, if its order is $p \geq 1$

Remark 1 *In the case of the Backward Differential formulae (BDF) we can write both conditions (3.4.32) and (3.4.33) as*

$$\sum_{v=0}^{k} \alpha_v v^s = s k^{s-1} \tag{3.4.34}$$

for $s = 0, \ldots, p$.

Lemma 1 *Let us assume (3.4.34) with* $k = p$. *Then (3.4.34) has unique solution.*

Proof. We will show that the matrix represented by (3.4.34) is regular. Because the matrix is of Vandermonde type we can compute determinant:

$$det \begin{pmatrix} 1 & 1 & 1 & 1 & \cdots & 1 \\ 0 & 1^1 & 2^1 & 3^1 & \cdots & k^1 \\ 0 & 1^2 & 2^2 & 3^2 & \cdots & k^2 \\ 0 & 1^3 & 2^3 & 3^3 & \cdots & k^3 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 1^k & 2^k & 3^k & \cdots & k^k \end{pmatrix} = \prod_{0 \le i < j \le k} (j - i) > 0 \tag{3.4.35}$$

□

Remark 2 *The coefficients of the BDF of the order* 2 *are*

$$\alpha_2 = \frac{3}{2}, \alpha_1 = -2, \alpha_0 = \frac{1}{2} \tag{3.4.36}$$

and the roots of characteristic polynomial ρ *are* 1 *and* $\frac{1}{3}$. *From this follows that BDF of the order* 2 *is D-stable.*
The coefficients of the BDF of the order 3 *are*

$$\alpha_3 = \frac{11}{6}, \alpha_2 = -3, \alpha_1 = \frac{3}{2}, \alpha_0 = -\frac{1}{3} \tag{3.4.37}$$

and the roots of characteristic polynomial ρ *are* 1, $\frac{7+i\sqrt{39}}{22}$ *and* $\frac{7-i\sqrt{39}}{22}$. *From this follows that BDF of the order* 3 *is D-stable.*
By a symbolic rather complicated calculation it is possible to show that BDF methods are D-stable up to order 6.

Lemma 2 *The system of equations (3.4.34) is equivalent to the system:*

$$\sum_{v=0}^{k} \alpha_v = 0 \tag{3.4.38}$$

$$\sum_{v=0}^{k} \alpha_v (k - v) = -1 \tag{3.4.39}$$

$$\sum_{v=0}^{k} \alpha_v (k - v)^s = 0 \tag{3.4.40}$$

for $s = 2, \ldots, k$.

Proof. We can see that matrix represented by the system (3.4.38)–(3.4.40) is regular, because similarly as in the previous case the matrix is of Vandermonde type. At first we will prove that (3.4.39) and (3.4.40) follows from (3.4.34). First equation (3.4.38) is one of the equations of (3.4.34). Second equation (3.4.39) we can divide into two parts and enumerate them by (3.4.32) and (3.4.33):

$$\sum_{v=0}^{k} \alpha_v (k-v) = k \sum_{v=0}^{k} \alpha_v - \sum_{v=0}^{k} \alpha_v v = 0 - 1 = -1. \tag{3.4.41}$$

The remaining equations (3.4.40) we will prove using (3.4.34).

$$\begin{aligned} \sum_{v=0}^{k} \alpha_v (k-v)^s &= \sum_{v=0}^{k} \alpha_v \sum_{i=0}^{s} \binom{s}{i} (-1)^i k^{s-i} v^i = \sum_{i=0}^{s} (-1)^i \binom{s}{i} k^{s-i} \sum_{v=0}^{k} \alpha_v v^i \\ &= \sum_{i=0}^{s} (-1)^i \binom{s}{i} k^{s-i} i k^{i-1} = k^{s-1} \sum_{i=0}^{s} (-1)^i \frac{s!}{i!(s-i)!} i \\ = k^{s-1} \sum_{i=1}^{s} (-1)^i \frac{s!}{i!(s-i)!} i &= -k^{s-1} \sum_{i=1}^{s} (-1)^{i-1} \frac{s(s-1)!}{(i-1)!(s-1-(i-1))!} \\ &= -sk^{s-1} \sum_{i=0}^{s-1} \binom{s-1}{i} (-1)^i = -sk^{s-1}(1-1)^{s-1} = 0 \end{aligned} \tag{3.4.42}$$

for $s \geq 2$.
Now we will prove that (3.4.34) follows from (3.4.38), (3.4.39) and (3.4.40). The equation (3.4.32) is exactly (3.4.38). The equations (3.4.33) we will prove by induction. As first step we will prove that (3.4.32) holds for $s = 1$.

$$\sum_{v=0}^{k} \alpha_v (k-v) = k \sum_{v=0}^{k} \alpha_v - \sum_{v=0}^{k} \alpha_v v = - \sum_{v=0}^{k} \alpha_v v = -1 \tag{3.4.43}$$

From that follows

$$\sum_{v=0}^{k} \alpha_v v^s = \sum_{v=0}^{k} \alpha_v v = 1 = sk^{s-1} \tag{3.4.44}$$

for $s = 1$. Then let us assume that (3.4.33) holds for $i = 1, \ldots, s-1$. Then

$$\begin{aligned} 0 = \sum_{v=0}^{k} \alpha_v (k-v)^s &= \sum_{v=0}^{k} \alpha_v \sum_{i=0}^{s} (-1)^i \binom{s}{i} k^{s-i} v^i = \sum_{i=0}^{s} (-1)^i \binom{s}{i} k^{s-i} \sum_{v=0}^{k} \alpha_v v^i \\ &= (-1)^s \sum_{v=0}^{k} \alpha_v v^s + \sum_{i=0}^{s-1} (-1)^i \binom{s}{i} k^{s-i} \sum_{v=0}^{k} \alpha_v v^i \end{aligned} \tag{3.4.45}$$

With the induction assumptions we will get from the second term:

$$\begin{aligned} \sum_{i=0}^{s-1} (-1)^i \binom{s}{i} k^{s-i} \sum_{v=0}^{k} \alpha_v v^i &= \sum_{i=0}^{s-1} (-1)^i \binom{s}{i} k^{s-i} i k^{i-1} \\ = k^{s-1} \sum_{i=0}^{s-1} (-1)^i \binom{s}{i} i &= -(-1)^s s k^{s-1} + k^{s-1} \sum_{i=0}^{s} (-1)^i \binom{s}{i} i \end{aligned} \tag{3.4.46}$$

Now it is sufficient to show that

$$\sum_{i=0}^{s}(-1)^i \begin{pmatrix} s \\ i \end{pmatrix} i = \sum_{i=1}^{s}(-1)^i \begin{pmatrix} s \\ i \end{pmatrix} i = \sum_{i=1}^{s}(-1)^i \frac{s!}{i!(s-i)!} i \tag{3.4.47}$$

$$= -s\sum_{i=1}^{s}(-1)^{i-1}\frac{(s-1)!}{(i-1)!(s-1-(i-1))!} = -s\sum_{i=1}^{s}(-1)^{i-1}\begin{pmatrix} s-1 \\ i-1 \end{pmatrix}$$

$$= -s\sum_{i=0}^{s-1}(-1)^i \begin{pmatrix} s-1 \\ i \end{pmatrix} = -s(1-1)^{s-1} = 0$$

for $s \geq 2$. From (3.4.45), (3.4.46) and (3.4.47) follows (3.4.33). □

Lemma 3 *Let (3.4.29) is BDF of order $k \geq 2$. Then*

$$\alpha_v = (-1)^{k-v} \begin{pmatrix} k \\ v \end{pmatrix} \frac{1}{k-v} \tag{3.4.48}$$

for $v = 0, \ldots, k-1$.

Proof. Because we have shown in Lemma 2 that system of equations (3.4.34) is equivalent to system (3.4.38), (3.4.39) and (3.4.40) and since α_k depends only on (3.4.38) we can prove our lemma by substituting to the (3.4.39) and (3.4.40).When we substitute to the (3.4.39) we get

$$\sum_{v=0}^{k}\alpha_v(k-v) = \sum_{v=0}^{k-1}\alpha_v(k-v) = \sum_{v=0}^{k-1}(-1)^{k-v}\begin{pmatrix} k \\ v \end{pmatrix}\frac{1}{k-v}(k-v) \tag{3.4.49}$$

$$= \sum_{v=0}^{k-1}(-1)^{k-v}\begin{pmatrix} k \\ v \end{pmatrix} = \sum_{v=0}^{k}(-1)^{k-v}\begin{pmatrix} k \\ v \end{pmatrix} - 1 = (1-1)^k - 1 = -1,$$

for $k \geq 1$. Now we will prove the rest by induction. We denote α_i^j the coefficient α_i of BDF of order j. As the first step we will prove that our α_v^j holds (3.4.40) for $s = 2$, $2 \leq j \leq k$.

$$\sum_{v=0}^{j}\alpha_v^j(j-v)^s = \sum_{v=0}^{j-1}\alpha_v^j(j-v)^2 = \sum_{v=0}^{j-1}(-1)^{j-v}\begin{pmatrix} j \\ v \end{pmatrix}\frac{1}{j-v}(j-v)^2 \tag{3.4.50}$$

$$= \sum_{v=0}^{j-1} -(-1)^{j-1-v}\frac{j(j-1)!}{v!(j-1-v)!} = -j\sum_{v=0}^{j-1}(-1)^{j-1-v}\begin{pmatrix} j-1 \\ v \end{pmatrix} = -j(1-1)^{j-1} = 0$$

Now let us assume that α_v^j holds (3.4.40) for $j = 2, \ldots, k-1$. Now we want to prove that α_v^k holds (3.4.40) for $2 \leq s \leq k$. We know that it holds for $s = 2$. We will assume that it holds for $s-1$. From this follows

$$\sum_{v=0}^{k}\alpha_v^k(k-v)^s = \sum_{v=0}^{k}\alpha_v^k(k-v)(k-v)^{s-1} = k\sum_{v=0}^{k}\alpha_v^k(k-v)^{s-1} - \sum_{v=0}^{k}\alpha_v^k(k-v)^{s-1}v \tag{3.4.51}$$

$$= 0 - \sum_{v=1}^{k-1} \alpha_v^k (k-v)^{s-1} v = - \sum_{v=1}^{k-1} (-1)^{k-v} \begin{pmatrix} k \\ v \end{pmatrix} (k-v)^{s-2} v$$

$$= - \sum_{v=1}^{k-1} (-1)^{k-1-(v-1)} \frac{k(k-1)!}{(v-1)!(k-1-(v-1))!} (k-1-(v-1))^{s-2}$$

$$= -k \sum_{v=0}^{k-2} (-1)^{k-1-v} \frac{(k-1)!}{v!(k-1-v)!} (k-1-v)^{s-2} = -k \sum_{v=0}^{k-2} \alpha_v^{k-1} (k-1-v)^{s-1} = 0$$

□

Lemma 4 *Let (3.4.29) is BDF of order $k \geq 2$. Then*

$$\alpha_k = \sum_{v=1}^{k} \frac{1}{v} \tag{3.4.52}$$

Proof. We will use the notation α_v^j for α_v of the BDF of order j. It is easy to compute that α_1^1 and α_2^2 satisfy our lemma. Now we want to show that

$$\alpha_{k+1}^{k+1} - \alpha_k^k = \frac{1}{k+1}, \tag{3.4.53}$$

which proves our lemma. From (3.4.38) follows

$$\alpha_{k+1}^{k+1} = - \sum_{v=0}^{k} \alpha_v^{k+1} = - \sum_{v=0}^{k} (-1)^{k+1-v} \begin{pmatrix} k+1 \\ v \end{pmatrix} \frac{1}{k+1-v} \tag{3.4.54}$$

$$= -(-1)^{k+1} \frac{1}{k+1} - \sum_{v=1}^{k} (-1)^{k+1-v} \begin{pmatrix} k+1 \\ v \end{pmatrix} \frac{1}{k+1-v}$$

$$= -(-1)^{k+1} \frac{1}{k+1} - \sum_{v=1}^{k} (-1)^{k-(v-1)} \frac{1}{v} \frac{(k+1)k!}{(v-1)!(k-(v-1))!} \frac{1}{k-(v-1)}$$

$$= -(-1)^{k+1} \frac{1}{k+1} - \sum_{v=0}^{k-1} (-1)^{k-v} \begin{pmatrix} k \\ v \end{pmatrix} \frac{1}{k-v} \frac{k+1}{v+1}$$

$$= -(-1)^{k+1} \frac{1}{k+1} - \sum_{v=0}^{k-1} \alpha_v^k \frac{k+1}{v+1}$$

$$\alpha_k^k = - \sum_{v=0}^{k-1} \alpha_v^k \tag{3.4.55}$$

Now we can compute $\alpha_{k+1}^{k+1} - \alpha_k^k$:

$$\alpha_{k+1}^{k+1} - \alpha_k^k = -(-1)^{k+1} \frac{1}{k+1} - \sum_{v=0}^{k-1} \alpha_v^k \left(\frac{k+1}{v+1} - 1 \right) \tag{3.4.56}$$

$$= -(-1)^{k+1} \frac{1}{k+1} - \sum_{v=0}^{k-1} \alpha_v^k \left(\frac{k-v}{v+1} \right)$$

$$= -(-1)^{k+1}\frac{1}{k+1} - \sum_{v=0}^{k-1}(-1)^{k-v}\left(\begin{array}{c} k \\ v \end{array}\right)\frac{1}{k-v}\left(\frac{k-v}{v+1}\right)$$

$$= -(-1)^{k+1}\frac{1}{k+1} - \sum_{v=0}^{k-1}(-1)^{k-v}\frac{k!}{(v+1)!(k-v)!}$$

$$= -\frac{1}{k+1}\left((-1)^{k+1} + \sum_{v=0}^{k-1}(-1)^{k+1-(v+1)}\frac{(k+1)!}{(v+1)!(k+1-(v+1))!}\right)$$

$$= -\frac{1}{k+1}\left((-1)^{k+1} + \sum_{v=1}^{k}(-1)^{k+1-v}\left(\begin{array}{c} k+1 \\ v \end{array}\right)\right)$$

$$= -\frac{1}{k+1}\sum_{v=0}^{k}(-1)^{k+1-v}\left(\begin{array}{c} k+1 \\ v \end{array}\right)$$

$$= \frac{1}{k+1} - \frac{1}{k+1}\sum_{v=0}^{k+1}(-1)^{k+1-v}\left(\begin{array}{c} k+1 \\ v \end{array}\right)$$

$$= \frac{1}{k+1} - \frac{1}{k+1}(-1+1)^{k+1} = \frac{1}{k+1} - 0 = \frac{1}{k+1}$$

□

Remark 3 *We can verify by simple calculation that for $k = 1$ holds Lemma 3 and Lemma 4 too. Later we can also find useful that $\alpha_k > 0$, which follows from Lemma 4.*

Lemma 5 *Let α_v are coefficients of BDF of order k. Then*

$$\sum_{v=0}^{k}|\alpha_v| \leq 2k\left(\begin{array}{c} k \\ \lfloor \frac{k}{2} \rfloor \end{array}\right) =: A, \tag{3.4.57}$$

where $\lfloor x \rfloor = \max\{n \in I\!N \cup \{0\} : n \leq x\}$.

Proof.

$$\sum_{v=0}^{k}|\alpha_v| = |\alpha_k| + \sum_{v=0}^{k-1}|\alpha_v| = \left|-\sum_{v=0}^{k-1}\alpha_v\right| + \sum_{v=0}^{k-1}|\alpha_v| \leq 2\sum_{v=0}^{k-1}|\alpha_v| \leq 2k\left(\begin{array}{c} k \\ \lfloor \frac{k}{2} \rfloor \end{array}\right) \tag{3.4.58}$$

□

Lemma 6 *Let us consider that $y : [a, b] \to R$ is the solution of the problem (3.4.28). Let $y \in C^{p+1}([a, b])$, where $p \geq 0$ is the order of the method (3.4.29). Then for $a \leq t_n \leq t_{n+k} \leq b$ holds*

$$\sum_{v=0}^{k}\alpha_v y(t_{n+v}) = \tau\sum_{v=0}^{k}\beta_v F(t_{n+v}, y(t_{n+v})) + O(\tau^{p+1}) \tag{3.4.59}$$

Remark 4 *The Lemma 6 estimates the local discretization error of the method (3.4.29).*

Proof. Let us set $t = t_n$. Then for $t, t + v\tau \in [a, b]$ holds

$$y(t + v\tau) = \sum_{s=0}^{p} \frac{y^{(s)}(t)v^s\tau^s}{s!} + \tau^{p+1}\frac{v^{p+1}}{(p+1)!}y^{(p+1)}(\psi_v), \tag{3.4.60}$$

where $\psi_v \in [t, t + v\tau] \subset [a, b]$, and

$$\tau y'(t + v\tau) = \sum_{s=1}^{p} \frac{y^{(s)}(t)v^{s-1}\tau^s}{(s-1)!} + \tau^{p+1}\frac{v^p}{p!}y^{(p+1)}(\tilde{\psi}_v), \tag{3.4.61}$$

where $\tilde{\psi}_v \in [t, t + v\tau] \subset [a, b]$.
Now we are ready to estimate the error.

$$\left|\sum_{v=0}^{k} \alpha_v y(t + v\tau) - \tau\sum_{v=0}^{k} \beta_v F(t + v\tau, y(t + v\tau))\right| = \left|\sum_{v=0}^{k} \alpha_v y(t + v\tau) - \sum_{v=0}^{k} \beta_v \tau y'(t + v\tau)\right| \tag{3.4.62}$$

$$= \left|y(t)\sum_{v=0}^{k} \alpha_v + \sum_{s=1}^{p} \tau^s y^{(s)}(t)\left(\sum_{v=0}^{k} \alpha_v \frac{v^s}{s!} - \sum_{v=0}^{k} \beta_v \frac{v^{s-1}}{(s-1)!}\right) + R_{p+1}\right| = |R_{p+1}|.$$

To finish the proof we only estimate the $|R_{p+1}|$.

$$|R_{p+1}| = \left|\sum_{v=0}^{k} \alpha_v \tau^{p+1}\frac{v^{p+1}}{(p+1)!}y^{(p+1)}(\psi_v) - \sum_{v=0}^{k} \beta_v \tau^{p+1}\frac{v^p}{p!}y^{(p+1)}(\tilde{\psi}_v)\right| \tag{3.4.63}$$

$$\leq \tau^{p+1} \max_{\psi\in[a,b]} |y^{(p+1)}(\psi)| \left|\sum_{v=0}^{k} \alpha_v \frac{v^{p+1}}{(p+1)!} - \sum_{v=0}^{k} \beta_v \frac{v^p}{p!}\right|.$$

□

Now we can look at the global error estimates.

Lemma 7 *Let the method (3.4.29) be D-stable and let γ_j, $j = 0, 1, \ldots$ are real coefficients such that*

$$\frac{1}{\alpha_k + \alpha_{k-1}z + \ldots + \alpha_0 z^k} = \sum_{j=0}^{\infty} \gamma_j z^j. \tag{3.4.64}$$

Then

$$\Gamma := \sup_{j=0,1,\ldots} |\gamma_j| < \infty. \tag{3.4.65}$$

Proof. Let us set

$$\tilde{\rho}(z) = \alpha_k + \alpha_{k-1}z + \ldots + \alpha_0 z^k = z^k \rho(\frac{1}{z}). \tag{3.4.66}$$

Because ρ has no roots outside of $\{z : |z| \leq 1\}$, all roots of $\tilde{\rho}$ are in the set $\{z : |z| \geq 1\}$. Let $z_1, \ldots, z_m$ are all roots of ρ, such that $|z_v| = 1$, $v = 1, \ldots, m$. These roots have multiplicity one. From that follows that $1/\tilde{\rho}(z)$ is a holomorphic function on the set $\{z : |z| < 1\}$ and on the set $\{z : |z| = 1\}$ has a finite number of poles z_v^{-1}, $v = 1, \ldots, m$ of multiplicity one. From that follows that there exists $\varepsilon > 0$ and constants A_v, $v = 1, \ldots, m$ such that function

$$f(z) = \frac{1}{\tilde{\rho}(z)} - \sum_{v=1}^{m} \frac{A_v}{z - z_v^{-1}} \tag{3.4.67}$$

is a holomorfic function on the set $\{z : |z| < 1 + \varepsilon\}$. Cauchy formula implies that the Taylor coefficients of function f are bounded and the Taylor coefficients of each term $\frac{A_v}{z - z_v^{-1}}$ are bounded too.

□

Remark 5 *Let us set $\gamma_j = 0$ for $j \leq -1$. Then we can multiply equation (3.4.64) by $\tilde{\rho}(z)$, we get*

$$1 = \frac{\tilde{\rho}(z)}{\tilde{\rho}(z)} = \tilde{\rho}(z) \sum_{j=0}^{\infty} \gamma_j z^j. \tag{3.4.68}$$

We can now compare the coefficients with the same powers of z. As the result we get these relations

$$m = 0 \qquad \sum_{v=0}^{k} \alpha_v \gamma_{m-k+v} = \alpha_k \gamma_m + \alpha_{k-1} \gamma_{m-1} + \ldots + \alpha_0 \gamma_{m-k} = \alpha_k \gamma_0 = 1 \tag{3.4.69}$$

$$m \geq 1 \qquad \sum_{v=0}^{k} \alpha_v \gamma_{m-k+v} = \alpha_k \gamma_m + \alpha_{k-1} \gamma_{m-1} + \ldots + \alpha_0 \gamma_{m-k} = 0$$

From this we can see that γ_j can be defined as the solution of the linear difference equation

$$\sum_{v=0}^{k} \alpha_v \gamma_{j+v} = 0 \tag{3.4.70}$$

for $j = -(k-1), \ldots, -1, 0, 1, \ldots$ with the initial condition

$$\begin{aligned} \gamma_{-(k-1)}, \ldots, \gamma_{-1} &= 0 \\ \gamma_0 &= \frac{1}{\alpha_k}. \end{aligned} \tag{3.4.71}$$

3.5 Discrete formulation

Now we are ready to introduce the discrete problem. To this end, we consider an uniform partition $t_s = s\tau$ $s = 0, \ldots, r$ of the time interval $[0, T]$ with time step $\tau = \frac{T}{r}$, the exact solution $u(t_s)$ will be approximated by an element $u^s \in S_h$, as test functions w we shall use functions $w_h \in S_h$ and to overcome the time derivative in (3.3.19) is used the backward differential formulae. In order to obtain a stable and efficient scheme, the forms a_h, J_h^σ and ℓ_h will be treated implicitly whereas the nonlinear form b_h will be treated using an explicit extrapolation. We will use extrapolation $\tilde{u}^{s+k}$ to the u^{s+k} such that we don't loose the order of accuracy. In this way we arrive at the following method.

Definition 4 *Let α_v are coefficients of some k step BDF method of order k. Then we define the* approximate solution *of problem (2.0.11) a) – c) as functions u_h^{s+k}, $t_{s+k} \in [0,T]$, satisfying the conditions*

$$a) \qquad u_h^{s+k} \in S_h, \tag{3.5.72}$$

$$b) \qquad \left(\frac{\sum_{v=0}^{k} \alpha_v u_h^{s+v}}{\tau}, w_h \right) + a_h(u_h^{s+k}, w_h) + b_h(\tilde{u}_h^{s+k}, w_h) + \varepsilon J_h^{\sigma}(u_h^{s+k}, w_h) = \ell_h(w_h)\,(t_{s+k})$$

$$\forall\, w_h \in S_h,\ s = 0,1,2,\dots,r-k,$$

$$\tilde{u}_h^{s+k} = -\sum_{v=1}^{k} (-1)^v \begin{pmatrix} k \\ v \end{pmatrix} u_h^{s+k-v}$$

$$c) \qquad (u_h^0, w_h) = (u^0, w_h) \quad \forall\, w_h \in S_h$$

The function u_h^s is called the approximate solution at time t_s.

Remark 6 *We can see that $\tilde{u}_h^{s+k}$ depends on $u_h^s, \dots, u_h^{s+k-1}$ and is independent of u_h^{s+k}.*

Remark 7 *The approximate solution u_h^0 at $t = 0$ given by (3.5.72), c) is the L^2-projection of the function u^0 from the initial condition (2.0.3). If we define the operator $\Pi^{L^2} : L^2(\Omega) \to S_h$ of the $L^2(\Omega)$-projection on S_h, i.e. for $w \in L^2(\Omega)$ we assume that $\Pi^{L^2} w \in S_h$ and*

$$\left(\Pi^{L^2} w, \varphi\right) = (w, \varphi) \quad \forall \varphi \in S_h, \tag{3.5.73}$$

then $u_h^0 = \Pi^{L^2} u^0$.

Remark 8 *Since (3.5.72) , a)-c) represents a k-step formula, we have to define the solution $u_h^1, \dots, u_h^{k-1}$ at time $t_1, \dots, t_{k-1}$. It can be done by a one step formula or either by k order Runge-Kutta scheme.*

The discrete problem (3.5.72), a) – c) is equivalent to a system of linear algebraic equations for each $t_{s+k} \in [0,T]$, which can be solved by a suitable solver. In what follows we shall be concerned with the analysis of method (3.5.72), a) – c).

Lemma 8 *The discrete problem (3.5.72) a) – c) has a unique solution.*

Proof. Problem (3.5.72) a) – c) can be rewritten in the following way.
Given $u_h^s, \dots, u_h^{s+k-1} \in S_h$, $\tau > 0$, we seek $u_h^{s+k} \in S_h$ such that

$$\boldsymbol{A}_h(u_h^{s+k}, w_h) = \boldsymbol{f}_h^s(w_h) \qquad \forall w_h \in S_h,\ s = 0,1,\dots,r-k \tag{3.5.74}$$

where

$$\begin{aligned} \boldsymbol{A}_h(u_h^{s+k}, w_h) &\equiv (\alpha_k u_h^{s+k}, w_h) + \tau \left(a_h(u_h^{s+k}, w_h) + \varepsilon J_h^{\sigma}(u_h^{s+k}, w_h) \right), \\ \boldsymbol{f}_h^s(w_h) &\equiv \left(-\sum_{v=0}^{k-1} \alpha_v u_h^{s+v}, w_h \right) + \tau \left(\ell_h(w_h)\,(t_{s+k}) - b_h(\tilde{u}_h^{s+k}, w_h) \right). \end{aligned} \tag{3.5.75}$$

Using the definition of forms (3.3.15) – (3.3.18) and (3.3.20), it is easy to observe that $\boldsymbol{A}_h$ is a bilinear form on the finite dimensional space S_h and $\boldsymbol{f}_h^s(w_h)$ is a linear functional on S_h (for a given $t_{s+k} \in [0, T]$). Moreover, the form $\boldsymbol{A}_h$ is coercive, since

$$a_h(w_h, w_h) \geq 0 \tag{3.5.76}$$

$$J_h^\sigma(w_h, w_h) \geq 0 \tag{3.5.77}$$

$$\begin{aligned} \boldsymbol{A}_h(w_h, w_h) &= \alpha_k(w_h, w_h) + \tau\left(a_h(w_h, w_h) + \varepsilon J_h^\sigma(w_h, w_h)\right) \\ &\geq \alpha_k \|w_h\|^2_{L^2(\Omega)} \forall w_h \in S_h. \end{aligned} \tag{3.5.78}$$

Hence, equation (3.5.74) has a unique solution $u_h^{s+k} \in S_h$. □

Chapter 4

Some auxiliary results

In this chapter we summarize some important results and properties which have been proved in [7] and [8].

4.1 Geometry of the mesh

Let us consider a system $\{\mathcal{T}_h\}_{h\in(0,h_0)}$, $h_0 > 0$, of partitions of the domain Ω, i.e. $\mathcal{T}_h = \{K_i\}_{i\in I_h}$, $I_h \subset Z^+$. For the sake of simplicity, we shall write I instead of I_h $(h \in (0,h_0))$ and the dependence of index sets $I, I_b, s(i), \gamma(i)$ and $S(i)$ on h will not be emphasized by the notation.

We shall assume that the system $\{\mathcal{T}_h\}_{h\in(0,h_0)}$ has the following properties:

(A1) Each element $K \in \mathcal{T}_h$, $h \in (0,h_0)$, is a *star-shaped* domain with respect to at least one point $x_K = (x_{K1},\dots,x_{Kd}) \in K^\circ$, where K° is the interior of K. We assume:

i) There exists a constant $\kappa > 0$ independent of K and h such that

$$\frac{\max_{x\in\partial K}|x-x_K|}{\min_{x\in\partial K}|x-x_K|} \le \kappa \quad \forall K \in \mathcal{T}_h,\ \forall h \in (0,h_0). \tag{4.1.1}$$

ii) The element K can be divided into a finite number of closed simplexes:

$$K = \bigcup_{S\in\boldsymbol{S}(K)} S. \tag{4.1.2}$$

There exists a positive constant C_2 independent of K, S and h such that

$$\frac{h_S}{\rho_S} \le C_2 \quad \forall S \in \boldsymbol{S}(K) \qquad \text{(shape regularity)}, \tag{4.1.3}$$

where h_S is the diameter of S, ρ_S is the radius of the largest d-dimensional ball inscribed into S and, moreover,

$$1 \le \frac{h_K}{h_S} \le \tilde\kappa < \infty \quad \forall S \in \boldsymbol{S}(K), \tag{4.1.4}$$

where $\tilde\kappa$ is a constant independent of K, S and h.

(A2) There exists a constant $C_3 > 0$ such that

$$h_{K_i} \leq C_3\, d(\Gamma_{ij}), \quad i \in I, j \in S(i), h \in (0, h_0). \tag{4.1.5}$$

Let us note that these properties can be verified, e.g. in the case of dual finite volumes constructed over a regular simplicial mesh.

4.2 Some important inequalities and estimates

Under the above assumptions, the following results can be established.

Lemma 9 (Multiplicative trace inequality) *There exists a constant $C_4 > 0$ independent of v, h and K such that*

$$\|w\|^2_{L^2(\partial K)} \leq C_4 \left(\|w\|_{L^2(K)}\, |w|_{H^1(K)} + h_K^{-1} \|w\|^2_{L^2(K)} \right), \tag{4.2.6}$$
$$K \in \mathcal{T}_h,\ w \in H^1(K),\ h \in (0, h_0).$$

Proof. See [8]. □

Lemma 10 (Inverse inequality) *There exists a constant $C_5 > 0$ independent of v, h and K such that*

$$|w|_{H^1(K)} \leq C_5 h_K^{-1} \|w\|_{L^2(K)} \quad \forall w \in P_p(K) \quad \forall K \in \mathcal{T}_h. \tag{4.2.7}$$

Proof. See [8]. □

Lemma 11 (Approximation properties of S_h) *There exist a constants $C_6 > 0$ independent of w and h and a linear mapping $\Pi : H^1(K) \to P_p(K)$, $p \geq 0$ such that*

$$\begin{aligned} \|\Pi w - w\|_{L^2(K)} &\leq C_6 h_K^{p+1} |w|_{H^{p+1}(K)}, \\ |\Pi w - w|_{H^1(K)} &\leq C_6 h_K^{p} |w|_{H^{p+1}(K)}, \end{aligned} \tag{4.2.8}$$

for all $w \in H^{p+1}(K)$, $K \in \mathcal{T}_h$ and $h \in (0, h_0)$.

Proof. If $\mathcal{T}_h$ is a simplicial mesh, then standard results from the finite element method can be employed (see, e. g. [3] or [2]). In our case, when general nonconvex elements are used, the approximation properties (4.2.8) are derived under assumptions (A1), (A2) in [16]. □

Remark 9 *The operator Π is not L^2- projector Π^{L^2} on S_h introduced in (3.5.73).*

It is clear that $\Pi w = w$ for $w \in S_h$. Moreover, we have

$$\|w - \Pi^{L^2} w\|_{L^2(\Omega)} \leq \|w - \Pi w\|_{L^2(\Omega)} \leq C_6 h^{p+1} |w|_{H^{p+1}(\Omega)},\ w \in H^{p+1}(\Omega), \tag{4.2.9}$$

as follows from (4.2.8).

4.3 Properties of the form b_h

Let assumptions (2.0.4), a), Assumptions (H) and Assumptions (A1), (A2) be satisfied. Then the form b_h has the following properties.

Lemma 12 *There exist constants $C_8 > 0$, $C_9 > 0$ and $C_{10} > 0$ independent of $h \in (0, h_0)$ such that*

$$
\begin{aligned}
|b_h(u, w) - b_h(\bar{u}, w)| \;&\leq\; C_8 \left(J_h^\sigma(w, w)^{1/2} + |w|_{H^1(\Omega, \mathcal{T}_h)} \right) \qquad (4.3.10)\\
&\quad \times \left(\|u - \bar{u}\|_{L^2(\Omega)} + \left(\sum_{i \in I} h_{K_i} \|u - \bar{u}\|^2_{L^2(\partial K_i)} \right)^{1/2} \right),\\
&\quad u,\ \bar{u},\ w \in H^1(\Omega, \mathcal{T}_h),\\
|b_h(u_h, w_h) - b_h(\bar{u}_h, w_h)| \;&\leq\; C_9 \left(J_h^\sigma(w_h, w_h)^{1/2} + |w_h|_{H^1(\Omega, \mathcal{T}_h)} \right) \|u_h - \bar{u}_h\|_{L^2(\Omega)}, \qquad (4.3.11)\\
&\quad u_h,\ \bar{u}_h,\ w_h \in S_h,\\
|b_h(u, w_h) - b_h(\Pi u, w_h)| \;&\leq\; C_{10} h^{p+1} \left(J_h^\sigma(w_h, w_h)^{1/2} + |w_h|_{H^1(\Omega, \mathcal{T}_h)} \right) |u|_{H^{p+1}(\Omega)}, \qquad (4.3.12)\\
&\quad u \in H^{p+1}(\Omega),\ w_h \in S_h,
\end{aligned}
$$

where Πu is the S_h-interpolant of u from Lemma 11.

Proof. In [8], Lemma 5 gives (4.3.10) and (4.3.11), Lemma 8 proves (4.3.12). □

Chapter 5

Error estimates

5.1 Preparation lemmas

Our goal is now to analyze the error estimates of the approximate solution $u_h^s,\ s = 0, 1, \ldots, r$ obtained by the method (3.5.72) under the assumptions that the exact solution u satisfies (2.0.13). We consider a uniform partition $t_s = s\tau,\ s = 0, 1, \ldots, r$, of the time interval $[0, T]$ with time step $\tau = T/r$, where $r > 1$ is an integer.

Let Πu^s be the S_h-interpolation of $u^s = u(t_s)$ $(s = 0, \ldots, r)$ from Lemma 11. We set

$$\xi^s = u_h^s - \Pi u^s \in S_h, \quad \eta^s = \Pi u^s - u^s \in H^{p+1}(\Omega, \mathcal{T}_h). \tag{5.1.1}$$

Then the error e_h^s can be expressed as

$$e_h^s \equiv u_h^s - u^s = \xi^s + \eta^s, \quad s = 0, \ldots, r. \tag{5.1.2}$$

The (3.5.72), b) say

$$\begin{aligned}
&\left(\sum_{v=0}^{k} \alpha_v u_h^{s+v}, w_h\right) + \tau \Big(a_h(u_h^{s+k}, w_h) + b_h(\tilde{u}_h^{s+k}, w_h) \\
&+\varepsilon J_h^\sigma(u_h^{s+k}, w_h) - \ell_h(w_h)\,(t_{s+k})\Big) = 0, \quad t_s, \ldots, t_{s+k} \in [0, T].
\end{aligned} \tag{5.1.3}$$

Moreover, setting $t := t_{s+k}$ in (3.3.27), we have

$$\begin{aligned}
&(u'(t_{s+k}), w_h) + a_h(u^{s+k}, w_h) + b_h(u^{s+k}, w_h) + \varepsilon J_h^\sigma(u^{s+k}, w_h) \\
&\quad -\ell_h(w_h)\,(t_{s+k}) = 0, \quad t_{s+k} \in [0, T],
\end{aligned} \tag{5.1.4}$$

where $u'(t_s) = \partial u/\partial t\,(t_s),\ t_s \in [0, T]$.

Multiplying (5.1.4) by τ, subtracting from (5.1.3) and using the linearity of the forms a_h and J_h^σ, we get

$$\begin{aligned}
&\left(\sum_{v=0}^{k} \alpha_v u_h^{s+v}, w_h\right) - \tau\,(u'(t_{s+k}), w_h) \\
&+\tau \Big(a_h(u_h^{s+k} - u^{s+k}, w_h) + b_h(\tilde{u}_h^{s+k}, w_h) - b_h(u^{s+k}, w_h) \\
&\quad +\varepsilon J_h^\sigma(u_h^{s+k} - u^{s+k}, w_h)\Big) = 0, \quad s = 0, \ldots, r-k.
\end{aligned} \tag{5.1.5}$$

Taking into account relations (5.1.1) and (5.1.2), from (5.1.5) we obtain

$$\begin{aligned} & \left(\sum_{v=0}^{k}\alpha_v\xi^{s+v},w_h\right)+\tau\left(a_h(\xi^{s+k},w_h)+\varepsilon J_h^\sigma(\xi^{s+k},w_h)\right) \\ = \; & \tau(u'(t_{s+k}),w_h)-\left(\sum_{v=0}^{k}\alpha_v u^{s+v},w_h\right)-\left(\sum_{v=0}^{k}\alpha_v\eta^{s+v},w_h\right) \\ & +\tau\left(b_h(u^{s+k},w_h)-b_h(\tilde{u}_h^{s+k},w_h)-a_h(\eta^{s+k},w_h)-\varepsilon J_h^\sigma(\eta^{s+k},w_h)\right), \end{aligned} \tag{5.1.6}$$

where $w_h \in S_h$. In what follows, we estimate the individual terms on the right-hand side of (5.1.6).

The Cauchy inequality implies that

$$J_h^\sigma(\eta^s,\xi^s) \;\leq\; (J_h^\sigma(\eta^s,\eta^s))^{1/2}\,(J_h^\sigma(\xi^s,\xi^s))^{1/2}, \quad s=0,\dots,r. \tag{5.1.7}$$

Lemma 13 *Let $f\in W^{n+1,1}(0,T)$. Then for $t,s\in[0,T]$ we have*

$$-f(t)+\sum_{j=0}^{n}f^{(j)}(s)\frac{\alpha^j}{j!}(-1)^j=(-1)^n\int_t^s\int_{z_1}^s\dots\int_{z_n}^s f^{(n+1)}(z_{n+1})\,\mathrm{d}z_{n+1}\dots\,\mathrm{d}z_1, \tag{5.1.8}$$

where $s-t=\alpha$.

Proof. Using notation $f=f(t), f_+=f(s)$ we will prove:

$$-f+\sum_{j=0}^{n}f_+^{(j)}\frac{\alpha^j}{j!}(-1)^j=(-1)^n\int_t^s\int_{z_1}^s\dots\int_{z_n}^s f^{(n+1)}(z_{n+1})\,\mathrm{d}z_{n+1}\dots\,\mathrm{d}z_1 \tag{5.1.9}$$

For $n=0$ it is obvious, because of

$$f_+-f=\int_t^s f'(z_1)\mathrm{d}z_1, \tag{5.1.10}$$

which is proved in [4].
We could also find useful this identity

$$\int_0^\alpha\int_0^{y_1}\dots\int_0^{y_{n-1}}1\,\mathrm{d}y_n\dots\,\mathrm{d}y_1=\frac{\alpha^n}{n!}. \tag{5.1.11}$$

Using induction we gain:

$$\begin{aligned} & -f+\sum_{j=0}^{n}f_+^{(j)}\frac{\alpha^j}{j!}(-1)^j=-f+\sum_{j=0}^{n-1}f_+^{(j)}\frac{\alpha^j}{j!}(-1)^j+(-1)^n f_+^{(n)}\frac{\alpha^n}{n!}= \\ & (-1)^{n-1}\int_t^s\int_{z_1}^s\dots\int_{z_{n-1}}^s f^{(n)}(z_n)\,\mathrm{d}z_n\dots\,\mathrm{d}z_1-(-1)^{n-1}f_+^{(n)}\frac{\alpha^n}{n!} \end{aligned} \tag{5.1.12}$$

We will use substitution $z_i = s - y_i$, $i = 1, \dots, n$.

$$(-1)^{n-1} \int_t^s \int_{z_1}^s \dots \int_{z_{n-1}}^s f^{(n)}(z_n)\, \mathrm{d}z_n \dots \mathrm{d}z_1 - (-1)^{n-1} f_+^{(n)} \frac{\alpha^n}{n!} = \tag{5.1.13}$$
$$(-1)^{n-1} \left(\int_0^\alpha \int_0^{y_1} \dots \int_0^{y_{n-1}} f^{(n)}(s - y_n)\, \mathrm{d}y_n \dots \mathrm{d}y_1 - f_+^{(n)} \frac{\alpha^n}{n!} \right)$$

Now we could use identity (5.1.11).

$$(-1)^{n-1} \left(\int_0^\alpha \int_0^{y_1} \dots \int_0^{y_{n-1}} f^{(n)}(s - y_n)\, \mathrm{d}y_n \dots \mathrm{d}y_1 - f_+^{(n)} \frac{\alpha^n}{n!} \right) = \tag{5.1.14}$$
$$(-1)^{n-1} \int_0^\alpha \int_0^{y_1} \dots \int_0^{y_{n-1}} f^{(n)}(s - y_n) - f_+^{(n)}\, \mathrm{d}y_n \dots \mathrm{d}y_1$$

To finish the proof we use the same substitution $z_i = s - y_i$, $i = 1, \dots, n$ and first induction step.

$$(-1)^{n-1} \int_0^\alpha \int_0^{y_1} \dots \int_0^{y_{n-1}} f^{(n)}(s - y_n) - f_+^{(n)}\, \mathrm{d}y_n \dots \mathrm{d}y_1 \tag{5.1.15}$$
$$= (-1)^{n-1} \int_t^s \int_{z_1}^s \dots \int_{z_{n-1}}^s f^{(n)}(z_n) - f_+^{(n)}\, \mathrm{d}z_n \dots \mathrm{d}z_1$$
$$= (-1)^n \int_t^s \int_{z_1}^s \dots \int_{z_n}^s f^{(n+1)}(z_{n+1})\, \mathrm{d}z_{n+1} \dots \mathrm{d}z_1$$

□

Lemma 14 *Under assumptions (2.0.13), for $t_s, \dots, t_{s+k} \in [0, T]$ we have*

$$\left| \left(\sum_{v=0}^k \alpha_v u^{s+v}, w_h \right) - \tau \left(u'(t_{k+1}), w_h \right) \right| \leq C_{11} \tau^{k+1} \|w_h\|_{L^2(\Omega)}, \tag{5.1.16}$$
$$\left\| \sum_{v=0}^k (-1)^v \begin{pmatrix} k \\ v \end{pmatrix} u^{s+k-v} \right\|_{L^2(\Omega)} \leq C_{12} \tau^k, \tag{5.1.17}$$
$$\left| \sum_{v=0}^k (-1)^v \begin{pmatrix} k \\ v \end{pmatrix} u^{s+k-v} \right|_{H^1(\Omega)} \leq C_{13} \tau^k, \tag{5.1.18}$$
$$|u^{s+v+1} - u^{s+v}|_{H^{p+1}(\Omega)} \leq C_{14} \tau, \tag{5.1.19}$$

with C_{11}, C_{12}, C_{13} and C_{14} depending on u and k, but independent of s and τ.

Proof.

i) Using (3.4.38) and (3.4.39) we get

$$\left| \left(\sum_{v=0}^k \alpha_v u^{s+v}, w_h \right) - \tau \left(u'(t_{k+1}), w_h \right) \right| \tag{5.1.20}$$

$$= \left| \left(\sum_{v=0}^{k-1} \alpha_v u^{s+v}, w_h \right) - \left(\sum_{v=0}^{k-1} \alpha_v u^{s+k}, w_h \right) - \tau \left(u'(t_{k+1}), w_h \right) \right|$$

$$= \left| \left(\sum_{v=0}^{k-1} \alpha_v u^{s+v}, w_h \right) - \left(\sum_{v=0}^{k-1} \alpha_v u^{s+k}, w_h \right) + \tau \sum_{v=0}^{k-1} \alpha_v (k-v) \left(u'(t_{k+1}), w_h \right) \right|$$

$$= \left| \sum_{v=0}^{k-1} \alpha_v \left((u^{s+v}, w_h) - \sum_{j=0}^{1} (u^{(j)}(t_{s+k}), w_h) \frac{\tau^j (k-v)^j}{j!} (-1)^j \right) \right|.$$

When we use (3.4.40) we have

$$0 = -\sum_{v=0}^{k-1} \alpha_v (k-v)^j = -\sum_{v=0}^{k-1} \alpha_v (u^{(j)}(t_{s+k}), w_h) \frac{\tau^j (k-v)^j}{j!} (-1)^j \tag{5.1.21}$$

and after summing (5.1.21) over $j = 2, \ldots, k$ we get

$$\begin{aligned} 0 &= -\sum_{j=2}^{k} \sum_{v=0}^{k-1} \alpha_v (u^{(j)}(t_{s+k}), w_h) \frac{\tau^j (k-v)^j}{j!} (-1)^j \\ &= -\sum_{v=0}^{k-1} \alpha_v \sum_{j=2}^{k} (u^{(j)}(t_{s+k}), w_h) \frac{\tau^j (k-v)^j}{j!} (-1)^j \end{aligned} \tag{5.1.22}$$

Putting (5.1.20) and (5.1.22) together we have

$$\begin{aligned} &\left| \left(\sum_{v=0}^{k} \alpha_v u^{s+v}, w_h \right) - \tau \left(u'(t_{k+1}), w_h \right) \right| \\ &= \left| \sum_{v=0}^{k-1} \alpha_v \left((u^{s+v}, w_h) - \sum_{j=0}^{k} (u^{(j)}(t_{s+k}), w_h) \frac{\tau^j (k-v)^j}{j!} (-1)^j \right) \right| \end{aligned} \tag{5.1.23}$$

Now we can estimate this term using Lemma 13 and Lemma 5.

$$\begin{aligned} &\left| \sum_{v=0}^{k-1} \alpha_v \left((u^{s+v}, w_h) - \sum_{j=0}^{k} (u^{(j)}(t_{s+k}), w_h) \frac{\tau^j (k-v)^j}{j!} (-1)^j \right) \right| \\ &\le \sum_{v=0}^{k-1} |\alpha_v| \left| \int_{t_{s+v}}^{t_{s+k}} \int_{z_1}^{t_{s+k}} \ldots \int_{z_k}^{t_{s+k}} (u^{(k+1)}(z_{k+1}), w_h) \, \mathrm{d}z_{k+1} \ldots \mathrm{d}z_1 \right| \\ &\le \sum_{v=0}^{k-1} |\alpha_v| \int_{t_{s+v}}^{t_{s+k}} \int_{z_1}^{t_{s+k}} \ldots \int_{z_k}^{t_{s+k}} \left| (u^{(k+1)}(z_{k+1}), w_h) \right| \, \mathrm{d}z_{k+1} \ldots \mathrm{d}z_1 \\ &\le A \frac{k^{k+1}}{(k+1)!} \tau^{k+1} \| u^{(k+1)} \|_{L^\infty(0,T;L^2(\Omega))} \| \boldsymbol{w}_h \|_{L^2(\Omega)} \end{aligned} \tag{5.1.24}$$

which proves (5.1.16) with

$$C_{11} = A \frac{k^{k+1}}{(k+1)!} \| u^{(k+1)} \|_{L^\infty(0,T;L^2(\Omega))}. \tag{5.1.25}$$

ii) Since $u \in W^{k,\infty}(0,T;H^{p+1}(\Omega)) \subset W^{k,\infty}(0,T;L^2(\Omega))$, we can write

$$\left\|\sum_{v=0}^{k}(-1)^v \begin{pmatrix} k \\ v \end{pmatrix} u^{s+k-v}\right\|_{L^2(\Omega)} = \left\|\sum_{v=0}^{k}(-1)^{k-v} \begin{pmatrix} k \\ v \end{pmatrix} u^{s+v}\right\|_{L^2(\Omega)} \tag{5.1.26}$$
$$= \left\|u^{s+k} + \sum_{v=0}^{k-1}\alpha_v (k-v) u^{s+v}\right\|_{L^2(\Omega)}.$$

Using (3.4.39) and (3.4.40) we get

$$\left\|u^{s+k} + \sum_{v=0}^{k-1}\alpha_v (k-v) u^{s+v}\right\|_{L^2(\Omega)} = \left\|\sum_{v=0}^{k-1}\alpha_v (k-v)(u^{s+v} - u^{s+k})\right\|_{L^2(\Omega)} \tag{5.1.27}$$
$$= \left\|\sum_{v=0}^{k-1}\alpha_v (k-v)(u^{s+v} - u^{s+k}) - \sum_{j=1}^{k-1} u^{(j)}(t_{s+k})\frac{\tau^j}{j!}(-1)^j \sum_{v=0}^{k-1}\alpha_v (k-v)^{j+1}\right\|_{L^2(\Omega)}$$
$$= \left\|\sum_{v=0}^{k-1}\alpha_v (k-v)(u^{s+v} - u^{s+k}) - \sum_{v=0}^{k-1}\alpha_v (k-v)\sum_{j=1}^{k-1} u^{(j)}(t_{s+k})\frac{\tau^j}{j!}(-1)^j (k-v)^j\right\|_{L^2(\Omega)}$$
$$= \left\|\sum_{v=0}^{k-1}\alpha_v (k-v)\left(u^{s+v} - \sum_{j=0}^{k-1} u^{(j)}(t_{s+k})\frac{\tau^j}{j!}(-1)^j (k-v)^j\right)\right\|_{L^2(\Omega)}.$$

This term we can estimate by Lemma 13 and Lemma 5.

$$\left\|\sum_{v=0}^{k-1}\alpha_v (k-v)\left(u^{s+v} - \sum_{j=0}^{k-1} u^{(j)}(t_{s+k})\frac{\tau^j}{j!}(-1)^j (k-v)^j\right)\right\|_{L^2(\Omega)} \tag{5.1.28}$$
$$= \left\|\sum_{v=0}^{k-1}\alpha_v (k-v)\left(\int_{t_{s+v}}^{t_{s+k}}\int_{z_1}^{t_{s+k}}\dots\int_{z_{k-1}}^{t_{s+k}} u^{(k)}(z_k)\,\mathrm{d}z_k\dots\mathrm{d}z_1\right)\right\|_{L^2(\Omega)}$$
$$\le k\sum_{v=0}^{k-1}|\alpha_v|\int_{t_{s+v}}^{t_{s+k}}\int_{z_1}^{t_{s+k}}\dots\int_{z_{k-1}}^{t_{s+k}} \|u^{(k)}(z_k)\|_{L^2(\Omega)}\,\mathrm{d}z_k\dots\mathrm{d}z_1$$
$$\le kA\frac{k^k}{k!}\tau^k\|u^{(k)}\|_{L^\infty(0,T;L^2(\Omega))},$$

which proves (5.1.17) with

$$C_{12} = A\frac{k^{k+1}}{k!}\|u^{(k)}\|_{L^\infty(0,T;L^2(\Omega))}. \tag{5.1.29}$$

iii) Since $u \in W^{k,\infty}(0,T;H^{p+1}(\Omega)) \subset W^{k,\infty}(0,T;H^1(\Omega))$, we have

$$\left|\sum_{v=0}^{k}(-1)^v \begin{pmatrix} k \\ v \end{pmatrix} u^{s+k-v}\right|_{H^1(\Omega)} = \left\|\sum_{v=0}^{k}(-1)^v \begin{pmatrix} k \\ v \end{pmatrix} \nabla u^{s+k-v}\right\|_{L^2(\Omega)} \tag{5.1.30}$$

Now we can use previous result for ∇u instead of u. Using

$$\frac{\partial}{\partial t}\left(\frac{\partial u}{\partial x_l}\right) = \frac{\partial}{\partial x_l}\left(\frac{\partial u}{\partial t}\right), \qquad l = 1, \ldots, d, \tag{5.1.31}$$

in the sense of distributions we get

$$\begin{aligned} \left|\sum_{v=0}^{k}(-1)^v \begin{pmatrix} k \\ v \end{pmatrix} u^{s+k-v}\right|_{H^1(\Omega)} &\leq A\frac{k^{k+1}}{k!}\tau^k \|\frac{\partial^k}{\partial t^k}\nabla u\|_{L^\infty(0,T;L^2(\Omega))} \\ = A\frac{k^{k+1}}{k!}\tau^k\|\nabla u^{(k)}\|_{L^\infty(0,T;L^2(\Omega))} &\leq A\frac{k^{k+1}}{k!}\tau^k\|u^{(k)}\|_{L^\infty(0,T;H^1(\Omega))} \end{aligned} \tag{5.1.32}$$

which is (5.1.18) with

$$C_{13} = A\frac{k^{k+1}}{k!}\|u^{(k)}\|_{L^\infty(0,T;H^1(\Omega))}. \tag{5.1.33}$$

iv) Taking into account property (5.1.31) and using a similar technique as in (5.1.32) we derive (5.1.19) with $C_{14} = \|u'\|_{L^\infty(0,T;H^{p+1}(\Omega))}$.

□

Lemma 15 *Under assumptions (2.0.13), for $t_s, t_{s+1} \in [0,T]$, $w_h \in S_h$ we have*

$$|(\eta^{s+1} - \eta^s, w_h)| \leq \tilde{C}_{16}\tau h^{p+1}\|w_h\|_{L^2(\Omega)}, \tag{5.1.34}$$

with $\tilde{C}_{16} = \tilde{C}_{16}(u)$.

Proof. The Cauchy inequality, relations (5.1.1), (4.2.8) and (5.1.19) imply that

$$\begin{aligned} |(\eta^{s+1} - \eta^s, w_h)| &\leq \|\eta^{s+1} - \eta^s\|_{L^2(\Omega)}\|w_h\|_{L^2(\Omega)} \\ &= \|\Pi(u^{s+1} - u^s) - (u^{s+1} - u^s)\|_{L^2(\Omega)}\|w_h\|_{L^2(\Omega)} \\ &\leq C_6 h^{p+1}|u^{s+1} - u^s|_{H^{p+1}(\Omega)}\|w_h\|_{L^2(\Omega)} \\ &\leq C_6 C_{14}\tau h^{p+1}\|w_h\|_{L^2(\Omega)}, \end{aligned} \tag{5.1.35}$$

which proves the lemma with $\tilde{C}_{16} := C_6 C_{14}$. □

Lemma 16 *Let (3.4.29) satisfies (3.4.32). Then under assumptions (2.0.13), for $t_s, \ldots, t_{s+k} \in [0,T], w_h \in S_h$ we have*

$$\left|\left(\sum_{v=0}^{k}\alpha_v \eta^{s+v}, w_h\right)\right| \leq C_{16}\tau h^{p+1}\|w_h\|_{L^2(\Omega)}, \tag{5.1.36}$$

with C_{16} depending on $\tilde{C}_{16}$ and k.

Proof. From (3.4.32) follows that 1 is a root of the characteristic polynomial $\rho(y)$. We set polynomial π such that :

$$\rho(y) = (y-1)\pi(y) \tag{5.1.37}$$

Now we should compute the coefficients of $\pi(y) = \sum\limits_{v=0}^{k-1} \beta_v y^v$.

$$\beta_0 = \pi(0) = \frac{\rho(0)}{-1} = -\alpha_0 = \sum_{v=1}^{k} \alpha_v \tag{5.1.38}$$

When we compute $v \geq 1$ derivations of $\rho(y)$ and enumerate them in 0 we get

$$\rho^{(v)}(y) = \sum_{j=0}^{v} \begin{pmatrix} v \\ j \end{pmatrix} \pi^{(v-j)}(y)(y-1)^{(j)} \tag{5.1.39}$$

$$v!\alpha_v = \rho^{(v)}(0) = -\pi^{(v)}(0) + v\pi^{(v-1)}(0) = -v!\beta_v + v!\beta_{v-1}$$

From this we can easily obtain by induction

$$\beta_v = \sum_{j=v+1}^{k} \alpha_j \tag{5.1.40}$$

Because $\pi(y)$ has property (5.1.37) we can write

$$\left| \left(\sum_{v=0}^{k} \alpha_v \eta^{s+v}, w_h \right) \right| = \left| \left(\sum_{v=0}^{k-1} \beta_v (\eta^{s+v+1} - \eta^{s+v}), w_h \right) \right| \leq \sum_{v=0}^{k-1} |\beta_v| \left| \left(\eta^{s+v+1} - \eta^{s+v}, w_h \right) \right| \tag{5.1.41}$$

When we also use (3.4.57) and (5.1.34) we get

$$\sum_{v=0}^{k-1} |\beta_v| \left| \left(\eta^{s+v+1} - \eta^{s+v}, w_h \right) \right| \leq kA\tilde{C}_{16} \tau h^{p+1} \|w_h\|_{L^2(\Omega)}, \tag{5.1.42}$$

which proves the lemma with $C_{16} := kA\tilde{C}_{16}$. □

Lemma 17 *There exist constants $C_{17} > 0$ and $C_{18} > 0$ independent of u, h, s, w_h and ε such that*

$$|a_h(\eta^s, w_h)| \leq C_{17} \varepsilon h^p |u^s|_{H^{p+1}(\Omega)} \left(J_h^\sigma(w_h, w_h)^{1/2} + |\boldsymbol{w}_h|_{H^1(\Omega, \mathcal{T}_h)} \right), \tag{5.1.43}$$

$$J_h^\sigma(\eta^s, \eta^s) \leq C_{18}\, h^{2p} |u^s|^2_{H^{p+1}(\Omega)}, \quad h \in (0, h_0),\ t_s \in [0, T]. \tag{5.1.44}$$

Proof. See [8], Lemma 9. □

Lemma 18 *Let us set*

$$\tilde{u}_h^{s+k} = -\sum_{v=1}^{k} (-1)^v \begin{pmatrix} k \\ v \end{pmatrix} u_h^{s+k-v} \tag{5.1.45}$$

for $h \in (0, h_0)$*,* $t_s, t_{s+k} \in [0, T]$*. Then we have*

$$\left| b_h(u^{s+k}, w_h) - b_h(\tilde{u}_h^{s+k}, w_h) \right| \le C_{19} \left(J_h^\sigma(w_h, w_h)^{1/2} + |w_h|_{H^1(\Omega,\mathcal{T}_h)} \right) \times \left(\sum_{v=0}^{k-1} \|\xi^{s+v}\|_{L^2(\Omega)} + h^{p+1} + \tau^k \right), \tag{5.1.46}$$

where $C_{19} = C_{19}(u)$ *is independent of* h*,* τ*,* s*.*

Proof. At first we set

$$\tilde{u}^{s+k} = -\sum_{v=1}^{k} (-1)^v \begin{pmatrix} k \\ v \end{pmatrix} u^{s+k-v}. \tag{5.1.47}$$

We can write

$$\begin{aligned} b_h(u^{s+k}, w_h) - b_h(\tilde{u}_h^{s+k}, w_h) = {} & b_h(u^{s+k}, w_h) - b_h(\tilde{u}^{s+k}, w_h) && (=: \Psi_1) \\ & + b_h(\tilde{u}^{s+k}, w_h) - b_h(\Pi\tilde{u}^{s+k}, w_h) && (=: \Psi_2) \\ & + b_h(\Pi\tilde{u}^{s+k}, w_h) - b_h(\tilde{u}_h^{s+k}, w_h) && (=: \Psi_3). \end{aligned}$$

We estimate the individual terms in (5.1.48). From (4.3.10) we have

$$|\Psi_1| \le C_8 \left(J_h^\sigma(w_h, w_h)^{1/2} + |w_h|_{H^1(\Omega,\mathcal{T}_h)} \right) \times \left(\left\| \sum_{v=0}^{k} (-1)^v \begin{pmatrix} k \\ v \end{pmatrix} u^{s+k-v} \right\|_{L^2(\Omega)} + \left(\sum_{i \in I} h_{K_i} \left\| \sum_{v=0}^{k} (-1)^v \begin{pmatrix} k \\ v \end{pmatrix} u^{s+k-v} \right\|_{L^2(\partial K_i)}^2 \right)^{1/2} \right). \tag{5.1.48}$$

Using (4.2.6), (5.1.17) and (5.1.18) we find that

$$\begin{aligned} & \sum_{i \in I} h_{K_i} \left\| \sum_{v=0}^{k} (-1)^v \begin{pmatrix} k \\ v \end{pmatrix} u^{s+k-v} \right\|_{L^2(\partial K_i)}^2 \\ \le {} & C_4 \sum_{i \in I} h_{K_i} \left\| \sum_{v=0}^{k} (-1)^v \begin{pmatrix} k \\ v \end{pmatrix} u^{s+k-v} \right\|_{L^2(K_i)} \left| \sum_{v=0}^{k} (-1)^v \begin{pmatrix} k \\ v \end{pmatrix} u^{s+k-v} \right|_{H^1(K_i)} \\ & + C_4 \sum_{i \in I} \left\| \sum_{v=0}^{k} (-1)^v \begin{pmatrix} k \\ v \end{pmatrix} u^{s+k-v} \right\|_{L^2(K_i)}^2 \le C_{21} \tau^{2k}, \end{aligned} \tag{5.1.49}$$

where $C_{21} := C_4(C_{12}C_{13}h_0 + C_{12}^2)$. Then (5.1.48), (5.1.49) and (5.1.17) give

$$|\Psi_1| \le C_8 \left(J_h^\sigma(w_h, w_h)^{1/2} + |w_h|_{H^1(\Omega,\mathcal{T}_h)} \right) \tau^k (\sqrt{C_{21}} + C_{12}). \tag{5.1.50}$$

Moreover, on the basis of (2.0.13) we can set

$$C_{20} = \|u\|_{L^\infty(0,T;H^{p+1}(\Omega))}. \tag{5.1.51}$$

From (4.3.12) and (4.3.11) from Lemma 12 we find that

$$\begin{aligned} |\Psi_2| &\le C_{10} h^{p+1} \left(J_h^\sigma(w_h, w_h)^{1/2} + |w_h|_{H^1(\Omega,\mathcal{T}_h)} \right) .(|\tilde{u}^{s+k}|_{H^{p+1}(\Omega)}), \qquad (5.1.52) \\ |\Psi_3| &\le C_9 \left(J_h^\sigma(w_h, w_h)^{1/2} + |w_h|_{H^1(\Omega,\mathcal{T}_h)} \right) .(\|\Pi\tilde{u}^{s+k} - \tilde{u}_h^{s+k}\|_{L^2(\Omega)}). \end{aligned}$$

By (5.1.48), (5.1.1), (5.1.50), (5.1.52) and (5.1.51),

$$\begin{aligned} \left| b_h(u^{s+k}, w_h) - b_h(\tilde{u}_h^{s+k}, w_h) \right| &\le C_{19} \left(J_h^\sigma(w_h, w_h)^{1/2} + |\, w_h|_{H^1(\Omega,\mathcal{T}_h)} \right) \qquad (5.1.53) \\ &\times \left(\sum_{v=0}^{k-1} \|\xi^{s+v}\|_{L^2(\Omega)} + h^{p+1} + \tau^k \right), \end{aligned}$$

with

$$C_{19} := \max \left(k \begin{pmatrix} k \\ \lfloor \frac{k}{2} \rfloor \end{pmatrix} C_9,\ k \begin{pmatrix} k \\ \lfloor \frac{k}{2} \rfloor \end{pmatrix} C_{10} C_{20},\ C_8 (C_{12} + \sqrt{C_{21}}) \right), \qquad (5.1.54)$$

which proves the lemma. □

Definition 5 *Let assumptions (2.0.4), a) - e), (H) from Section 3.3 and (A1)-(A2) from Section 4.1 be satisfied. Let u be the exact solution of problem (2.0.11) satisfying (2.0.13). Let* $t_s = s\tau,\ s = 0, 1, \ldots, r,\ \tau = T/r$, *be a time partition of* $[0, T]$ *and let* $u_h^s,\ s = 0, \ldots, r$ *be the approximate solution defined by (3.5.72). We set*

$$\begin{aligned} e_h^s &= u_h^s - u^s, \quad s = 0, 1, \ldots, r, \qquad (5.1.55) \\ \|e\|^2_{h,\tau,L^\infty(L^2)} &= \max_{s=0,\ldots,r} \|e_h^s\|^2_{L^2(\Omega)} \\ \|e\|^2_{h,\tau,L^2(H^1)} &= \tau\varepsilon \sum_{s=0}^{r} \left(|e_h^s|^2_{H^1(\Omega,\mathcal{T}_h)} + J_h^\sigma(e_h^s, e_h^s) \right) \end{aligned}$$

Lemma 19 *Let assumptions (2.0.4), a) - e), (H) from Section 3.3 and (A1)-(A2) from Section 4.1 be satisfied. Let u be the exact solution of problem (2.0.11) satisfying (2.0.13). Let* $t_s = s\tau,\ s = 0, 1, \ldots, r,\ \tau = T/r$, *be a time partition of* $[0, T]$ *and let* $u_h^s,\ s = 0, \ldots, r$ *be the approximate solution defined by (3.5.72), where* $k \ge 1$. *Then if we set* $\xi^s = u_h^s - \Pi u^s \in S_h$ *there exists constant K independent of h,* τ*, s and* ε *such that*

$$\begin{aligned} 2 \left(\sum_{v=0}^{k} \alpha_v \xi^{s+v}, w_h \right) + 2\tau \left(a_h(\xi^{s+k}, w_h) + \varepsilon J_h^\sigma(\xi^{s+k}, w_h) \right) \qquad (5.1.56) \\ \le \tau \|w_h\|^2_{L^2(\Omega)} + \tau \frac{K}{\varepsilon} \sum_{v=0}^{k-1} \|\xi^{s+v}\|^2_{L^2(\Omega)} \\ + \tau\varepsilon \left(|w_h|^2_{H^1(\Omega,\mathcal{T}_h)} + J_h^\sigma(w_h, w_h) \right) + \tau K q(\varepsilon, h, \tau, k), \end{aligned}$$

where

$$q(\varepsilon, h, \tau, k) \equiv h^{2p}\left(\varepsilon + h^2 + h^2/\varepsilon\right) + \tau^{2k}\left(1 + 1/\varepsilon\right) \tag{5.1.57}$$

Proof. As in (5.1.1), we set $\xi^s = u_h^s - \Pi u^s \in S_h$, $\eta^s = \Pi u^s - u^s$, $s = 0, \ldots, r$. Then (5.1.2) holds: $e_h^s = u_h^s - u^s = \xi^s + \eta^s$. When we multiply (5.1.6) by 2 we get

$$\begin{aligned} & 2\left(\sum_{v=0}^{k} \alpha_v \xi^{s+v}, w_h\right) + 2\tau\left(a_h(\xi^{s+k}, w_h) + \varepsilon J_h^\sigma(\xi^{s+k}, w_h)\right) \qquad (5.1.58) \\ = \; & 2\tau(u'(t_{s+k}), w_h) - 2\left(\sum_{v=0}^{k} \alpha_v u^{s+v}, w_h\right) - 2\left(\sum_{v=0}^{k} \alpha_v \eta^{s+v}, w_h\right) \\ & +2\tau\left(b_h(u^{s+k}, w_h) - b_h(\tilde{u}_h^{s+k}, w_h) - a_h(\eta^{s+k}, w_h) - \varepsilon J_h^\sigma(\eta^{s+k}, w_h)\right) =: \mathrm{RHS}. \end{aligned}$$

With the aid of Lemmas 14 – 18 and (5.1.7) we estimate the right-hand side of (5.1.58) denoted by RHS:

$$\begin{aligned} |\mathrm{RHS}| \leq \; & 2\left(C_{11}\tau^{k+1} + C_{16}\tau h^{p+1}\right) \|w_h\|_{L^2(\Omega)} \qquad (5.1.59) \\ & +2\tau\Big\{C_{17}\varepsilon h^p |u^{s+k}|_{H^{p+1}(\Omega)}\left(J_h^\sigma(w_h, w_h)^{1/2} + |w_h|_{H^1(\Omega,\mathcal{T}_h)}\right) \\ & +\sqrt{C_{18}}\varepsilon h^p |u^{s+k}|_{H^{p+1}(\Omega)} J_h^\sigma(w_h, w_h)^{1/2} \\ & +C_{19}\left(J_h^\sigma(w_h, w_h)^{1/2} + |w_h|_{H^1(\Omega,\mathcal{T}_h)}\right)\left(\sum_{v=0}^{k-1} \|\xi^{s+v}\|_{L^2(\Omega)} + h^{p+1} + \tau^k\right)\Big\}. \end{aligned}$$

Adding some nonnegative terms to the right-hand side of inequality (5.1.59) and using the notation (5.1.51), we have

$$\begin{aligned} |\mathrm{RHS}| \leq \; & 2\left(C_{11}\tau^{k+1} + C_{16}\tau h^{p+1}\right) \|w_h\|_{L^2(\Omega)} \qquad (5.1.60) \\ & +2\tau\left(J_h^\sigma(w_h, w_h)^{1/2} + |w_h|_{H^1(\Omega,\mathcal{T}_h)}\right) \\ & \times\left(C_{22}\varepsilon h^p + C_{19}\left(h^{p+1} + \tau^k + \sum_{v=0}^{k-1} \|\xi^{s+v}\|_{L^2(\Omega)}\right)\right), \end{aligned}$$

where $C_{22} = C_{20}(C_{17} + \sqrt{C_{18}})$.
Now we want to use Young's inequality in the form:

$$ab \leq \frac{\alpha}{4}a^2 + \frac{b^2}{\alpha}. \tag{5.1.61}$$

From (5.1.61) follows for $j = 1, 2, \ldots$

$$\left(\sum_{i=1}^{j} a_i\right)^2 \leq j \sum_{i=1}^{j} a_i^2 \tag{5.1.62}$$

Now with the aid of (5.1.61) we estimate both terms from (5.1.60). First we use (5.1.61) with $a = C_{11}\tau^k + C_{16}h^{p+1}$, $b = \|w_h\|_{L^2(\Omega)}$ and $\alpha = 2$ and then (5.1.62)with $a_1 = C_{11}\tau^k$, $a_2 = C_{16}h^{p+1}$ and $j = 2$:

$$2\left(C_{11}\tau^{k+1} + C_{16}\tau h^{p+1}\right)\|w_h\|_{L^2(\Omega)} = 2\tau\left(C_{11}\tau^k + C_{16}h^{p+1}\right)\|w_h\|_{L^2(\Omega)} \quad (5.1.63)$$
$$\leq \tau\|w_h\|^2_{L^2(\Omega)} + \tau\left(C_{11}\tau^k + C_{16}h^{p+1}\right)^2 \leq \tau\|w_h\|^2_{L^2(\Omega)} + 2C_{11}^2\tau^{2k+1} + 2C_{16}^2\tau h^{2(p+1)}.$$

The second term we estimate with (5.1.61) $a = J_h^\sigma(w_h, w_h)^{1/2} + |w_h|_{H^1(\Omega,\mathcal{T}_h)}$,
$b = C_{22}\varepsilon h^p + C_{19}\left(h^{p+1} + \tau^k + \sum_{v=0}^{k-1}\|\xi^{s+v}\|_{L^2(\Omega)}\right)$ and $\alpha = \varepsilon$

$$2\tau\left(J_h^\sigma(w_h, w_h)^{1/2} + |\ w_h|_{H^1(\Omega,\mathcal{T}_h)}\right) \quad (5.1.64)$$
$$\times\left(C_{22}\varepsilon h^p + C_{19}\left(h^{p+1} + \tau^k + \sum_{v=0}^{k-1}\|\xi^{s+v}\|_{L^2(\Omega)}\right)\right)$$
$$\leq \frac{1}{2}\tau\varepsilon\left(J_h^\sigma(w_h, w_h)^{1/2} + |w_h|_{H^1(\Omega,\mathcal{T}_h)}\right)^2 \quad (5.1.65)$$
$$+\frac{2\tau}{\varepsilon}\left(C_{22}\varepsilon h^p + C_{19}\left(h^{p+1} + \tau^k + \sum_{v=0}^{k-1}\|\xi^{s+v}\|_{L^2(\Omega)}\right)\right)^2. \quad (5.1.66)$$

The term (5.1.65) we estimate with (5.1.62) $a_1 = J_h^\sigma(w_h, w_h)^{1/2}$, $a_2 = |w_h|_{H^1(\Omega,\mathcal{T}_h)}$ and $j = 2$

$$\frac{1}{2}\tau\varepsilon\left(J_h^\sigma(w_h, w_h)^{1/2} + |w_h|_{H^1(\Omega,\mathcal{T}_h)}\right)^2 \leq \tau\varepsilon\left(J_h^\sigma(w_h, w_h) + |w_h|^2_{H^1(\Omega,\mathcal{T}_h)}\right) \quad (5.1.67)$$

The term (5.1.66) we estimate first with (5.1.62) with $a_1 = C_{22}\varepsilon h^p$,
$a_2 = C_{19}\left(h^{p+1} + \tau^k + \sum_{v=0}^{k-1}\|\xi^{s+v}\|_{L^2(\Omega)}\right)$ and $j = 2$ and then with (5.1.62) $a_1 = h^{p+1}$, $a_2 = \tau^k$,
$\sum_{i=3}^{k+2} a_i = \sum_{v=0}^{k-1}\|\xi^{s+v}\|^2_{L^2(\Omega)}$, and $j = k + 2$:

$$\frac{2\tau}{\varepsilon}\left(C_{22}\varepsilon h^p + C_{19}\left(h^{p+1} + \tau^k + \sum_{v=0}^{k-1}\|\xi^{s+v}\|_{L^2(\Omega)}\right)\right)^2 \quad (5.1.68)$$
$$\leq 4\tau\left(\varepsilon C_{22}^2 h^{2p} + \frac{1}{\varepsilon}C_{19}^2\left(h^{p+1} + \tau^k + \sum_{v=0}^{k-1}\|\xi^{s+v}\|_{L^2(\Omega)}\right)^2\right)$$
$$\leq 4\tau\left(\varepsilon C_{22}^2 h^{2p} + \frac{k+2}{\varepsilon}C_{19}^2\left(h^{2p+2} + \tau^{2k} + \sum_{v=0}^{k-1}\|\xi^{s+v}\|^2_{L^2(\Omega)}\right)\right)$$

Taking (5.1.58), (5.1.60), (5.1.63), (5.1.64), (5.1.67) and (5.1.68) together we get:

$$2\left(\sum_{v=0}^{k}\alpha_v\xi^{s+v}, w_h\right) + 2\tau\left(a_h(\xi^{s+k}, w_h) + \varepsilon J_h^\sigma(\xi^{s+k}, w_h)\right) \quad (5.1.69)$$

$$\leq \tau \|w_h\|^2_{L^2(\Omega)} + 2C_{11}^2 \tau^{2k+1} + 2C_{16}^2 \tau h^{2(p+1)} + \tau \frac{4(k+2)C_{19}^2}{\varepsilon} \sum_{v=0}^{k-1} \|\xi^{s+v}\|^2_{L^2(\Omega)}$$

$$+\tau\varepsilon \left(J_h^\sigma(w_h, w_h) + |w_h|^2_{H^1(\Omega,\mathcal{T}_h)} \right) + 4\tau\varepsilon C_{22}^2 h^{2p} + \tau \frac{4(k+2)C_{19}^2}{\varepsilon} h^{2p+2} + \tau^{2k+1} \frac{4(k+2)C_{19}^2}{\varepsilon}.$$

When we use (5.1.57), we get

$$2 \left(\sum_{v=0}^{k} \alpha_v \xi^{s+v}, w_h \right) + 2\tau \left(a_h(\xi^{s+k}, w_h) + \varepsilon J_h^\sigma(\xi^{s+k}, w_h) \right) \tag{5.1.70}$$

$$\leq \tau \|w_h\|^2_{L^2(\Omega)} + \tau \frac{K}{\varepsilon} \sum_{v=0}^{k-1} \|\xi^{s+v}\|^2_{L^2(\Omega)}$$

$$+\tau\varepsilon \left(|w_h|^2_{H^1(\Omega,\mathcal{T}_h)} + \varepsilon J_h^\sigma(w_h, w_h) \right) + \tau K q(\varepsilon, h, \tau, k),$$

with

$$K = \max \left(2C_{11}^2, 2C_{16}^2, 4(k+2)C_{19}^2, 4C_{22}^2 \right). \tag{5.1.71}$$

□

Because there exists unique sequence γ_j satisfying (3.4.70), we know from Lemma 7 that this sequence γ_j is bounded. To prove our result we need to know that γ_j is also nonnegative. This we will prove in the next lemma for $k = 2, 3$.

Lemma 20 *Let α_v are coefficients of k step BDF of order $k = 2, 3$. Let γ_j is the sequence defined in Lemma 7 by (3.4.64). Then $\gamma_j \geq 0$ for $j = 0, 1, \ldots$.*

Proof. We know that γ_j defined by (3.4.64) for $j = 0, 1, \ldots$ and $\gamma_j = 0$ for $j = -(k-2), \ldots, -1$ satisfy (3.4.70). Because all the coefficients of difference equation (3.4.70) are real constants α_v, the solution $\gamma_j \in \mathbb{R}$. For $k = 2$ it is easy to see that

$$\gamma_j = 1 - \frac{1}{3^{j+1}} \geq 0 \quad j = 0, 1, \ldots \ . \tag{5.1.72}$$

For $k = 3$ we can compute that γ_j is the solution of

$$\frac{11}{6}\gamma_{j+3} - 3\gamma_{j+2} + \frac{3}{2}\gamma_{j+1} - \frac{1}{3}\gamma_j = 0 \tag{5.1.73}$$

for $j = -2, -1, 0, 1, \ldots$ with the initial condition

$$\gamma_{-2} = \gamma_{-1} = 0 \tag{5.1.74}$$

$$\gamma_0 = \frac{6}{11}.$$

We set $\delta_j = \gamma_{j-1} - \gamma_{j-2}$, which is the solution of

$$\frac{11}{6}\delta_{j+2} - \frac{7}{6}\delta_{j+1} + \frac{1}{3}\delta_j = 0 \tag{5.1.75}$$

for $j = 0, 1, \dots$ with the initial condition

$$\delta_0 = 0 \qquad (5.1.76)$$
$$\delta_1 = \frac{6}{11}.$$

When we solve (5.1.75) with the initial condition (5.1.76) we get

$$\delta_j = \frac{-6i}{\sqrt{39}} \left(\frac{7 + i\sqrt{39}}{22} \right)^j + \frac{6i}{\sqrt{39}} \left(\frac{7 - i\sqrt{39}}{22} \right)^j. \qquad (5.1.77)$$

Now we can easily estimate $|\delta_j|$

$$|\delta_j| \leq 2\frac{6}{\sqrt{39}} \left| \frac{7 + i\sqrt{39}}{22} \right|^j \leq 2 \left(\sqrt{\frac{2}{11}} \right)^j. \qquad (5.1.78)$$

If we compute $\gamma_1, \gamma_2, \gamma_3, \gamma_4$ we can realize that all of them are positive and $\gamma_4 > 1$. Then for every $j \geq 5$ we can estimate γ_j in the following way.

$$\gamma_j = \gamma_4 + \sum_{i=6}^{j+1} \delta_i \geq \gamma_4 - \sum_{i=6}^{j+1} |\delta_i| > 1 - 2\sum_{i=6}^{j+1} \left(\sqrt{\frac{2}{11}} \right)^i > 1 - 2\frac{8}{1331} \sum_{i=0}^{\infty} \left(\sqrt{\frac{2}{11}} \right)^i \qquad (5.1.79)$$
$$= 1 - \frac{16}{1331} \frac{1}{1 - \sqrt{\frac{2}{11}}} > 1 - \frac{16}{1331} 2 > 0$$

□

5.2 Error estimation of the second order

Theorem 1 *Let assumptions (2.0.4), a) - e), (H) from Section 3.3 and (A1)-(A2) from Section 4.1 be satisfied. Let u be the exact solution of problem (2.0.11) satisfying (2.0.13). Let $t_s = s\tau$, $s = 0, 1, \dots, r$, $\tau = T/r$, be a time partition of $[0, T]$ and let u_h^s, $s = 0, \dots, r$ be the approximate solution defined by (3.5.72) with $k = 2$ and let $\tau \leq 1$. Then there exists constant $\tilde{C} = O(\exp(3\Gamma T(1 + 2K/\varepsilon)))$ such that*

$$\|e\|^2_{h,\tau,L^\infty(L^2)} \leq \tilde{C} \left(h^{2p} \left(\varepsilon + h^2 + h^2/\varepsilon \right) + \tau^4 (1 + 1/\varepsilon) + \max(\|e_h^0\|^2_{L^2(\Omega)}, \|e_h^1\|^2_{L^2(\Omega)}) \right) \qquad (5.2.80)$$

where K is defined in Lemma 19.

Proof. As in (5.1.1), we set $\xi^s = u_h^s - \Pi u^s \in S_h$, $\eta^s = \Pi u^s - u^s$, $s = 0, \dots, r$. Then (5.1.2) holds: $e_h^s = u_h^s - u^s = \xi^s + \eta^s$. It is easy to see that

$$\|e_h^s\|^2_{L^2(\Omega))} \leq \tilde{C} \left(h^{2p} \left(\varepsilon + h^2 + h^2/\varepsilon \right) + \tau^4 (1 + 1/\varepsilon) + \max(\|e^0\|^2_{L^2(\Omega)}, \|e^1\|^2_{L^2(\Omega)}) \right) \qquad (5.2.81)$$

for $s = 0, 1$. So we should only estimate e_h^s for $s = 2, \dots, r$. Putting $w_h = \xi^{s+2}$ in (5.1.56) and using the relations

$$a_h(\xi^{s+2}, \xi^{s+2}) = \varepsilon|\xi^{s+2}|^2_{H^1(\Omega,\mathcal{T}_h)} \qquad (5.2.82)$$

and

$$
\begin{gathered}
2\left(\tfrac{3}{2}\xi^{s+2}-2\xi^{s+1}+\tfrac{1}{2}\xi^{s},\xi^{s+2}\right)=\tfrac{3}{2}\|\xi^{s+2}\|^2_{L^2(\Omega)}-2\|\xi^{s+1}\|^2_{L^2(\Omega)}+\tfrac{1}{2}\|\xi^{s}\|^2_{L^2(\Omega)} \quad (5.2.83)\\
+2\|\xi^{s+2}-\xi^{s+1}\|^2_{L^2(\Omega)}-\tfrac{1}{2}\|\xi^{s+2}-\xi^{s}\|^2_{L^2(\Omega)}\geq\tfrac{3}{2}\|\xi^{s+2}\|^2_{L^2(\Omega)}-2\|\xi^{s+1}\|^2_{L^2(\Omega)}+\tfrac{1}{2}\|\xi^{s}\|^2_{L^2(\Omega)}\\
+2\|\xi^{s+2}-\xi^{s+1}\|^2_{L^2(\Omega)}-\|\xi^{s+2}-\xi^{s+1}\|^2_{L^2(\Omega)}-\|\xi^{s+1}-\xi^{s}\|^2_{L^2(\Omega)}=\\
\tfrac{3}{2}\|\xi^{s+2}\|^2_{L^2(\Omega)}-2\|\xi^{s+1}\|^2_{L^2(\Omega)}+\tfrac{1}{2}\|\xi^{s}\|^2_{L^2(\Omega)}+\|\xi^{s+2}-\xi^{s+1}\|^2_{L^2(\Omega)}-\|\xi^{s+1}-\xi^{s}\|^2_{L^2(\Omega)},
\end{gathered}
$$

for $s=0,\dots,r-2$ we get

$$
\begin{aligned}
&\frac{3}{2}\|\xi^{s+2}\|^2_{L^2(\Omega)}-2\|\xi^{s+1}\|^2_{L^2(\Omega)}+\frac{1}{2}\|\xi^{s}\|^2_{L^2(\Omega)}+\|\xi^{s+2}-\xi^{s+1}\|^2_{L^2(\Omega)} \quad (5.2.84)\\
&\quad-\|\xi^{s+1}-\xi^{s}\|^2_{L^2(\Omega)}+2\tau\varepsilon\left(|\xi^{s+2}|^2_{H^1(\Omega,\mathcal{T}_h)}+J_h^\sigma(\xi^{s+2},\xi^{s+2})\right)\\
&\quad\leq\tau\|\xi^{s+2}\|^2_{L^2(\Omega)}+\tau\frac{K}{\varepsilon}\sum_{v=0}^{1}\|\xi^{s+v}\|^2_{L^2(\Omega)}\\
&\quad+\tau\varepsilon\left(|\xi^{s+2}|^2_{H^1(\Omega,\mathcal{T}_h)}+\varepsilon J_h^\sigma(\xi^{s+2},\xi^{s+2})\right)+\tau Kq(\varepsilon,h,\tau,2).
\end{aligned}
$$

We want to estimate $\|\xi^m\|^2_{L^2(\Omega)}$,where $m=2,\dots,r$ is arbitrary, but fixed. For simplicity we modify (5.2.84) by adding some non-negative terms to RHS and omitting some non-negative terms from LHS

$$
\begin{aligned}
&\frac{3}{2}\|\xi^{s+2}\|^2_{L^2(\Omega)}-2\|\xi^{s+1}\|^2_{L^2(\Omega)}+\frac{1}{2}\|\xi^{s}\|^2_{L^2(\Omega)}+\|\xi^{s+2}-\xi^{s+1}\|^2_{L^2(\Omega)}-\|\xi^{s+1}-\xi^{s}\|^2_{L^2(\Omega)} \quad (5.2.85)\\
&\quad\leq\tau\|\xi^{s+2}\|^2_{L^2(\Omega)}+\tau\frac{K}{\varepsilon}\sum_{v=0}^{1}\|\xi^{s+v}\|^2_{L^2(\Omega)}+\tau Kq(\varepsilon,h,\tau,2).
\end{aligned}
$$

where $s=0,\dots,m-2$.

First we multiply (5.2.85) by

$$
\gamma_j=1-\left(\frac{1}{3}\right)^{j+1} \quad (5.2.86)
$$

where $j=m-s-2$. This sequence γ_j satisfies condition (3.4.69) and is bounded because of Lemma 7. We can see that $\gamma_j\geq 0$ for $j=0,1,\dots$.
After summing over $s=0,\dots,m-2$ we get

$$
\begin{aligned}
&\|\xi^m\|^2_{L^2(\Omega)}+(\frac{1}{2}\gamma_{m-3}-2\gamma_{m-2})\|\xi^1\|^2_{L^2(\Omega)}+\frac{1}{2}\gamma_{m-2}\|\xi^0\|^2_{L^2(\Omega)}+\frac{2}{3}\|\xi^m-\xi^{m-1}\|^2_{L^2(\Omega)} \quad (5.2.87)\\
&+2\sum_{j=0}^{m-3}\left(\frac{1}{3}\right)^{j+2}\|\xi^{m-j-1}-\xi^{m-j-2}\|^2_{L^2(\Omega)}-\gamma_{m-2}\|\xi^1-\xi^0\|^2_{L^2(\Omega)}\leq\tau\gamma_0\|\xi^m\|^2_{L^2(\Omega)}
\end{aligned}
$$

$$+\tau \sum_{s=2}^{m-3} \gamma_{m-2-s} \|\xi^{s+2}\|^2_{L^2(\Omega)} + \tau \frac{K}{\varepsilon} \sum_{s=0}^{m-2} \gamma_{m-2-s} \left(\|\xi^{s+1}\|^2_{L^2(\Omega)} + \|\xi^s\|^2_{L^2(\Omega)} \right)$$
$$+\tau K \sum_{s=0}^{m-2} \gamma_{m-2-s} q(\varepsilon, h, \tau, 2).$$

Using $|\gamma_j| \leq \Gamma$,putting some terms together and omitting some non-negative terms on the left side implies

$$(1 - \tau\gamma_0)\|\xi^m\|^2_{L^2(\Omega)} \leq \tau\Gamma \left(1 + \frac{2K}{\varepsilon}\right) \sum_{s=0}^{m-1} \|\xi^s\|^2_{L^2(\Omega)} \tag{5.2.88}$$
$$+4\Gamma\|\xi^1\|^2_{L^2(\Omega)} + \frac{3}{2}\Gamma\|\xi^0\|^2_{L^2(\Omega)} + T\Gamma K q(\varepsilon, h, \tau, 2).$$

If we set

$$X = \Gamma \frac{2K/\varepsilon + 1}{1 - \tau\gamma_0}, Y = \Gamma \frac{\frac{3}{2}\|\xi^0\|^2_{L^2(\Omega)} + 4\|\xi^1\|^2_{L^2(\Omega)}}{1 - \tau\gamma_0}, Z = \frac{T\Gamma K}{1 - \tau\gamma_0} \tag{5.2.89}$$

we can write

$$\|\xi^m\|^2_{L^2(\Omega)} \leq \tau X \sum_{s=0}^{m-1} \|\xi^s\|^2_{L^2(\Omega)} + Y + Zq(\varepsilon, h, \tau, 2). \tag{5.2.90}$$

It is simple to see that for $s = 0, 1$

$$\|\xi^s\|^2_{L^2(\Omega)} \leq (Y + Zq(\varepsilon, h, \tau, 2))(1 + \tau X)^s. \tag{5.2.91}$$

We want to prove this estimate for $s = m$. We will prove it by induction.
Let's assume that this estimate holds for $s = 0, \ldots, m-1$. Then we can substitute to (5.2.90) and derive

$$\|\xi^m\|^2_{L^2(\Omega)} \leq Y + Zq(\varepsilon, h, \tau, 2) + \tau X(Y + Zq(\varepsilon, h, \tau, 2)) \sum_{s=0}^{m-1} (1 + \tau X)^s \tag{5.2.92}$$
$$= Y + Zq(\varepsilon, h, \tau, 2) + \tau X(Y + Zq(\varepsilon, h, \tau, 2)) \frac{(1 + \tau X)^m - 1}{\tau X}$$
$$= (Y + Zq(\varepsilon, h, \tau, 2))(1 + \tau X)^m.$$

Using

$$\frac{1}{1 - \tau\gamma_0} \leq 3, \tag{5.2.93}$$
$$(1 + \tau X)^s \leq e^{s\tau X} \leq e^{TX}, \tag{5.2.94}$$
$$\|\xi^s\|^2_{L^2(\Omega)} \leq 2\|e_h^s\|^2_{L^2(\Omega)} + 2\|\eta^s\|^2_{L^2(\Omega)} \tag{5.2.95}$$

we gain

$$\|\xi^m\|^2_{L^2(\Omega)} \le (Y + Zq(\varepsilon,h,\tau,2))e^{TX} \tag{5.2.96}$$

$$\le (\frac{9}{2}\Gamma\|\xi^0\|^2_{L^2(\Omega)} + 12\Gamma\|\xi^1\|^2_{L^2(\Omega)} + 3T\Gamma Kq(\varepsilon,h,\tau,2))e^{3T\Gamma(2K/\varepsilon+1)}$$

$$\le (9\Gamma\|e_h^0\|^2_{L^2(\Omega)} + 24\Gamma\|e_h^1\|^2_{L^2(\Omega)} + 9\Gamma\|\eta^0\|^2_{L^2(\Omega)}$$

$$+24\Gamma\|\eta^1\|^2_{L^2(\Omega)} + 3T\Gamma Kq(\varepsilon,h,\tau,2))e^{3T\Gamma(2K/\varepsilon+1)}$$

By (5.1.1), (4.2.8) and (5.1.51),

$$\|\eta^s\|^2_{L^2(\Omega)} \quad \le \quad C_6^2 h^{2(p+1)} |u^s|^2_{H^{p+1}(\Omega)} \le C_6^2 C_{20}^2 h^{2(p+1)}. \tag{5.2.97}$$

Now, using (5.1.55), (5.1.1), (5.2.96) and (5.2.97), we find that

$$\begin{aligned} \|e\|^2_{h,\tau,L^\infty(L^2)} &\le 2\max_{s=0,\dots,r}\left(\|\xi^s\|^2_{L^2(\Omega)} + \|\eta^s\|^2_{L^2(\Omega)}\right) \\ &\le 2(9\Gamma\|e_h^0\|^2_{L^2(\Omega)} + 24\Gamma\|e_h^1\|^2_{L^2(\Omega)} + C_{25a}h^{2(p+1)} \\ &\quad +3T\Gamma Kq(\varepsilon,h,\tau,2))e^{3\Gamma T(2K/\varepsilon+1)} \end{aligned} \tag{5.2.98}$$

with $C_{25a} = C_6^2 C_{20}^2(1+33\Gamma)$, which implies estimate (5.2.80). □

5.3 Error estimation of the third order

Lemma 21 *Let assumptions (2.0.4), a) - e), (H) from Section 3.3 and (A1)-(A2) from Section 4.1 be satisfied. Let u be the exact solution of problem (2.0.11) satisfying (2.0.13). Let $t_s = s\tau$, $s = 0,1,\dots,r$, $\tau = T/r$, be a time partition of $[0,T]$ and let u_h^s, $s = 0,\dots,r$ be the approximate solution defined by (3.5.72) with $k = 3$. Then for $m = 3,\dots,r$*

$$\sum_{s=2}^{m-1}\|\xi^s - \xi^{s-1}\|^2_{L^2(\Omega)} \le \tau\left(3 + \frac{9K}{4\varepsilon}\right)\sum_{s=0}^{m-1}\|\xi^s\|^2_{L^2(\Omega)} + \frac{3}{4}TKq(\varepsilon,h,\tau,3) \tag{5.3.99}$$

$$+\frac{23}{4}\|\xi^2\|^2_{L^2(\Omega)} + \frac{29}{4}\|\xi^1\|^2_{L^2(\Omega)} + \frac{3}{2}\|\xi^0\|^2_{L^2(\Omega)} + \frac{3}{4}\tau\varepsilon\left(|\xi^2|^2_{H^1(\Omega,\mathcal{T}_h)} + J_h^\sigma(\xi^2,\xi^2)\right),$$

where K is defined in Lemma 19.

Proof. When we set $w_h = \xi^{s+3} - \xi^{s+2}$ in (5.1.56) with the notation $y^s = \xi^s - \xi^{s-1}$ for $s = 1,\dots,r$, we get

$$2\left(\frac{11}{6}\xi^{s+3} - \frac{18}{6}\xi^{s+2} + \frac{9}{6}\xi^{s+1} - \frac{2}{6}\xi^s, y^{s+3}\right) + 2\tau\left(a_h(\xi^{s+3},y^{s+3}) + \varepsilon J_h^\sigma(\xi^{s+3},y^{s+3})\right) \tag{5.3.100}$$

$$\le \tau\|y^{s+3}\|^2_{L^2(\Omega)} + \tau\frac{K}{\varepsilon}\sum_{v=0}^{2}\|\xi^{s+v}\|^2_{L^2(\Omega)}$$

$$+\tau\varepsilon\left(|y^{s+3}|^2_{H^1(\Omega,\mathcal{T}_h)} + J_h^\sigma(y^{s+3},y^{s+3})\right) + \tau Kq(\varepsilon,h,\tau,3).$$

When we rewrite (5.3.100) using

$$2\left(\frac{11}{6}\xi^{s+3}-\frac{18}{6}\xi^{s+2}+\frac{9}{6}\xi^{s+1}-\frac{2}{6}\xi^{s},y^{s+3}\right)=2\left(\frac{11}{6}y^{s+3}-\frac{7}{6}y^{s+2}+\frac{2}{6}y^{s+1},y^{s+3}\right) \tag{5.3.101}$$

$$\begin{aligned}
&= 2\|y^{s+3}\|^2_{L^2(\Omega)}+\frac{5}{6}\|y^{s+3}\|^2_{L^2(\Omega)}-\frac{7}{6}\|y^{s+2}\|^2_{L^2(\Omega)}+\frac{2}{6}\|y^{s+1}\|^2_{L^2(\Omega)}\\
&\quad +\frac{7}{6}\|y^{s+3}-y^{s+2}\|^2_{L^2(\Omega)}-\frac{2}{6}\|y^{s+3}-y^{s+1}\|^2_{L^2(\Omega)}\\
&\geq \frac{17}{6}\|y^{s+3}\|^2_{L^2(\Omega)}-\frac{7}{6}\|y^{s+2}\|^2_{L^2(\Omega)}+\frac{2}{6}\|y^{s+1}\|^2_{L^2(\Omega)}\\
&\quad +\frac{3}{6}\|y^{s+3}-y^{s+2}\|^2_{L^2(\Omega)}-\frac{4}{6}\|y^{s+2}-y^{s+1}\|^2_{L^2(\Omega)}\\
&\geq \frac{17}{6}\|y^{s+3}\|^2_{L^2(\Omega)}-\frac{9}{6}\|y^{s+2}\|^2_{L^2(\Omega)}+\frac{3}{6}\|y^{s+3}-y^{s+2}\|^2_{L^2(\Omega)}-\frac{3}{6}\|y^{s+2}-y^{s+1}\|^2_{L^2(\Omega)}
\end{aligned}$$

and

$$\begin{aligned}
2\tau\left(a_h(\xi^{s+3},y^{s+3})+\varepsilon J_h^\sigma(\xi^{s+3},y^{s+3})\right)&=\tau\left(a_h(\xi^{s+3},\xi^{s+3})+\varepsilon J_h^\sigma(\xi^{s+3},\xi^{s+3})\right) \qquad (5.3.102)\\
-\tau\left(a_h(\xi^{s+2},\xi^{s+2})+\varepsilon J_h^\sigma(\xi^{s+2},\xi^{s+2})\right)&+\tau\left(a_h(y^{s+3},y^{s+3})+\varepsilon J_h^\sigma(y^{s+3},y^{s+3})\right)
\end{aligned}$$

and (5.2.82), we get

$$\begin{aligned}
&\frac{17}{6}\|y^{s+3}\|^2_{L^2(\Omega)}-\frac{9}{6}\|y^{s+2}\|^2_{L^2(\Omega)}+\frac{3}{6}\|y^{s+3}-y^{s+2}\|^2_{L^2(\Omega)}-\frac{3}{6}\|y^{s+2}-y^{s+1}\|^2_{L^2(\Omega)} \qquad (5.3.103)\\
&\quad +\tau\varepsilon\left(|\xi^{s+3}|^2_{H^1(\Omega,\mathcal{T}_h)}+J_h^\sigma(\xi^{s+3},\xi^{s+3})-|\xi^{s+2}|^2_{H^1(\Omega,\mathcal{T}_h)}+J_h^\sigma(\xi^{s+2},\xi^{s+2})\right)\\
&\quad \leq 2\tau\left(\|\xi^{s+3}\|^2_{L^2(\Omega)}+\|\xi^{s+2}\|^2_{L^2(\Omega)}\right)+\tau\frac{K}{\varepsilon}\sum_{v=0}^{2}\|\xi^{s+v}\|^2_{L^2(\Omega)}+\tau Kq(\varepsilon,h,\tau,3).
\end{aligned}$$

After summing over $s=0,\dots,m-4$ we get

$$\begin{aligned}
&\frac{8}{6}\sum_{s=0}^{m-4}\|y^{s+3}\|^2_{L^2(\Omega)}+\frac{9}{6}\|y^{m-1}\|^2_{L^2(\Omega)}-\frac{9}{6}\|y^{2}\|^2_{L^2(\Omega)}+\frac{3}{6}\|y^{m-1}-y^{m-2}\|^2_{L^2(\Omega)} \qquad (5.3.104)\\
&-\frac{3}{6}\|y^{2}-y^{1}\|^2_{L^2(\Omega)}+\tau\varepsilon\left(|\xi^{m-1}|^2_{H^1(\Omega,\mathcal{T}_h)}+J_h^\sigma(\xi^{m-1},\xi^{m-1})-|\xi^{2}|^2_{H^1(\Omega,\mathcal{T}_h)}+J_h^\sigma(\xi^{2},\xi^{2})\right)\\
&\quad \leq 2\tau\sum_{s=0}^{m-4}\left(\|\xi^{s+3}\|^2_{L^2(\Omega)}+\|\xi^{s+2}\|^2_{L^2(\Omega)}\right)+3\tau\frac{K}{\varepsilon}\sum_{s=0}^{m-2}\|\xi^{s}\|^2_{L^2(\Omega)}+TKq(\varepsilon,h,\tau,3)\\
&\quad \leq \tau\left(4+3\frac{K}{\varepsilon}\right)\sum_{s=0}^{m-1}\|\xi^{s}\|^2_{L^2(\Omega)}+TKq(\varepsilon,h,\tau,3).
\end{aligned}$$

Omitting some non-negative terms on the left-hand side of (5.3.104) and moving the initial terms to right we get

$$\frac{8}{6}\sum_{s=2}^{m-1}\|\xi^{s}-\xi^{s-1}\|^2_{L^2(\Omega)}=\frac{8}{6}\sum_{s=2}^{m-1}\|y^{s}\|^2_{L^2(\Omega)} \tag{5.3.105}$$

$$\leq \frac{8}{6}\|y^2\|^2_{L^2(\Omega)} + \tau\left(4 + 3\frac{K}{\varepsilon}\right)\sum_{s=0}^{m-1}\|\xi^s\|^2_{L^2(\Omega)} + TKq(\varepsilon, h, \tau, 3)$$
$$+\frac{9}{6}\|y^2\|^2_{L^2(\Omega)} + \frac{3}{6}\|y^2 - y^1\|^2_{L^2(\Omega)} + \tau\varepsilon\left(|\xi^2|^2_{H^1(\Omega,\mathcal{T}_h)} + J_h^\sigma(\xi^2, \xi^2)\right)$$
$$\leq \tau\left(4 + 3\frac{K}{\varepsilon}\right)\sum_{s=0}^{m-1}\|\xi^s\|^2_{L^2(\Omega)} + TKq(\varepsilon, h, \tau, 3)$$
$$+\frac{23}{6}\|y^2\|^2_{L^2(\Omega)} + \|y^1\|^2_{L^2(\Omega)} + \tau\varepsilon\left(|\xi^2|^2_{H^1(\Omega,\mathcal{T}_h)} + J_h^\sigma(\xi^2, \xi^2)\right)$$
$$\leq \tau\left(4 + 3\frac{K}{\varepsilon}\right)\sum_{s=0}^{m-1}\|\xi^s\|^2_{L^2(\Omega)} + TKq(\varepsilon, h, \tau, 3)$$
$$\frac{23}{3}\|\xi^2\|^2_{L^2(\Omega)} + \frac{29}{3}\|\xi^1\|^2_{L^2(\Omega)} + 2\|\xi^0\|^2_{L^2(\Omega)} + \tau\varepsilon\left(|\xi^2|^2_{H^1(\Omega,\mathcal{T}_h)} + J_h^\sigma(\xi^2, \xi^2)\right)$$

which implies the estimate (5.3.99). □

Theorem 2 *Let assumptions (2.0.4), a) - e), (H) from Section 3.3 and (A1)-(A2) from Section 4.1 be satisfied. Let u be the exact solution of problem (2.0.11) satisfying (2.0.13). Let $t_s = s\tau$, $s = 0, 1, \ldots, r$, $\tau = T/r$, be a time partition of $[0, T]$ and let u_h^s, $s = 0, \ldots, r$ be the approximate solution defined by (3.5.72) with $k = 3$ and let $\tau \leq 1$. Then there exists constant $\tilde{C} = O\left(\exp(T(30 + 117K/4\varepsilon))\right)$ such that*

$$\|e\|^2_{h,\tau,L^\infty(L^2)} \leq \tilde{C}\left(h^{2p}\left(\varepsilon + h^2 + h^2/\varepsilon\right) + \tau^6\left(1 + 1/\varepsilon\right)\right) \tag{5.3.106}$$
$$+\tilde{C}\left(\max_{s=0,1,2}(\|e_h^s\|^2_{L^2(\Omega)}) + \tau\varepsilon\max_{s=0,1,2}(|e_h^s|^2_{H^1(\Omega,\mathcal{T}_h)} + J_h^\sigma(e_h^s, e_h^s))\right),$$

where K is defined in Lemma 19.

Proof. It is easy to see that for $s = 0, 1, 2$ holds (5.3.106). So from this moment we should estimate $\|e_h^s\|^2_{L^2(\Omega)}$ for $s = 3, \ldots, r$. When we use notation (5.1.1) and (5.1.2), then to obtain the estimate for $\|e_h^s\|^2_{L^2(\Omega)}$, we can estimate $\|\xi^s\|^2_{L^2(\Omega)}$ and $\|\eta^s\|^2_{L^2(\Omega)}$ induvidually. First we will use (5.1.56) and we substitute $w_h = \xi^{s+3}$. Then we get

$$2\left(\frac{11}{6}\xi^{s+3} - \frac{18}{6}\xi^{s+2} + \frac{9}{6}\xi^{s+1} - \frac{2}{6}\xi^s, \xi^{s+3}\right) + 2\tau\left(a_h(\xi^{s+3}, \xi^{s+3}) + \varepsilon J_h^\sigma(\xi^{s+3}, \xi^{s+3})\right) \tag{5.3.107}$$
$$\leq \tau\|\xi^{s+3}\|^2_{L^2(\Omega)} + \tau\frac{K}{\varepsilon}\sum_{v=0}^{2}\|\xi^{s+v}\|^2_{L^2(\Omega)}$$
$$+\tau\varepsilon\left(|\xi^{s+3}|^2_{H^1(\Omega,\mathcal{T}_h)} + J_h^\sigma(\xi^{s+3}, \xi^{s+3})\right) + \tau Kq(\varepsilon, h, \tau, 3).$$

With the relations

$$a_h(\xi^{s+3}, \xi^{s+3}) = \varepsilon|\xi^{s+3}|^2_{H^1(\Omega,\mathcal{T}_h)} \tag{5.3.108}$$

and

$$
\begin{aligned}
&2\left(\frac{11}{6}\xi^{s+3}-\frac{18}{6}\xi^{s+2}+\frac{9}{6}\xi^{s+1}-\frac{2}{6}\xi^{s},\xi^{s+3}\right) \\
&=\frac{11}{6}\|\xi^{s+3}\|^2_{L^2(\Omega)}-\frac{18}{6}\|\xi^{s+2}\|^2_{L^2(\Omega)}+\frac{9}{6}\|\xi^{s+1}\|^2_{L^2(\Omega)}-\frac{2}{6}\|\xi^{s}\|^2_{L^2(\Omega)} \\
&+\frac{18}{6}\|\xi^{s+3}-\xi^{s+2}\|^2_{L^2(\Omega)}-\frac{9}{6}\|\xi^{s+3}-\xi^{s+1}\|^2_{L^2(\Omega)}+\frac{2}{6}\|\xi^{s+3}-\xi^{s}\|^2_{L^2(\Omega)} \\
&\geq\frac{11}{6}\|\xi^{s+3}\|^2_{L^2(\Omega)}-\frac{18}{6}\|\xi^{s+2}\|^2_{L^2(\Omega)}+\frac{9}{6}\|\xi^{s+1}\|^2_{L^2(\Omega)}-\frac{2}{6}\|\xi^{s}\|^2_{L^2(\Omega)} \\
&+\frac{18}{6}\|\xi^{s+3}-\xi^{s+2}\|^2_{L^2(\Omega)}-\frac{18}{6}\|\xi^{s+3}-\xi^{s+2}\|^2_{L^2(\Omega)}-\frac{18}{6}\|\xi^{s+2}-\xi^{s+1}\|^2_{L^2(\Omega)} \\
&\geq\frac{11}{6}\|\xi^{s+3}\|^2_{L^2(\Omega)}-\frac{18}{6}\|\xi^{s+2}\|^2_{L^2(\Omega)}+\frac{9}{6}\|\xi^{s+1}\|^2_{L^2(\Omega)}-\frac{2}{6}\|\xi^{s}\|^2_{L^2(\Omega)} \\
&-\frac{18}{6}\|\xi^{s+2}-\xi^{s+1}\|^2_{L^2(\Omega)}
\end{aligned}
\tag{5.3.109}
$$

we get

$$
\begin{aligned}
&\frac{11}{6}\|\xi^{s+3}\|^2_{L^2(\Omega)}-\frac{18}{6}\|\xi^{s+2}\|^2_{L^2(\Omega)}+\frac{9}{6}\|\xi^{s+1}\|^2_{L^2(\Omega)}-\frac{2}{6}\|\xi^{s}\|^2_{L^2(\Omega)} \\
&+\tau\varepsilon\left(|\xi^{s+3}|^2_{H^1(\Omega,\mathcal{T}_h)}+J_h^\sigma(\xi^{s+3},\xi^{s+3})\right) \\
&\leq\tau\|\xi^{s+3}\|^2_{L^2(\Omega)}+\tau\frac{K}{\varepsilon}\sum_{v=0}^{2}\|\xi^{s+v}\|^2_{L^2(\Omega)}+3\|\xi^{s+2}-\xi^{s+1}\|^2_{L^2(\Omega)}+\tau Kq(\varepsilon,h,\tau,3).
\end{aligned}
\tag{5.3.110}
$$

When we omit non-negative term $\tau\varepsilon\left(|\xi^{s+3}|^2_{H^1(\Omega,\mathcal{T}_h)}+J_h^\sigma(\xi^{s+3},\xi^{s+3})\right)$ we finally get

$$
\begin{aligned}
&\frac{11}{6}\|\xi^{s+3}\|^2_{L^2(\Omega)}-\frac{18}{6}\|\xi^{s+2}\|^2_{L^2(\Omega)}+\frac{9}{6}\|\xi^{s+1}\|^2_{L^2(\Omega)}-\frac{2}{6}\|\xi^{s}\|^2_{L^2(\Omega)} \\
&\leq\tau\|\xi^{s+3}\|^2_{L^2(\Omega)}+\tau\frac{K}{\varepsilon}\sum_{v=0}^{2}\|\xi^{s+v}\|^2_{L^2(\Omega)}+3\|\xi^{s+2}-\xi^{s+1}\|^2_{L^2(\Omega)}+\tau Kq(\varepsilon,h,\tau,3).
\end{aligned}
\tag{5.3.111}
$$

There exist the sequence γ_j such that γ_j is the solution of

$$
\begin{aligned}
\gamma_{-2},\gamma_{-1}&=0 \\
\gamma_0&=\frac{6}{11} \\
\frac{11}{6}\gamma_{j+3}-\frac{18}{6}\gamma_{j+2}+\frac{9}{6}\gamma_{j+1}-\frac{2}{6}\gamma_j&=0
\end{aligned}
\tag{5.3.112}
$$

for $j=-2,-1,0,1,\ldots$. From Lemma 7 we know that there exist constant Γ such that $|\gamma_j|\leq\Gamma$ for $j=0,1,\ldots$. From Lemma 20 follows that $\gamma_j\geq 0$ for $j=0,1,\ldots$. We multiply (5.3.111) with γ_{m-3-s} and after summing over $s=0,\ldots,m-3$ using the property of γ_j (5.3.112) we get

$$
\|\xi^m\|^2_{L^2(\Omega)}+\left(-\frac{18}{6}\gamma_{m-3}+\frac{9}{6}\gamma_{m-4}-\frac{2}{6}\gamma_{m-5}\right)\|\xi^2\|^2_{L^2(\Omega)}
\tag{5.3.113}
$$

$$
\begin{aligned}
&+\left(\frac{9}{6}\gamma_{m-3}-\frac{2}{6}\gamma_{m-4}\right)\|\xi^1\|^2_{L^2(\Omega)}-\frac{2}{6}\gamma_{m-3}\|\xi^0\|^2_{L^2(\Omega)}\\
\leq{}& \gamma_0\tau\|\xi^m\|^2_{L^2(\Omega)}+\tau\sum_{s=0}^{m-4}\gamma_{m-3-s}\|\xi^{s+3}\|^2_{L^2(\Omega)}+\tau\frac{K}{\varepsilon}\sum_{s=0}^{m-3}\gamma_{m-3-s}\sum_{v=0}^{2}\|\xi^{s+v}\|^2_{L^2(\Omega)}\\
&+3\sum_{s=0}^{m-3}\gamma_{m-3-s}\|\xi^{s+2}-\xi^{s+1}\|^2_{L^2(\Omega)}+\tau K\sum_{s=0}^{m-3}\gamma_{m-3-s}q(\varepsilon,h,\tau,3).
\end{aligned}
$$

Omitting some non-negative terms from the left-hand side of (5.3.113) and estimating terms on the right-hand side we have

$$
\begin{aligned}
\|\xi^m\|^2_{L^2(\Omega)} \leq{}& \tau\gamma_0\|\xi^m\|^2_{L^2(\Omega)}+\tau\Gamma\left(1+\frac{3K}{\varepsilon}\right)\sum_{s=0}^{m-1}\|\xi^s\|^2_{L^2(\Omega)} \\
&+3\Gamma\sum_{s=2}^{m-1}\|\xi^s-\xi^{s-1}\|^2_{L^2(\Omega)}+T\Gamma Kq(\varepsilon,h,\tau,3)\\
&+\frac{20}{6}\Gamma\|\xi^2\|^2_{L^2(\Omega)}+\frac{2}{6}\Gamma\|\xi^1\|^2_{L^2(\Omega)}+\frac{2}{6}\Gamma\|\xi^0\|^2_{L^2(\Omega)}
\end{aligned}
\tag{5.3.114}
$$

When we estimate (5.3.114) with (5.3.99) we get

$$
\begin{aligned}
\|\xi^m\|^2_{L^2(\Omega)} \leq{}& \tau\gamma_0\|\xi^m\|^2_{L^2(\Omega)}+\tau\Gamma\left(10+\frac{39K}{4\varepsilon}\right)\sum_{s=0}^{m-1}\|\xi^s\|^2_{L^2(\Omega)} \\
&+\frac{13}{4}T\Gamma Kq(\varepsilon,h,\tau,3)+\frac{9}{4}\Gamma\tau\varepsilon\left(|\xi^2|^2_{H^1(\Omega,\mathcal{T}_h)}+J_h^\sigma(\xi^2,\xi^2)\right)\\
&+\frac{247}{12}\Gamma\|\xi^2\|^2_{L^2(\Omega)}+\frac{265}{12}\Gamma\|\xi^1\|^2_{L^2(\Omega)}+\frac{29}{6}\Gamma\|\xi^0\|^2_{L^2(\Omega)}.
\end{aligned}
\tag{5.3.115}
$$

If we set

$$
X=\Gamma\frac{\frac{39K}{4\varepsilon}+10}{1-\tau\gamma_0}, Z=\frac{13}{4}\frac{T\Gamma K}{1-\tau\gamma_0}
\tag{5.3.116}
$$

$$
Y=\Gamma\frac{\frac{9}{4}\tau\varepsilon\left(|\xi^2|^2_{H^1(\Omega,\mathcal{T}_h)}+J_h^\sigma(\xi^2,\xi^2)\right)+\frac{247}{12}\|\xi^2\|^2_{L^2(\Omega)}+\frac{265}{12}\|\xi^1\|^2_{L^2(\Omega)}+\frac{29}{6}\|\xi^0\|^2_{L^2(\Omega)}}{1-\tau\gamma_0}
$$

we can write

$$
\|\xi^m\|^2_{L^2(\Omega)}\leq\tau X\sum_{s=0}^{m-1}\|\xi^s\|^2_{L^2(\Omega)}+Y+Zq(\varepsilon,h,\tau,3).
\tag{5.3.117}
$$

It is simple to see that for $s=0,1,2$

$$
\|\xi^s\|^2_{L^2(\Omega)}\leq(Y+Zq(\varepsilon,h,\tau,3))(1+\tau X)^s.
\tag{5.3.118}
$$

We want to prove this estimate for $s = m$. We will prove it by induction. Let's assume that this estimate holds for $s = 0, \dots, m-1$. Then we can substitute to (5.3.117) and derive

$$\begin{aligned}\|\xi^m\|^2_{L^2(\Omega)} &\le Y + Zq(\varepsilon,h,\tau,3) + \tau X(Y + Zq(\varepsilon,h,\tau,3))\sum_{s=0}^{m-1}(1+\tau X)^s \\ &= Y + Zq(\varepsilon,h,\tau,3) + \tau X(Y + Zq(\varepsilon,h,\tau,3))\frac{(1+\tau X)^m - 1}{\tau X} \\ &= (Y + Zq(\varepsilon,h,\tau,3))(1+\tau X)^m.\end{aligned} \tag{5.3.119}$$

Using

$$\frac{1}{1-\tau\gamma_0} \le 3, \tag{5.3.120}$$

$$(1+\tau X)^s \le e^{s\tau X} \le e^{TX} \tag{5.3.121}$$

$$\|\xi^s\|^2_{L^2(\Omega)} \le 2\|e_h^s\|^2_{L^2(\Omega)} + 2\|\eta^s\|^2_{L^2(\Omega)} \tag{5.3.122}$$

$$\begin{aligned}\tau\varepsilon\left(|\xi^s|^2_{H^1(\Omega,\mathcal{T}_h)} + J_h^\sigma(\xi^s,\xi^s)\right) &\le 2\tau\varepsilon\left(|e_h^s|^2_{H^1(\Omega,\mathcal{T}_h)} + J_h^\sigma(e_h^s,e_h^s)\right) \\ &\quad + 2\tau\varepsilon\left(|\eta^s|^2_{H^1(\Omega,\mathcal{T}_h)} + J_h^\sigma(\eta^s,\eta^s)\right)\end{aligned} \tag{5.3.123}$$

we gain

$$\begin{aligned}\|\xi^m\|^2_{L^2(\Omega)} &\le (Y + Zq(\varepsilon,h,\tau,3))e^{TX} \\ &\le (\frac{29}{2}\Gamma\|\xi^0\|^2_{L^2(\Omega)} + \frac{265}{4}\Gamma\|\xi^1\|^2_{L^2(\Omega)} + \frac{247}{4}\Gamma\|\xi^2\|^2_{L^2(\Omega)} \\ &+ \frac{27}{4}\Gamma\tau\varepsilon\left(|\xi^2|^2_{H^1(\Omega,\mathcal{T}_h)} + J_h^\sigma(\xi^2,\xi^2)\right) + \frac{39}{4}T\Gamma Kq(\varepsilon,h,\tau,3))e^{T\Gamma(117K/4\varepsilon+30)}\end{aligned} \tag{5.3.124}$$

By (5.1.1), (4.2.8) ,(5.1.43) , (5.1.44) and (5.1.51),

$$\|\eta^s\|^2_{L^2(\Omega)} \le C_6^2 h^{2(p+1)}|u^s|^2_{H^{p+1}(\Omega)} \le C_6^2 C_{20}^2 h^{2(p+1)} \tag{5.3.125}$$

$$\begin{aligned}\tau\varepsilon\left(|\eta^s|^2_{H^1(\Omega,\mathcal{T}_h)} + J_h^\sigma(\eta^s,\eta^s)\right) &\le \tau\varepsilon\left(C_6^2 h^{2p}|u^s|^2_{H^{p+1}(\Omega,\mathcal{T}_h)} + C_{18}\, h^{2p}|u^s|^2_{H^{p+1}(\Omega)}\right) \\ &\le \tau\varepsilon C_{27} h^{2p},\end{aligned} \tag{5.3.126}$$

where $C_{27} = C_{20}^2(C_6^2 + C_{18})$. Now, using (5.1.55), (5.1.1), (5.3.124) , (5.3.125) and (5.3.126), we find that

$$\begin{aligned}\|e\|^2_{h,\tau,L^\infty(L^2)} &\le 2\max_{s=0,\dots,r}\left(\|\xi^s\|^2_{L^2(\Omega)} + \|\eta^s\|^2_{L^2(\Omega)}\right) \\ &\le 2C_{25}h^{2(p+1)} + 2(\frac{29}{2}\Gamma\|\xi^0\|^2_{L^2(\Omega)} + \frac{265}{4}\Gamma\|\xi^1\|^2_{L^2(\Omega)}\end{aligned} \tag{5.3.127}$$

$$
\begin{aligned}
&+\frac{247}{4}\Gamma\|\xi^2\|^2_{L^2(\Omega)} + \frac{27}{4}\Gamma\tau\varepsilon\left(|\xi^2|^2_{H^1(\Omega,\mathcal{T}_h)} + J_h^\sigma(\xi^2,\xi^2)\right)\\
&+\frac{39}{4}T\Gamma K q(\varepsilon,h,\tau,3))e^{T\Gamma(117K/4\varepsilon+30)}\\
\leq\ & 2(29\Gamma\|e_h^0\|^2_{L^2(\Omega)} + \frac{265}{2}\Gamma\|e_h^1\|^2_{L^2(\Omega)} + \frac{247}{2}\Gamma\|e_h^2\|^2_{L^2(\Omega)}\\
&+\frac{27}{2}\Gamma\tau\varepsilon\left(|e_h^2|^2_{H^1(\Omega,\mathcal{T}_h)} + J_h^\sigma(e_h^2,e_h^2)\right) + C_{25a}h^{2(p+1)}\\
&+\frac{27}{2}C_{27}\Gamma\tau\varepsilon h^{2p} + \frac{39}{4}T\Gamma K q(\varepsilon,h,\tau,3))e^{T\Gamma(117K/4\varepsilon+30)}
\end{aligned}
$$

with $C_{25} = C_6^2 C_{20}^2$ and $C_{25a} = C_{25}(1+285\Gamma)$, which implies estimate (5.3.106). □

5.4 Main result

Theorem 3 *Let assumptions (2.0.4), a) - e), (H) from Section 3.3 and (A1)-(A2) from Section 4.1 be satisfied. Let u be the exact solution of problem (2.0.11) satisfying (2.0.13). Let $t_s = s\tau$, $s = 0,1,\ldots,r$, $\tau = T/r$, be a time partition of $[0,T]$ and let u_h^s, $s = 0,\ldots,r$ be the approximate solution defined by (3.5.72) with $k = 2,3$ and let $\tau \leq 1$. Let*

$$
\max_{s=0,\ldots,k-1}\|e_h^s\|^2_{L^2(\Omega)} \leq C_n(h^{2(p+1)} + \tau^{2k}) \tag{5.4.128}
$$

$$
\max_{s=0,\ldots,k-1}(|e_h^s|^2_{H^1(\Omega,\mathcal{T}_h)} + J_h^\sigma(e_h^s,e_h^s),) \leq C_{nn}(h^{2p} + \tau^{2k}), \tag{5.4.129}
$$

where C_n and C_{nn} are independent of τ,ε,h. Then there exist constants $\tilde{C}$ and $\hat{C}$ such that

$$
\|e\|^2_{h,\tau,L^\infty(L^2)} \leq \tilde{C}\left(h^{2p}\left(\varepsilon + h^2 + \frac{h^2}{\varepsilon}\right) + \tau^{2k}\left(1+\frac{1}{\varepsilon}\right)\right), \tag{5.4.130}
$$

$$
\|e\|^2_{h,\tau,L^2(H^1)} \leq \hat{C}\left(h^{2p}\left(\varepsilon + h^2 + \frac{h^2}{\varepsilon} + 1 + \frac{h^2}{\varepsilon^2}\right) + \tau^{2k}\left(\varepsilon + 1 + \frac{1}{\varepsilon} + \frac{1}{\varepsilon^2}\right)\right) \tag{5.4.131}
$$

Proof. As in (5.1.1), we set $\xi^s = u_h^s - \Pi u^s \in S_h$, $\eta^s = \Pi u^s - u^s$, $s = 0,\ldots,r$. Then (5.1.2) holds: $e_h^s = u_h^s - u^s = \xi^s + \eta^s$.

i) For $k = 2$ (5.4.130) is a direct consequence of Theorem 1 and (5.4.128).

$$
\begin{aligned}
\|e\|^2_{h,\tau,L^\infty(L^2)} \leq\ & 2(33\Gamma\max(\|e_h^0\|^2_{L^2(\Omega)},\|e_h^1\|^2_{L^2(\Omega)}) + C_{25a}h^{2(p+1)}\\
&+3T\Gamma K q(\varepsilon,h,\tau,2))e^{3\Gamma T(2K/\varepsilon+1)}\\
\leq\ & 2(33\Gamma C_n(h^{2(p+1)} + \tau^4) + C_{25a}h^{2(p+1)}\\
&+3T\Gamma K q(\varepsilon,h,\tau,2))e^{3\Gamma T(2K/\varepsilon+1)},
\end{aligned} \tag{5.4.132}
$$

which implies the estimate (5.4.130) for $k = 2$, where $\tilde{C} = O(\exp(3\Gamma T(1+2K/\varepsilon)))$. For $k = 3$ (5.4.130) is a direct consequence of Theorem 2, (5.4.128) and (5.4.129).

$$
\|e\|^2_{h,\tau,L^\infty(L^2)} \leq 2(29\Gamma\|e_h^0\|^2_{L^2(\Omega)} + \frac{265}{2}\Gamma\|e_h^1\|^2_{L^2(\Omega)} + \frac{247}{2}\Gamma\|e_h^2\|^2_{L^2(\Omega)} \tag{5.4.133}
$$

$$
\begin{aligned}
& +\frac{27}{2}\Gamma\tau\varepsilon\left(|e_h^2|^2_{H^1(\Omega,\mathcal{T}_h)} + J_h^\sigma(e_h^2, e_h^2)\right) + C_{25a}h^{2(p+1)} \\
& +\frac{27}{2}C_{27}\Gamma\tau\varepsilon h^{2p} + \frac{39}{4}T\Gamma K q(\varepsilon, h, \tau, 3))e^{T\Gamma(117K/4\varepsilon+30)} \\
\leq\ & (570\Gamma C_n(h^{2(p+1)} + \tau^6) + 2C_{25a}h^{2(p+1)} + 27C_{nn}\tau\varepsilon\Gamma(h^{2p} + \tau^6) \\
& +27C_{27}\tau\varepsilon\Gamma h^{2p} + \frac{39}{2}T\Gamma K q(\varepsilon, h, \tau, 3))e^{\Gamma T(117K/4\varepsilon+30)},
\end{aligned}
$$

which implies the estimate (5.4.130) for $k = 3$, where $\tilde{C} = O\left(\exp(\Gamma T(30 + 117K/4\varepsilon))\right)$.

ii) Now let as derive (5.4.131). When we substitute $w_h = \xi^{s+k}$ in Lemma 19 and use (5.2.82) we have

$$
\begin{aligned}
& 2\left(\sum_{v=0}^{k}\alpha_v\xi^{s+v}, \xi^{s+k}\right) + \tau\varepsilon\left(|\xi^{s+k}|^2_{H^1(\Omega,\mathcal{T}_h)} + J_h^\sigma(\xi^{s+k}, \xi^{s+k})\right) \\
& \leq \tau\|\xi^{s+k}\|^2_{L^2(\Omega)} + \tau\frac{K}{\varepsilon}\sum_{v=0}^{k-1}\|\xi^{s+v}\|^2_{L^2(\Omega)} + \tau K q(\varepsilon, h, \tau, k),
\end{aligned} \tag{5.4.134}
$$

For $k = 2$ using the same step as in (5.2.83) we get

$$
\begin{aligned}
& \frac{3}{2}\|\xi^{s+2}\|^2_{L^2(\Omega)} - 2\|\xi^{s+1}\|^2_{L^2(\Omega)} + \frac{1}{2}\|\xi^s\|^2_{L^2(\Omega)} + \|\xi^{s+2} - \xi^{s+1}\|^2_{L^2(\Omega)} \\
& -\|\xi^{s+1} - \xi^s\|^2_{L^2(\Omega)} + \tau\varepsilon\left(|\xi^{s+2}|^2_{H^1(\Omega,\mathcal{T}_h)} + J_h^\sigma(\xi^{s+2}, \xi^{s+2})\right) \\
& \leq \tau\|\xi^{s+2}\|^2_{L^2(\Omega)} + \tau\frac{K}{\varepsilon}\sum_{v=0}^{1}\|\xi^{s+v}\|^2_{L^2(\Omega)} + \tau K q(\varepsilon, h, \tau, 2).
\end{aligned} \tag{5.4.135}
$$

Now after summing (5.4.135) over $s = 0, \ldots, r-2$ we obtain

$$
\begin{aligned}
& \frac{3}{2}\|\xi^r\|^2_{L^2(\Omega)} - \frac{1}{2}\|\xi^{r-1}\|^2_{L^2(\Omega)} - \frac{3}{2}\|\xi^1\|^2_{L^2(\Omega)} + \frac{1}{2}\|\xi^0\|^2_{L^2(\Omega)} + \|\xi^r - \xi^{r-1}\|^2_{L^2(\Omega)} \\
& -\|\xi^1 - \xi^0\|^2_{L^2(\Omega)} + \tau\varepsilon\sum_{s=2}^{r}\left(|\xi^s|^2_{H^1(\Omega,\mathcal{T}_h)} + J_h^\sigma(\xi^s, \xi^s)\right) \\
& \leq \tau\sum_{s=2}^{r}\|\xi^s\|^2_{L^2(\Omega)} + \tau\frac{K}{\varepsilon}\sum_{s=1}^{r-1}\left(\|\xi^s\|^2_{L^2(\Omega)} + \|\xi^{s-1}\|^2_{L^2(\Omega)}\right) + TKq(\varepsilon, h, \tau, 2).
\end{aligned} \tag{5.4.136}
$$

From that follows

$$
\begin{aligned}
& \tau\varepsilon\sum_{s=2}^{r}\left(|\xi^s|^2_{H^1(\Omega,\mathcal{T}_h)} + J_h^\sigma(\xi^s, \xi^s)\right) \leq \frac{1}{2}\|\xi^{r-1}\|^2_{L^2(\Omega)} + \frac{7}{2}\|\xi^1\|^2_{L^2(\Omega)} + \frac{3}{2}\|\xi^0\|^2_{L^2(\Omega)} \\
& +\tau\sum_{s=2}^{r}\|\xi^s\|^2_{L^2(\Omega)} + \tau\frac{2K}{\varepsilon}\sum_{s=0}^{r-1}\|\xi^s\|^2_{L^2(\Omega)} + TKq(\varepsilon, h, \tau, 2).
\end{aligned} \tag{5.4.137}
$$

When we use

$$
\|\xi^s\|^2_{L^2(\Omega)} \leq 2\|e\|^2_{h,\tau,L^\infty(L^2)} + 2C_{25}h^{2(p+1)}, \tag{5.4.138}
$$

where $C_{25} = C_6^2 C_{20}^2$, we get

$$\tau\varepsilon \sum_{s=2}^{r} \left(|\xi^s|^2_{H^1(\Omega,\mathcal{T}_h)} + J_h^\sigma(\xi^s,\xi^s) \right) \tag{5.4.139}$$
$$\leq \left(11 + 2T + 4T\frac{K}{\varepsilon} \right) \left(\|e\|^2_{h,\tau,L^\infty(L^2)} + C_{25}h^{2(p+1)} \right) \quad + \quad TKq(\varepsilon,h,\tau,2).$$

Then (5.4.139), (5.4.129) and

$$|\xi^s|^2_{H^1(\Omega,\mathcal{T}_h)} + J_h^\sigma(\xi^s,\xi^s) \leq 2|e_h^s|^2_{H^1(\Omega,\mathcal{T}_h)} + 2J_h^\sigma(e_h^s,e_h^s) + 2C_{27}h^{2p} \tag{5.4.140}$$

give

$$\tau\varepsilon \sum_{s=0}^{r} \left(|\xi^s|^2_{H^1(\Omega,\mathcal{T}_h)} + J_h^\sigma(\xi^s,\xi^s) \right) \leq \left(11 + 2T + 4T\frac{K}{\varepsilon} \right) \left(\|e\|^2_{h,\tau,L^\infty(L^2)} + C_{25}h^{2(p+1)} \right) \tag{5.4.141}$$
$$+TKq(\varepsilon,h,\tau,2) + 4C_{nn}(h^{2p}+\tau^4) + 4\tau\varepsilon C_{27}h^{2p}.$$

Using (5.4.141), (5.4.132) and notation $C_{hx} = \max(11+2T, 4TK)$,
$C_p = 33\Gamma C_n + C_{25a}$ and $C_{pp} = 4C_{nn} + 4C_{27}$ we get

$$\tau\varepsilon \textstyle\sum_{s=0}^{r} \left(|\xi^s|^2_{H^1(\Omega,\mathcal{T}_h)} + J_h^\sigma(\xi^s,\xi^s) \right) \leq \tag{5.4.142}$$
$$C_{hx}(1+\tfrac{1}{\varepsilon})(33\Gamma C_n(h^{2(p+1)}+\tau^4) + C_{25a}h^{2(p+1)} + 3T\Gamma Kq(\varepsilon,h,\tau,2))e^{3\Gamma T(2K/\varepsilon+1)}$$
$$+C_{hx}(1+\tfrac{1}{\varepsilon})C_{25}h^{2(p+1)} + TKq(\varepsilon,h,\tau,2) + 4C_{nn}(h^{2p}+\tau^4) + 4\tau\varepsilon C_{27}h^{2p}$$
$$\leq 2C_{hx}(1+\tfrac{1}{\varepsilon})(33\Gamma C_n(h^{2(p+1)}+\tau^4) + C_{25a}h^{2(p+1)}$$
$$+(3\Gamma+1)TKq(\varepsilon,h,\tau,2) + 4C_{nn}(h^{2p}+\tau^4) + 4\tau\varepsilon C_{27}h^{2p})e^{3\Gamma T(2K/\varepsilon+1)}$$
$$\leq 2C_{hx}(1+\tfrac{1}{\varepsilon})(C_p(h^{2(p+1)}+\tau^4) + C_{pp}(1+\varepsilon)(h^{2p}+\tau^4) + (3\Gamma+1)TKq(\varepsilon,h,\tau,2))e^{3\Gamma T(2K/\varepsilon+1)}$$

From (5.1.1), (4.2.8), (5.1.44) and (5.1.51) we have

$$\tau\varepsilon \sum_{k=0}^{r} \left(|\eta^k|^2_{H^1(\Omega,\mathcal{T}_h)} + J_h^\sigma(\eta^k,\eta^k) \right) \tag{5.4.143}$$
$$\leq \tau\varepsilon \left(C_6^2 h^{2p} \sum_{k=0}^{r} |u^k|^2_{H^{p+1}(\Omega,\mathcal{T}_h)} + \sum_{k=0}^{r} C_{18}\, h^{2p} |u^k|^2_{H^{p+1}(\Omega)} \right)$$
$$\leq \varepsilon C_{27} h^{2p}(T+\tau).$$

Finally, using (5.1.55), (5.1.1), (5.1.7), Young's inequality, (5.4.142), (5.4.143) and the inequalities $\tau \leq 1$, we have

$$\|e\|^2_{h,\tau,L^2(H^1)} \leq 2\tau\varepsilon \sum_{k=0}^{r} \left(|\xi^k|^2_{H^1(\Omega,\mathcal{T}_h)} + J_h^\sigma(\xi^k,\xi^k) + |\eta^k|^2_{H^1(\Omega,\mathcal{T}_h)} + J_h^\sigma(\eta^k,\eta^k) \right) \tag{5.4.144}$$
$$\leq 4C_{hx}(1+\frac{1}{\varepsilon})(C_p(h^{2(p+1)}+\tau^4) + C_{pp}(1+\varepsilon)(h^{2p}+\tau^4)$$
$$+(3\Gamma+1)TKq(\varepsilon,h,\tau,2))e^{3\Gamma T(2K/\varepsilon+1)} + 2\varepsilon C_{27}h^{2p}(T+\tau)$$
$$\leq 4C_{hx}(1+\frac{1}{\varepsilon})(C_p(h^{2(p+1)}+\tau^4) + C_{pp}(1+\varepsilon)(h^{2p}+\tau^4)$$
$$+2\varepsilon C_{27}h^{2p}(T+1) + (3\Gamma+1)TKq(\varepsilon,h,\tau,2))e^{3\Gamma T(2K/\varepsilon+1)}$$

Now, assertion (5.4.131) for $k = 2$ of the theorem follows from (5.1.57) and (5.4.144) with $\tilde{C} = O\left(\exp(3\Gamma T(1 + 2K/\varepsilon))\right)$.

For $k = 3$ using the same step as in (5.3.109) we get from (5.4.134)

$$\begin{aligned}\frac{11}{6}\|\xi^{s+3}\|^2_{L^2(\Omega)} - \frac{18}{6}\|\xi^{s+2}\|^2_{L^2(\Omega)} + \frac{9}{6}\|\xi^{s+1}\|^2_{L^2(\Omega)} - \frac{2}{6}\|\xi^{s}\|^2_{L^2(\Omega)} \\ + \tau\varepsilon\left(|\xi^{s+3}|^2_{H^1(\Omega,\mathcal{T}_h)} + J_h^\sigma(\xi^{s+3},\xi^{s+3})\right) \\ \leq \tau\|\xi^{s+3}\|^2_{L^2(\Omega)} + \tau\frac{K}{\varepsilon}\sum_{v=0}^{2}\|\xi^{s+v}\|^2_{L^2(\Omega)} + 3\|\xi^{s+2} - \xi^{s+1}\|^2_{L^2(\Omega)} + \tau K q(\varepsilon, h, \tau, 3)\end{aligned} \tag{5.4.145}$$

Now after summing (5.4.135) over $s = 0, \ldots, r-3$ we obtain

$$\begin{aligned}\frac{11}{6}\|\xi^{r}\|^2_{L^2(\Omega)} - \frac{7}{6}\|\xi^{r-1}\|^2_{L^2(\Omega)} + \frac{2}{6}\|\xi^{r-2}\|^2_{L^2(\Omega)} - \frac{11}{6}\|\xi^{2}\|^2_{L^2(\Omega)} + \frac{7}{2}\|\xi^{1}\|^2_{L^2(\Omega)} \\ - \frac{2}{6}\|\xi^{0}\|^2_{L^2(\Omega)} + \tau\varepsilon\sum_{s=3}^{r}\left(|\xi^{s}|^2_{H^1(\Omega,\mathcal{T}_h)} + J_h^\sigma(\xi^{s},\xi^{s})\right) \leq \tau\sum_{s=3}^{r}\|\xi^{s}\|^2_{L^2(\Omega)} \\ + \tau\frac{K}{\varepsilon}\sum_{s=0}^{r-3}\sum_{v=0}^{2}\|\xi^{s+v}\|^2_{L^2(\Omega)} + 3\sum_{s=2}^{r-1}\|\xi^{s} - \xi^{s-1}\|^2_{L^2(\Omega)} + TKq(\varepsilon, h, \tau, 3).\end{aligned} \tag{5.4.146}$$

When we use (5.3.99) omitting some non-negative terms on the left-hand side we have

$$\begin{aligned}\tau\varepsilon\sum_{s=3}^{r}\left(|\xi^{s}|^2_{H^1(\Omega,\mathcal{T}_h)} + J_h^\sigma(\xi^{s},\xi^{s})\right) \leq \frac{7}{6}\|\xi^{r-1}\|^2_{L^2(\Omega)} + \frac{229}{12}\|\xi^{2}\|^2_{L^2(\Omega)} + \frac{73}{4}\|\xi^{1}\|^2_{L^2(\Omega)} \\ + \frac{29}{6}\|\xi^{0}\|^2_{L^2(\Omega)} + \tau\sum_{s=3}^{r}\|\xi^{s}\|^2_{L^2(\Omega)} + \tau\frac{3K}{\varepsilon}\sum_{s=0}^{r-1}\|\xi^{s}\|^2_{L^2(\Omega)} + \frac{13}{4}TKq(\varepsilon, h, \tau, 3) \\ + \tau\left(9 + \frac{27K}{4\varepsilon}\right)\sum_{s=0}^{r-1}\|\xi^{s}\|^2_{L^2(\Omega)} + \frac{9}{4}\tau\varepsilon\left(|\xi^{2}|^2_{H^1(\Omega,\mathcal{T}_h)} + J_h^\sigma(\xi^{2},\xi^{2})\right)\end{aligned} \tag{5.4.147}$$

When we use

$$\|\xi^{s}\|^2_{L^2(\Omega)} \leq 2\|e\|^2_{h,\tau,L^\infty(L^2)} + 2C_{25}h^{2(p+1)}, \tag{5.4.148}$$

where $C_{25} = C_6^2 C_{20}^2$, we get

$$\begin{aligned}\tau\varepsilon\sum_{s=3}^{r}\left(|\xi^{s}|^2_{H^1(\Omega,\mathcal{T}_h)} + J_h^\sigma(\xi^{s},\xi^{s})\right) \leq \frac{9}{4}\tau\varepsilon\left(|\xi^{2}|^2_{H^1(\Omega,\mathcal{T}_h)} + J_h^\sigma(\xi^{2},\xi^{2})\right) \\ + \left(\frac{260}{3} + 20T + T\frac{39K}{2\varepsilon}\right)\left(\|e\|^2_{h,\tau,L^\infty(L^2)} + C_{25}h^{2(p+1)}\right) + \frac{13}{4}TKq(\varepsilon, h, \tau, 3).\end{aligned} \tag{5.4.149}$$

Then (5.4.149), (5.4.129) and

$$|\xi^{s}|^2_{H^1(\Omega,\mathcal{T}_h)} + J_h^\sigma(\xi^{s},\xi^{s}) \leq 2|e_h^s|^2_{H^1(\Omega,\mathcal{T}_h)} + 2J_h^\sigma(e_h^s,e_h^s) + 2C_{27}h^{2p} \tag{5.4.150}$$

give

$$\tau\varepsilon\sum_{s=0}^{r}\left(|\xi^s|^2_{H^1(\Omega,\mathcal{T}_h)}+J_h^\sigma(\xi^s,\xi^s)\right)\le\left(\frac{260}{3}+20T+T\frac{39K}{2\varepsilon}\right)\left(\|e\|^2_{h,\tau,L^\infty(L^2)}+C_{25}h^{2(p+1)}\right)$$
$$+\frac{13}{4}TKq(\varepsilon,h,\tau,3)+\frac{21}{2}C_{nn}(h^{2p}+\tau^6)+\frac{21}{2}\tau\varepsilon C_{27}h^{2p}. \quad (5.4.151)$$

Using (5.4.151), (5.4.133) and notation $C_{hx}=\max(\frac{260}{3}+20T,\frac{39TK}{2})$, $C_p=570\Gamma C_n+2C_{25a}$ and $C_{pp}=(27\Gamma+\frac{21}{2})(C_{nn}+C_{27})$ we get

$$\begin{aligned}\tau\varepsilon\sum_{s=0}^{r}\left(|\xi^s|^2_{H^1(\Omega,\mathcal{T}_h)}+J_h^\sigma(\xi^s,\xi^s)\right) &\le C_{hx}(1+\frac{1}{\varepsilon})(570\Gamma C_n(h^{2(p+1)}+\tau^6) \qquad (5.4.152)\\ &+2C_{25a}h^{2(p+1)}\\ &+27C_{nn}\tau\varepsilon\Gamma(h^{2p}+\tau^6)+27C_{27}\tau\varepsilon\Gamma h^{2p}\\ &+\frac{39}{2}T\Gamma Kq(\varepsilon,h,\tau,3))e^{\Gamma T(117K/4\varepsilon+30)}\\ &+C_{hx}(1+\frac{1}{\varepsilon})C_{25}h^{2(p+1)}+\frac{13}{4}TKq(\varepsilon,h,\tau,3)\\ &+\frac{21}{2}C_{nn}(h^{2p}+\tau^6)+\frac{21}{2}\tau\varepsilon C_{27}h^{2p}\\ &\le 2C_{hx}(1+\frac{1}{\varepsilon})(570\Gamma C_n(h^{2(p+1)}+\tau^6)+2C_{25a}h^{2(p+1)}\\ &+(27\Gamma+\frac{21}{2})(C_{nn}+C_{27})(1+\varepsilon)(h^{2p}+\tau^6)\\ &+(\frac{39}{2}\Gamma+\frac{13}{4})TKq(\varepsilon,h,\tau,3))e^{\Gamma T(117K/4\varepsilon+30)}\\ &\le 2C_{hx}(1+\frac{1}{\varepsilon})(C_p(h^{2(p+1)}+\tau^6)+C_{pp}(1+\varepsilon)(h^{2p}+\tau^6)\\ &+(\frac{39}{2}\Gamma+\frac{13}{4})TKq(\varepsilon,h,\tau,3))e^{\Gamma T(117K/4\varepsilon+30)}\end{aligned}$$

Finally, using (5.1.55), (5.1.1), (5.1.7), Young's inequality, (5.4.152), (5.4.143) and the inequalities $\tau\le 1$, we have

$$\begin{aligned}\|e\|^2_{h,\tau,L^2(H^1)} &\le 2\tau\varepsilon\sum_{k=0}^{r}\left(|\xi^k|^2_{H^1(\Omega,\mathcal{T}_h)}+J_h^\sigma(\xi^k,\xi^k)+|\eta^k|^2_{H^1(\Omega,\mathcal{T}_h)}+J_h^\sigma(\eta^k,\eta^k)\right) \qquad (5.4.153)\\ &\le 4C_{hx}(1+\frac{1}{\varepsilon})(C_p(h^{2(p+1)}+\tau^6)+C_{pp}(1+\varepsilon)(h^{2p}+\tau^6)\\ &\quad+(\frac{39}{2}\Gamma+\frac{13}{4})TKq(\varepsilon,h,\tau,3))e^{\Gamma T(117K/4\varepsilon+30)}+2\varepsilon C_{27}h^{2p}(T+\tau)\\ &\le 4C_{hx}(1+\frac{1}{\varepsilon})(C_p(h^{2(p+1)}+\tau^6)+C_{pp}(1+\varepsilon)(h^{2p}+\tau^6)\\ &\quad+2\varepsilon C_{27}h^{2p}(T+1)+(\frac{39}{2}\Gamma+\frac{13}{4})TKq(\varepsilon,h,\tau,3))e^{\Gamma T(117K/4\varepsilon+30)}\end{aligned}$$

Now, assertion (5.4.131) for $k = 3$ of the theorem follows from (5.1.57) and (5.4.153) with $\tilde{C} = O\left(\exp(\Gamma T(30 + 117K/4\varepsilon))\right)$.

$\square$

Chapter 6

Numerical results

In this chapter we want to compute experimental orders of convergence and verify the orders of convergence that we have derived theoreticaly in Chapter 5. We solve the problem (2.0.1)–(2.0.3) with $N = 2$, $\Omega = (0,1)^2$, $f_s(u) = \frac{1}{2}u^2$, $s = 1,2$, $T = 1$, $\varepsilon = 0.01$ and the functions u_D, u_0 and g are chosen in such a way that

$$u(x_1, x_2, t) = 16\frac{e^{10t} - 1}{e^{10} - 1} x_1(1 - x_1)x_2(1 - x_2). \tag{6.0.1}$$

The computations were carried out on a triangular mesh having 591 elements with piecewise cubic approximation in space, because we want to restrain the discretization error in space, and with BDF method of order $k = 1,2,3$ in time for different time steps. In the form b_h we use the numerical flux

$$\begin{aligned} H(u_1, u_2, \boldsymbol{n}) &= \sum_{s=1}^{2} f_s(u_1) n_s \quad \text{if } A > 0 \\ H(u_1, u_2, \boldsymbol{n}) &= \sum_{s=1}^{2} f_s(u_2) n_s \quad \text{if } A \le 0, \end{aligned} \tag{6.0.2}$$

where

$$A = \sum_{s=1}^{2} f_s' \left(\frac{u_1 + u_2}{2} \right) n_s \quad \boldsymbol{n} = (n_1, n_2). \tag{6.0.3}$$

To solve our problem we use programmes DGFEM and LASPACK. For time step τ we set the error $e(\tau)$. We define the *local experimental order of convergence* for $L^\infty(L^2)$–norm and for $L^2(H^1)$–seminorm by

$$\begin{aligned} \Phi_{L^2} &= \frac{\log\left(\frac{\|e(\tau_1)\|_{h,\tau,L^\infty(L^2)}}{\|e(\tau_2)\|_{h,\tau,L^\infty(L^2)}} \right)}{\log\left(\frac{\tau_1}{\tau_2}\right)} \\ \Phi_{H^1} &= \frac{\log\left(\frac{\|e(\tau_1)\|_{h,\tau,L^2(H^1)}}{\|e(\tau_2)\|_{h,\tau,L^2(H^1)}} \right)}{\log\left(\frac{\tau_1}{\tau_2}\right)}. \end{aligned} \tag{6.0.4}$$

τ	$\lVert e\rVert_{h,\tau,L^\infty(L^2)}$	Φ_{L^2}	$\lVert e\rVert_{h,\tau,L^2(H^1)}$	Φ_{H^1}
$\frac{1}{20}$	1.452 E−01	-	6.712 E−01	-
$\frac{1}{40}$	6.997 E−02	1.054	3.218 E−01	1.061
$\frac{1}{80}$	3.431 E−02	1.028	1.574 E−01	1.032
$\frac{1}{160}$	1.698 E−02	1.014	7.779 E−02	1.016
$\frac{1}{320}$	8.449 E−03	1.007	3.867 E−02	1.008
$\frac{1}{640}$	4.213 E−03	1.004	1.928 E−02	1.004

Table 6.1: Errors of the first order scheme for different time steps τ.

τ	$\lVert e\rVert_{h,\tau,L^\infty(L^2)}$	Φ_{L^2}	$\lVert e\rVert_{h,\tau,L^2(H^1)}$	Φ_{H^1}
$\frac{1}{20}$	3.474 E−02	-	1.679 E−01	-
$\frac{1}{40}$	9.964 E−03	1.802	4.819 E−02	1.801
$\frac{1}{80}$	2.701 E−03	1.883	1.309 E−02	1.880
$\frac{1}{160}$	7.062 E−04	1.936	3.431 E−03	1.932
$\frac{1}{320}$	1.808 E−04	1.966	8.839 E−04	1.957
$\frac{1}{640}$	4.575 E−05	1.982	2.411 E−04	1.874

Table 6.2: Errors of the second order scheme for different time steps τ.

τ	$\lVert e\rVert_{h,\tau,L^\infty(L^2)}$	Φ_{L^2}	$\lVert e\rVert_{h,\tau,L^2(H^1)}$	Φ_{H^1}
$\frac{1}{20}$	1.066 E−02	-	5.432 E−02	-
$\frac{1}{40}$	1.759 E−03	2.600	9.045 E−03	2.586
$\frac{1}{80}$	2.558 E−04	2.781	1.329 E−03	2.767
$\frac{1}{160}$	3.461 E−05	2.886	2.025 E−04	2.714
$\frac{1}{320}$	4.543 E−06	2.930	9.531 E−05	1.087
$\frac{1}{640}$	8.374 E−07	2.440	9.239 E−05	0.045

Table 6.3: Errors of the third order scheme for different time steps τ.

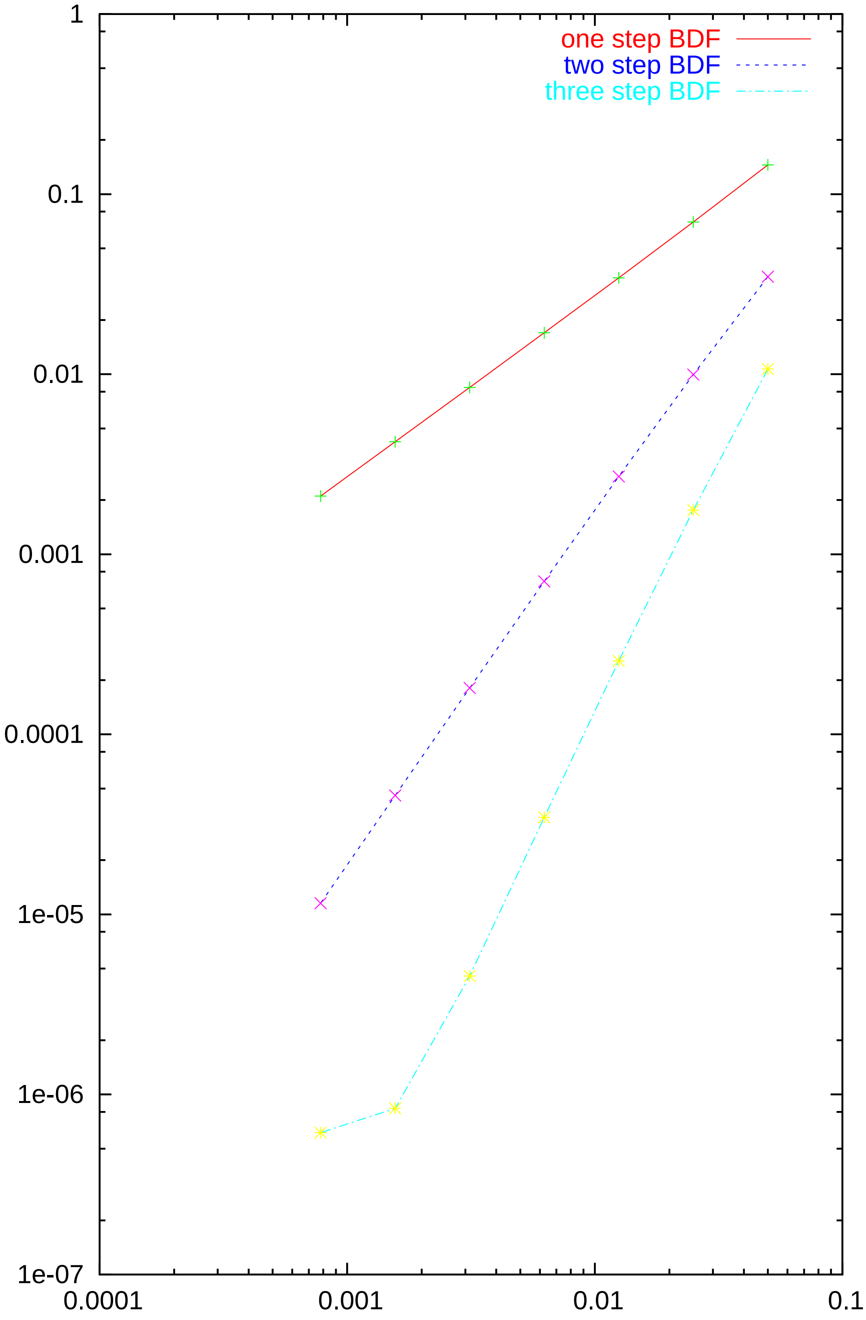

Figure 6.1: Progress of error in L^2–norm for different time steps τ.

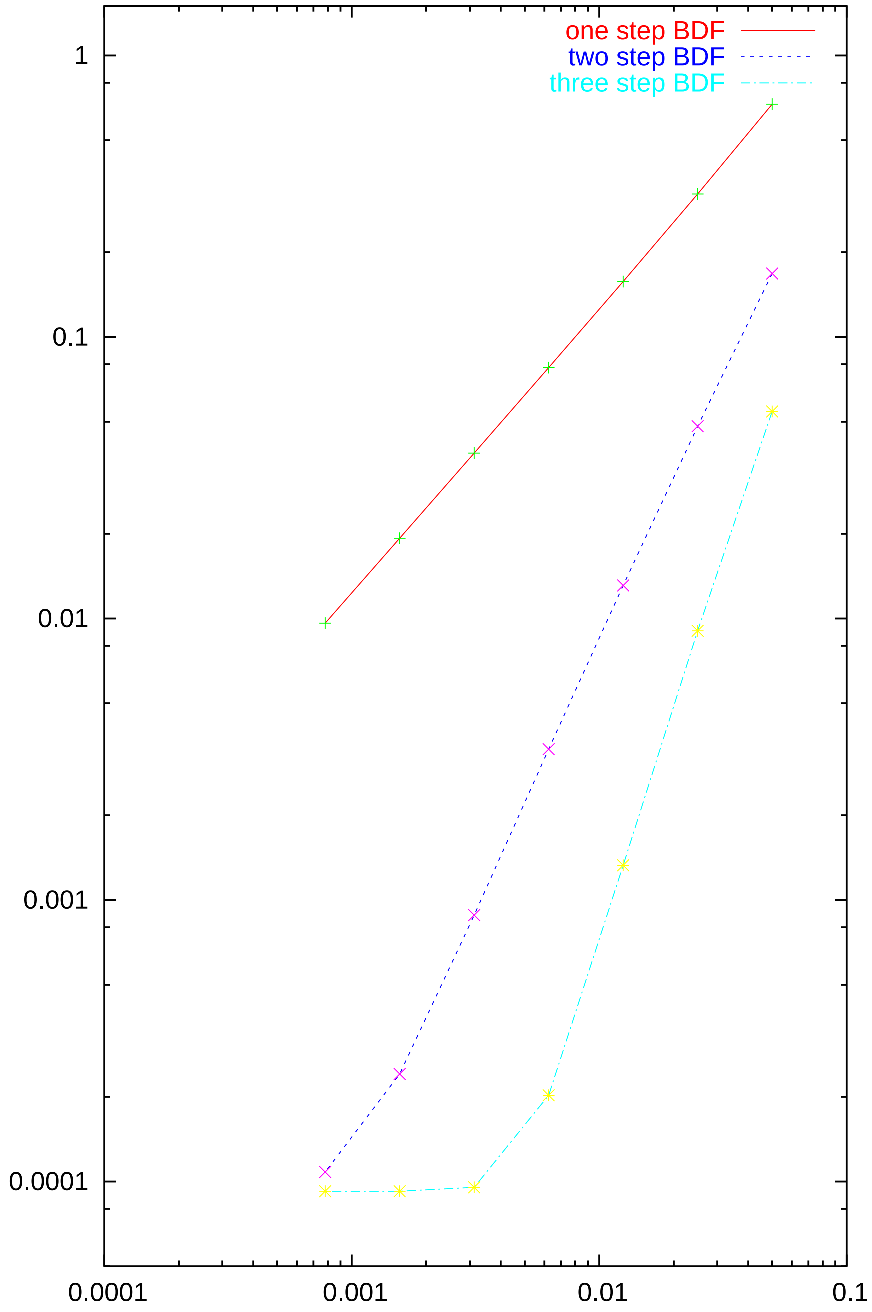

Figure 6.2: Progress of error in H^1–semininorm for different time steps τ.

Errors and experimental orders of convergence of the schemes are presented in Tables 6.1–6.3 and Figures 6.1–6.2. From the tables and graphs representing the progress of error we can come to conclusion that the numerical orders of convergence satisfies the theoretically derived orders. For the smallest time steps we get worse results, because the influence of error in space starts to play an important role. When we compare the three used methods we should mention that almost the same time for computation is needed in all three cases.

Chapter 7

Conclusion

We presented a higher order numerical method for the solution of nonstacionary nonlinear convection-diffusion problems, which is based on the discretization by the discontinuous Galerkin finite element method in space and the semi-implicit backward differential formulae in time. We have derived a priori error estimates, namely $O(h^p + \tau^k)$ in $L^\infty(0, T; L^2(\Omega))$–norm and in $L^2(0, T; H^1(\Omega))$–seminorm for $k = 2, 3$. This estimates are suboptimal in space with respect to approximation property of finite element space S_h. The obtained results show that the DGFEM is a challenging method for the numerical solution of nonstacionary nonlinear convection-diffusion problems. It is sufficiently accurate and robust and yields sufficiently precise approximations of solutions with steep gradients. There are also several items to do. In the estimates we obtained the constants exponentially depending on $\frac{1}{\varepsilon}$. One of these items is development of estimates avoiding this uncomfortable blow up behaviour with respect to $\varepsilon \to 0+$. Another issue is derivation of optimal error estimates in space or derivation of a posteriori error estimates.

Bibliography

[1] D. N. Arnold, F. Brezzi, B. Cockburn, and L. D. Marini. Unified analysis of discontinuous Galerkin methods for elliptic problems. *SIAM J. Numer. Anal.*, **39**(5):1749–1779, 2002.

[2] S. Brenner and R. L. Scott. *The Mathematical Theory of Finite Element Methods.* Spriger, New York, 1994.

[3] P. G. Ciarlet. *The Finite Elements Method for Elliptic Problems.* North-Holland, Amsterdam, New York, Oxford, 1979.

[4] V. Dolejší, M. Feistauer, J. Felcman, and A. Kliková. Error estimates for barycentric finite volumes combined with nonconforming finite elements applied to nonlinear convection–diffusion problems. *Appl. Math.*, **47**(4):301–340, 2002.

[5] V. Dolejší, M. Feistauer. Error estimates of the discontinuous Galerkin method for nonlinear nonstationary convection-diffusion problems. Numer. Func. Anal. Optimiz. **26**(25-26): 2709-2733, 2005.

[6] V. Dolejší, M. Feistauer, J. Hozman: Analysis of Semi-Implicit DGFEM for Nonlinear Convection-Diffusion Problems, Comput. Methods Appl. Mech. Engrg. (submitted).

[7] V. Dolejší, M. Feistauer, and C. Schwab. A finite volume discontinuous Galerkin scheme for nonlinear convection–diffusion problems. *Calcolo*, 39:1–40, 2002.

[8] V. Dolejší, M. Feistauer, and V. Sobotíková. Analysis of the Discontinuous Galerkin Method for Nonlinear Convection–Diffusion Problems, *Comput. Methods Appl. Mech. Eng.* **194**: 2709-2733, 2005.

[9] M. Feistauer. *Discrete Methods for the Solution of Differential Equations.* SPN, Prague, 1981. (in Czech)

[10] M. Feistauer, J. Felcman, M. Lukáčová-Medvidová, and G. Warnecke. Error estimates of a combined finite volume – finite element method for nonlinear convection – diffusion problems. *SIAM J. Numer. Anal.*, **36**(5):1528–1548, 1999.

[11] A. Kufner, O. John, and S. Fučík. *Function Spaces.* Academia, Prague, 1977.

[12] P. L. Lions. *Mathematical Topics in Fluid Mechanics.* Oxford Science Publications, 1996.

[13] K. Rektorys. *The Method of Discretization in Time and Partial Differential Equations.* Reidel, Dodrecht, 1982.

[14] B. Rivière and M. F. Wheeler. A discontinuous Galerkin method applied to nonlinear parabolic equations. In B. Cockburn, editor, *Discontinuous Galerkin methods. Theory, computation and applications.*, volume 11 of *Lect. Notes Comput. Sci. Eng.*, pages 231–244. Berlin, Springer, 2000.

[15] T. Skalický. *LASPack.* Dresden University of Technology, version 1.12.2 edition, 1995. www.netlib.org/laspack.

[16] R. Verfürth. A note on polynomial approximation in Sobolev spaces. M^2AN, 33:715–719, 1999.