Numer. Math. 61, 335–357 (1992)

Numerische Mathematik

Streamline diffusion finite element method for quasilinear elliptic problems

G. Lube

Technische Universität "Otto von Guericke", Fakultät für Mathematik, Institut für Analyis, Postfach 4120, O-3010 Magdeburg, Federal Republic of Germany

Received June 5, 1990 / Revised version received March 4, 1991

Summary. We extend the analysis of the streamline diffusion finite element method to quasilinear elliptic problems of second order. An existence theorem and error estimates are given in the case of branches of nonsingular solutions following a recent abstract approach in [12, 13, 26].

Mathematics Subject Classification (1991): 65N30

1 Introduction

We consider finite element approximations of quasilinear boundary value problems

(1.1) $$N_\varepsilon(u) \equiv -\varepsilon \Delta u + B(x, u) \cdot \nabla u + C(x, u) = 0, \qquad x \in \Omega$$

(1.2) $$u = 0, \qquad x \in \partial\Omega$$

in a bounded domain $\Omega \subset \mathbb{R}^2$ with Lipschitz continuous boundary $\partial\Omega$ and with a parameter ε, $0 < \varepsilon_2 \leqq \varepsilon \leqq \varepsilon_1 \leqq 1$. ν denotes the outward pointing normal vector on $\partial\Omega$. (1.1) and (1.2) describe diffusion-convection-reaction problems where the convection-reaction dominated case $\varepsilon \ll 1$ is of special interest: In general, the solutions u of (1.1) and (1.2) and u_0 of the corresponding first order limit problem ($\varepsilon = 0$) with boundary conditions on "inflow parts" $\Gamma_- \equiv \{x \in \partial\Omega : \exists\, \nu(x), B(x, 0) \cdot \nu(x) < 0\}$ only are very close with the following exceptions: There are "boundary layers" of u at $\partial\Omega \backslash \Gamma_-$ and non-smoothness of u_0 in Ω results in "interior layer" behaviour of u. For a more detailed asymptotic analysis we refer to [14] and [32].

Standard Galerkin-Bubnov finite element methods for (somewhat more general) quasilinear elliptic problems in case of $\varepsilon = O(1)$ were considered for instance in [5–10]. For $\varepsilon \ll 1$ such methods are known to generate unphysical oscillations of discrete solutions unless the mesh is refined. Artificial diffusion or upwind finite element methods, often used as a remedy, have limited order of accuracy and tend to smear sharp fronts (at least in directions perpendicular to the streamlines, i.e. the characteristics of the limit problem with $\varepsilon=0$). The approach with exponentially fitted Galerkin schemes [3, 27, 28] seems too complicated in case of a complex geometry of $\Omega\subset\mathbb{R}^N$, $N\geqq 2$ and is not straightforward in the quasilinear case (1.1), (1.2).

In this paper, we prefer the streamline diffusion method (SDFEM for short below [1, 2, 11, 15–18, 20–25, 29–31]), as the approach most used in recent time. For linear problems (1.1), (1.2) with $B(x,u)=b(x)$, $C(x,u)=c(x)u-f(x)$, SDFEM can be characterized either as a Galerkin-Petrov method with basis functions w_h and test functions $w_h+\delta b\cdot\nabla w_h$ or as a Galerkin method with anisotropic artificial diffusion $\delta\int_\Omega \nabla u_h(b^\perp b)\nabla w_h\,dx$. It has the following properties: The discretization error is quasi-optimal for the "streamline derivative" $b\cdot\nabla$ and nearly optimal in the L_2-norm [1, 2, 17, 24]. Furthermore boundary and interior layers are well localized and affect the accuracy only locally [18, 24]. Local and uniform in ε L_∞-estimates are given for a modified streamline method in [18, 25]. For some asymptotically fitted variants of SDFEM we refer to [20, 21].

SDFEM has been introduced for problems (1.1), (1.2) with nonlinear convection-reaction terms in [31] with no rigorous finite element analysis. In [23] we have started such an analysis. In this paper, which is a much more detailed and extended version of the note [22], we consider existence and convergence results for branches $\{(\lambda, u(\lambda)):\ \lambda\in\Lambda,\ u\in W_0^{1,2}(\Omega)\}$ of nonsingular solutions of (1.1), (1.2) where $\lambda=\varepsilon^{-1}\in\Lambda\equiv[\lambda_1,\lambda_2]\subset\mathbb{R}^+$, $\lambda_i=\varepsilon_i^{-1}$. The analysis is based on an abstract result in [12, 13, 26] and on techniques frequently used in streamline diffusion methods. This paper is organized as follows: In Sect. 2 we give notation and main assumptions concerning the continuous problem. The discrete problem is presented in Sect. 3. Section 4 contains the main result of existence and convergence of branches of discrete solutions. Several auxiliary results for the proof are collected in Sect. 5. In Sect. 6 we give a numerical example.

2 Notation. Main assumptions

Let $\Omega\subset\mathbb{R}^2$ be a bounded domain with a Lipschitz continuous boundafy $\partial\Omega$. For subdomains $G\subseteq\Omega$ with Lipschitz continuous boundary ∂G and $1\leqq p\leqq\infty$ let be $L_p(G)\equiv\{u=u(x):\ \|u\|_{0,p,G}<\infty\}$ with

$$\|u\|_{0,p,G}\equiv\Big(\int_G |u(x)|^p\,dx\Big)^{1/p},\quad 1\leqq p<\infty;\qquad \|u\|_{0,\infty,G}\equiv\sup_{x\in G} ess\,|u(x)|.$$

For $k\in\mathbb{N}$, $1\leqq p\leqq\infty$ let $W^{k,p}(G)$ denote the Sobolev space of all functions $u\in L_p(G)$ whose derivatives $D^\alpha u$ (in the sense of distributions) of order $|\alpha|\leqq k$ again belong to $L_p(G)$. $W^{k,p}(G)$ is equipped with the usual norm $\|\cdot\|_{k,p,G}$ and seminorm $|\cdot|_{k,p,G}$. $W_0^{k,p}(G)$ is the closure of $C_0^\infty(G)$ in the norm of $W^{k,p}(G)$. $(\cdot,\cdot)_G$ denotes the inner product in $L_2(G)$. If there is no doubt we omit the index Ω.

Now we assume that the coefficients $B=(B^j)_{j=1,2}$, C and their derivatives satisfy a growth condition

(H.1) $|B^j(x,s)|+|D_u B^j(x,s)|+|D_{uu} B^j(x,s)|+|\operatorname{div} B(x,s)|$
$+|C(x,s)|+|D_u C(x,s)|+|D_{uu} C(x,s)|\leqq G(|s|)$, $x\in\Omega$ a.e.,

$\forall s\in\mathbb{R}$ with a nondecreasing function $G:\mathbb{R}^+\to\mathbb{R}^+$.

(H.1) is satisfied, for instance, for equations of Burger's type (cf. Sect. 6) or various diffusion-convection-reaction models with a nonlinear reaction rate. Further we denote by $B(u)$, $C(u)$, $B_u(u)$, $H(u, \nabla u)$ and $\tilde{H}(u, \nabla u)$ the Nemyzki operators associated with $B(x, u)$, $C(x, u)$, $D_u B(x, u)$, $B(x, u)\cdot\nabla u + C(x, u)$ and $D_u B(x, u)\cdot\nabla u + D_u C(x, u)$, respectively, by $B(u)\colon u \to B(\cdot, u(\cdot))$ etc. They are well-defined for $u \in W^{1,2}(\Omega) \cap L_\infty(\Omega)$.

The variational formulation of (1.1), (1.2) in $X \cap L_\infty(\Omega)$, $X = W_0^{1,2}(\Omega)$ reads now:

$(\mathcal{N})$ Find $u \in X \cap L_\infty(\Omega)$ s.t. $\forall w \in X$:

$$a_0(u, w) \equiv \varepsilon(\nabla u, \nabla w) + \langle H(u, \nabla u), w\rangle = 0.$$

The operator $N_\varepsilon \in (X \cap L_\infty(\Omega) \to (X \cap L_\infty(\Omega))^*)$ associated with $(\mathcal{N})$ by $\langle N_\varepsilon(u), w\rangle \equiv a_0(u, w)$ is differentiable in the sense of Frechèt:

$$\langle D_u N_\varepsilon(u)\cdot v, w\rangle \equiv \varepsilon(\nabla v, \nabla w) + \langle B(u)\cdot\nabla v + \tilde{H}(u, \nabla u) v, w\rangle \tag{2.1}$$

where $\langle\cdot,\cdot\rangle$ denotes the scalar product between $X \cap L_\infty(\Omega)$ and the corresponding dual space.

Let $\lambda \equiv \varepsilon^{-1} \in \Lambda \equiv [\lambda_1, \lambda_2]$, $\lambda_i = \varepsilon_i^{-1}$. The main assumptions on $(\mathcal{N})$ are:

(H.2) There exists a branch $\{(\lambda, u(\lambda))\colon \lambda \in \Lambda,\ u(\lambda) \in X \cap L_\infty(\Omega)\}$ of nonsingular solutions of $(\mathcal{N})$, i.e. $D_u N_\varepsilon(u)$ is an isomorphism of $X \cap L_\infty(\Omega)$ for all $\lambda \in \Lambda$.

(H.3) $\forall \lambda \in \Lambda$: $\|u(\lambda)\|_{1,2,\Omega} + \|u(\lambda)\|_{0,\infty,\Omega} \leqq M$.

(H.4) The mapping $\phi \to \Delta\phi$ is an isomorphism from $W^{2,r}(\Omega) \cap X$ onto $L_r(\Omega)$ for all $r \in [1; 2]$.

We give some remarks concerning our assumptions.

Remarks.

2.1. *(H.2)* is equivalent to the hypothesis that the linearized (in u) Dirichlet problem is uniquely solvable in X for all $\lambda \in \Lambda$.

2.2. Sufficient conditions for existence of solutions of quasi-linear elliptic problems in $X \cap L_\infty(\Omega)$ are given for instance in [4, 19]. Often one can prove *(H.3)* by means of (weak) maximum principles. On the other hand, $u \in L_\infty(\Omega)$ follows from *(H.1)*, *(H.2)*, *(H.4)* – cf. Lemma 2.1 below.

2.3. Sufficient for *(H.4)* is that Ω is a convex polygon or has a more smooth boundary $\partial\Omega \subset \mathscr{C}^{1,1}$ ([13], Remark I.1.2).

2.4. The case of nonhomogeneous Dirichlet boundary conditions could be handled via a splitting $u = u_{bc} + \tilde{u}$, $\tilde{u} \in X$ (cf. [8]).

Further, *(H.1)*–*(H.4)* imply better regularity property of solutions of $(\mathcal{N})$.

Lemma 2.1. *Under the assumptions (H.1), (H.2), (H.4) any solution u of $(\mathcal{N})$ belongs to $X \cap L_\infty(\Omega) \cap W^{2,2}(\Omega)$.*

Proof. *(H.1)*, *(H.2)* imply $H(u, \nabla u) \in L_{\tilde{r}}(\Omega)$, $\tilde{r} \in [1, 2)$ arbitrary, using compact imbedding of X and [13], Cor. I.1.1. Now *(H.4)* implies $u \in W^{2,\tilde{r}}(\Omega)$ and Sobolev's imbedding theorem (with appropriate values $\tilde{r}$) yields $u \in L_\infty(\Omega)$ and hence $H(u, \nabla u) \in L_2(\Omega)$. *(H.4)* completes the proof. □

Remark 2.5. As a conclusion of Lemma 2.1, Eq. (1.1) is satisfied in the sense of $L_2(\Omega)$. Similarly it holds $\tilde{H}(u, \nabla u) \in L_2(\Omega)$.

In Sect. 4 we assume sometimes some more regularity of solutions of $(\mathcal{N})$ according to

(H.5) $\forall \lambda \in \Lambda: u \in W^{k+1,2}(\Omega), k \in \mathbb{N}$.

3 Discrete problem

Now we consider finite element approximations of problem $(\mathcal{N})$ using piecewise polynomial basis functions of degree $k \in \mathbb{N}$. Let Ω be (for simplicity) a convex polygon. Further let $T_h = \{\tau_i\}_{i=1}^{I(h)}$ be a finite element partition of Ω with $\bar{\Omega} = \bigcup_{i=1}^{I(h)} \bar{\tau}_i$ by means of elements $\tau_i \in T_h$ of diameter h_i. For the case of a general domain Ω, we refer to [6] and [8]. We assume that T_h is quasi-regular such that each element τ_i contains a ball with radius Ch where $C>0$ is independent on $h = \max_{i=1,\dots,I(h)} h_i$ (see also Remark 3.2). As a conclusion one can prove for the discrete space of piecewise polynomial shape functions of Lagrangian type

$$V_h := \{w_h \in X: w_h|_{\tau_i} \in P_k(\tau_i) \; \forall \tau_i \in T_h\} \tag{3.1}$$

local inverse and interpolation properties [6]:

(A.1) $\forall v \in V_h: |v|_{m,q,\tau_i} \leqq C h_i^{l-m+\min(0;2(1/q-1/p))} |v|_{l,p,\tau_i}$
$0 \leqq l \leqq m \leqq k+1,\ 1 \leqq p \leqq q \leqq \infty,\ \forall \tau_i \in T_h$

(A.2) $\forall v \in W^{l+1,p}(\tau_i): \exists \pi_h v \in V_h$ s.t.
$\|v - \pi_h v\|_{m,q,\tau_i} \leqq C h_i^{l-m+2(1/q-1/p)} \|v\|_{l+1,p,\tau_i}$
$0 \leqq l \leqq k;\ m \leqq l+1,\ 1 \leqq p \leqq q \leqq \infty,\ \forall \tau_i \in T_h$.

In detail, we find from *(A.1)* and Sobolev's imbedding theorem that

$$\forall v \in V_h: \; \|v\|_{0,\infty,\Omega} \leqq C h^{-\kappa} \|v\|_{1,2,\Omega} \text{ with } \kappa > 0 \text{ arbitrary.} \tag{3.2}$$

Now we introduce the form

$$a_\delta(\psi; v, w) := a_0(v, w) + \sum_{i=1}^{I(h)} (N_\varepsilon(v), \delta_i B(\psi) \cdot \nabla w)_{\tau_i} \tag{3.3}$$

with "upwind" parameters δ_i such that with sufficiently small constants $K_i \leqq K$ (not depending on ε, h – cf. proof of Lemma 5.4)

(H.6) $\forall \lambda = \varepsilon^{-1} \in \Lambda$: (i) $0 \leqq \varepsilon \delta_i \leqq K_i h_i^2,\ i = 1, \dots, I(h)$
(ii) $\delta_i \leqq C h_i$ (C independent on ε, h).

The discrete problem of the SDFEM reads now

$(\mathcal{N}_h)$ Find $u_h \in V_h$ s.t. $\forall w \in V_h$:

$$a_\delta(u_h; u_h, w) = 0.$$

Remark 3.1. $(\mathcal{N}_h)$ corresponds to a Petrov-Galerkin method with test functions $w_h+\delta_i B(u_h)\cdot\nabla w_h$, $w_h\in V_h$ or to introducing artificial viscosity acting in the "streamline direction" $B(\cdot, u_h)$. In [15] and [17] SDFEM is characterized as a Galerkin/least-squares method where the residual of (1.1) is weighted with $\delta_i B(u_h)\cdot\nabla w_h$ elementwise and added to the Galerkin formulation.

Remark 3.2. A more careful analysis shows that the restriction on T_h to be quasi-uniform is not necessary. For details in the linear case of $(\mathcal{N})$ we refer to [11]. We state the assumption of a quasi-uniform mesh here only for simplicity.

4 Existence and convergence of discrete solutions

The main result of this paper is the following existence and convergence theorem for the SDFEM $(\mathcal{N}_h)$.

Theorem

1. Let *(H.1)–(H.4), (H.6)* be satisfied. Then there exists for all $\lambda=\varepsilon^{-1}\in\Lambda$ a (sufficiently small) neighbourhood $\mathcal{O}$ of the origin in X and for all h, $0<h\leqq h_0(\Lambda)$ a unique branch $\{(\lambda, u_h(\lambda)): \lambda\in\Lambda, u_h\in V_h\}$ of nonsingular solutions of $(\mathcal{N}_h)$ s.t.

$$\forall\lambda\in\Lambda:\ u(\lambda)-u_h(\lambda)\in\mathcal{O}. \tag{4.1}$$

2. Moreover, if it additionally holds *(H.5)*, then we have:

$$\forall\lambda\in\Lambda:\ |||(u-u_h)(\lambda)|||_{\varepsilon,\delta}\leqq T(\Lambda)\, h^k(\sqrt{\varepsilon}+\sqrt{h}) \tag{4.2}$$

where

$$|||v|||^2_{\varepsilon,\delta}\equiv\varepsilon\,|v|^2_{1,2,\Omega}+\sum_i \delta_i\,\|B(u_h)\cdot\nabla v\|^2_{0,2,\tau_i}. \tag{4.3}$$

For the proof of the first part of this theorem we make use of an result in [13] on approximations of branches $\{(\lambda, u(\lambda)): \lambda\in\Lambda, u\in X\}$ of nonsingular solutions of the abstract nonlinear equation

$$F(\lambda, u(\lambda))\equiv u+T G_0(\lambda, u)=0. \tag{4.4}$$

Let be $\Lambda\subset\mathbb{R}$ compact, X, Y-Banach spaces, $V\subset X$ a closed subspace, $T\in\mathcal{L}(Y, X)$ and $G_0: \Lambda\times X\mapsto Y$ a $\mathcal{C}^1$-mapping. (4.4) is approximated by means of the discrete equation

$$F_h(\lambda, u_h(\lambda))\equiv u_h+T_h G_h(\lambda, u_h)=0 \tag{4.5}$$

where $V_h\subset X$ denotes a closed subspace and $Y\subset Y_h$ with continuous imbedding. Further let be $T_h\in\mathcal{L}(Y_h, V_h)$, $G_h: \Lambda\times V_h\to Y_h$ a $\mathcal{C}^1$-mapping, $\|\cdot\|_X$ and $\|\cdot\|_h$ norms on X and Y_h, respectively. $B(z,\rho)$ denotes the ball $\{v\in X: \|v-z\|_X\leqq\rho\}$. Then it holds ([13], IV, Sect. 3.4).

Lemma 4.1. *Let the following assumptions (i)–(vi) be valid:*

(4.6) (i) $\exists \pi_h \in \mathscr{L}(V, V_h)$: $\lim\limits_{h\to 0} \|v-\pi_h v\|_X = 0 \quad \forall v \in V$

(4.7) (ii) $\|T_h\|_{\mathscr{L}(Y_h, V_h)} \leqq \text{const.}$

(4.8) (iii) $\lim\limits_{h\to 0} \|T_h - T\|_{\mathscr{L}(Y,X)} = 0$

(4.9) (iv) $\lim\limits_{h\to 0} \sup\limits_{\lambda\in\Lambda} \|G_h(\lambda, \pi_h u(\lambda)) - G_0(\lambda, u(\lambda))\|_h = 0$

(4.10) (v) $\forall w_h \in V_h \ \exists D_u G_h(\lambda, w_h) \in \mathscr{L}(V_h, Y_h)$:
$\lim\limits_{h\to 0} \sup\limits_{\lambda\in\Lambda} \|D_u G_h(\lambda, \pi_h u(\lambda)) - D_u G_0(\lambda, u(\lambda))\|_{\mathscr{L}(V_h, Y_h)} = 0$

(4.11) (vi) $\forall u_1, u_2 \in B(\pi_h u(\lambda); \rho) \cap V_h$:

$$\|G_h(\lambda, u_1) - G_h(\lambda, u_2) - D_u G_h(\lambda, \pi_h u)\cdot(u_1 - u_2)\|_h \leqq L_h(\lambda, \rho, \|\pi_h u\|_X)\, \|u_1 - u_2\|_X$$

with a continuous, monotonically increasing (in each variable) function L_h s.t.

$$\lim_{h\to 0} \sup_{\lambda\in\Lambda} L_h(\lambda, 0, \mu) = 0 \qquad \forall \mu \in \mathbb{R}^+.$$

Then there exists a neighbourhood $\mathcal{O}$ of the origin in X and for $\forall h$, $0 < h \leqq h_0(\Lambda)$ a unique $\mathscr{C}^0$-mapping u_h: $\Lambda \mapsto V_h$ s.t.

1. $F_h(\lambda, u_h(\lambda)) = 0$
2. $\forall \lambda \in \Lambda$: $u_h(\lambda) - u(\lambda) \in \mathcal{O}$
3. $\|u_h(\lambda) - u(\lambda)\|_X \leqq \tilde{T}(\Lambda)\{\|(I-\pi_h)\, u(\lambda)\|_X$
$+ \|(T - T_h)\, G_0(\lambda, u(\lambda))\|_X + \|G_h(\lambda, \pi_h u(\lambda)) - G_0(\lambda, u(\lambda))\|_h\}$.

Proof of Theorem (1). We set $\Lambda = [\lambda_1, \lambda_2]$ with $\lambda = \varepsilon^{-1}$, $\lambda_i = \varepsilon_i^{-1}$, $X = W_0^{1,2}(\Omega)$ with norm $\|\cdot\|_X = |\cdot|_{1,2,\Omega}$, $X^* = W^{-1,2}(\Omega)$, $\langle\cdot,\cdot\rangle$ scalar product between X^* and X, $Y = L_2(\Omega)$, $V = X \cap W^{2,2}(\Omega)$. *(H.1), (H.2)* imply that

$$H(v, \nabla v) \in L_r(\Omega) \subset X^*, \forall r \in [1, 2], \forall v \in X.$$

Hence with

$$g(v; w) := \langle \lambda H(v, \nabla v), w\rangle \qquad \forall v, w \in X, \ \forall \lambda \in \Lambda$$

$g(v; \cdot)$ is a linear continuous functional on X and there exists an operator G_0: $\Lambda \times X \to X^*$ with

$$\langle G_0(\lambda, v), w\rangle := \langle \lambda H(v, \nabla v), w\rangle \qquad \forall w \in X. \tag{4.12}$$

With the inverse Laplace operator $T \in \mathscr{L}(X^*, X)$ solving

$$(\nabla(Tf^*), \nabla w) = \langle f^*, w\rangle \tag{4.13}$$

we reformulate problem $(\mathcal{N})$ as

$$u\in X:\ (\nabla u, \nabla w)+\langle G_0(\lambda, u), w\rangle=0 \qquad \forall w\in X$$

or

$$(\mathcal{N}^*) \qquad F(\lambda, u(\lambda)):=u+TG_0(\lambda, u)=0.$$

With respect to *(H.4)* we have $T\in\mathcal{L}(Y, V)$.

Similarly we define for the discrete problem

$$(4.14)\quad g_h(\psi; v, w):=\langle\lambda H(v, \nabla v), w\rangle+\sum_i \delta_i(-\Delta v+\lambda H(v, \nabla v), B(\psi)\cdot\nabla w)_{\tau_i}$$

$$\forall \psi, v\in X\cap L_\infty(\Omega), \qquad \forall w\in V_h.$$

It holds

Lemma 4.2. *Under the assumptions (H.1)–(H.4) and (H.6) we conclude for* $\psi, v\in X\cap L_\infty(\Omega)$ *with* $\|\psi\|_{0,\infty,\Omega}, \|v\|_{0,\infty,\Omega}\leqq M$ *and* $w\in V_h$ *that*

$$|\sum_i \delta_i(-\Delta v+\lambda H(v, \nabla v), B(\psi)\cdot\nabla w)_{\tau_i}|\leqq hG(M)C\lambda\{\lambda+(1+\lambda h)\|v\|_X\}\|w\|_X.$$

Proof. *(H.1)* and an inverse inequality imply that

$$\begin{aligned}\delta_i\,&\|-\Delta v+\lambda H(v, \nabla v)\|_{0,2,\tau_i}\\ &\leqq G(M)\{\delta_i h_i^{-1}|v|_{1,2,\tau_i}+\lambda\delta_i(|v|_{1,2,\tau_i}+1)\}\\ &\leqq \delta G(M)\{\lambda+(h^{-1}+\lambda)|v|_{1,2,\tau_i}\}.\end{aligned}$$

Thus we find that

$$\begin{aligned}|\sum_i &\delta_i(-\Delta v+\lambda H(v, \nabla v), B(\psi)\cdot\nabla w)_{\tau_i}|\\ &\leqq(\sum_i \delta_i^2\|-\Delta v+\lambda H(v, \nabla v)\|_{0,2,\tau_i}^2)^{1/2}G(M)(\sum_i\|\nabla w\|_{0,2,\tau_i}^2)^{1/2}\\ &\leqq C\delta G(M)(\sum_i\{(h^{-1}+\lambda)|v|_{1,2,\tau_i}+\lambda\}^2)^{1/2}\|w\|_X\\ &\leqq C\delta G(M)(\sum_i\lambda^2+\sum_i(h^{-2}+\lambda^2)|v|_{1,2,\tau_i}^2)^{1/2}\|w\|_X\\ &\leqq CG(M)\delta(h^{-2}\lambda^2+(h^{-2}+\lambda^2)\|v\|_X^2)^{1/2}\|w\|_X\\ &\leqq CG(M)\delta(h^{-1}\lambda+(h^{-1}+\lambda)\|v\|_X)\|w\|_X\\ &\leqq CG(M)\lambda h(\lambda+(1+h\lambda)\|v\|_X)\|w\|_X \quad \text{(by } (H.6)). \quad \square\end{aligned}$$

As a conclusion, $g_h(\psi; v, \cdot)$ is a linear continuous functional on V_h and there exists an operator G_h: $\Lambda\times(X\cap L_\infty(\Omega))\to X^*$ with

$$(4.15)\qquad \langle G_h(\lambda, v), w\rangle:=g_h(v; v, w) \qquad \forall w\in V_h.$$

Remark 4.1. It follows from (4.12), (4.15) and Lemma 4.2 that

(4.16) $$\lim_{h\to 0}\langle G_h(\lambda, v), w\rangle = \langle G_0(\lambda, v), w\rangle \qquad \forall\,\lambda\in\Lambda,\ \forall\, v\in X\cap L_\infty(\Omega),\ \forall\, w\in V_h.$$

Let $T_h\in\mathscr{L}(X^*, V_h)$ be the discrete inverse Laplace operator solving

(4.17) $$(\nabla(T_h f^*), \nabla w) = \langle f^*, w\rangle \qquad \forall\, w\in V_h.$$

Hence $(\mathscr{N}_h)$ is reformulated as

$(\mathscr{N}_h^*)$ $$F_h(\lambda, u_h(\lambda)) := u_h + T_h\, G_h(\lambda, u_h) = 0.$$

With $Y_h = Y$ it holds $T_h\in\mathscr{L}(Y_h, V_h)$. Further we define a norm on Y_h

$$\|\cdot\|_h := \sup_{w\in V_h}\frac{\langle\cdot, w\rangle}{\|w\|_X}.$$

Now we prove properties (i)–(vi) of Lemma 4.1.

(i) From interpolation property *(A.2)* and regularity according to *(H.6)* we find

(4.18) $$\|v-\pi_h v\|_X \leqq C\,h^k\,|v|_{k+1,2},$$

from which, in particular, (4.6) follows.

(ii) The usual linear theory for the discrete Poisson problem (4.17) yields $\|T_h f^*\|_X \leqq \|f^*\|_h$ implying (4.7).

(iii) As another consequence of (4.13), (4.17) we find by Cea's Lemma, (4.18) with $k=1$ and *(H.4)*

(4.19) $$\begin{aligned}\|(T-T_h)f^*\|_X &\leqq \inf_{w\in V_h}\|Tf^*-w\|_X\\ &\leqq C\,h\,|Tf^*|_{2,2,\Omega} \leqq C\,h\,\|f^*\|_{0,2,\Omega}\end{aligned}$$

from which (4.8) follows.

(iv) For proving (4.9) we recognize Remark 2.4, such that $-\Delta u + \lambda H(u, \nabla u) = 0$ in $L_2(\Omega)$. Consequently

(4.20) $$\begin{aligned}&\langle G_h(\lambda, \pi_h u), w\rangle - \langle G_0(\lambda, u), w\rangle\\ &\quad = \langle G_h(\lambda, \pi_h u), w\rangle - \langle G_h(\lambda, u), w\rangle.\end{aligned}$$

Lemma 5.1 yields

(4.21) $$\|G_h(\lambda, \pi_h u(\lambda)) - G_h(\lambda, u(\lambda))\|_{\mathscr{L}(V_h, Y_h)} \leqq C_1(\Lambda)\,h^{k-1}(h+\delta)$$

which implies (4.9). Notice that $\lim_{h\to 0}\delta = 0$ follows from *(H.6)* and compactness of Λ.

(v) For the derivative $D_u G_h(\lambda, z)\cdot v$ we find

$$(4.22)\quad \langle D_u G_h(\lambda, z)\cdot v, w\rangle = \lambda\{(B(z)\cdot\nabla v, w)+(\tilde{H}(z,\nabla z)v, w)\} + \sum_i \delta_i\{(-\Delta v+\lambda B(z)\cdot\nabla v+\lambda\tilde{H}(z,\nabla z)v, B(z)\cdot\nabla w)_{\tau_i} + ([-\Delta z+\lambda H(z,\nabla z)]\,v, D_u B(z)\cdot\nabla w)_{\tau_i}\}.$$

Using again Remark 2.4 we conclude from Lemma 5.2

$$(4.23)\quad \|D_u G_h(\lambda, \pi_h u)-D_u G_0(\lambda, u)\|_{\mathscr{L}(V_h, Y_h)} = \|D_u G_h(\lambda, \pi_h u)-D_u G_h(\lambda, u)\|_{\mathscr{L}(V_h, Y_h)} \leqq C_2(\Lambda)\, h^k(1+\delta h^{-1-\kappa})$$

with κ from (3.2). (4.23) implies (4.10) according to *(H.6)* and to compactness of Λ.

(vi) The analysis of Lemma 5.3 shows that for all $u_1, u_2\in B(\pi_h u;\rho)\cap V_h$ with $\rho\leqq C h^{\bar\kappa}$, $\bar\kappa>\kappa$

$$(4.24)\quad \|G_h(\lambda, u_1)-G_h(\lambda, u_2)-D_u G_h(\lambda, \pi_h u)\cdot(u_1-u_2)\|_h \leqq L_h(\lambda, \rho, \|\pi_h u\|_X)\,\|u_1-u_2\|_X$$

with

$$(4.25)\quad \lim_{h\to 0}\sup_{\lambda\in\Lambda} L_h(\lambda, 0, \mu) = \lim_{h\to 0} C[G(3M)]^2\{\delta h^{-1-\kappa}(h+\mu)(1+\lambda_2)+\lambda_2\,\delta h^{-\kappa}(1+\mu)\}=0$$

according to *(H.6)*. Thus (4.11) is valid.

Now Lemma 4.1 implies assertion (1) of the Theorem. More precisely, Lemma 4.1 (iii), (4.18), (4.21) and the (compared with (4.20)) improved inequality

$$(4.26)\quad \|(T-T_h)f^*\|_X\leqq C h^k\,|Tf^*|_{k+1,2,\Omega}$$

yield

$$(4.27)\quad \forall\lambda\in\Lambda:\ \|u(\lambda)-u_h(\lambda)\|_X\leqq C_3(\Lambda)\,h^k$$

where $C_3(\Lambda)$ depends on $\tilde{T}(\Lambda)$ from Lemma 4.1, $|u|_{k+1,2,\Omega}$, $C_1(\Lambda)$.

As a next step, we improve (4.27) by means of standard arguments in SDFEM-analysis. Note that we have the following estimates from *(H.1)*, (4.27) and (3.2)

$$(4.28)\quad \|u_h\|_X\leqq\|u\|_X+\|u-u_h\|_X\leqq M+C_3(\Lambda)\,h^k\leqq 2M\quad\forall h:\ 0<h\leqq h_0(\Lambda)$$

$$(4.29)\quad \|u_h\|_{0,\infty,\Omega}\leqq C h^{-\kappa}\,\|u_h\|_X\leqq 2CM\,h^{-\kappa}\equiv M_h.$$

Remark 4.2. It is open whether it holds a uniform (in h) estimate for $\|u_h\|_{0,\infty,\Omega}$. In the linear case one can prove $\|u_h\|_{0,\infty,\Omega}\leqq C h^{-1/4}\,|\log h|^{3/4}$ [18]. Unfortunately, a discrete maximum principle is not valid and the discrete solution has sometimes local oscillations in the neighbourhood of discontinuities.

Proof of Theorem (2). ($\mathcal{N}$) together with Remark 2.4 and ($\mathcal{N}_h$) imply that with $\vartheta=\pi_h u-u_h$ it holds

$$(4.30)\qquad V=[a_0(\pi_h u,\vartheta)-a_0(u_h,\vartheta)]+\sum_i \delta_i(N_\varepsilon(\pi_h u)-N_\varepsilon(u_h), B(u_h)\cdot\nabla\vartheta)_{\tau_i}$$
$$=VI=[a_0(\pi_h u,\vartheta)-a_0(u,\vartheta)]+\sum_i \delta_i(N_\varepsilon(\pi_h u)-N_\varepsilon(u), B(u_h)\cdot\nabla\vartheta)_{\tau_i}.$$

From Lemma 5.4 and Lemma 5.5 we conclude that

$$(4.31)\qquad V\geqq \tfrac{1}{2}|||\vartheta|||^2_{\varepsilon,\delta}-CM\,G(M_h)\{1+MG(M_h)\delta h^{-2\kappa}\}\,|\vartheta|^2_{1,2}$$

and

$$(4.32)\qquad VI\leqq \tfrac{1}{4}|||\vartheta|||^2_{\varepsilon,\delta}+C_4(\Lambda)\,h^{2k}(\varepsilon+h).$$

(4.30)–(4.32) imply

$$|||\vartheta|||^2_{\varepsilon,\delta}\leqq C_4(\Lambda)\,h^{2k}(\varepsilon+h)+C_5(\Lambda)\,\varepsilon\,|\vartheta|^2_{1,2}$$

with

$$(4.33)\qquad C_5(\Lambda):=C\sup_{\lambda\in\Lambda}\lambda M\,G(M_h)(1+MG(M_h)\delta h^{-2\kappa}).$$

Recognizing now (4.27) and *(A.2)* we find from triangle inequality that

$$|\vartheta|_{1,2}\leqq|u-u_h|_{1,2}+|\pi_h u-u|_{1,2}\leqq C_6(\Lambda)\,h^k$$

with

$$(4.34)\qquad C_6(\Lambda):=C_3(\Lambda)+C\,|u|_{k+1,2},$$

hence by (4.33), (4.34) with $[\hat{T}(\Lambda)]^2=C_4(\Lambda)+C_5(\Lambda)[C_6(\Lambda)]^2$

$$(4.35)\qquad |||\vartheta|||_{\varepsilon,\delta}\leqq\hat{T}(\Lambda)\,h^k(\sqrt{\varepsilon}+\sqrt{h}).$$

The triangle inequality together with (4.35) and *(A.2)* concludes the proof. □

5 Auxiliary results

In this section, we summarize some auxiliary results which are needed to prove the main result. We start with the first part of the Theorem (cf. Lemma 5.1–5.3). Note that by *(H.3)* and for sufficiently small h it holds

$$\|\pi_h u\|_{0,\infty,\Omega}\leqq\|u\|_{0,\infty,\Omega}+\|u-\pi_h u\|_{0,\infty,\Omega}\leqq 2M,\qquad 0<h\leqq h_0(\Lambda).$$

As a first step, we show continuity of the operator G_h defined by (cf. (4.15))

$$\langle G_h(\lambda,v),w\rangle:=\langle G_0(\lambda,v),w\rangle+\sum_i\delta_i(-\Delta v+\lambda H(v,\nabla v), B(v)\cdot\nabla w)_{\tau_i}.$$

Lemma 5.1. *Under the assumptions (H.1)–(H.6) it holds*

$$\|G_h(\lambda, \pi_h u(\lambda)) - G_h(\lambda, u(\lambda))\|_{\mathscr{L}(V_h, Y_h)} \leqq C_1(\Lambda)\, h^{k-1}(h+\delta).$$

Proof. We split as follows

$$\text{(5.1)}\qquad \langle G_h(\lambda, \pi_h u), w\rangle - \langle G_h(\lambda, u), w\rangle = \{\langle G_0(\lambda, \pi_h u), w\rangle - \langle G_0(\lambda, u), w\rangle + \sum_i \delta_i \{(-\Delta \pi_h u + \lambda H(\pi_h u, \nabla \pi_h u), B(\pi_h u)\cdot \nabla w)_{\tau_i} - (-\Delta u + \lambda H(u, \nabla u), B(u)\cdot\nabla w)_{\tau_i}\} \equiv I_1 + I_2.$$

The first term can be estimated using first Hölder's inequality, mean value theorem, *(H.1)*, the estimate

$$|\pi_h u|_{1,2,\Omega} \leqq |u - \pi_h u|_{1,2,\Omega} + |u|_{1,2,\Omega} \leqq 2\,\|u\|_X \quad \forall h\colon\ 0 < h \leqq h_0(\Lambda)$$

and secondly Sobolev's imbedding theorem:

$$\begin{aligned}|I_1| &\leqq |([B(\pi_h u) - B(u)]\cdot\nabla \pi_h u + B(u)\cdot\nabla(\pi_h u - u) + [C(\pi_h u) - C(u)], w)_\Omega| \\ &\leqq C\,G(2M)\{\|\pi_h u - u\|_{0,6}\, |\pi_h u|_{1,2}\, \|w\|_{0,3} \\ &\quad + |\pi_h u - u|_{1,2}\, |w|_{1,2} + \|\pi_h u - u\|_{0,2}\, |w|_{1,2}\} \\ &\leqq C\,G(2M)\{\|u - \pi_h u\|_X\, \|u\|_X + \|u - \pi_h u\|_X\}\, \|w\|_X.\end{aligned}$$

When using the mean value theorem, $D_u B$ and $D_u C$ are evaluated at a convex combination of u and $\pi_h u$ and such it will be bounded in the L_∞-norm by $2M$.

By *(A.2)* it follows

$$|I_1| \leqq C\,G(2M)\, h^k (1 + \|u\|_X)\, |u|_{k+1,2}\, \|w\|_X \leqq C_{11}(\Lambda)\, h^k\, \|w\|_X$$

with

$$\text{(5.2)}\qquad C_{11}(\Lambda) := C\,G(2M) \sup_{\lambda\in\Lambda} (1 + |u|_{1,2})\, |u|_{k+1,2}.$$

Similarly we find for the second term in (5.1)

$$\begin{aligned}|I_2| &\leqq |\sum_i \delta_i(-\Delta(\pi_h u - u), B(\pi_h u)\cdot\nabla w)_{\tau_i} + (-\Delta u, [B(\pi_h u) - B(u)]\cdot\nabla w)_{\tau_i} \\ &\quad + \lambda(H(\pi_h u, \nabla \pi_h u) - H(u, \nabla u), B(\pi_h u)\cdot\nabla w)_{\tau_i} \\ &\quad + \lambda(H(u, \nabla u), [B(\pi_h u) - B(u)]\cdot\nabla w)_{\tau_i}\}| \\ &\leqq C\,\delta\, G(2M)\{(\sum_i |u - \pi_h u|^2_{2,2,\tau_i})^{1/2}\, |w|_{1,2} + |u|_{2,2}\, \|u - \pi_h u\|_{0,\infty}\, |w|_{1,2}\} \\ &\quad + \lambda_2\, |\sum_i \delta_i \{([B(\pi_h u) - B(u)]\cdot\nabla \pi_h u + B(u)\cdot\nabla(\pi_h u - u) \\ &\quad + [C(\pi_h u) - C(u)], B(\pi_h u)\cdot\nabla w)_{\tau_i} \\ &\quad + (B(u)\cdot\nabla u + C(u), [B(\pi_h u) - B(u)]\cdot\nabla w)_{\tau_i}\}| \\ &\leqq C\,\delta\, G(2M)(h^{k-1}\, |u|_{k+1,2} + |u|_{2,2}\, h^k\, |u|_{k+1,2})\, |w|_{1,2} \\ &\quad + \lambda_2\, C\,\delta\, [G(2M)]^2 (|\pi_h u|_{1,3}\, \|u - \pi_h u\|_{0,6}\, |w|_{1,2} \\ &\quad + |u - \pi_h u|_{1,2}\, |w|_{1,2} + \|u - \pi_h u\|_{0,2}\, |w|_{1,2} \\ &\quad + |u|_{1,6}\, \|u - \pi_h u\|_{0,3}\, |w|_{1,2} + \|u - \pi_h u\|_{0,2}\, |w|_{1,2}).\end{aligned}$$

Repeated use of *(A.1)*, *(A.2)* and Sobolev's imbedding theorem yield

$$|\pi_h u|_{1,3} \leqq C h^{-1/3} |\pi_h u|_{1,2} \leqq C h^{-1/3} \|u\|_X, \qquad 0 < h \leqq h_0(\Lambda)$$
$$\|u - \pi_h u\|_{0,6} \leqq C |u - \pi_h u|_{1,2} \leqq C h^k |u|_{k+1,2}; \qquad |u|_{1,6} \leqq C |u|_{2,2}$$

and thus with *(A.2)*

$$\begin{aligned} (5.3) \qquad |I_2| &\leqq C \delta G(2M) |u|_{k+1,2} (1 + h |u|_{2,2}) h^{k-1} \|w\|_X \\ &\quad + [G(2M)]^2 \lambda_2 C \delta \|w\|_X (h^{k-1/3} |u|_{1,2} |u|_{k+1,2} + h^k |u|_{k+1,2} \\ &\quad + |u|_{2,2} h^k |u|_{k+1,2}) \\ &\leqq C_{12}(\Lambda) \delta h^{k-1} \|w\|_X \end{aligned}$$

with

$$C_{12}(\Lambda) := C \sup_{\lambda \in \Lambda} |u|_{k+1,2} \{ G(2M)(1 + h |u|_{2,2}) + [G(2M)]^2 \lambda_2 [h^{1-1/3} |u|_{1,2} + h + h |u|_{2,2}]) \}.$$

Summarizing (5.1)–(5.3) we find the assertion of Lemma 5.1 with $C_1(\Lambda) := C_{11}(\Lambda) + C_{12}(\Lambda)$. □

Lemma 5.2 is a continuity result for the Frèchet derivative $D_u G_h(\lambda, \cdot)$ given by

$$\begin{aligned} \langle D_u G_h(\lambda, z) \cdot v, w \rangle &= \underbrace{\lambda \{ \langle B(z) \cdot \nabla v, w \rangle + \langle \tilde{H}(z, \nabla z) v, w \rangle \}}_{= \langle D_u G_0(\lambda, z) \cdot v, w \rangle} \\ &\quad + \sum_i \delta_i \{ (-\Delta v + \lambda B(z) \cdot \nabla v + \lambda \tilde{H}(z, \nabla z) v, B(z) \cdot \nabla w)_{\tau_i} \\ &\quad + ([-\Delta z + \lambda H(z, \nabla z)] v, B_u(z) \cdot \nabla w)_{\tau_i} \}. \end{aligned}$$

Lemma 5.2. *Under the assumptions (H.1)–(H.6) it holds*

$$\|D_u G_h(\lambda, \pi_h u) - D_u G_h(\lambda, u)\|_{\mathscr{L}(V_h, Y_h)} \leqq C_2(\Lambda) h^k (1 + \delta h^{-1-\kappa}).$$

Proof. We split as follows

$$\begin{aligned} (5.4) \qquad & \langle [D_u G_h(\lambda, \pi_h u) - D_u G_h(\lambda, u)] v, w \rangle \\ &= \langle [D_u G_0(\lambda, \pi_h u) - D_u G_0(\lambda, u)] v, w \rangle \\ &\quad + \sum_i \delta_i \{ (-\Delta v, [B(\pi_h u) - B(u)] \cdot \nabla w)_{\tau_i} \\ &\quad + (-\Delta(\pi_h u - u) v, D_u B(\pi_h u) \cdot \nabla w)_{\tau_i} \\ &\quad + (-\Delta u \cdot v, [D_u B(\pi_h u) - D_u B(u)] \cdot \nabla w)_{\tau_i} \} \\ &\quad + \lambda \sum_i \delta_i \{ ([B(\pi_h u) - B(u)] \cdot \nabla v, B(\pi_h u) \cdot \nabla w)_{\tau_i} \\ &\quad + (B(u) \cdot \nabla v, [B(\pi_h u) - B(u)] \cdot \nabla w)_{\tau_i} \} \\ &\quad + \lambda \sum_i \delta_i \{ (\tilde{H}(\pi_h u, \nabla \pi_h u) v, B(\pi_h u) \cdot \nabla w)_{\tau_i} - (\tilde{H}(u, \nabla u) v, B(u) \cdot \nabla w)_{\tau_i} \\ &\quad + \lambda \sum_i \delta_i \{ (\tilde{H}(\pi_h u, \nabla \pi_h u) v, D_u B(\pi_h u) \cdot \nabla w)_{\tau_i} \\ &\quad - (\tilde{H}(u, \nabla u) v, D_u B(u) \cdot \nabla w)_{\tau_i} \} \\ &\equiv II_1 + II_2 + II_3 + II_4 + II_5. \end{aligned}$$

With similar arguments as in the proof of Lemma 5.1 and using (3.2) we find

$$(5.5)\qquad \begin{aligned} |II_1| &= \lambda\,|([B(\pi_h u)-B(u)]\cdot\nabla v, w)+(([D_u B(\pi_h u)-D_u B(u)]\cdot\nabla\pi_h u \\ &\quad + D_u B(u)\cdot\nabla(\pi_h u-u)+[D_u C(\pi_h u)-D_u C(u)])\,v, w)| \\ &\leqq \lambda_2 C G(2M)(\|\pi_h u-u\|_{0,6}\,|v|_{1,2}\,\|w\|_{0,3} \\ &\quad +|\pi_h u|_{1,2}\,\|u-\pi_h u\|_{0,6}\,\|v\|_{0,6}\,\|w\|_{0,6} \\ &\quad +|u-\pi_h u|_{1,2}\,\|v\|_{0,6}\,\|w\|_{0,3}+\|u-\pi_h u\|_{0,3}\,\|v\|_{0,3}\,\|w\|_{0,3}) \\ &\leqq C G(2M)\,\lambda_2\,\|v\|_X\,\|w\|_X(h^k\,|u|_{k+1,2}+(1+|u|_{1,2})\,h^k\,|u|_{k+1,2}) \\ &\leqq C_{21}(\Lambda)\,h^k\,\|v\|_X\,\|w\|_X; \end{aligned}$$

with

$$C_{21}(\Lambda):=\sup_{\lambda\in\Lambda} C\,G(2M)\,\lambda(1+|u|_{1,2})\,|u|_{k+1,2}$$

$$(5.6)\qquad \begin{aligned} |II_2| &\leqq C\,\delta\,G(2M)\{h^{-1}\,|v|_{1,2}\,\|u-\pi_h u\|_{0,\infty}\,|w|_{1,2} \\ &\quad +(\sum_i |u-\pi_h u|^2_{2,2,\tau_i})^{1/2}\,h^{-\kappa}\,|v|_{1,2}\,|w|_{1,2} \\ &\quad +|u|_{2,2}\,h^{-\kappa}\,|v|_{1,2}\,\|u-\pi_h u\|_{0,\infty}\,|w|_{1,2} \\ &\leqq C\,\delta\,G(2M)(h^{k-1}+h^{k-1-\kappa}+h^{k-\kappa}\,|u|_{2,2})\,|u|_{k+1,2}\,\|v\|_X\,\|w\|_X \\ &\leqq C_{22}(\Lambda)\,\delta\,h^{k-1-\kappa}\,\|v\|_X\,\|w\|_X, \end{aligned}$$

with

$$C_{22}(\Lambda):=\sup_{\lambda\in\Lambda} G(2M)\,C(1+h^\kappa+h\,|u|_{2,2})\,|u|_{k+1,2}$$

$$(5.7)\qquad \begin{aligned} |II_3| &\leqq \lambda\,C\,\delta[G(2M)]^2\,\|u-\pi_h u\|_{0,\infty}\,|v|_{1,2}\,|w|_{1,2} \\ &\leqq C_{23}(\Lambda)\,\delta\,h^k\,\|v\|_X\,\|w\|_X; \end{aligned}$$

with

$$C_{23}(\Lambda):=\sup_{\lambda\in\Lambda}\lambda[G(2M)]^2\,|u|_{k+1,2}$$

$$(5.8)\qquad \begin{aligned} |II_4| &\leqq \lambda\sum_i \delta_i\,|(v[D_u B(\pi_h u)-D_u B(u)]\cdot\nabla\pi_h u,\ B(\pi_h u)\cdot\nabla w)_{\tau_i} \\ &\quad +(v[D_u B(u)\cdot\nabla\pi_h u],\ [B(\pi_h u)-B(u)]\cdot\nabla w)_{\tau_i} \\ &\quad +(v[D_u B(u)\cdot\nabla(\pi_h u-u)],\ B(u)\cdot\nabla w)_{\tau_i} \\ &\quad +\lambda\sum_i \delta_i\,|(v[D_u C(\pi_h u)-D_u C(u)],\ B(\pi_h u)\cdot\nabla w)_{\tau_i} \\ &\quad +(v\,D_u C(u),\ [B(\pi_h u)-B(u)]\cdot\nabla w)_{\tau_i}| \\ &\leqq \lambda_2\,\delta\,C[G(2M)]^2(|\pi_h u|_{1,2}\,\|u-\pi_h u\|_{0,\infty}\,\|v\|_{0,\infty}\,|w|_{1,2} \\ &\quad +|u-\pi_h u|_{1,2}\,\|v\|_{0,\infty}\,|w|_{1,2}+\|u-\pi_h u\|_{0,6}\,\|v\|_{0,3}\,|w|_{1,2}) \\ &\leqq C\,\lambda_2\,\delta[G(2M)]^2(|u|_{1,2}\,h^{k-\kappa}+h^{k-\kappa})\,|u|_{k+1,2}\,\|v\|_X\,\|w\|_X \\ &\leqq C_{24}(\Lambda)\,h^{k-\kappa}\,\delta\,\|v\|_X\,\|w\|_X \end{aligned}$$

with

$$C_{24}(\Lambda):=\sup_{\lambda\in\Lambda} C\,\lambda[G(2M)]^2\,|u|_{k+1,2}(|u|_{1,2}+1).$$

The same estimate is valid for term $|II_5|$, hence summarizing (5.4)–(5.8) the assertion of Lemma 5.2 follows with $C_2(\Lambda) := \sum_{i=1}^{4} C_{2i}(\Lambda)$. □

Lemma 5.3 shows essentially that the operator G_h is locally Lipschitz continuous.

Lemma 5.3. *Under the assumptions (H.1)–(H.6) it holds for* u_1, $u_2 \in B(\pi_h u; \rho) \cap V_h$ *with* $\rho \leqq C h^{\bar\kappa}$, $\bar\kappa > \kappa$

$$\|G_h(\lambda, u_1) - G_h(\lambda, u_2) - D_u G_h(\lambda, \pi_h u)\cdot(u_1 - u_2)\|_h \leqq L_h(\lambda, \rho, \|\pi_h u\|_X)\, \|u_1 - u_2\|_X$$

with

$$\lim_{h\to 0} \sup_{\lambda\in\Lambda} L_h(\lambda, 0, \mu) = 0 \qquad \forall \mu \in \mathbb{R}^+.$$

Proof. We split as follows with $e = u_1 - u_2$, $z = \pi_h u$

$$\begin{aligned}
(5.9)\qquad & |\langle G_h(\lambda, u_1), w\rangle - \langle G_h(\lambda, u_2), w\rangle - \langle D_u G_h(\lambda, z)\cdot(u_1 - u_2), w\rangle| \\
&\leqq |\langle G_0(\lambda, u_1), w\rangle - \langle G_0(\lambda, u_2), w\rangle - \langle D_u G_0(\lambda, z)\cdot(u_1 - u_2), w\rangle| \\
&\quad + |\sum_i \delta_i \{(-\Delta u_1 + \lambda H(u_1, \nabla u_1), B(u_1)\cdot\nabla w)_{\tau_i} \\
&\quad - (-\Delta u_2 + \lambda H(u_2, \nabla u_2), B(u_2)\cdot\nabla w)_{\tau_i} \\
&\quad - (-\Delta e + \lambda B(z)\cdot\nabla e + \lambda \tilde H(z, \nabla z) e, B(z)\cdot\nabla w)_{\tau_i} \\
&\quad - ([-\Delta z + \lambda H(z, \nabla z)]\, e, D_u B(z)\cdot\nabla w)_{\tau_i}\}| \equiv III + IV.
\end{aligned}$$

For using *(H.1)* we need a L_∞-bound for elements in the ball $B(\pi_h u; \rho) \cap V_h$:

$$\begin{aligned}
\|v\|_{0,\infty,\Omega} &\leqq \|\pi_h u\|_{0,\infty,\Omega} + \|\pi_h u - v\|_{0,\infty,\Omega} \leqq 2M + C h^{-\kappa} \|\pi_h u - v\|_{1,2,\Omega} \\
&\leqq 2M + C h^{\bar\kappa - \kappa} \leqq 3M \qquad \forall h:\ 0 < h \leqq h_0(\Lambda).
\end{aligned}$$

With similar arguments as in the proof of (5.2) we find

$$\begin{aligned}
III &\equiv \lambda |(B(u_1)\cdot\nabla u_1 - B(u_2)\cdot\nabla u_2 + C(u_1) - C(u_2), w) \\
&\quad - (B(z)\cdot\nabla(u_1 - u_2), w) \\
&\quad - ([D_u B(z)\cdot\nabla z + D_u C(z)](u_1 - u_2), w)| \\
&\leqq \lambda |([B(u_1) - B(u_2)]\cdot\nabla u_1, w) \\
&\quad + ([B(u_2) - B(z)]\cdot\nabla(u_1 - u_2), w) \\
&\quad - ([D_u B(z)\cdot\nabla z](u_1 - u_2), w)| \\
&\quad + \lambda_2 |(C(u_1) - C(u_2) - D_u C(z)(u_1 - u_2), w)| \\
&\equiv III_1 + III_2
\end{aligned}$$

$$\begin{aligned}
III_1 &\leqq \lambda |([D_u B(z_1)\cdot\nabla u_1]\, e, w) - ([D_u B(z)\cdot\nabla z]\, e, w) \\
&\quad + ([D_u B(z_2)\cdot\nabla e](u_2 - z), w)| \\
&\quad (\text{with } z_1 := \vartheta_1 u_1 + (1 - \vartheta_1) u_2,\ z_2 := \vartheta_2 u_2 + (1 - \vartheta_2) z,\ 0 < \vartheta_i < 1,\ i = 1, 2) \\
&\leqq \lambda |(([D_u B(z_1) - D_u B(z)]\cdot\nabla u_1)\, e, w) \\
&\quad + ([D_u B(z)\cdot\nabla(u_1 - z)]\, e, w) + ([D_u B(z_2)\cdot\nabla e](u_2 - z), w)| \\
&\leqq \lambda C G(3M) \{|u_1|_{1,2}\, \|z_1 - z_2\|_{0,6}\, \|e\|_{0,6}\, \|w\|_{0,6} \\
&\quad + |u_1 - z|_{1,2}\, \|e\|_{0,4}\, \|w\|_{0,4} + |e|_{1,2}\, \|u_2 - z\|_{0,4}\, \|w\|_{0,4}\}.
\end{aligned}$$

Sobolev's imbedding theorem and the estimates $\|u_i - z\|_X \leqq \rho$, $\|u_i\|_X \leqq \|z\|_X + \rho$ imply

$$(5.10)\qquad III_1 \leqq C G(3M)\, \lambda \rho (1 + \|z\|_X + \rho)\, \|e\|_X\, \|w\|_X.$$

On the other hand it is similarly

(5.11) $$III_2 \leqq C\,G(3M)\,\lambda\,\rho\,\|e\|_X\,\|w\|_X,$$

hence summarizing (5.10), (5.11)

(5.12) $$III \leqq L_h^1(\lambda, \rho, \|z\|_X)\,\|e\|_X\,\|w\|_X, \qquad L_h^1 := C\,G(3M)\,\lambda\,\rho(1+\|z\|_X+\rho).$$

Further we have

(5.13) $$\begin{aligned} IV \leqq & \Big|\sum_i \delta_i\{(-\Delta(u_1-u_2), B(u_1)\cdot\nabla w)_{\tau_i} \\ & +(-\Delta u_2, [B(u_1)-B(u_2)]\cdot\nabla w)_{\tau_i} \\ & +(\Delta e, B(z)\cdot\nabla w)_{\tau_i}+(\Delta z\cdot e, D_u B(z)\cdot\nabla w)_{\tau_i}\}\Big| \\ & +\lambda\Big|\sum_i \delta_i(H(u_1,\nabla u_1)-H(u_2,\nabla u_2), B(u_1)\cdot\nabla w)_{\tau_i}\Big| \\ & +\lambda\Big|\sum_i \delta_i(H(u_2,\nabla u_2), [B(u_1)-B(u_2)]\cdot\nabla w)_{\tau_i}\Big| \\ & +\lambda\Big|\sum_i \delta_i(-B(z)\cdot\nabla e, B(z)\cdot\nabla w)_{\tau_i}\Big| \\ & +\lambda\Big|\sum_i \delta_i(-\tilde{H}(z,\nabla z)e, B(z)\cdot\nabla w)_{\tau_i}\Big| \\ & +\lambda\Big|\sum_i \delta_i(-H(z,\nabla z)e, D_u B(z)\cdot\nabla w)_{\tau_i}\Big| \equiv \sum_{i=1}^{6} IV_i \end{aligned}$$

with

(5.14) $$\begin{aligned} IV_1 \leqq & \Big|\sum_i \delta_i\{(-\Delta e, B(u_1)\cdot\nabla w)_{\tau_i} \\ & +(\Delta e, B(z)\cdot\nabla w)_{\tau_i}+(\Delta z\cdot e, D_u B(z)\cdot\nabla w)_{\tau_i} \\ & -(\Delta u_2, [D_u B(z_1)\cdot\nabla w]\,e)_{\tau_i}\}\Big| \\ \leqq & C\,G(3M)\,\delta\,h^{-1}\,|w|_{1,2} \\ & \cdot\{|e|_{1,2}+|z|_{1,2}\,h^{-\kappa}\,|e|_{1,2}+|u_2|_{1,2}\,h^{-\kappa}\,|e|_{1,2}\} \\ \leqq & C\,G(3M)\,\delta\,h^{-1-\kappa}(h^\kappa+\|z\|_X+\rho)\,\|e\|_X\,\|w\|_X, \end{aligned}$$

(5.15) $$\begin{aligned} IV_2 \leqq & \lambda\,\delta\,|(B(u_1)\cdot\nabla(u_1-u_2)+[B(u_1)-B(u_2)]\cdot\nabla u_2| \\ & +[C(u_1)-C(u_2)], B(u_1)\cdot\nabla w)| \\ \leqq & C\,G(3M)\,\lambda\,\delta\,|w|_{1,2} \\ & \cdot(|e|_{1,2}+h^{-\kappa}\,|e|_{1,2}\,|u|_{1,2}+G(3M)\,\|e\|_{0,2}) \\ \leqq & C\,G(3M)\,\lambda\,\delta\,h^{-\kappa}(h^\kappa+\|z\|_X+\rho\,G(3M))\,\|e\|_X\,\|w\|_X, \end{aligned}$$

(5.16) $$\begin{aligned} IV_3 \leqq & \lambda\,\delta\,|(B(u_2)\cdot\nabla u_2+C(u_2), [D_u B(z_1)\cdot\nabla w]\,e)| \\ \leqq & C\,\lambda\,\delta[G(3M)]^2(|u_2|_{1,2}+1)\,h^{-\kappa}\,|e|_{1,2}\,|w|_{1,2} \\ \leqq & C\,\lambda\,\delta[G(3M)]^2\,h^{-\kappa}(1+\|z\|_X+\rho)\,\|e\|_X\,\|w\|_X, \end{aligned}$$

(5.17) $$IV_4 \leqq C\,\lambda\,\delta[G(3M)]^2\,\|e\|_X\,\|w\|_X,$$

(5.18) $$\begin{aligned} IV_5 \leqq & \lambda\,\delta\,|([D_u B(z)\cdot\nabla z+D_u C(z)]\,e, B(z)\cdot\nabla w)| \\ \leqq & C\,\lambda\,\delta[G(3M)]^2\,h^{-\kappa}(1+|z|_{1,2})\,|e|_{1,2}\,|w|_{1,2} \\ \leqq & C\,\lambda\,\delta[G(3M)]^2\,h^{-\kappa}(1+\|z\|_X)\,\|e\|_X\,\|w\|_X, \end{aligned}$$

(5.19) $$\begin{aligned} IV_6 \leqq & \lambda\,\delta\,|([B(z)\cdot\nabla z+C(z)]\,e, D_u B(z)\cdot\nabla w)| \\ \leqq & C[G(3M)]^2\,\lambda\,\delta\,h^{-\kappa}\,\|e\|_X(1+\|z\|_X)\,\|w\|_X. \end{aligned}$$

(5.13)–(5.19) yield

$$IV \leqq L_h^2(\lambda, \rho, \|z\|_X)\, \|e\|_X\, \|w\|_X, \tag{5.20}$$

with

$$L_h^2 := C[G(3M)]^2 \{\delta h^{-1-\kappa}(h^\kappa + \|z\|_X + \rho)(1+\lambda) + \lambda \delta h^{-\kappa}(1+\|z\|_X)\},$$

hence from (5.9), (5.12), (5.20) it follows

$$III + IV \leqq L_h(\lambda, \rho, \|z\|_X)\, \|e\|_X\, \|w\|_X, \tag{5.21}$$

$$L_h := L_h^1 + L_h^2 := C\, G(3M)\{\lambda \rho (1+\|z\|_X+\rho) + \delta h^{-1-\kappa} G(3M) \cdot (h^\kappa + \|z\|_X + \rho)(1+\lambda) + G(3M)\lambda \delta h^{-\kappa}(1+\|z\|_X)\}.$$

(5.21) together with *(H.6)* implies

$$\lim_{h\to 0} \sup_{\lambda \in A} L_h(\lambda, 0, \mu) = \lim_{h \to 0} C[G(3M)]^2 \{\delta h^{-1-\kappa}(h^\kappa + \mu)(1+\lambda_2) + \lambda_2 \delta h^{-\kappa}(1+\mu)\} = 0. \quad \square$$

Thus the auxiliary results for the first part of the Theorem and for the estimate (4.27) are complete. Note that as a consequence of *(H.3)* and (4.27) we have the following estimates

$$\|u_h\|_X \leqq 2M, \qquad \|u_h\|_{0,\infty,\Omega} \leqq 2CMh^{-\kappa} \equiv M_h. \tag{5.22}$$

As a last point, we consider auxiliary results from standard SDFEM analysis to prove Theorem (2).

Lemma 5.4. *Under the assumptions (H.1)–(H.6) it holds with* $\vartheta = \pi_h u - u_h$

$$V = a_0(\pi_h u, \vartheta) - a_0(u_h, \vartheta) + \sum_i \delta_i (N_\varepsilon(\pi_h u) - N_\varepsilon(u_h), B(u_h)\cdot \nabla \vartheta)_{\tau_i}$$

$$\geqq \tfrac{1}{2} |||\vartheta|||_{\varepsilon,\delta}^2 - CM\, G(M_h)(1 + M\, G(M_h)\delta h^{-2\kappa})\, |\vartheta|_{1,2}^2.$$

Proof. Hölder's inequality yields

$$\begin{aligned} &|\sum_i \delta_i(-\varepsilon \Delta \vartheta, B(u_h)\cdot \nabla \vartheta)_{\tau_i} \\ &\leqq \sum_i \varepsilon \sqrt{\delta_i}\, \|\Delta \vartheta\|_{0,2,\tau_i} \sqrt{\delta_i}\, \|B(u_h)\cdot\nabla\vartheta\|_{0,2,\tau_i} \\ &\leqq \Big(\sum_i \varepsilon^2 \delta_i \|\Delta\vartheta\|_{0,2,\tau_i}^2\Big)^{1/2} \Big(\sum_i \delta_i \|B(u_h)\cdot\nabla\vartheta\|_{0,2,\tau_i}^2\Big)^{1/2} \\ &\leqq \tfrac{1}{4} \sum_i \delta_i \|B(u_h)\cdot\nabla\vartheta\|_{0,2,\tau_i}^2 \\ &\quad + \sum_i C_I^2\, \varepsilon(\varepsilon \delta_i h_i^{-2})\, \|\nabla\vartheta\|_{0,2,\tau_i}^2 \qquad \text{(by } (A.1)) \\ &\leqq \tfrac{1}{4} \sum_i \delta_i \|B(u_h)\cdot\nabla\vartheta\|_{0,2,\tau_i}^2 + \tfrac{1}{4}\varepsilon |\vartheta|_{1,2}^2 \quad \text{(by } (H.6)) \end{aligned} \tag{5.23}$$

where C_I is the constant in the inverse inequality *(A.1)* and we assumed that $4C_I^2 K \leqq 1$ in *(H.6)*. Thus

$$(5.24)\qquad V=\varepsilon|\vartheta|_{1,2}^2-\sum_i \varepsilon\delta_i(\Delta\vartheta, B(u_h)\cdot\nabla\vartheta)_{\tau_i} + \underbrace{\sum_i (H(\pi_h u, \nabla\pi_h u)-H(u_h, \nabla u_h), \vartheta+\delta_i B(u_h)\cdot\nabla\vartheta)_{\tau_i}}_{=:V_1}$$
$$\geqq \varepsilon|\vartheta|_{1,2}^2-\tfrac{1}{4}|||\vartheta|||_{\varepsilon,\delta}^2+V_1.$$

In more detail, we have

$$(5.25)\qquad V_1=\sum_i (B(u_h)\cdot\nabla\vartheta+(B(\pi_h u)-B(u_h))\cdot\nabla\pi_h u + C(\pi_h u)-C(u_h), \vartheta+\delta_i B(u_h)\cdot\nabla\vartheta)_{\tau_i}$$
$$=\sum_i \delta_i\|B(u_h)\cdot\nabla\vartheta\|_{0,2,\tau_i}^2+\underbrace{(B(u_h)\cdot\nabla\vartheta, \vartheta)}_{=:V_2} + \underbrace{\sum_i ([B(\pi_h u)-B(u_h)]\cdot\nabla\pi_h u+[C(\pi_h u)-C(u_h)], \vartheta+\delta_i B(u_h)\cdot\nabla\vartheta)_{\tau_i}}_{=V_3}.$$

Integration by parts and *(H.1)*, (5.21) yield

$$(5.26)\qquad |V_2|=|-\tfrac{1}{2}(\vartheta, \vartheta\operatorname{div}B(u_h))|\leqq C\,G(M_h)\,\|\vartheta\|_{0,2}^2.$$

Using *(H.1)*, standard inequalities and Sobolev's imbedding theorem, we find

$$(5.27)\qquad |V_3|\leqq G(M_h)(\|\vartheta\|_{0,4}|\pi_h u|_{1,2}\|\vartheta\|_{0,4}+\|\vartheta\|_{0,2}^2) + G(M_h)\sum_i \sqrt{\delta_i}(\|\vartheta\|_{0,\infty}|\pi_h u|_{1,2,\tau_i}+\|\vartheta\|_{0,2,\tau_i})\sqrt{\delta_i}\,\|B(u_h)\cdot\nabla\vartheta\|_{0,2,\tau_i}$$
$$\leqq C\,G(M_h)(|u|_{1,2}\|\vartheta\|_{1,2}^2+\|\vartheta\|_{0,2}^2)+\tfrac{1}{4}\sum_i \delta_i\|B(u_h)\cdot\nabla\vartheta\|_{0,2,\tau_i}^2 + C[G(M_h)]^2\,\delta(h^{-2\kappa}|\vartheta|_{1,2}^2 M^2+\|\vartheta\|_{0,2}^2)$$
$$\leqq \tfrac{1}{4}\sum_i \delta_i\|B(u_h)\cdot\nabla\vartheta\|_{0,2,\tau_i}^2+C\,G(M_h) \cdot\{M\|\vartheta\|_{1,2}^2+G(M_h)M^2\delta h^{-2\kappa}|\vartheta|_{1,2}^2\}$$
$$\leqq \tfrac{1}{4}\sum_i \delta_i\|B(u_h)\cdot\nabla\vartheta\|_{0,2,\tau_i}^2 + C\,M\,G(M_h)\|\vartheta\|_{1,2}^2(1+M\,G(M_h)\,\delta h^{-2\kappa}).$$

The assertion follows from (5.24)–(5.27) and Friedrich's inequality:

$$V\geqq \varepsilon|\vartheta|_{1,2}^2-\tfrac{1}{4}|||\vartheta|||_{\varepsilon,\delta}^2+\sum_i \delta_i\|B(u_h)\cdot\nabla\vartheta\|_{0,2,\tau_i}^2-|V_2|-|V_3|$$
$$\geqq \tfrac{1}{2}|||\vartheta|||_{\varepsilon,\delta}^2-C\,M\,G(M_h)\,\|\vartheta\|_{1,2}^2(1+M\,G(M_h)\,\delta h^{-2\kappa}).\quad\square$$

Lemma 5.5. *Under the assumptions (H.1)–(H.6) it holds with* $\vartheta=\pi_h u-u_h$

$$VI=[a_0(\pi_h u,\vartheta)-a_0(u,\vartheta)]+\sum_i \delta_i(N_\varepsilon(\pi_h u)-N_\varepsilon(u), B(u_h)\cdot\nabla\vartheta)_{\tau_i}$$
$$=\tfrac{1}{4}\,|||\vartheta|||^2_{\varepsilon,\delta}+C_4(\Lambda)\,h^{2k}(\varepsilon+h).$$

Proof. With $\eta=\pi_h u-u$ we find

$$\begin{aligned}(5.28)\quad VI=&\,\varepsilon(\nabla\eta,\nabla\vartheta)+(H(\pi_h u,\nabla\pi_h u)-H(u,\nabla u),\vartheta)\\&+\sum_i\delta_i\{(-\varepsilon\Delta\eta, B(u_h)\cdot\nabla\vartheta)_{\tau_i}\\&+(H(\pi_h u,\nabla\pi_h u)-H(u,\nabla u), B(u_h)\cdot\nabla\vartheta)_{\tau_i}\}\\ \leqq&\,\tfrac{1}{8}\varepsilon|\vartheta|^2_{1,2}+C\varepsilon|\eta|^2_{1,2}+\tfrac{1}{8}\sum_i\delta_i\|B(u_h)\cdot\nabla\vartheta\|^2_{0,2,\tau_i}\\&+C\varepsilon^2\sum_i\delta_i|\eta|^2_{2,2,\tau_i}+\underbrace{|(B(\pi_h u)\cdot\nabla\eta,\vartheta)|}_{=:VI_1}\\&+\underbrace{|([B(\pi_h u)-B(u)]\cdot\nabla u+[C(\pi_h u)-C(u)],\vartheta)|}_{=:VI_2}\\&+\underbrace{|\sum_i\delta_i(B(\pi_h u)\cdot\nabla\eta+[B(\pi_h u)-B(u)]\cdot\nabla u+[C(\pi_h u)-C(u)], B(u_h)\cdot\nabla\vartheta)_{\tau_i}|}_{=:VI_3}.\end{aligned}$$

Integration by parts, *(H.1)* and Friedrich's inequality yield

$$\begin{aligned}(5.29)\quad |VI_1|&=|-(\eta, B(\pi_h u)\cdot\nabla\vartheta)-(\eta,\vartheta\,\mathrm{div}\,B(\pi_h u))|\\&\leqq C\,G(2M)\,\|\eta\|_{0,2}(|\vartheta|_{1,2}+\|\vartheta\|_{0,2})\\&\leqq C\,G(2M)\,\|\eta\|_{0,2}\,|\vartheta|_{1,2}.\end{aligned}$$

Next we find by *(H.1)*, Hölder's inequality, Sobolev's imbedding theorem and Friedrich's inequality that

$$\begin{aligned}(5.30)\quad |VI_2|&\leqq G(2M)\{\|\eta\|_{0,2}\,|u|_{1,3}\,\|\vartheta\|_{0,6}+\|\eta\|_{0,2}\,\|\vartheta\|_{0,2}\}\\&\leqq C\,G(2M)\,\|\eta\|_{0,2}(|u|_{2,2}+1)\,|\vartheta|_{1,2}\\&\leqq C\,\|\eta\|^2_{0,2}+C[G(2M)]^2(1+|u|_{2,2})^2\,|\vartheta|^2_{1,2},\end{aligned}$$

$$\begin{aligned}(5.31)\quad |VI_3|&\leqq\sum_i(\sqrt{\delta_i}\,\|B(u_h)\cdot\nabla\vartheta\|_{0,2,\tau_i})\,G(2M)\\&\quad\cdot\sqrt{\delta_i}(|\eta|_{1,2,\tau_i}+\|\eta\|_{0,6,\tau_i}\,|u|_{1,3,\tau_i}+\|\eta\|_{0,2,\tau_i})\\&\leqq\tfrac{1}{8}\sum_i\delta_i\,\|B(u_h)\cdot\nabla\vartheta\|^2_{0,2,\tau_i}\\&\quad+C[G(2M)]^2\cdot\delta(|\eta|^2_{1,2}+|u|^2_{2,2}\,|\eta|^2_{1,2}+\|\eta\|^2_{0,2}).\end{aligned}$$

Summarizing (5.28)–(5.31) we arrive at

$$VI \leqq \tfrac{1}{4} |||\vartheta|||_{\varepsilon,\delta}^2 + C[G(2M)]^2 (1+|u|_{2,2}^2)\,|\vartheta|_{1,2}^2 + C\{\varepsilon |\eta|_{1,2}^2 + \varepsilon^2 \sum_i \delta_i |\eta|_{2,2,\tau_i}^2 + \|\eta\|_{0,2}^2 + [G(2M)]^2\,\delta(1+|u|_{2,2}^2)\,|\eta|_{1,2}^2\}$$

and by *(A.2)*

(5.32) $$VI \leqq \tfrac{1}{4} |||\vartheta|||_{\varepsilon,\delta}^2 + C_{41}(\Lambda)\,\varepsilon\,|\vartheta|_{1,2}^2 + C_{42}(\Lambda)\,h^{2k}(\varepsilon+h+\delta)$$

where

(5.33) $$C_{41}(\Lambda) := C[G(2M)]^2(1+|u|_{2,2}^2)\,\lambda$$

(5.34) $$C_{42}(\Lambda) := C\,|u|_{k+1,2}^2 (1+[G(2M)]^2(1+|u|_{2,2}^2)).$$

By means of (4.27), triangle inequality and *(A.2)* it follows

(5.35) $$|\vartheta|_{1,2} \leqq |\eta|_{1,2} + |u-u_h|_{1,2} \leqq C h^k |u|_{k+1,2} + C_3(\Lambda)\,h^k = C_{43}(\Lambda)\,h^k \quad \text{with} \quad C_{43}(\Lambda) := C_3(\Lambda) + C\,|u|_{k+1,2}.$$

Hence with *(A.6)* we arrive at

(5.36) $$VI \leqq \tfrac{1}{4} |||\vartheta|||_{\varepsilon,\delta}^2 + C_4(\Lambda)\,h^{2k}(\varepsilon+h)$$

where

(5.37) $$C_4(\Lambda) := C_{41}(\Lambda)\cdot[C_{43}(\Lambda)]^2 + C_{42}(\Lambda)$$

which concludes the proof. □

6 Numerical example

We mention that the proof of Theorem (1) in Sect. 4 justifies Newton's method for $(\mathcal{N}_h)$ in a neighbourhood of the nonsingular branch $\{(\lambda, u(\lambda)\colon \lambda\in\Lambda,\ u\in X\}$ of solutions of $(\mathcal{N})$. In order to find the solution, we studied in [23] simple $1D$-problems with parameter continuation (in $\lambda=\varepsilon^{-1}$) in connection with Newton's method [13]. Some numerical $2D$-results for $B(x,u)\equiv b(x)$ can be found in [22] and [31].

For nonlinear equations of Burger's type

(6.1) $$-\varepsilon\Delta u + B(x,u)\cdot\nabla u = f \quad \text{in } \Omega\subset\mathbb{R}^2$$

we studied also the simple iteration technique (frequently used for approximating the Navier-Stokes equations)

(6.2) Find $u_h^{m+1}\in V_h$, $m\in\mathbb{N}_0$ s.t. $\forall v\in V_h$

$$\varepsilon(\nabla u_h^{m+1}, \nabla v)_\Omega + (B(u_h^m)\cdot\nabla u_h^{m+1}, v)_\Omega + \sum_i \delta_i(-\varepsilon\Delta u_h^{m+1} + B(u_h^m)\cdot\nabla u_h^{m+1} - f,\ B(u_h^m)\cdot\nabla v)_{\tau_i} = (f, v)_\Omega.$$

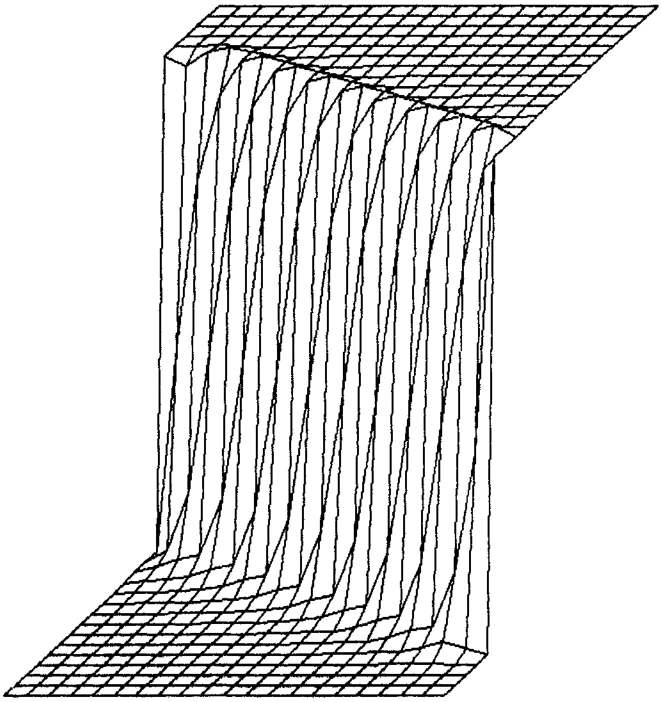

Fig. 1. Discrete solution of (6.2)–(6.4), $\varepsilon = 10^{-6}$, $h = \frac{1}{20}$

As an example we consider [cf. Giles/Rose (1984), J. Comput. Phys. **56**, 513–529]

$$(6.3) \qquad -\varepsilon \Delta u + u \frac{\partial u}{\partial x_1} + 0.4 \frac{\partial u}{\partial x_2} = 0 \quad \text{in } \Omega = (0, 1)^2$$

with

$$(6.4) \qquad u = \begin{cases} 0.8, & \text{if } x_1 \leqq 0.3 + 0.5 x_2,\ 0 \leqq x_2 \leqq 1 \\ -0.4, & \text{if } x_1 > 0.3 + 0.5 x_2,\ 0 \leqq x_2 \leqq 1 \end{cases}$$

$$\frac{\partial u}{\partial n} = 0, \quad \text{if } x_2 = 1,\ 0 < x_1 < 1$$

with an interior layer at $x_1 = 0.3 + 0.5 x_2$. With piecewise linear approximation ($k = 1$) and on an uniform 20×20 mesh on Ω, a converged solution of (6.2) was obtained for $\varepsilon = 10^{-6}$ after 30 iteration steps. The discrete solution is presented in Figs. 1–3.

7 Remarks. Open problems

7.1. The constant $T(\Lambda)$ in our Theorem in Sect. 4 strongly depends on ε_2, $0 < \varepsilon_2 \leqq \varepsilon$. In the special case of strict monotone problems $(\mathcal{N})$ such that

$$(7.1) \qquad \exists \gamma > 0\ \forall v \in X\colon\ \langle D_u N_\varepsilon(v) \cdot w, w \rangle \geqq \gamma \{\varepsilon |w|_{1,2}^2 + \|w\|_{0,2}^2\} \qquad \forall w \in X$$

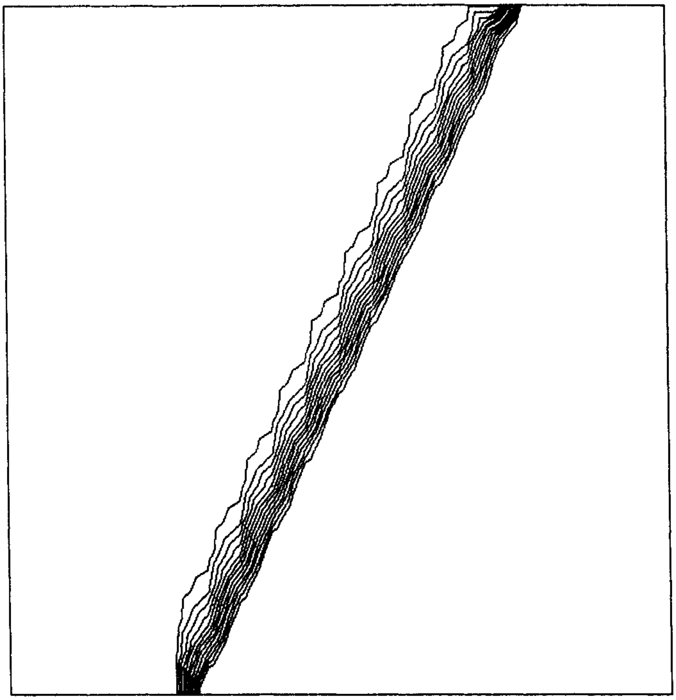

Fig. 2. Discrete solution of (6.2) (6.4), isolines

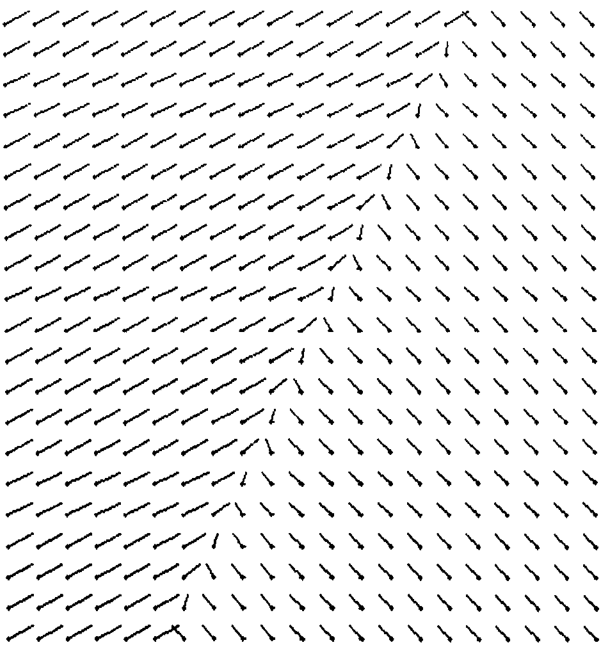

Fig. 3. Example (6.3), (6.4), Vector plot $B(u) = (u;\ 0.4)$ after 30 iteration steps in (6.2)

it is shown in [23] that even for $\varepsilon \geqq 0$

(7.2) $$|||u-u_h|||_{\varepsilon,\delta} \leqq C h^k(\sqrt{\varepsilon}+\sqrt{h})\,|u|_{k+1,2}$$

as in the linear case $B(x, u)=b(x)$, $C(x, u)=c(x)\,u-f(x)$. Unfortunately, (7.1) implies $B(x, u) \equiv b(x)$. It is conjectured that (7.2) remains valid in the somewhat more general case where in (7.1) v is replaced by the solution u of $(\mathcal{N})$.

7.2. In the genuinely quasilinear case $(\mathcal{N})$ our analysis obviously does not work in the limit case $\varepsilon=0$. It is an open problem whether the analysis of an improved SDFEM variant ("shock capturing streamline diffusion method") introduced in [29] and [30] for time-dependent conservation laws and the incompressible Navier-Stokes equations can be used as a remedy. It is essentially based on special (compensated) compactness arguments in $L_\infty(\Omega)$.

7.3. The analysis in Sects. 4 and 5 remains partially correct in the $3D$-case. Note that the basic estimate (3.2) then reads

(3.2)* $$\forall v_h \in V_h\colon\ \|v_h\|_{0,\infty,\Omega} \leqq C h^{-1/2}\,\|v_h\|_{1,2,\Omega}.$$

With some modifications in the proof of Lemmas 5.1–5.3, Theorem (1) and estimate (4.27) remain valid. Further it holds instead of (4.29)

(4.29)$^+$ $$\|u_h\|_{0,\infty,\Omega} \leqq M_h := 2CM h^{-1/2}.$$

Consequently the corresponding estimate of $|||u-u_h|||_{\varepsilon,\delta}$ in Theorem (2) holds only with loss of some powers of h depending on the growth condition *(H.1)*.

Acknowledgement. The author would like to thank the referees for many helpful comments.

References

1. Axelsson, O. (1984): On the numerical solution of convection dominated convection-diffusion problems. Teubner-Texte zur Mathematik, Bd. 63. Teubner, Leipzig
2. Axelsson, O., Eijkhout, V., Polman, B., Vassilevski, P. (1987): Iterative solution of singular perturbation second order boundary value problems by use of incomplete block-matrix factorization methods. Cath. Univ. Nijmegen, Report 8709
3. Bagaev, B.M. (1986): Solution of a parabolic equation with a small parameter multiplying the time derivative [Russ.]. Čisl. Met. Mech. Splošn. Sredy **17**, 13–26
4. Boccardo, L., Murat, L., Puel, J.P. (1988): Quelques proprietes des operateurs elliptiques quasi lineaires, C.R.A.S., t. **307**, Ser. I, 749–752
5. Ciarlet, P. (1978): The Finite Element Method for Elliptic Problems. North Holland, Amsterdam
6. Ciarlet, P., Schultz, M.H., Varga, R.S. (1969): Numerical methods of higher order accuracy for nonlinear boundary value problems. V. Monotone operator theory. Numer. Math. **13**, 51–77
7. Dobrowolski, M., Rannacher, R. (1980): Finite element methods for nonlinear elliptic systems of second order. Math. Nachr. **94**, 155–172
8. Feistauer, M., Sobotikova, V. (1990): Finite element approximation of nonlinear elliptic problems with discontinuous coefficients, Math. Modell. Numer. Anal. **24**, 457–500
9. Feistauer, M., Ženišek, A. (1987): Finite element solution of nonlinear elliptic problems. Numer. Math. **50**, 451–475
10. Feistauer, M., Ženišek, A. (1988): Compactness method in the finite element theory of nonlinear elliptic problems. Numer. Math. **52**, 147–163

11. Franca, L.P., Frey, S.L., Hughes, T.J.R. (1991): Stabilized finite element methods. I. Application to the advective-diffusive model. Comp. Meth. Appl. Mech. Engrg. (to appear)
12. Girault, V., Raviart, P.A. (1982): An analysis of upwind schemes for the Navier-Stokes equations. SIAM J. Numer. Anal. **19**, 312–333
13. Girault, V., Raviart, P.A. (1986): Finite Element Methods for Navier-Stokes Equations. Theory and Algorithms, Springer, Berlin Heidelberg New York
14. Howes, F.A. (1981): Perturbed boundary value problems whose reduced solutions are nonsmooth. Indiana Univ. Math. J. **90**, 267–280
15. Hughes, T.J.R. (1987): Recent progress in the development and understanding of SUPG methods with special reference to the compressible Euler and Navier-Stokes equations. Int. J. Numer. Meth. Fluids **7**, 1261–1275
16. Hughes, T.J.R., Brooks, A. (1979): A multidimensional upwind scheme with no crosswind diffusion. In: T.J.R. Hughes, ed., Finite Element Methods for Convection Dominated Flows, AMD, Vol. 34. ASME, New York, pp. 19–35
17. Hughes, T.J.R., Franca, L.P., Hulbert, G.M. (1989): A new finite element formulation for computational fluid dynamics, VIII. The Galerkin/Least-squares method for advective-diffusive equations. Comput. Meth. Appl. Mech. Engrg. **73**, 173–189
18. Johnson, C., Schatz, A.H., Wahlbin, L. (1987): Crosswind smear and pointwise errors in streamline diffusion finite element methods. Math. Comput. **49**, 25–38
19. Ladyshenskaya, O.A., Uralceva, N.N. (1973): Lineare und quasilineare Gleichungen elliptischen Typs [Russ.]. Nauka, Moskau
20. Lube, G. (1988): Uniform in ε discretization error estimates for convection dominated convection-diffusion problems. Math. Modell. Numer. Anal. **22**, 477–498
21. Lube, G. (1992): An asymptotically fitted finite element method for convection dominated convection-reaction-diffusion problems. ZAMM (to appear)
22. Lube, G., Weiß, D. (1991): Streamline diffusion finite element method for quasilinear elliptic problems. ZAMM **71**, T671–T674
23. Lube, G., Friedrich, L. (1989): Finite element methods of streamline diffusion type for singularly perturbed quasilinear elliptic boundary value problems. Wiss. Z. TU Magdeburg **33**, 21–30
24. Nävert, U. (1982): A finite element method for convection-diffusion problems. Ph.D. Thesis, Dept. of Comp. Science. Chalmers Univ. Göteborg
25. Niijima, K. (1990): Pointwise error estimates for a streamline diffusion finite element scheme. Numer. Math. **56**, 707–719
26. Rappaz, J. (1984): Approximation of a nondifferentiable nonlinear problem related to MHD equilibria. Numer. Math. **45**, 117–133
27. Schieweck, F. (1986): Eine exponentiell angepaßte Finite-Element-Methode für singulär gestörte elliptische Randwertaufgaben. Diss. A, TH Magdeburg
28. Sechin, A.Yu. (1986): On the approximate solution of two-dimensional singularly perturbed convection-diffusion equations. Soviet J. Numer. Anal. Math. Modell. **1**
29. Szepessy, A. (1989): Convergence of the streamline diffusion finite element method for conservation laws. Ph.D. Thesis, Dept. of Math., Chalmers Univ. Göteborg
30. Szepessy, A., Hansbo, P. (1990): A velocity-pressure streamline diffusion finite element method for the incompressible Navier-Stokes equations. Comput. Meth. Appl. Mech. Engrg. **84**, 175–192
31. Tezduyar, T.E., Park, Y.J. (1986): Discontinuity-capturing finite element formulation for nonlinear convection-diffusion-reaction systems. Comput. Meth. Appl. Mech. Engrg. **59**, 307–325
32. Tobiska, L. (1984): Die asymptotische Lösung singulär gestörter elliptischer Randwertaufgaben. Diss. B, TH Magdeburg