SPECTRAL APPROXIMATION OF A NONLINEAR ELASTIC LIMITING STRAIN MODEL

ENDRE SÜLI

ABSTRACT. We construct a numerical algorithm for the approximate solution of a nonlinear elastic limiting strain model based on the Fourier spectral method. The existence and uniqueness of the numerical solution is proved. Assuming that the weak solution to the boundary-value problem possesses suitable Sobolev regularity, the sequence of numerical solutions is shown to converge to the weak solution of the problem at an optimal rate. The numerical method represents a finite-dimensional system of nonlinear equations. An iterative method is proposed for the approximate solution of this system equations, and is shown to converge, at a linear rate, to the unique solution of the numerical method.

KEYWORDS: Fourier spectral method, convergence, nonlinear elasticity, implicit constitutive theory

1. INTRODUCTION

During the past decade there has been considerable progress in developing implicit constitutive models for the description of nonlinear responses of materials (see, for example, [11], [12]). In the field of solid mechanics, one of the main achievements of implicit constitutive theory is in providing a theoretical background for nonlinear models involving the linearized strain. In particular, within the realm of implicit constitutive theory, it is possible to have models in which the linearized strain is in all circumstances a bounded function, even when the stress is very large. This subclass of implicit constitutive models, proposed by Rajagopal in [12], are referred to as limiting strain models, and have the potential to be useful in modelling stress concentration effects in instances when the gradient of the displacement is relatively small (e.g. in modeling brittle materials near crack tips or notches, or concentrated loads inside the body or on its boundary). Models with limiting finite strain are also found to be useful in describing the response of various soft tissues that exhibit the phenomenon of finite extensibility. For example, Rajagopal’s limiting-strain elastic models stemming from implicit constitutive theory seem to provide good description of Fung’s experimental data concerning the passive response of biological tissues, which indicate that the stress/strain response of the tissue is, to a good approximation, exponential (see, for example, [6]).

As has been indicated in reference [3] and the survey article [1], limiting strain models have been thus far studied in several situations. In the case of special deformations such as shearing, compressions, torsion, etc., Rajagopal himself, and Bustamante and Rajagopal aimed to assess whether the models exhibit the expected responses (cf. [4], [14], [13]). In the case of anti-plane strain (stress) problems, considered in domains with nonconvex cross-sections (including thus the domains with V-notches or cracks), the resulting scalar problem in two space dimensions has been analyzed by methods of asymptotic analysis in [15], by performing systematic computational tests in [9], and by analytical methods of modern theory of nonlinear partial differential equations in [2]; the last result establishes the existence of weak solutions in nonconvex domains for values of the model parameter \( r \) in the range \( r \in (0, 2) \), see equation (2) below, and in convex domains for the range \( r \in (0, \infty) \). A detailed computational study of the complete problem in planar domains was performed in Ortiz et al. [10]. The recent paper [3] has been the first one with focus on the mathematical analysis of general boundary-value problems (which include systems of \( \frac{1}{2}d(d + 3) \) time-independent nonlinear partial differential equations of first order), featuring in limiting strain models, in bounded subsets of \( \mathbb{R}^d \), \( d \geq 2 \); the existence of a weak solution was shown,
in the case of periodic boundary conditions, for all values of the model parameter $r$ in the range $(0, \frac{2}{3})$ and the existence of a renormalized solution was established for all values of $r \in (0, \infty)$. The subsequent paper [1] surveys the physical background and the mathematical analysis of boundary-value problems associated with models with limiting small strain, and presents the first analytical result concerning the existence of weak solutions in general three-dimensional domains.

The analysis of numerical algorithms for limiting strain models is currently lacking. The present paper is a first step in the direction of rigorous analysis of a numerical method for a limiting strain model. The numerical algorithm considered here is posed in the context of the paper [3], i.e., in an axiparallel parallelopipedal domain subject to periodic boundary conditions, as this is the only setting involving the complete nonlinear system of equations in the model for which existence of a solution of any kind has been shown for the complete range $r \in (0, \infty)$ of the model parameter $r$.

# 2. Formulation of the problem and summary of the main results

As has been explained above, we shall consider a domain of a special form: namely an axiparallel parallelepiped, with spatially periodic boundary conditions in the various co-ordinate directions. This essential simplification helps us to introduce not only the concept of weak solution to the problem under consideration, but also the concept of a renormalized solution. The spatially periodic setting also helps us to construct the solution via a specific numerical method, namely the Fourier spectral method. The proof of existence of weak and renormalized solutions to the model in this geometry presented in [3] is therefore, at the same time, a proof of the convergence of the sequence of numerical approximations to the unknown analytical solution. This simplified setting allows us to provide a fairly complete picture regarding numerical analysis for a nontrivial example of a strain-limiting nonlinear elastic model.

The problem under consideration here is the following: suppose that $\Omega = (0, 2\pi)^d$, with $d \geq 2$, and $f$ is a given $d$-component vector-function (the load-vector), which is $2\pi$-periodic in each of the $d$ co-ordinate directions. Our objective is to construct a Fourier spectral approximation $(S_N, u_N)$ to $(S, u)$, where $S$ is the stress tensor and $u$ is the displacement, which belong to suitable function spaces consisting of $d \times d$ matrix functions and $d$-component vector functions, respectively, that are $2\pi$-periodic in each co-ordinate direction, such that

$$-\text{div } S = f$$

and

$$D(u) = F(S).$$

Here $F \in C^1(\mathbb{R}^{d \times d}; \mathbb{R}^{d \times d})$ is defined by

$$F(S) = \frac{S}{(1 + |S|^r)^{\frac{1}{r}}}, \quad S \in \mathbb{R}^{d \times d},$$

where $r > 0$, and $|\cdot|$ denotes the Frobenius norm on $\mathbb{R}^{d \times d}$, defined by $|X|^2 := X : X = \text{tr}(X^T X)$. It is a straightforward matter to show that the function $F$ has the following properties:

(P1) $F(0) = 0$ and $|F(A)| \leq 1$ for all $A \in \mathbb{R}^{d \times d}$;

(P2) There exist two constants $c_a = c_a(r) > 0$ and $c_b \geq 1$ such that the following inequalities hold:

(P2a)

$$(F(A) - F(B)) : (A - B) \geq c_a \frac{|A - B|^2}{(1 + |A| + |B|)^{r+1}} \quad \forall A, B \in \mathbb{R}^{d \times d},$$

and

$$(F(A) : A \geq c_a \frac{|A|^2}{1 + |A|} \quad \forall A \in \mathbb{R}^{d \times d};$$

(P2b)

$$|F(A) - F(B)| \leq c_b |A - B| \quad \forall A, B \in \mathbb{R}^{d \times d}.$$
Lemma 1. For any \( y \geq 0 \) and \( r > 0 \), we have that
\[
\min(1, 2^{-1+\frac{1}{r}}) (1 + y) \leq (1 + y^r)^\frac{1}{r} \leq \max(1, 2^{-1+\frac{1}{r}}) (1 + y).
\]

Lemma 2. Let \( r > 0 \) and consider the mapping
\[X \in \mathbb{R}^{d \times d} \mapsto F(X) := X(1 + |X|^r)^{-\frac{1}{r}} \in \mathbb{R}^{d \times d}.
\]
Then, for each \( A, B \in \mathbb{R}^{d \times d} \), we have that
\[|F(A) - F(B)| \leq 2|A - B|,
\]
and
\[(F(A) - F(B)) : (A - B) \geq \min(1, 2^{1-r^{-1}}) |A - B|^2 (1 + |A| + |B|)^{-r^{-1}}.
\]

Thanks to Lemma 2, (P2b) holds with \( c_b = 2 \) and the first inequality in (P2a) holds with \( c_a = \min(1, 2^{1-r^{-1}}) \); thanks to Lemma 1, the second inequality in (P2a) holds with \( c_b = \min(1, 2^{1-r^{-1}}) \).

The next lemma collects some elementary but helpful results concerning the function \( F \) and related functions that will arise in our analysis.

Lemma 3. The following statements hold:
(a) Suppose that \( \alpha > 0 \). The function \( t \in [0, \infty) \mapsto (1 + t)^{-\alpha} \in (0, 1] \) is Lipschitz continuous, with Lipschitz constant \( \alpha \).
(b) Suppose that \( \mu \in (0, 1] \); then, the function \( x \in \mathbb{R}^{d \times d} \mapsto |x|^\mu \in [0, \infty) \) is Hölder-continuous; in particular,
\[|x|^\mu - |y|^\mu | \leq \frac{1}{\mu} |x| - |y| |x|^\mu \leq \frac{1}{\mu} |x - y|^\mu \quad \forall x, y \in \mathbb{R}^{d \times d}.
\]
(c) Suppose that \( \mu \in (1, \infty) \) and let \( B(0, R) \) be the closed ball in \( \mathbb{R}^{d \times d} \) with radius \( R > 0 \), centred at the origin; then, the function \( x \in B(0, R) \mapsto |x|^\mu \in [0, \infty) \) is Lipschitz-continuous; in particular,
\[|x|^\mu - |y|^\mu | \leq \mu R^{\alpha-1} |x| - |y| \leq \mu R^{\alpha-1} |x - y| \quad \forall x, y \in B(0, R).
\]
(d) The composition of a \( (0, 1] \)-valued Lipschitz-continuous function defined on \([0, \infty)\) and a \((0, \infty)\)-valued Hölder continuous function defined on \( B(0, R) \), with Hölder exponent \( \min(1, r) \), is a \((0, 1]\)-valued Hölder-continuous function defined on \( B(0, R) \), with exponent \( \min(1, r) \).
In particular, for any \( \alpha > 0 \) and \( r > 0 \), the function \( x \in B(0, R) \mapsto (1 + |x|^r)^{-\alpha} \in (0, 1] \) is Hölder continuous, with exponent \( \min(1, r) \).
(e) Suppose that \( p > d^2 \); then, \( W^{1,p}(B(0, R)) \subseteq C^{0,\alpha}(B(0, R)) \) with \( \alpha = 1 - \frac{d^2}{p} \). In particular, for any \( \varepsilon \in (0, 1) \), the function
\[x \in B(0, R) \mapsto \frac{x}{|x|} \in B(0, R^{1-\varepsilon})
\]
belongs to \( W^{1,p}(B(0, R)) \) for \( p \in \left[1, \frac{d^2}{\varepsilon}\right) \), and hence to \( C^{0,\delta}(B(0, R)) \) for \( \delta \in (0, 1 - \varepsilon) \).

The paper is structured as follows. In section 3 we formulate the numerical approximation of the problem and recall from [3] various results concerning the existence and uniqueness of weak solutions for the range \( r \in (0, \frac{2}{3}) \) and the existence of a renormalized solution for the range \( r \in (0, \infty) \). The existence proofs are based on various weak compactness arguments and are omitted here as they do not directly relate to the topic of the present paper. For the sake of completeness of our discussion of the numerical method here, we have however included the proof, from [3], of the existence and uniqueness of a solution to the numerical approximation of the boundary-value problem under consideration. In section 4 we assume that the pair \((S, D(u))\) has additional regularity beyond that of a weak solution, i.e., that it belongs to the Sobolev space \([H^s(\Omega)]^{d \times d} \times [H^s(\Omega)]^{d \times d}, \) with \( s > \frac{d}{2} \), and use a fixed point argument to prove that the numerical method exhibits optimal order convergence in the \( L^2 \) norm. The numerical method represents a finite-dimensional system of nonlinear equations. In section 5 an iterative method is proposed for the approximate solution of this system equations, and is shown to converge to the unique solution.
3. Definition of the Approximation: Existence and Uniqueness of Solutions

Consider the domain $\Omega := (0, 2\pi)^d$ in $\mathbb{R}^d$, $d \geq 2$. All function spaces consisting of real-valued $2\pi$-periodic functions (by which we mean $2\pi$-periodic in each of the $d$ coordinate directions) will be labelled with the subscript $\#$: subspaces of these, consisting of $2\pi$-periodic functions whose integral over $\Omega$ is equal to 0, will be labelled with the subscript $\ast$; in order to avoid notational clutter we shall not use the symbols $\#$ and $\ast$ in the various norm signs. It will be clear from the argument of the norm which of the symbols $\#$ or $\ast$ is intended. For example, $L^p_\#(\Omega)$ will denote the Lebesgue space of all real-valued $2\pi$-periodic functions $v$ such that $|v|^p$ is integrable of $\Omega$, equipped with the norm $\|v\|_{L^p_\#(\Omega)}$. It is understood that the usual modification is made when $p = \infty$. Spaces of $d$-component vector functions, where each component belongs to a certain function space $X$, will be denoted by $[X]^d$, while spaces of $d \times d$ component matrix functions each of whose components is an element of $X$ will be signified by $[X]^{d \times d}$. Letting $C^\infty_\#(\Omega)$ denote the linear space consisting of the restriction to $\Omega$ of all $2\pi$-periodic $C^\infty$ functions defined on $\mathbb{R}^d$, we note that $C^\infty_\#(\Omega)$ is dense in $L^p_\#(\Omega)$ for all $p \in [1, \infty)$; analogously, $C^\infty_\ast(\Omega)$ is dense in $L^p_\ast(\Omega)$ for $1 \leq p < \infty$. The Sobolev space $W^{1,p}_\#(\Omega)$, $1 \leq p < \infty$, will be defined as the closure of $C^\infty_\#(\Omega)$ in the Sobolev norm $\| \cdot \|_{W^{1,p}_\#(\Omega)}$, where

$$ \|v\|_{W^{1,p}_\#(\Omega)} := \left( \|v\|_{L^p_\#(\Omega)}^p + \|\nabla v\|_{L^p_\#(\Omega)}^p \right)^{\frac{1}{p}}; $$

here, $\|v\|_{L^p_\#(\Omega)} := \|\nabla v\|_{L^p_\#(\Omega)}$, where $|\nabla v|$ denotes the Euclidean norm of $\nabla v$. Analogously, $W^{1,p}_\ast(\Omega)$, $1 \leq p < \infty$, will be defined as the closure of $C^\infty_\ast(\Omega)$ in the Sobolev norm $\| \cdot \|_{W^{1,p}_\ast(\Omega)}$.

\[ \begin{align*} \text{We further define} \\
H_{\#}(\text{div}; \Omega) := \{ v \in [L^2_\#(\Omega)]^d : \text{such that div } v \in L^2_\#(\Omega) \}, \end{align*} \]

equipped with the norm

$$ \|v\|_{H(\text{div}; \Omega)} := \left( \|v\|_{L^2_\#(\Omega)}^2 + \|\text{div } v\|_{L^2_\#(\Omega)}^2 \right)^{\frac{1}{2}}. $$

For $s > 0$, the (potentially, fractional-order, ) Hilbertian Sobolev spaces of periodic functions $H^s_\#(\Omega)$ and $H^s_\ast(\Omega)$, are defined analogously, through closure of $C^\infty_\#(\Omega)$ in the norm of $H^s_\ast(\Omega)$.

We shall require the following periodic version of Korn’s inequality [3].

**Lemma 4** (Korn’s inequality in $L^p$). Let $p \in (1, \infty)$, $d \geq 2$ and $\Omega := (0, 2\pi)^d$. There exists a positive constant $c_p$ such that the following inequalities hold:

$$ \|\nabla v\|_{L^p_\#(\Omega)} \leq c_p \left( \|D(v)\|_{L^p_\#(\Omega)} + \|\text{div } v\|_{L^p_\#(\Omega)} \right) \quad \forall v \in [W^{1,p}_\#(\Omega)]^d, $$

and, hence, also, with a possibly different constant $c_p$,

$$ \|\nabla v\|_{L^p_\#(\Omega)} \leq c_p \|D(v)\|_{L^p_\#(\Omega)} \quad \forall v \in [W^{1,p}_\#(\Omega)]^d. $$

Let, further, $D^{\text{dev}}(v) := D(v) - \frac{1}{d}(\text{div } v)I$ denote the deviatoric part of $D(v)$, where $I$ is the identity matrix in $\mathbb{R}^{d \times d}$; then, there exists a positive constant $c_p$ such that

$$ \|\nabla v\|_{L^p_\#(\Omega)} \leq c_p \|D^{\text{dev}}(v)\|_{L^p_\#(\Omega)} \quad \forall v \in [W^{1,p}_\#(\Omega)]^d. $$

Besides being dependent on $p$, the constant $c_p$ also depends on $d$, but we do not explicitly indicate that. In each case, the left-hand side of the inequality can be further bounded below by $C_p \|v\|_{W^{1,p}_\#(\Omega)}$, where $C_p$ is another positive constant dependent on $p$ and $d$, but independent of $v$. 

of the numerical method. We conclude, in section 6, with a summary of the main results of the paper and indications of some relevant open problems.
3.1. Construction of the numerical method. Let
\[ \Sigma_N \subset H_{\# \text{sym}}(\div; \Omega) := \{ S \in [L^2_{\#}(\Omega)]^{d \times d} : S = S^T, \; \div S \in [L^2_{\#}(\Omega)]^d \}, \]
equipped with norm
\[ \| S \|_{H(\div; \Omega)} := \left( \| S \|^2_{L^2(\Omega)} + \| \div S \|^2_{L^2(\Omega)} \right)^{\frac{1}{2}}, \]
and
\[ V_N \subset [W^{1,2}_{\#}(\Omega)]^d := \left\{ v \in [W^{1,2}_{\#}(\Omega)]^d : \int_\Omega v(x) \, dx = 0 \right\} \]
be a pair of finite-dimensional spaces consisting of, respectively, \( \mathbb{R}^{d \times d} \)-valued and \( \mathbb{R}^d \)-valued functions, whose components are 2\pi-periodic real-valued trigonometric polynomials of degree \( N \), \( N \geq 1 \), in each of the \( d \)-coordinate directions. The pair of spaces \((\Sigma_N, V_N)\) satisfies the following inf-sup condition: let \( b(v,T) := -(v, \div T) \); then, there exists a positive constant \( c_{\text{inf-sup}} \), independent of \( N \), such that
\[ \inf_{v_N \in V_N \setminus \{0\}} \sup_{T_N \in \Sigma_N \setminus \{0\}} \frac{b(v_N, T_N)}{\|v_N\|_{L^2(\Omega)} \|T_N\|_{H(\div; \Omega)}} \geq c_{\text{inf-sup}}. \]
For a short proof of (3) we refer to the Appendix in [3], where it is shown that \( c_{\text{inf-sup}} \geq \frac{1}{2} \).

Suppose that \( f \in [L^1_\#(\Omega)]^d \); in order to avoid trivialities, it will be assumed throughout that \( f \neq 0 \) (and therefore \( S \neq 0 \)). We consider the following discrete problem: find \((S_N, u_N) \in \Sigma_N \times V_N\) such that
\[ -(\div S_N, v_N) = (f, v_N) \quad \forall v_N \in V_N, \]
\[ D_N := F(S_N), \]
\[ (D(u_N), T_N) = (D_N, T_N) \quad \forall T_N \in \Sigma_N. \]

We are now ready to embark on the proof of existence and uniqueness of a solution to the discrete problem (4)–(6).

3.2. Existence and uniqueness of solutions to the numerical method. Theorem 1 below, guaranteeing the existence and uniqueness of a solution to the discrete problem (4)–(6), was established in [3]; for the sake of completeness of our analysis of the discretization, and for the convenience of the reader, we include its proof here. It relies on the following corollary of Brouwer’s fixed point theorem (cf. Girault & Raviart [8], Corollary 1.1, p.279).

**Lemma 5.** Let \( \mathcal{H} \) be a finite-dimensional Hilbert space whose inner product is denoted by \((\cdot, \cdot)_{\mathcal{H}}\) and the corresponding norm by \( \| \cdot \|_{\mathcal{H}} \). Let \( \mathcal{F} \) be a continuous mapping from \( \mathcal{H} \) into \( \mathcal{H} \) with the following property: there exists a \( \mu > 0 \) such that \( (\mathcal{F}(v), v)_{\mathcal{H}} > 0 \) for all \( v \in \mathcal{H} \) with \( \|v\|_{\mathcal{H}} = \mu \). Then, there exists an element \( u \in \mathcal{H} \) such that \( \|u\|_{\mathcal{H}} \leq \mu \) and \( \mathcal{F}(u) = 0 \).

**Theorem 1.** Suppose that \( f \in [L^1_\#(\Omega)]^d \) and \( N \geq 1 \). Then, the discrete problem (4)–(6) has a unique solution \((S_N, u_N) \in \Sigma_N \times V_N\).

**Proof.** Assuming for the moment the existence of a solution \((S_N, u_N) \in \Sigma_N \times V_N\) to (4)–(6), we shall show that the solution must be unique. Suppose otherwise, that there exist \((S_N^1, u_N^1) \in \Sigma_N \times V_N\) that solve (4)–(6) for \( i = 1, 2 \). Hence,
\[ -(\div(S_N^1 - S_N^2), v_N) - (D(u_N^1 - u_N^2), T_N) + (F(S_N^1) - F(S_N^2), T_N) = 0 \]
for all \((T_N, v_N) \in \Sigma_N \times V_N\). We take \( T_N = S_N^1 - S_N^2 \) and \( v_N = u_N^1 - u_N^2 \), and note that, after partial integration in the first term,
\[ -(\div(S_N^1 - S_N^2), u_N^1 - u_N^2) - (D(u_N^1 - u_N^2), S_N^1 - S_N^2) \]
\[ = (S_N^1 - S_N^2, \div(u_N^1 - u_N^2)) - (D(u_N^1 - u_N^2), S_N^1 - S_N^2) \]
\[ = (S_N^1 - S_N^2, D(u_N^1 - u_N^2)) - (D(u_N^1 - u_N^2), S_N^1 - S_N^2) = 0. \]

Consequently,
\[ (F(S_N^1) - F(S_N^2), S_N^1 - S_N^2) = 0. \]
Property (P2a) then implies that \( S_N^L \equiv S_N^J \) on \( \Omega \), and hence \( D_N^L \equiv D_N^J \) on \( \Omega \), which yields that \( D(u_N^1 - u_N^2) \equiv 0 \) on \( \Omega \). By Korn’s inequality stated in Lemma 4, we then have that \( u_N^1 - u_N^2 \equiv 0 \) on \( \Omega \), thus completing the proof of uniqueness of the solution to discrete problem (4)–(6).

Next we prove the existence of a solution to (4)–(6). First we choose any \( \hat{S}_N \in \Sigma_N \) such that \(-(\text{div} \hat{S}_N, v_N) = (f, v_N)\) for all \( v_N \in V_N \), and let \( S_{N,0} := \hat{S}_N - \hat{S}_N \). The existence of such an \( S_N \) will be shown below; for the time being, we shall proceed by taking the existence of such an \( \hat{S}_N \) for granted. Clearly, \(-(\text{div} S_{N,0}, v_N) = 0\) for all \( v_N \in V_N \), which then motivates us to define

\[
\Sigma_{N,0} := \{ T_N \in \Sigma_N : -(\text{div} T_N, v_N) = 0 \text{ for all } v_N \in V_N \}.
\]

As \( 0 \in \Sigma_{N,0} \), the set \( \Sigma_{N,0} \) is nonempty. Problem (4)–(6) can be therefore restated in the following equivalent form: find \((S_{N,0}, u_N) \in \Sigma_{N,0} \times V_N\) such that

\[
(D(u_N), T_N) = \left( F(S_{N,0} + \hat{S}_N), T_N \right) \quad \forall T_N \in \Sigma_N.
\]

Now, for \( T_N \in \Sigma_{N,0}, (D(u_N), T_N) = (\nabla v_N, T_N) = -(v_N, \text{div} T_N) = -(\text{div} T_N, v_N) = 0 \) for all \( v_N \in V_N \). Hence, (7) indicates that we should seek \( S_{N,0} \in \Sigma_{N,0} \) such that

\[
(F(S_{N,0} + \hat{S}_N), T_N) = 0 \quad \forall T_N \in \Sigma_{N,0}.
\]

Let us consider the nonlinear operator \( \mathfrak{F} : \Sigma_{N,0} \to \Sigma_{N,0} \), defined on the finite-dimensional Hilbert space \( \Sigma_{N,0} \), equipped with the inner product and norm of \([L^2_{\#}(\Omega)]^{d \times d}\), by

\[
\mathfrak{F}(U_N) := P_N F(U_N + \hat{S}_N), \quad U_N \in \Sigma_{N,0},
\]

where \( P_N \) denotes the orthogonal projector in \([L^2_{\#}(\Omega)]^{d \times d}\) onto \( \Sigma_{N,0} \).

Thanks to property (P2b), we then have that

\[
\| \mathfrak{F}(U_N) - \mathfrak{F}(\hat{U}_N) \|_{L^2(\Omega)} \leq c_0 \| U_N^1 - U_N^2 \|_{L^2(\Omega)} \quad \forall U_N^1, U_N^2 \in \Sigma_{N,0},
\]

and therefore \( \mathfrak{F} : \Sigma_{N,0} \to \Sigma_{N,0} \) is (globally) Lipschitz continuous on \( \Sigma_{N,0} \).

Note further that by (P2a) and (P1),

\[
\mathfrak{F}(U_N, U_N) = \left( F(U_N + \hat{S}_N), U_N \right) = \left( F(U_N + \hat{S}_N), U_N + S_N \right) - \left( F(U_N + \hat{S}_N), S_N \right) \geq c_0 \int_{\Omega} \frac{|U_N + \hat{S}_N|^2}{1 + |U_N + \hat{S}_N|} \, dx - \| S_N \|_{L^1(\Omega)}
\]

\[
\geq \frac{1}{2} c_0 \int_{\Omega} \frac{|U_N|^2}{1 + |U_N + \hat{S}_N|} \, dx - \| S_N \|_{L^1(\Omega)}
\]

\[
\geq \frac{1}{2} c_0 \int_{\Omega} \frac{|U_N|^2}{1 + |U_N + \hat{S}_N|} \, dx - c_0 \int_{\Omega} \frac{|S_N|^2}{1 + |U_N + \hat{S}_N|} \, dx - \| S_N \|_{L^1(\Omega)}. \]

As \( |U_N + \hat{S}_N| \leq |U_N| + |\hat{S}_N| \leq \| U_N \|_{L^\infty(\Omega)} + \| \hat{S}_N \|_{L^\infty(\Omega)} \), it follows by the Nikol’skii inequality \( \| U_N \|_{L^\infty(\Omega)} \leq C_{\text{inv}} N^2 \| U_N \|_{L^2(\Omega)} \) that for any \( U_N \in \Sigma_{N,0} \) such that \( \| U_N \|_{L^2(\Omega)} = \mu > 0 \), we have that

\[
\mathfrak{F}(U_N, U_N) \geq \frac{c_0 \mu^2}{2(1 + C_{\text{inv}} N^2 \mu + \| \hat{S}_N \|_{L^\infty(\Omega)})} - \| \Omega \| \| \hat{S}_N \|_{L^\infty(\Omega)} - \| \Omega \| \| \hat{S}_N \|_{L^\infty(\Omega)}.
\]

For \( N \geq 1 \) fixed (and therefore \( \| \hat{S}_N \|_{L^\infty(\Omega)} \) also fixed), the expression on the right-hand side of the last displayed inequality is a continuous function of \( \mu \in (0, \infty) \), which converges to \( +\infty \) as \( \mu \to +\infty \); thus, there exists a \( \mu_0 = \mu_0(d, N, \| \hat{S}_N \|_{L^\infty(\Omega)}) \), such that \( \mathfrak{F}(U_N, U_N) > 0 \) for all \( U_N \in \Sigma_{N,0} \) satisfying \( \| U_N \|_{L^2(\Omega)} = \mu \), for \( \mu > \mu_0 \).

By taking \( H = \Sigma_{N,0} \), equipped with the inner product and norm of \([L^2_{\#}(\Omega)]^{d \times d}\), we deduce from Lemma 5 the existence of an \( S_{N,0} \in \Sigma_{N,0} \) that solves (8), and thus, recalling that \( S_N = S_{N,0} + \hat{S}_N \), we have also shown the existence of an \( S_N \in \Sigma_N \) such that \(-(\text{div} S_N, v_N) = (f, v_N)\) for all \( v_N \in V_N \).

Having shown the existence of \( S_N \in \Sigma_N \), we return to (7) in order to show the existence of a \( u_N \in V_N \) such that

\[
(D(u_N), T_N) = (F(S_N), T_N) \quad \forall T_N \in \Sigma_N.
\]
Equivalently, we wish to show the existence of a $u_N \in V_N$ such that
\begin{equation}
    b(u_N, T_N) = \ell(T_N) \quad \forall T_N \in \Sigma_N,
\end{equation}
where
\begin{equation}
    b(v_N, T_N) := -(v_N, \text{div} T_N) \quad \text{and} \quad \ell(T_N) := (F(S_N), T_N).
\end{equation}

We note that $\ell(T_N) = 0$ for all $T_N \in \Sigma_{N,0}$, i.e., $\ell \in (\Sigma_{N,0})^d$ (the annihilator of $\Sigma_{N,0}$).

The existence of a unique $u_N \in V_N$ satisfying (9) then follows, thanks to the inf-sup condition (3), from the fundamental theorem of the theory of mixed variational problems stated in Lemma 4.1(ii) on p.40 of Girault & Raviart [7].

At the very beginning of our proof of existence of solutions we postulated the existence of an $\hat{S}_N \in \Sigma_N$ such that $-(\text{div} \hat{S}_N, v_N) = (f, v_N)$ for all $v_N \in V_N$. Part (iii) of Lemma 4.1 on p.40 of Girault & Raviart [7] implies, again thanks to the inf-sup condition (3), the existence of an $\hat{S}_N \in \Sigma_N$ such that $b(v_N, \hat{S}_N) = (f, v_N)$ for all $v_N \in V_N$; i.e., $-(\text{div} \hat{S}_N, v_N) = (f, v_N)$ for all $v_N \in V_N$. Thus we have proved both the existence and the uniqueness of solutions to the discrete problem (4)–(6).

\begin{remark}
The statement in the final paragraph of the proof above, that $\hat{S}_N \in \Sigma_N$, can be refined: in fact, $\hat{S}_N \in \Sigma_{N,0}^\perp$, where $\Sigma_{N,0}^\perp$ is the orthogonal complement of $\Sigma_{N,0}$ in $\Sigma_N$ with respect to the $[L^2(\Omega)]^d \times (\Omega)$ inner product.

The regularity hypothesis, that $f \in [L^1(\Omega)]^d$, is only used in the final paragraph of the proof. We note in particular that in order to apply Part (iii) of Lemma 4.1 on p.40 of [7], it is not necessary to demand that $f \in [L^2(\Omega)]^d$. Indeed, the Nikol’skii inequality $\|v_N\|_{L^\infty(\Omega)} \leq C_{\text{inv}} N^{\frac{d}{2}} \|v_N\|_{L^2(\Omega)}$ for any $v_N \in V_N$, implies that
\begin{equation}
    |(f, v_N)| \leq C_{\text{inv}} N^{\frac{d}{2}} \|f\|_{L^1(\Omega)} \|v_N\|_{L^2(\Omega)},
\end{equation}
and hence $v_N \mapsto (f, v_N)$ is a bounded linear functional on (the Hilbert space) $V_N$, equipped with the $[L^2(\Omega)]^d$ norm, as is required in Part (iii) of Lemma 4.1 on p.40 of [7].
\end{remark}

3.3. Convergence of the sequence of numerical solutions. Next we will address the question of convergence of the sequence of approximate solutions generated by (4)–(6). To this end, we define the function space
\begin{equation}
    D_{1,\infty}^1(\Omega) := \left\{ w \in [L^1(\Omega)]^d : D(w) \in [L^\infty(\Omega)]^{d \times d}, \int_\Omega w(x) \, dx = 0 \right\}.
\end{equation}

Trivially, $V_N \subset D_{1,\infty}^1(\Omega)$ for each $N \geq 1$. As, by Hölder’s inequality, $\|D(w)\|_{L^p(\Omega)} < \infty$ for any $w \in D_{1,\infty}^1(\Omega)$ and any $p \in [1, \infty)$, Korn’s inequality (cf. Lemma 4) implies that the seminorm $w \in D_{1,\infty}^1(\Omega) \mapsto \|D(w)\|_{L^1(\Omega)}$ is in fact a norm on $D_{1,\infty}^1(\Omega)$. Furthermore (cf. [3]), $[C^\infty_{\text{per}}(\Omega)]^d$ is weak-* dense in $D_{1,\infty}^1(\Omega)$ against $[L^\infty(\Omega)]^{d \times d}$, in the sense that for each $v \in D_{1,\infty}^1(\Omega)$ there exists a sequence $\{v_n\}_{n \geq 1} \subset [C^\infty_{\text{per}}(\Omega)]^d$ such that
\begin{equation}
    \int_\Omega T(x) : D(v_n(x)) \, dx \to \int_\Omega T(x) : D(v(x)) \, dx \quad \forall T \in [L^1(\Omega)]^{d \times d}.
\end{equation}

We recall the following result from [3] concerning the convergence of the sequence of approximate solutions generated by (4)–(6) to a weak solution of the boundary-value problem.

\begin{theorem}
Suppose that $f \in [W^{1, t}(\Omega)]^d$ for some $t > 1$; then, there exists a unique pair $(S, u) \in [L^1(\Omega)]^{d \times d} \times D_{1,\infty}^1(\Omega)$, such that
\begin{equation}
    (S, D(v)) = (f, v) \quad \forall v \in D_{1,\infty}^1(\Omega),
\end{equation}
and
\begin{equation}
    D(u) = F(S) \quad \text{with} \quad \begin{cases} r \in (0, 1] & \text{if } d = 2, \\ r \in (0, \frac{2}{d}] & \text{if } d > 2. \end{cases}
\end{equation}

Furthermore, the sequence of (uniquely defined) solution pairs $(S_N, u_N) \in \Sigma_N \times V_N$, $N \geq 1$, generated by (4)–(6), converges to $(S, u)$ in the following sense:
Theorem 3. \( \leq c \) example, Theorem 1.1 in [5]): suppose that
\( (t, \text{omal bound in the } L^2 \) the unique weak solution to the problem (cf. Theorem 5.1 in [3]) for any
\( \text{Proof. We begin by rewriting (4)–(6) in the following form: find } (S, u) \) and \( c \) a positive constant
\( \text{Consider } \hat{\theta} \) it follows that (13), (14) can be rewritten in the following form:
\( \text{Thus, by letting } S_{N,0} := S_N - \hat{S}_N, \)
we deduce that
\( S_{N,0} \in \Sigma_{N,0} := \{ T_N \in \Sigma_N : (\text{div } T_N, v_N) = 0 \quad \forall v_N \in V_N \}. \)
It follows that (13), (14) can be rewritten in the following form:
\( (\text{div } S_{N,0}, v_N) = 0 \quad \forall v_N \in V_N, \)
\( (F(S_{N,0} + \hat{S}_N), T_N) + (u_N, \text{div } T_N) = 0 \quad \forall T_N \in \Sigma_N. \)
Hence, in particular,
\( (F(S_{N,0} + \hat{S}_N), T_N) = 0 \quad \forall T_N \in \Sigma_{N,0}, \)

4. Error analysis of the numerical method

The proof of the next theorem will rely on the following classical approximation result (cf., for example, Theorem 1.1 in [5]): suppose that \( T \in [H^s_{\#}(\Omega)]^{d \times d} \); then, there exists a positive constant \( c_1 = c_1(s, d) \), independent of \( N \), such that

\[
\| T - P_NT \|_{H^s(\Omega)} \leq c_1 N^{s' - s} \| T \|_{H^s(\Omega)} \quad \forall N \geq 1,
\]
where \( 0 \leq s' \leq s \).

\*Theorem 3. Suppose that \((S, D(u)) \in [H^s_{\#}(\Omega)]^{d \times d} \times [H^s_{\#}(\Omega)]^{d \times d}, \) where \( s > \frac{d}{2} \). Then, there exists a positive constant \( c_0 \), independent of \( N \), and a positive integer \( N_* \) such that

\[
\| S - S_N \|_{L^2(\Omega)} \leq (c_1 + c_4) N^{-s} \| S \|_{H^s(\Omega)} + \| D(u) \|_{H^s(\Omega)} \quad \forall N \geq N_*,
\]
and

\[
\| D(u) - D(u_N) \|_{L^2(\Omega)} \leq c_5(c_1 + c_4) N^{-s} \| S \|_{H^s(\Omega)} + \| D(u) \|_{H^s(\Omega)} \quad \forall N \geq N_*,
\]

\*Proof. We begin by rewriting (4)–(6) in the following form: find \((S_N, u_N) \in \Sigma_N \times V_N \) such that

\[
-(\text{div } S_N, v_N) = (f, v_N) \quad \forall v_N \in V_N,
\]

\[
(F(S_N), T_N) - (D(u_N), T_N) = 0 \quad \forall T_N \in \Sigma_N.
\]
Consider \( \hat{S}_N := P_NS \), the orthogonal projection in \([L^2_{\#}(\Omega)]^{d \times d} \) of \( S \) onto \( \Sigma_N \). Clearly,

\[
-(\text{div } \hat{S}_N, v_N) = -(\text{div } P_NS, v_N) = (P_NS, \nabla v_N) = (P_NS, D(v_N)) = (S, D(v_N)) = (S, \nabla v_N) = -(\text{div } S, v_N) = (f, v_N) \quad \forall v_N \in V_N.
\]

Thus, by letting
\( S_{N,0} := S_N - \hat{S}_N, \)
we deduce that
\( S_{N,0} \in \Sigma_{N,0} := \{ T_N \in \Sigma_N : (\text{div } T_N, v_N) = 0 \quad \forall v_N \in V_N \}. \)
and therefore
\[
(F(S_{N,0} + \hat S_N) - F(\hat S_N), T_N) = -(F(\hat S_N), T_N)
\]
\[
= (F(S) - F(\hat S_N), T_N) = (F(S), T_N)
\]
\[
= (F(S) - F(\hat S_N), T_N) - (D(u), T_N)
\]
\[
= (F(S) - F(\hat S_N), T_N) - (D(u) - P_N D(u), T_N) - (P_N D(u), T_N)
\]
\[
= (F(S) - F(\hat S_N), T_N) - (D(u) - P_N D(u), T_N) - (D(u), T_N)
\]
\[
= (F(S) - F(\hat S_N), T_N) - (D(u) - P_N D(u), T_N) + (u, \text{div} T_N)
\]
\[
= (F(S) - F(\hat S_N), T_N) - (D(u) - P_N D(u), T_N) \quad \forall T_N \in \Sigma_{N,0}.
\]

Now, for \( S \) and \( u \) fixed, consider the linear functional \( \ell : \Sigma_N \to \mathbb{R} \), defined by
\[
\ell(T_N) := (F(S) - F(\hat S_N), T_N) - (D(u) - P_N D(u), T_N), \quad T_N \in \Sigma_N.
\]

We then deduce from (18) that
\[
\ell(T_N) := (F(S_{N,0} + \hat S_N) - F(\hat S_N), T_N) = \ell(T_N) \quad \forall T_N \in \Sigma_{N,0}.
\]

Thanks to (P2b) and (10), we have that
\[
|\ell(T_N)| \leq \left( c_0 \|S - \hat S_N\|_{L^2(\Omega)} + \|D(u) - P_N D(u)\|_{L^2(\Omega)} \right) \|T_N\|_{L^2(\Omega)}
\]
\[
\leq \left( c_0 c_1 N^{-\frac{s}{2}} \|S\|_{H^s(\Omega)} + c_1 N^{-s} \|D(u)\|_{H^{-s}(\Omega)} \right) \|T_N\|_{L^2(\Omega)}
\]
\[
\leq c_0 c_1 N^{-\frac{s}{2}} \left( \|S\|_{H^s(\Omega)} + \|D(u)\|_{H^{-s}(\Omega)} \right) \|T_N\|_{L^2(\Omega)} \quad \forall T_N \in \Sigma_N.
\]

Our objective is to prove that there exist a \( c_* > 0 \), independent of \( N \), and \( N_* \in \mathbb{N} \), such that for each \( N \geq N_* \) there exists a unique \( S_{N,0} \in \Sigma_{N,0} \) such that (19) holds and
\[
\|S_{N,0}\|_{L^2(\Omega)} \leq c_* N^{-s} \left( \|S\|_{H^s(\Omega)} + \|D(u)\|_{H^{-s}(\Omega)} \right).
\]

We shall use a fixed point theorem to this end. In order to define the fixed point mapping, we begin by noting that, by Lemma 3.2 in [3],
\[
(F(A) - F(B)) : C = \int_0^1 G(\theta A + (1 - \theta)B; A - B, C) \, d\theta,
\]
where, for \( \alpha, \beta, \gamma \in \mathbb{R}^{d \times d} \),
\[
G(\gamma; \alpha, \beta) := \frac{\alpha : \beta}{(1 + |\gamma|)^\frac{s}{2}} - (\alpha : \gamma)(\beta : \gamma) \frac{|\gamma|^{r - 2}}{(1 + |\gamma|)^{1 + \frac{s}{2}}}
\]

Note that
\[
|G(\gamma; \alpha, \beta)| \leq \frac{2 |\alpha| |\beta|}{(1 + |\gamma|)^\frac{s}{2}} \quad \forall \alpha, \beta, \gamma \in \mathbb{R}^{d \times d},
\]
\[
G(\gamma; \alpha, \alpha) \geq \frac{|\alpha|^2}{(1 + |\gamma|)^{1 + \frac{s}{2}}} \quad \forall \alpha, \gamma \in \mathbb{R}^{d \times d}.
\]

We define the set
\[
\mathcal{B}_{N,0} := \{ T_N \in \Sigma_{N,0} : \|T_N\|_{L^2(\Omega)} \leq c_* N^{-s} \left( \|S\|_{H^s(\Omega)} + \|D(u)\|_{H^{-s}(\Omega)} \right) \}.
\]

As \( 0 \in \mathcal{B}_{N,0} \), the set \( \mathcal{B}_{N,0} \) is nonempty, regardless of the choice of \( c_* > 0 \); also, \( \mathcal{B}_{N,0} \) is a closed subset of the finite-dimensional linear space \( \Sigma_{N,0} \).

Let us rewrite (19) as follows: find \( S_{N,0} \in \Sigma_{N,0} \) such that
\[
\int_0^1 G(\theta (S_{N,0} + \hat S_N) + (1 - \theta)\hat S_N; S_{N,0}, T_N) \, d\theta \, dx = \ell(T_N) \quad \forall T_N \in \Sigma_{N,0}.
\]

Equivalently, we can write this as follows: find \( S_{N,0} \in \Sigma_{N,0} \) such that
\[
\int_0^1 G(\theta \hat S_N + \theta S_{N,0}; S_{N,0}, T_N) \, d\theta \, dx = \ell(T_N) \quad \forall T_N \in \Sigma_{N,0}.
\]
Motivated by this equivalent restatement of (19), we consider the following mapping: to each \( \varphi \in \mathcal{B}_{N,0} \) we assign \( S_{N,\varphi} \in \Sigma_{N,0} \) such that

\[
(24) \quad \int_{\Omega} \int_0^1 G(\tilde{S}_N + \theta \varphi; S_{N,\varphi}, T_N) \, d\theta \, dx = \ell(T_N) \quad \forall T_N \in \Sigma_{N,0}.
\]

It follows from (23) that, for \( \tilde{S}_N \in \Sigma_N \) and \( \varphi \in \mathcal{B}_{N,0} \) fixed, (24) has at most one solution \( S_{N,\varphi} \in \Sigma_{N,0} \). Since \( \Sigma_{N,0} \) is a finite-dimensional linear space and (24) is a linear problem, the uniqueness of the solution implies its existence. Thus we deduce that the mapping \( \varphi \in \mathcal{B}_{N,0} \mapsto S_{N,\varphi} \in \Sigma_{N,0} \) is correctly defined. Next we will show that there exists a constant \( c_* > 0 \), independent of \( N \), and \( N_* \in \mathbb{N} \), such that if \( \varphi \in \mathcal{B}_{N,0} \) with \( N \geq N_* \), then \( S_{N,\varphi} \in \mathcal{B}_{N,0} \), in fact.

Note that by (23), (24) and (20),

\[
\frac{\|S_{N,\varphi}\|^2_{L^2(\Omega)}}{(1 + (\|\tilde{S}_N\|_{L^\infty(\Omega)} + \|\varphi\|_{L^\infty(\Omega)})^r)^{1 + \frac{1}{r}} + 1} \leq \int_{\Omega} \int_0^1 \frac{|S_{N,\varphi}|^2}{(1 + (\|\tilde{S}_N + \theta \varphi\|_{L^\infty(\Omega)})^r)^{1 + \frac{1}{r}} + 1} \, d\theta \, dx
\]

\[
\leq \int_{\Omega} \int_0^1 G(\tilde{S}_N + \theta \varphi; S_{N,\varphi}, \varphi) \, d\theta \, dx = \ell(S_{N,\varphi}) \leq c_* N^{-r} (\|S\|_{H^r(\Omega)} + \|D(u)\|_{H^r(\Omega)}) \|S_{N,\varphi}\|_{L^2(\Omega)}.
\]

Thus we deduce that

\[
(25) \quad \|S_{N,\varphi}\|_{L^2(\Omega)} \leq c_* N^{-r} (\|S\|_{H^r(\Omega)} + \|D(u)\|_{H^r(\Omega)})(1 + (\|\tilde{S}_N\|_{L^\infty(\Omega)} + \|\varphi\|_{L^\infty(\Omega)})^r)^{1 + \frac{1}{r}}.
\]

In order to prove that \( S_{N,\varphi} \in \mathcal{B}_{N,0} \) for a suitable \( c_* > 0 \) and all \( N \geq N_* \), with a certain positive integer \( N_* \), our aim is to show that, for a suitable constant \( c_* > 0 \), independent of \( N \), and a suitable positive integer \( N_* \),

\[
(26) \quad c_* N^{-r} (\|S\|_{H^r(\Omega)} + \|D(u)\|_{H^r(\Omega)}) \leq c_* N^{-r} (\|S\|_{H^r(\Omega)} + \|D(u)\|_{H^r(\Omega)}) \quad \forall N \geq N_*.
\]

This is equivalent to showing that, for a suitable constant \( c_* > 0 \), independent of \( N \), and a suitable positive integer \( N_* \),

\[
(27) \quad c_* N^{-r} (\|S\|_{H^r(\Omega)} + \|\varphi\|_{L^\infty(\Omega)})^{1 + \frac{1}{r}} \leq c_* \quad \forall N \geq N_*.
\]

We shall derive a sufficient condition for (27) to hold by replacing \( \|\tilde{S}_N\|_{L^\infty(\Omega)} \) and \( \|\varphi\|_{L^\infty(\Omega)} \) in (27) by upper bounds on them.

First note that

\[
\|\tilde{S}_N\|_{L^\infty(\Omega)} = \|P_N S\|_{L^\infty(\Omega)} \leq \|S\|_{L^\infty(\Omega)} + \|S - P_N S\|_{L^\infty(\Omega)}.
\]

As, by hypothesis, \( s > \frac{d}{2} \), there exists an \( s' \in \left( \frac{d}{2}, s \right) \). By Sobolev embedding, and using the approximation property (10) of the projector \( P_N \), we have that

\[
\|S - P_N S\|_{L^\infty(\Omega)} \leq C(s', d) \|S - P_N S\|_{H^{s'}(\Omega)} \leq c_1 C(s', d) N^{s' - s} \|S\|_{H^r(\Omega)}.
\]

As \( s > s' \), there exists a positive integer \( N_* \) such that

\[
c_1 C(s', d) N^{s' - s} \|S\|_{H^r(\Omega)} \leq \|S\|_{L^\infty(\Omega)} \quad \forall N \geq N_*.
\]

For example, we can take

\[
N_* := \left( \frac{c_1 C(s', d) \|S\|_{H^r(\Omega)}}{\|S\|_{L^\infty(\Omega)}} \right)^{1 - s'}.r
\]

Hence,

\[
\|\tilde{S}_N\|_{L^\infty(\Omega)} \leq 2 \|S\|_{L^\infty(\Omega)} \quad \forall N \geq N_*.
\]

Since by the Nikol’skii inequality \( \|T_N\|_{L^\infty(\Omega)} \leq C_{\text{inv}} N^{\frac{d}{2}} \|T_N\|_{L^2(\Omega)} \) for any \( T_N \in \Sigma_{N,0} \), it follows that a sufficient condition for (27) to hold is that

\[
(28) \quad c_* N^{\frac{d}{2}} (\|S\|_{L^\infty(\Omega)} + C_{\text{inv}} N^{\frac{d}{2}} \|\varphi\|_{L^2(\Omega)})^{r + \frac{1}{r}} \leq c_* \quad \forall N \geq N_*.
\]
where \( N_s \geq N_{**} \) is a positive integer, to be chosen below.

We define
\[
c_s := c_{\epsilon 1} (1 + (2 \| S \|_{L^\infty(\Omega)} + C_{\mathrm{uv}} (\| S \|_{H^s(\Omega)} + \| D(u) \|_{H^s(\Omega)}))^\gamma)\frac{1}{1+s}.
\]

With this definition of \( c_s \), (28) becomes equivalent to the inequality
\[
N_s^\frac{1}{2} \| \varphi \|_{L^2(\Omega)} \leq \| S \|_{H^s(\Omega)} + \| D(u) \|_{H^s(\Omega)} \quad \forall N \geq N_s.
\]

As \( \varphi \in B_{N,0} \), a sufficient condition for (29) to hold is that
\[
c_s N_s^{-\frac{1}{2}} \leq 1 \quad \forall N \geq N_s.
\]

Since \( s > \frac{d}{2} \), there exists an \( N_s \geq N_{**} \), such that this inequality holds; for example, one can take
\[
N_s := \max \left( \left[ c_s \frac{2}{s+1} \right], N_{**} \right).
\]

With \( c_s \) and \( N_s \) thus defined, (30) holds; and, therefore, (29), (28), (27) all hold, and, since (27) is equivalent to (26), it follows that (26) also holds. Having shown the existence of \( c_s \) and \( N_s \) such that (26) holds, it follows from (25) that
\[
\| S_N,\varphi \|_{L^2(\Omega)} \leq c_s N_s^{-\frac{1}{2}} (\| S \|_{H^s(\Omega)} + \| D(u) \|_{H^s(\Omega)}) \quad \forall N \geq N_s.
\]

Hence, \( S_{N,\varphi} \in B_{N,0} \) for all \( N \geq N_s \). As the function \( \varphi \mapsto S_{N,\varphi} \) maps the bounded closed ball \( B_{N,0} \) contained in the finite-dimensional linear space \( \Sigma_{N,0} \) into itself, Brouwer’s fixed point theorem will imply the existence of a fixed point \( S_{N,\varphi} \in B_{N,0} \) for this mapping, once we have shown the continuity of this mapping.

To this end, we consider \( \varphi_1, \varphi_2 \in B_{N,0} \) and the associated \( S_{N,\varphi_1}, S_{N,\varphi_2} \in B_{N,0}, N \geq N_s \), defined, for \( i = 1, 2 \), by
\[
\int_0^\frac{1}{2} \int_0^1 G(\tilde{S}_N + \theta \varphi_i; S_{N,\varphi_i}, T_N) \, d\theta \, dx = \ell(T_N) \quad \forall T_N \in \Sigma_{N,0}.
\]

We thus have that
\[
\int_0^\frac{1}{2} \int_0^1 G(\tilde{S}_N + \theta \varphi_1; S_{N,\varphi_1} - S_{N,\varphi_2}, T_N) \, d\theta \, dx
\]
\[
= \int_0^\frac{1}{2} \int_0^1 G(\tilde{S}_N + \theta \varphi_2; S_{N,\varphi_2}, T_N) \, d\theta \, dx - \int_0^\frac{1}{2} \int_0^1 G(\tilde{S}_N + \theta \varphi_1; S_{N,\varphi_1}, T_N) \, d\theta \, dx.
\]

By taking \( T_N = S_{N,\varphi_1} - S_{N,\varphi_2} \) we deduce from Lemma 2 that
\[
\| S_{N,\varphi_1} - S_{N,\varphi_2} \|_{L^2(\Omega)}^2
\]
\[
\leq \int_0^\frac{1}{2} \int_0^1 \left| G(\tilde{S}_N + \theta \varphi_2; S_{N,\varphi_2}, S_{N,\varphi_1} - S_{N,\varphi_2}) - G(\tilde{S}_N + \theta \varphi_1; S_{N,\varphi_1}, S_{N,\varphi_1} - S_{N,\varphi_2}) \right| \, d\theta \, dx.
\]

For \( \alpha, \beta, \gamma \in \mathbb{R}^{d \times d} \), we choose \( \varepsilon \in \{ \max\{0, 1 - \frac{d}{2}\}, 1 \} \) and rewrite \( G(\gamma; \alpha, \beta) \) as follows:
\[
G(\gamma; \alpha, \beta) := \frac{\alpha : \beta}{(1 + |\gamma|^r)^\frac{1}{2}} = \left( \frac{\alpha : \gamma}{|\gamma|^{r}} \right) \left( \frac{\beta : \gamma}{|\gamma|^{r}} \right) \frac{|\gamma|^{r-2+2\varepsilon}}{(1 + |\gamma|^r)^{1+\frac{2}{s}}}.
\]

Note that with such an \( \varepsilon \), one has \( r - 2 + 2\varepsilon > 0 \). The functions
\[
\gamma \mapsto \frac{1}{(1 + |\gamma|^r)^\frac{1}{2}}, \quad \gamma \mapsto \frac{\gamma}{|\gamma|^r}, \quad \gamma \mapsto |\gamma|^{r-2+2\varepsilon}, \quad \gamma \mapsto \frac{1}{(1 + |\gamma|^r)^{1+\frac{2}{s}}}
\]
are Hölder-continuous on any bounded ball \( B(0,R) \) in \( \mathbb{R}^{d \times d} \) of radius \( R \); the Hölder exponents \( \delta_i \), \( i = 1, 2, 3, 4 \), of these four functions are, respectively,
\[
\delta_1 = \min(1, r), \quad \delta_2 < 1 - \varepsilon, \quad \delta_3 = \min(1, r - 2 + 2\varepsilon), \quad \delta_4 = \min(1, r).
\]

These statements follow from Lemma 3, parts (d); (c); (b) and (c); and (d), respectively.
Let $\delta_0 = \min(\delta_1, \delta_2, \delta_3, \delta_4)$; clearly, $\delta_0 \in (0, 1)$. Let $\delta \in (0, \delta_0]$. Hence,
\[
\int_\Omega \int_0^1 \left| G(\hat{S}_N + \theta \varphi_2; S_{N, \varphi_2}, s_{N, \varphi_1} - s_{N, \varphi_2}) - G(\hat{S}_N + \theta \varphi_1; S_{N, \varphi_2}, s_{N, \varphi_1} - s_{N, \varphi_2}) \right| \, d\theta \, dx
\leq C(r, \varepsilon, \|S_{N, \varphi_2}\|_{L^\infty(\Omega)}, \|\varphi_1\|_{L^\infty(\Omega)}, \|\varphi_2\|_{L^\infty(\Omega)}) \int_\Omega |\varphi_1 - \varphi_2| \delta \|S_{N, \varphi_1} - s_{N, \varphi_2}\| \, dx.
\]
Thus we deduce that
\[
\|s_{N, \varphi_1} - s_{N, \varphi_2}\|_{L^2(\Omega)} \leq C(r, \varepsilon, \|\hat{S}_N\|_{L^\infty(\Omega)}, \|S_{N, \varphi_2}\|_{L^\infty(\Omega)}, \|\varphi_1\|_{L^\infty(\Omega)}, \|\varphi_2\|_{L^\infty(\Omega)}) \|\varphi_1 - \varphi_2\|_{L^2(\Omega)}^2,
\]
for all $\varphi_1, \varphi_2 \in B_{N, 0}$. As $\delta \in (0, 1)$, it follows by Hölder’s inequality that
\[
\|s_{N, \varphi_1} - s_{N, \varphi_2}\|_{L^2(\Omega)} \leq C(r, \varepsilon, \|\hat{S}_N\|_{L^\infty(\Omega)}, \|S_{N, \varphi_2}\|_{L^\infty(\Omega)}, \|\varphi_1\|_{L^\infty(\Omega)}, \|\varphi_2\|_{L^\infty(\Omega)}) \|\varphi_1 - \varphi_2\|_{L^2(\Omega)}^2,
\]
for all $\varphi_1, \varphi_2 \in B_{N, 0}$. We note that, for $N \geq N_*$, we have that
\[
\|\hat{S}_N\|_{L^\infty(\Omega)} \leq 2\|S\|_{L^\infty(\Omega)},
\]
\[
\|S_{N, \varphi_2}\|_{L^\infty(\Omega)} \leq C_{\text{inv}} \varepsilon \frac{N^{\frac{d}{2} - s}}{(\|S\|_{H^s(\Omega)} + \|D(u)\|_{H^s(\Omega)})},
\]
\[
\|\varphi_i\|_{L^\infty(\Omega)} \leq C_{\text{inv}} \varepsilon \frac{N^{\frac{d}{2} - s}}{(\|S\|_{H^s(\Omega)} + \|D(u)\|_{H^s(\Omega)})}, \quad i = 1, 2.
\]
Hence, for $(S, D(u)) \in [H^s_0(\Omega)]^{d \times d} \times [H^s_0(\Omega)]^{d \times d}$ fixed, with $s > \frac{d}{2}$,
\[
\|S_{N, \varphi_1} - S_{N, \varphi_2}\|_{L^2(\Omega)} \leq C(r, \varepsilon) \|\varphi_1 - \varphi_2\|_{L^2(\Omega)}^2 \quad \forall \varphi_1, \varphi_2 \in B_{N, 0}, \quad N \geq N_*.
\]
This implies the (Hölder) continuity of the map $\varphi \in B_{N, 0} \mapsto S_{N, \varphi} \in B_{N, 0}$ for $N \geq N_*$. Hence, $\varphi \mapsto S_{N, \varphi}$ maps the bounded closed ball $B_{N, 0}$ contained in the finite-dimensional linear space $\Sigma_{N, 0}$ continuously into itself; Brouwer’s fixed point theorem therefore implies the existence of a fixed point $S_{N, *}$ in $B_{N, 0}$ for this mapping; i.e.,
\[
\int_\Omega \int_0^1 G(\hat{S}_N + \theta S_{N,*}; s_{N,*}, T_N) \, d\theta \, dx = \ell(T_N) \quad \forall T_N \in \Sigma_{N, 0}.
\]
Since the uniqueness of the fixed point is not guaranteed by Brouwer’s fixed point theorem, it is not clear at this stage whether $S_{N,*}$ is equal to $S_{N,0}$. In order to show that this is the case, we proceed as follows. First note that (32) is equivalent to
\[
(F(S_{N,*} + \hat{S}_N), T_N) = 0 \quad \forall T_N \in \Sigma_{N, 0}.
\]
Recall from (17) that, on the other hand,
\[
(F(S_{N,0} + \hat{S}_N), T_N) = 0 \quad \forall T_N \in \Sigma_{N, 0}.
\]
It follows from the last two equations, and setting $T_N = (S_{N,*} + \hat{S}_N) - (S_{N,0} + \hat{S}_N) = S_{N,*} - S_{N,0} \in \Sigma_{N, 0}$, that
\[
(F(S_{N,*} + \hat{S}_N) - F(S_{N,0} + \hat{S}_N), (S_{N,*} + \hat{S}_N) - (S_{N,0} + \hat{S}_N)) = 0.
\]
By Lemma 2, with $A = S_{N,*} + \hat{S}_N$, $B = S_{N,0} + \hat{S}_N$, this then implies that
\[
\int_\Omega \min \left(1, 2^{-\frac{1}{r}} \right) \frac{|(S_{N,*} - S_{N,0})|^2}{1 + |S_{N,*} + \hat{S}_N| + |S_{N,0} + \hat{S}_N|} \, dx \leq 0.
\]
Hence, $|S_{N,*} - S_{N,0}|^2 = 0$ a.e. on $\Omega$, whereby $S_{N,*} = S_{N,0}$ a.e. on $\Omega$. Since both $S_{N,*}$ and $S_{N,0}$ are trigonometric polynomials, it follows that $S_{N,*}(x) = S_{N,0}(x)$ for all $x \in \Omega$.
Thus we have finally shown that there exists a unique $S_{N,0} \in B_{N, 0}$, with
\[
S_{N,0} := S - \hat{S} = S_N - PN S,
\]
such that (17) holds. Now, by the triangle inequality and (10), and because \( S_{N,0} \in \mathfrak{B}_{N,0} \), we have that

\[
\| S - S_N \|_{L^2(\Omega)} \leq \| S - P_N S \|_{L^2(\Omega)} + \| S_{N,0} \|_{L^2(\Omega)}
\]

(33)

\[
\leq c_1 N^{-s} \| S \|_{H^s(\Omega)} + c_2 N^{-s} (\| S \|_{H^s(\Omega)} + \| D(u) \|_{H^1(\Omega)})
\]

(34)

Further, by (14), (P2b) and (33), we have that, for all \( N \geq N_* \).

\[
\| P_N D(u) - D(u_N) \|_{L^2(\Omega)} = \sup_{T_N \in \Sigma_N \setminus \{0\}} \frac{(P_N D(u) - D(u_N), T_N)}{\|T_N\|_{L^2(\Omega)}}
\]

= \sup_{T_N \in \Sigma_N \setminus \{0\}} \frac{(D(u) - D(u_N), T_N)}{\|T_N\|_{L^2(\Omega)}}

\[
= \sup_{T_N \in \Sigma_N \setminus \{0\}} \frac{(F(S) - F(S_N), T_N)}{\|T_N\|_{L^2(\Omega)}}
\]

\[
\leq \| F(S) - F(S_N) \|_{L^2(\Omega)} \leq c_6 \| S - S_N \|_{L^2(\Omega)}
\]

(35)

From (35), by the triangle inequality and noting that \( c_6 \geq 1 \), it follows that, for all \( N \geq N_* \),

\[
\| D(u) - D(u_N) \|_{L^2(\Omega)} \leq \| D(u) - P_N D(u) \|_{L^2(\Omega)} + \| P_N D(u) - D(u_N) \|_{L^2(\Omega)}
\]

\[
\leq c_6 (c_1 + c_2) N^{-s} \| S \|_{H^s(\Omega)} + (c_1 + c_2) N^{-s} \| D(u) \|_{H^s(\Omega)}
\]

\[
\leq c_6 (c_1 + c_2) N^{-s} (\| S \|_{H^s(\Omega)} + \| D(u) \|_{H^s(\Omega)})
\]

That completes the proof. \( \square \)

**Remark 2.** We note that by Korn's inequality (cf. Lemma 4),

\[
\| u - u_N \|_{H^s(\Omega)} \leq \text{Const.} N^{-s} (\| S \|_{H^s(\Omega)} + \| D(u) \|_{H^s(\Omega)})
\]

For each \( N \geq 1 \), the numerical method (4)–(6) is a finite-dimensional system of nonlinear equations. In the next section we propose an iterative method for the solution of the discrete problem (4)–(6) and we explore its convergence, with \( N \) kept fixed.

### 5. Iterative Solution of the Finite-dimensional Nonlinear System

We consider the following iterative method for the solution of (4)–(6): let \( S^k_N \) := 0; for \( k = 1, 2, \ldots \), we define \( (S^k_N, u^k_N) \in \Sigma_N \times V_N \) as the solution of the following problem

\[
-(\text{div} \, S^k_N, v_N) = (f, v_N) \quad \forall v_N \in V_N,
\]

(36)

\[
(S^k_N, T_N) - \lambda D(u^k_N), T_N) = (S^{k-1}_N, T_N) - \lambda (F(S^{k-1}_N), T_N) \quad \forall T_N \in \Sigma_N,
\]

(37)

where \( \lambda > 0 \) is a parameter, to be fixed below.

We begin by noting that this iteration is correctly defined, in the sense that, for each \( k \in \mathbb{N} \), there exists a unique pair \( (S^k_N, u^k_N) \in \Sigma_N \times V_N \) satisfying (36), (37). To this end, let \( S^0_{N,0} := S^k_N - S^{k-1}_N \), and note that

\[
(\text{div} \, S^0_{N,0}, v_N) = 0 \quad \forall v_N \in V_N,
\]

(38)

\[
(S^0_{N,0}, T_N) - \lambda (D(u^0_N), T_N) = -\lambda (F(S^{k-1}_N), T_N) \quad \forall T_N \in \Sigma_N.
\]

Hence, \( S^k_{N,0} \in \Sigma_{N,0} \), and therefore,

\[
(S^k_{N,0}, T_N) = -\lambda (F(S^{k-1}_N), T_N) \quad \forall T_N \in \Sigma_{N,0}.
\]

Consequently, \( S^k_{N,0} \) is uniquely defined as the orthogonal projection of \( -\lambda F(S^{k-1}_N) \) onto the finite-dimensional linear subspace \( \Sigma_{N,0} \) of \( \Sigma_N \), with respect to the inner product of \([L^2_#(\Omega)]^{d \times d}\), which
then uniquely defines $S_N^k = S_N^{k-1} + S_{N,0}^k \in \Sigma_N$. For $S_N^k$ thus fixed, we rewrite (37) as follows:
\[-(u_N^k, \text{div } T_N) = \frac{1}{\lambda} (s_N^k - s_N^{k-1}, T_N) + (F(s_N^{k-1}), T_N) \quad \forall T_N \in \Sigma_N.\]
By introducing the bilinear form $b(v,T) := -(v, \text{div } T)$ on $V_N \times \Sigma_N$ and the linear functional $\ell(T) := \frac{1}{2} (s_N^k - s_N^{k-1}, T) + (F(s_N^{k-1}), T)$ on $\Sigma_N$, the proof of existence of a unique solution $u_N^k$ to the problem $b(u_N^k, T_N) = \ell(T_N)$ for all $T_N \in \Sigma_N$ proceeds analogously as in the case of problem (9): the bilinear form $b(\cdot,\cdot)$ satisfies the inf-sup condition (3), and the linear functional $\ell \in (\Sigma_N,0)^0$ (the annihilator of $\Sigma_N$). The existence of a unique solution $u_N^k$ satisfying $b(u_N^k, T_N) = \ell(T_N)$ for all $T_N \in \Sigma_N$ therefore follows from the fundamental theorem of the theory of mixed variational problems, stated in Lemma 4.1(ii) on p.40 of Girault & Raviart [7].

Next, we will show that, for each fixed $N \geq 1$, $(S_N^k, u_N^k) \to (S_N, u_N)$ as $k \to \infty$.

**Theorem 4.** Let
\[c_a := \min(1, 2^{-\frac{1}{2}}), \quad c_b := 1 + (2 + C_{\text{inv}} N^{\frac{1}{2}} |\Omega|^\frac{1}{2}) \|S_N\|_{L^\infty(\Omega)}, \quad c_0 := \frac{c_a}{c_b},\]
and let $\lambda \in (0, \frac{1}{2} c_0)$. Then,
\[L^2 := 1 - 2c_0 \lambda + 4\lambda^2 \in (0, 1),\]
and for each $k \geq 1$,
\[\|S_N - S_N^k\|_{L^2(\Omega)}^2 + \lambda^2 \|D(u_N - u_N^k)\|_{L^2(\Omega)}^2 \leq L^2 k \|S_N\|_{L^\infty(\Omega)}^2 \quad \forall k \geq 1.\]

**Proof.** We subtract (36), (37) from (13), (14), respectively; hence,
\[(\text{div } (S_N - S_N^k), v_N) = 0 \quad \forall v_N \in V_N,\]
\[\begin{align*}
(S_N - S_N^k, T_N) &= (S_N - S_N^{k-1}, T_N) - \lambda (F(S_N) - F(S_N^{k-1}), T_N) \\
&\quad + \lambda (D(u_N - u_N^k), T_N) \quad \forall T_N \in \Sigma_N.
\end{align*}\]
Equation (40) implies that $S_N - S_N^k \in \Sigma_N$; thus, by taking $T_N = S_N - S_N^k$ in (41), we have that
\[\|S_N - S_N^k\|_{L^2(\Omega)}^2 = (S_N - S_N^{k-1}, S_N - S_N^k) - \lambda (F(S_N) - F(S_N^{k-1}), S_N - S_N^k).\]
Next, we take $T_N = S_N - S_N^{k-1}$ in (41); hence,
\[\|S_N - S_N^k\|_{L^2(\Omega)}^2 = (S_N - S_N^{k-1}, S_N - S_N^{k-1}) - \lambda (F(S_N) - F(S_N^{k-1}), S_N - S_N^{k-1}).\]
Finally, we take $T_N = P_N(F(S_N) - F(S_N^{k-1}))$ in (41); thus,
\[\begin{align*}
(S_N - S_N^k, F(S_N) - F(S_N^{k-1})) &= (S_N - S_N^{k-1}, P_N(F(S_N) - F(S_N^{k-1}))) \\
&\quad - \lambda (F(S_N) - F(S_N^{k-1}), P_N(F(S_N) - F(S_N^{k-1}))) \\
&\quad + \lambda (D(u_N - u_N^k), P_N(F(S_N) - F(S_N^{k-1}))) \\
&= (S_N - S_N^{k-1}, F(S_N) - F(S_N^{k-1})) \\
&\quad - \lambda (F(S_N) - F(S_N^{k-1}), P_N(F(S_N) - F(S_N^{k-1}))) \\
&\quad + \lambda (D(u_N - u_N^k), F(S_N) - F(S_N^{k-1})).
\end{align*}\]
Substitution of (43) and (44) into (42) yields
\[\begin{align*}
\|S_N - S_N^k\|_{L^2(\Omega)}^2 &= \|S_N - S_N^{k-1}\|_{L^2(\Omega)}^2 - \lambda (F(S_N) - F(S_N^{k-1}), S_N - S_N^{k-1}) \\
&\quad - \lambda (S_N - S_N^{k-1}, F(S_N) - F(S_N^{k-1})) \\
&\quad + \lambda^2 (F(S_N) - F(S_N^{k-1}), P_N(F(S_N) - F(S_N^{k-1}))) \\
&\quad - \lambda^2 (D(u_N - u_N^k), F(S_N) - F(S_N^{k-1})).
\end{align*}\]
We shall transform the final term in (45) by taking $T_N = D(u_N - u_N^k)$ in (41):

\[ \lambda \|D(u_N - u_N^k)\|_{L^2(\Omega)}^2 = (S_N - S_N^k, D(u_N - u_N^k)) - (S_N - S_N^{-1}, D(u_N - u_N^k)) \]

(46)

\[ + \lambda (F(S_N) - F(S_N^{-1}), D(u_N - u_N^k)). \]

As the first two terms on the right-hand side of (46) are both equal to 0 and $\lambda > 0$, it follows that

\[ (D(u_N - u_N^k), F(S_N) - F(S_N^{-1})) = \|D(u_N - u_N^k)\|_{L^2(\Omega)}^2. \]

Substituting (47) into (45), we arrive at the following identity:

\[
\begin{align*}
|S_N - S_N^k|_{L^2(\Omega)}^2 + \lambda^2 \|D(u_N - u_N^k)\|_{L^2(\Omega)}^2 \\
= |S_N - S_N^{-1}|_{L^2(\Omega)}^2 - 2\lambda (F(S_N) - F(S_N^{-1}), S_N - S_N^{-1}) \\
+ \lambda^2 (F(S_N) - F(S_N^{-1}), P_N(F(S_N) - F(S_N^{-1}))).
\end{align*}
\]

(48)

As $|F(A) - F(B)| \leq 2|A - B|$ (cf. Lemma 2), it follows that

\[
|S_N - S_N^{-1}|_{L^2(\Omega)}^2 - 2\lambda (F(S_N) - F(S_N^{-1}), S_N - S_N^{-1}) + \lambda^2 (F(S_N) - F(S_N^{-1}), P_N(F(S_N) - F(S_N^{-1}))) \leq 0.
\]

(49)

We focus our attention on the second term on the right-hand side of (49).

Thanks to Lemma 2,

\[
(F(S_N) - F(S_N^{k-1}), S_N - S_N^{k-1}) \geq c_0 \int_\Omega \frac{|S_N - S_N^{k-1}|^2}{(1 + |S_N| + |S_N^{k-1}|)^{\alpha+1}} \mathrm{d}x,
\]

(50)

\[
\geq \frac{c_0}{1 + \|S_N\|_{L^\infty(\Omega)} + \|S_N^{k-1}\|_{L^\infty(\Omega)}} \|S_N - S_N^{k-1}\|_{L^2(\Omega)}^2,
\]

where $c_0 = \min(1, 2^{\alpha-\frac{1}{2}})$. As $S_N^0 := 0$, there exists a positive constant $c_0$, independent of $k$ (but possibly dependent on $N$), such that

\[
1 + \|S_N\|_{L^\infty(\Omega)} + \|S_N^0\|_{L^\infty(\Omega)} \leq c_0.
\]

Suppose, for induction, that we have already shown that

\[
1 + \|S_N\|_{L^\infty(\Omega)} + \|S_N^m\|_{L^\infty(\Omega)} \leq c_0 \quad \forall m \in \{0, \ldots, k-1\},
\]

(51)

for some $k \geq 1$. It then follows from (50) and (51) that

\[
(F(S_N) - F(S_N^{k-1}), S_N - S_N^{k-1}) \geq c_0 \|S_N - S_N^{k-1}\|_{L^2(\Omega)}^2,
\]

with $c_0 := \min(1, 2^{\alpha-\frac{1}{2}})$. Substituting this into the right-hand side of (49) we deduce that

\[
\|S_N - S_N^k\|_{L^2(\Omega)}^2 + \lambda^2 \|D(u_N - u_N^k)\|_{L^2(\Omega)}^2 \leq (1 - 2c_0\lambda + 4\lambda^2) \|S_N - S_N^{k-1}\|_{L^2(\Omega)}^2.
\]

(52)

Let us choose

\[
\lambda \in \left(0, \frac{1}{2} c_0\right).
\]

Then,

\[
L^2 := 1 - 2c_0\lambda + 4\lambda^2 \in (0, 1).
\]

Consequently, (52) yields

\[
\|S_N - S_N^k\|_{L^2(\Omega)}^2 + \lambda^2 \|D(u_N - u_N^k)\|_{L^2(\Omega)}^2 \leq L^2 \|S_N - S_N^{k-1}\|_{L^2(\Omega)}^2, \quad L \in (0, 1).
\]

(53)

In order to complete the inductive step, it remains to show that (51) holds for all $m \in \{0, \ldots, k\}$, $k \geq 1$. To this end, we note that (53) implies that

\[
\|S_N - S_N^k\|_{L^2(\Omega)} \leq L^k \|S_N - S_N^0\|_{L^2(\Omega)} = L^k \|S_N\|_{L^2(\Omega)}.
\]

(54)
Thus, by the Nikol’skii inequality $\|T_N\|_{L^\infty(\Omega)} \leq C_{\text{inv}} N^{\frac{1}{2}} \|T_N\|_{L^2(\Omega)}$, $T_N \in \Sigma_N$, we have that

$$
\|S_N\|_{L^\infty(\Omega)} \leq \|S_N - S_N^k\|_{L^\infty(\Omega)} + \|S_N\|_{L^\infty(\Omega)} 
\leq C_{\text{inv}} N^{\frac{1}{2}} L^2 \|S_N\|_{L^2(\Omega)} + \|S_N\|_{L^\infty(\Omega)}.
$$

(55)

Hence,

$$
1 + \|S_N\|_{L^\infty(\Omega)} + \|S_N^k\|_{L^\infty(\Omega)} \leq 1 + 2\|S_N\|_{L^\infty(\Omega)} + C_{\text{inv}} N^{\frac{1}{2}} L^2 \|S_N\|_{L^2(\Omega)} 
\leq 1 + (2 + C_{\text{inv}} N^{\frac{1}{2}} |\Omega|^\frac{1}{2}) \|S_N\|_{L^\infty(\Omega)} 
\leq 1 + (2 + C_{\text{inv}} N^{\frac{1}{2}} |\Omega|^\frac{1}{2}) \|S_N\|_{L^\infty(\Omega)}.
$$

(56)

Thus we define

$$
c_0 := 1 + (2 + C_{\text{inv}} N^{\frac{1}{2}} |\Omega|^\frac{1}{2}) \|S_N\|_{L^\infty(\Omega)}
$$

to deduce that, with this definition of $c_0$, (51) holds with $k-1$ replaced by $k$, which then completes the inductive step. In particular, this implies that (53), and therefore also (54), holds for all $k \geq 1$.

Thus, from (53) and (54) we deduce that

$$
\|S_N - S_N^k\|_{L^2(\Omega)}^2 + \lambda^2 \|D(u_N - u_N^k)\|_{L^2(\Omega)}^2 \leq L^{2k} \|S_N\|_{L^2(\Omega)}^2 \quad \forall k \geq 1,
$$

where $L \in (0, 1)$, and hence $(S_N^k, D(u_N^k)) \to (S_N, D(u_N))$ as $k \to \infty$; thus, by Korn’s inequality, also $(S_N^k, u_N^k) \to (S_N, u_N)$ as $k \to \infty$. \hfill \Box

**Remark 3.** Some remarks are in order at this point. As a function of $\lambda$, $L^2 = 1 - 2c_0 \lambda + 4\lambda^2$ is minimized for $\lambda = \frac{1}{2} c_0$, yielding $L^2 = 1 - \frac{c_0^2}{4}$ (assuming that $c_0 \in (0, 2)$, which can always be achieved by choosing $c_0 > \frac{1}{4} c_0$).

Our next remark concerns the choice of $c_0$. As $C_{\text{inv}} L^k N^{\frac{1}{2}} |\Omega|^\frac{1}{2} \to 0$ when $k \to \infty$, there exists a positive integer $k_0 = k_0(N)$ such that $C_{\text{inv}} L^k N^{\frac{1}{2}} |\Omega|^\frac{1}{2} \leq 1$ for all $k \geq k_0$. For example, one can take

$$
k_0 := \left[ \frac{\log C_{\text{inv}} |\Omega|^\frac{1}{2} + \frac{d}{2} \log N}{\log \frac{1}{2}} \right] + 1.
$$

Using this refined upper bound in (56) allows us to redefine $c_0$ as $c_0 := 1 + 3\|S_N\|_{L^\infty(\Omega)}$. In fact, since we know from the proof of Theorem 3 that $\|S_N\|_{L^\infty(\Omega)} \leq 2\|S\|_{L^\infty(\Omega)}$ for all $N \geq N_*$, with $N_*$ as defined in the proof of Theorem 3, we can further redefine $c_0$ as

$$
c_0 := 1 + \frac{1}{2} c_0 + 6\|S\|_{L^\infty(\Omega)},
$$

thus rendering $c_0 := \frac{c_0}{2} \in (0, 2)$ independent of $N$, and thereby $\lambda = \frac{1}{2} c_0$ and $L^2 = 1 - \frac{c_0^2}{4}$ become independent of $N$. In other words, once $N \geq N_*$ and $k \geq k_0(N) \sim \frac{d}{4} \log N$, the asymptotic rate of convergence of the iterative method (36), (37) is independent of $N$, provided that $(S, D(u)) \in [H^m_0(\Omega)]^{d \times d} \times [H^m_0(\Omega)]^{d \times d}$ with $s > \frac{d}{2}$.

6. Conclusions

This paper provides an initial step towards the rigorous mathematical analysis of numerical approximations to nonlinear elastic limiting strain models. We have constructed a spectral approximation of the model problem under consideration and have shown that the spectral method exhibits optimal order convergence. We have also proposed an iterative method for the numerical solution of the finite-dimensional system of nonlinear equations featuring in the method and have shown that the iterations converge, at a linear rate, to the unique solution of the numerical method. Our aim in future work will be to extend the results developed here to limiting strain models in general multidimensional domains, in the spirit of the PDE analysis developed in the paper [1], using a finite element method.
References


Mathematical Institute, University of Oxford, Woodstock Road, Oxford OX2 6GG, United Kingdom
E-mail address: Endre.Suli@maths.ox.ac.uk