Computer Assisted Mechanics and Engineering Sciences, 8: 425–437, 2001.


# General aspects of Trefftz method and relations to error estimation of finite element approximations

Stephan Ohnimus, Marcus Rüter and Erwin Stein
*Institute for Structural and Computational Mechanics (IBNM), University of Hannover*
*Appelstr. 9A, 30167 Hannover, Germany*

(Received June 30, 2000)

In this paper a guaranteed upper bound of the global discretization error in linear elastic finite element approximations is presented, based on a generalized Trefftz functional. Therefore, the general concept of complementary energy functionals and the corresponding approximation methods of Ritz, Trefftz, the method of orthogonal projection and the hypercircle method are briefly outlined. Furthermore, it is shown how to use a generalized Trefftz functional to solve a Neumann problem in linear elasticity. Based on an implicit a posteriori error estimator within the finite element method, using equilibrated local Neumann problems, the generalized Trefftz functional yields a computable guaranteed upper bound of the discretization error without multiplicative constants.

## 1. Introduction

The finite element method has been well established as an important tool for the numerical solution of many kinds of boundary value problems in applied mathematics and engineering sciences. As a matter of fact, the great success of this approach, a Ritz or Galerkin method applied to finite subdomains, is based on a variational principle. While Ritz method is based on the minimization of the total potential energy of a system, the dual idea naturally arises to maximize the corresponding complementary energy. This method is strongly correlated with the name of Trefftz [45] who presented this theory in 1926 on the 2nd International Congress on Applied Mechanics held in Zurich, Switzerland. Trefftz himself called his pioneering work "a counterpart to Ritz method", but nowadays this approach is simply called "Trefftz method" in honour of Erich Trefftz.

Meanwhile, a vast of literature has been devoted to Trefftz method applied to second order elliptic boundary value problems, where many sophisticated methods, especially within the finite element method (FEM) and the boundary integral equation method (BIEM), were developed. Particularly, many authors dealt with the development of the so-called hybrid Trefftz (HT) finite elements with relations to the hybrid stress method as introduced by Pian [30]. The HT finite elements were initiated by Jirousek [18] and considerably developed by many different authors like Zieliński and Zienkiewicz [51], Peters et al. [29], Qin [32] and Freitas et al. [13]. Furthermore, an a posteriori error indicator for HT elements was presented by Jirousek and Venkatesh [19]. A central point of Trefftz method is that the functions have to fulfill the governing PDE in the domain, but they are not subjected to any boundary condition. This leads to the Trefftz boundary element method (TBEM), as presented by Herrera [15], Jin et al. [17], Kita and Kamiya [22] and others, in contrast to BIEM, where a Green's function as fundamental solution is used. A coupling of finite elements with boundary elements was shown by Brink et al. [7].

An important property of Trefftz method, however, is the bounding property of Trefftz functionals in convex analysis which makes the method very attractive, particularly in view of the a posteriori error analysis of finite element approximations. In an early work of Cooperman [12] Trefftz method

is already applied to obtain local bounds of the approximated solution in linear elasticity. More recently, different approaches based on complementary functionals have been developed. Kelly [21], for instance, presented an a posteriori error estimator based on a complementary energy functional evaluated elementwise. Oden et al. [26] made use of both, the lower bound property of the energy functional and the upper bound property of the corresponding complementary energy functional, to derive an error estimator in the primal as well as in the dual approximation. The dual theory was also used by Han [14] to estimate the error caused by the linearization of nonlinear problems. As pointed out by Repin and Xanthis [34, 35] it can be rather cumbersome to compute an approximated solution of the complementary problem. In order to bypass this numerical difficulty, they developed an error estimator for both, linear and nonlinear variational problems, without directly solving the complementary problem.

The present paper is organized as follows. In Section 2 the complementary energy functional associated with the potential energy of an elastic body is derived. Moreover, it is shown how to compute numerical solutions of the corresponding maximization and minimization principles using the methods of Ritz and Trefftz as well as the method of orthogonal projection and the hypercircle method. In what follows, we deal with Neumann boundary value problems in linear elasticity. A generalized Trefftz functional is applied in Section 3 to compute solutions of Neumann problems using a maximization principle. In Section 4, the finite element method is introduced, and an a posteriori error estimator based on equilibrated local Neumann problems is presented. Finally, we present in Section 5 a computable guaranteed upper bound of the discretization error using a generalized Trefftz functional.

## 2. General concept of complementary variational principles

### 2.1. The linear elasticity problem

Within this section we outline the general concept of finding solutions of the linear elasticity problem by minimizing a quadratic energy functional and, conversely, by maximizing the corresponding quadratic complementary energy functional, see e.g. Mikhlin [24], Arthurs [3], Nowinski [25] and Reddy [33]. Following standard usage, let the elastic body be given by the closure of a bounded open set $\Omega \subset \mathbb{R}^3$ with piecewise smooth and Lipschitz continuous boundary $\Gamma$ such that $\Gamma = \bar{\Gamma}_D \cup \bar{\Gamma}_N$ and $\Gamma_D \cap \Gamma_N = \emptyset$, where $\Gamma_D$ and $\Gamma_N$ denote the measurable Dirichlet and the Neumann boundary, respectively. Material points of the body are denoted by $\mathbf{x} \in \bar{\Omega} = \Omega \cup \Gamma$. Furthermore, the total potential energy of the elastic body $\Pi : [H^1(\Omega)]^3 \to \mathbb{R}$ is given by

$$\Pi(\mathbf{u}) = \frac{1}{2} a(\mathbf{u}, \mathbf{u}) - l(\mathbf{u}), \tag{1}$$

where we introduced the continuous, symmetric, positive semidefinite bilinear form $a : [H^1(\Omega)]^3 \times [H^1(\Omega)]^3 \to \mathbb{R}$ and the potential of conservative exterior loads, given by the continuous linear form $l : [H^1(\Omega)]^3 \to \mathbb{R}$, defined by

$$a(\mathbf{u}, \mathbf{u}) = 2 \int_\Omega W_s(\mathbf{x}, \boldsymbol{\varepsilon}) \, \mathrm{d}V \qquad \text{and} \qquad l(\mathbf{u}) = \int_\Omega \bar{\mathbf{f}} \cdot \mathbf{u} \, \mathrm{d}V + \int_{\Gamma_N} \bar{\mathbf{t}} \cdot \mathbf{u} \, \mathrm{d}S, \tag{2}$$

respectively. Here, $\mathbf{u} : \bar{\Omega} \to \mathbb{R}^3$ is the displacement field, $W_s : \bar{\Omega} \times \mathbb{M}^3 \to \mathbb{R}$ ($\mathbb{M}^n$ denotes the set of all real square tensors of order $n$) is the convex strain-energy per unit volume of the system, $\boldsymbol{\varepsilon} = \frac{1}{2}(\nabla \mathbf{u} + (\nabla \mathbf{u})^{\mathrm{T}})$ is the strain tensor, $\bar{\mathbf{f}} \in [L_2(\Omega)]^3$ are prescribed body forces and $\bar{\mathbf{t}} \in [L_2(\Omega)]^3$ are prescribed surface tractions. For the sake of simplicity, let us restrict our considerations to the case of isotropy. Then, the strain-energy per unit volume is given by $W_s(\mathbf{x}, \cdot) = \frac{1}{2} \boldsymbol{\varepsilon} : \mathbb{C} : \boldsymbol{\varepsilon}$, where $\mathbb{C} = \partial^2_{\boldsymbol{\varepsilon}} W_s$ is the positive definite, constant, fourth-order material stiffness tensor depending only on the two Lamé parameters $\lambda$ and $\mu$. Now, the basic problem of the energetic equilibrium formulation consists of finding the minimum of the energy functional (1) (Dirichlet's principle)

$$\Pi(\hat{\mathbf{u}}) = \min_{\mathbf{u} \in \mathcal{V} + \mathbf{u}_0} \Pi(\mathbf{u}) \tag{3}$$

for a given $\mathbf{u}_0 \in [H^1(\Omega)]^3$, such that $\mathbf{u}_0|_{\Gamma_D} = \bar{\mathbf{u}} \in [H^{\frac{1}{2}}(\Gamma_D)]^3$, over the linear manifold

$$\mathcal{V} + \mathbf{u}_0 = \left\{ \mathbf{v} \in \left[H^1(\Omega)\right]^3 \; ; \; \mathbf{v} = \mathbf{w} + \mathbf{u}_0 \, , \; \mathbf{w} \in \mathcal{V} \right\}, \tag{4}$$

where the subspace $\mathcal{V} \subset [H^1(\Omega)]^3$ is given by

$$\mathcal{V} = \left\{ \mathbf{v} \in \left[H^1(\Omega)\right]^3 \; ; \; \mathbf{v}|_{\Gamma_D} = \mathbf{0} \right\}. \tag{5}$$

We point out that the bilinear form $a$ is positive definite for all $\mathbf{u} \in \mathcal{V}$ and $\mathbf{u} \neq \mathbf{0}$. As a consequence, $a$ generates a norm in $\mathcal{V}$, the so-called energy norm $\|\cdot\|_E = \sqrt{a(\cdot,\cdot)}$. Thus, $\mathcal{V}$ is a closed subspace of $[H^1(\Omega)]^3$. Indeed, it is easy to see that the energy norm is generally a seminorm in the entire Sobolev space $[H^1(\Omega)]^3$, denoted by $|\cdot|_E$.

A well-known theorem states that $\hat{\mathbf{u}} \in \mathcal{V} + \mathbf{u}_0$ is a solution of the minimization problem (3) if and only if the first variation of the energy functional $\Pi(\mathbf{u})$,

$$a(\hat{\mathbf{u}}, \mathbf{v}) = l(\mathbf{v}) \qquad \forall \mathbf{v} \in \mathcal{V}, \tag{6}$$

vanishes for arbitrary $\mathbf{v} \in \mathcal{V}$. Furthermore, due to the lemma of Lax–Milgram, the solution $\hat{\mathbf{u}} \in \mathcal{V} + \mathbf{u}_0$ of the variational problem (6) exists and is unique under the conditions that $a$ is $\mathcal{V}$-elliptic, i.e. $a(\mathbf{v}, \mathbf{v}) \geq \alpha \|\mathbf{v}\|_{\mathcal{V}}^2$ for all $\mathbf{v} \in \mathcal{V}$ with a positive constant $\alpha$ as well as $a$ and $l$ are continuous as postulated above.

It can be easily verified that the variational form (6) is the weak formulation of the boundary value problem of linear elasticity which consists of finding a solution $\hat{\mathbf{u}} \in [C^2(\Omega)]^3 \cap [C^1(\bar{\Omega})]^3$ such that

$$-\nabla \cdot \boldsymbol{\sigma}(\hat{\mathbf{u}}) = \bar{\mathbf{f}} \qquad \text{in } \Omega, \tag{7a}$$

$$\boldsymbol{\sigma}(\hat{\mathbf{u}}) = \mathbb{C} : \boldsymbol{\varepsilon}(\hat{\mathbf{u}}) \qquad \text{in } \Omega, \tag{7b}$$

$$\hat{\mathbf{u}} = \bar{\mathbf{u}} \qquad \text{on } \Gamma_D, \tag{7c}$$

$$\boldsymbol{\sigma}(\hat{\mathbf{u}}) \cdot \mathbf{n} = \bar{\mathbf{t}} \qquad \text{on } \Gamma_N, \tag{7d}$$

where $\boldsymbol{\sigma} = \partial_{\boldsymbol{\varepsilon}} W_s$ is the symmetric stress tensor and $\mathbf{n}$ is the unit outward normal.

### 2.2. The complementary energy functional

In what follows, we want to derive the corresponding complementary energy functional to $\Pi(\mathbf{u})$. As a point of departure, it proves convenient to introduce the orthogonal complement $\mathcal{V}^\perp$ of $\mathcal{V}$ in $[H^1(\Omega)]^3$, i.e.

$$\mathcal{V}^\perp = \left\{ \mathbf{v} \in \left[H^1(\Omega)\right]^3 \; ; \quad a(\mathbf{v}, \mathbf{u}) = 0 \quad \forall \mathbf{u} \in \mathcal{V} \right\}. \tag{8}$$

More precisely, $\mathcal{V}^\perp$ consists of all functions $\mathbf{v} \in [H^1(\Omega)]^3$ fulfilling the homogeneous PDE, $-\nabla \cdot \boldsymbol{\sigma}(\hat{\mathbf{u}}) = \mathbf{0}$ in $\Omega$, and the homogeneous Neumann boundary condition $\boldsymbol{\sigma}(\hat{\mathbf{u}}) \cdot \mathbf{n} = \mathbf{0}$ on $\Gamma_N$ without any Dirichlet boundary condition. Moreover, the closed subspaces $\mathcal{V}$ and $\mathcal{V}^\perp$ are orthogonal with respect to the inner product $a$, and it follows that $\mathcal{V} \cap \mathcal{V}^\perp = \{\mathbf{0}\}$. From the well-known projection theorem we conclude that

$$\left[H^1(\Omega)\right]^3 = \mathcal{V} \oplus \mathcal{V}^\perp. \tag{9}$$

In other words, every $\mathbf{v} \in [H^1(\Omega)]^3$ has the unique representation $\bar{\mathbf{v}} + \mathbf{v}^\perp$, with $\bar{\mathbf{v}} \in \mathcal{V}$ and $\mathbf{v}^\perp \in \mathcal{V}^\perp$. Let us now consider the following variational problem: find a solution $\mathbf{w} \in [H^1(\Omega)]^3$ such that

$$a(\mathbf{w}, \mathbf{v}) = l(\mathbf{v}) \qquad \forall \mathbf{v} \in \mathcal{V}. \tag{10}$$

Obviously, the solution $\mathbf{w}$ is not unique, since the bilinear form $a$ is positive semidefinite. But from (6) we conclude that the variational problem (10) has at least one solution, namely $\mathbf{w} = \hat{\mathbf{u}}$. Now, let $\mathbf{w}_0 \in [H^1(\Omega)]^3$ be any solution of (10) with the unique representation $\mathbf{w}_0 = \bar{\mathbf{w}}_0 + \mathbf{w}_0^\perp$, where $\bar{\mathbf{w}}_0 \in \mathcal{V}$ and $\mathbf{w}_0^\perp \in \mathcal{V}^\perp$, due to (9). By substituting $\mathbf{w} = \mathbf{w}_0$ in (10), the variational formulation reduces to

$$a(\bar{\mathbf{w}}_0, \mathbf{v}) = l(\mathbf{v}) \qquad \forall \mathbf{v} \in \mathcal{V}. \tag{11}$$

Likewise, every $\mathbf{w} \in \mathcal{V}^\perp + \mathbf{w}_0$ leads to the same variational formulation (11) and is therefore a solution of (10). Furthermore, from (11) and (6) we see that if $\mathbf{w} \in \mathcal{V}^\perp + \mathbf{w}_0$ and $\mathbf{w} \in \mathcal{V} + \mathbf{u}_0$, then $\mathbf{w} = \hat{\mathbf{u}}$, or, more concisely, the solution is the intersection of the two manifolds $\mathcal{V}^\perp + \mathbf{w}_0 \cap \mathcal{V} + \mathbf{u}_0 = \hat{\mathbf{u}}$.

Following Velte [47], let us introduce the energy seminorm

$$|\mathbf{u} - \mathbf{w}|_E^2 = a(\mathbf{u}, \mathbf{u}) - 2a(\mathbf{u}, \mathbf{w}) + a(\mathbf{w}, \mathbf{w}) \geq 0 \tag{12}$$

which holds for arbitrary $\mathbf{u}, \mathbf{w} \in [H^1(\Omega)]^3$. In the special case that $\mathbf{u} \in \mathcal{V} + \mathbf{u}_0$ and $\mathbf{w} \in \mathcal{V}^\perp + \mathbf{w}_0$, the variational formulation (10) can be expressed with $\mathbf{v} = \mathbf{u} - \mathbf{u}_0 \in \mathcal{V}$ by

$$a(\mathbf{w}, \mathbf{u}) = l(\mathbf{u}) + a(\mathbf{w}, \mathbf{u}_0) - l(\mathbf{u}_0). \tag{13}$$

Thus, substituting of (13) in (12) results in

$$|\mathbf{u} - \mathbf{w}|_E^2 = 2\Pi(\mathbf{u}) - 2\overset{*}{\Pi}(\mathbf{w}) \geq 0 \qquad \forall \mathbf{u} \in \mathcal{V} + \mathbf{u}_0\,, \quad \forall \mathbf{w} \in \mathcal{V}^\perp + \mathbf{w}_0\,, \tag{14}$$

with the complementary energy functional

$$\overset{*}{\Pi}(\mathbf{w}) = -\frac{1}{2} a(\mathbf{w}, \mathbf{w}) + a(\mathbf{w}, \mathbf{u}_0) - l(\mathbf{u}_0) \tag{15}$$

which implies $\Pi(\mathbf{u}) \geq \overset{*}{\Pi}(\mathbf{w})$. In particular, this inequality must be fulfilled on the one hand for a fixed $\mathbf{u}$ and all admissible $\mathbf{w}$ and on the other hand for a fixed $\mathbf{w}$ and all admissible $\mathbf{u}$. Furthermore, the solution $\hat{\mathbf{u}}$ of the variational problem (6) fulfills $\Pi(\hat{\mathbf{u}}) = \overset{*}{\Pi}(\hat{\mathbf{u}})$. We thus finally arrive at the important result

$$\min_{\mathbf{u} \in \mathcal{V} + \mathbf{u}_0} \Pi(\mathbf{u}) = \Pi(\hat{\mathbf{u}}) = \overset{*}{\Pi}(\hat{\mathbf{u}}) = \max_{\mathbf{w} \in \mathcal{V}^\perp + \mathbf{w}_0} \overset{*}{\Pi}(\mathbf{w}). \tag{16}$$

### 2.3. The methods of Ritz and Trefftz

We now turn to the numerical solution of minimization and maximization problems as developed in the preceding section. First recall *Ritz method*, see Ritz [37], which is based on the minimization problem (3). It turns out that this problem is a special case of (14) in which $\mathbf{w} = \hat{\mathbf{u}}$, since we have

$$|\mathbf{u} - \hat{\mathbf{u}}|_E^2 = 2\Pi(\mathbf{u}) - 2\Pi(\hat{\mathbf{u}}) \geq 0 \qquad \forall \mathbf{u} \in \mathcal{V} + \mathbf{u}_0\,. \tag{17}$$

Apparently, it is generally impossible to give a closed-form solution of this minimization problem. Therefore, Ritz-method is based on solving the minimization problem (3) in a finite-dimensional manifold $\mathcal{V}_h + \mathbf{u}_0$ with finite-dimensional subspace $\mathcal{V}_h \subset \mathcal{V}$. The resulting approximated solution $\hat{\mathbf{u}}_h$ is the orthogonal projection of the exact solution $\hat{\mathbf{u}}$ on $\mathcal{V}_h + \mathbf{u}_0$.

Alternatively, based on (14), it is evident that an approximated solution $\hat{\mathbf{u}}_h$ of the minimization problem (3) can be computed by solving the maximization problem of the complementary energy functional (15). This is the basic idea of *Trefftz method*, see Trefftz [45]. Analogously to (17), the maximization problem is of the form

$$|\mathbf{w} - \hat{\mathbf{u}}|_E^2 = 2\overset{*}{\Pi}(\hat{\mathbf{u}}) - 2\overset{*}{\Pi}(\mathbf{w}) \geq 0 \qquad \forall \mathbf{w} \in \mathcal{V}^\perp + \mathbf{w}_0 \tag{18}$$

which we solve in a finite-dimensional manifold $\mathcal{V}_h^\perp + \mathbf{w}_0$ with subspace $\mathcal{V}_h^\perp \subset \mathcal{V}^\perp$. Note that the solution is not unique, but it exists at least one solution of the maximization problem. However, a unique solution is obtained by choosing a basis for $\mathcal{V}_h^\perp$ such that the resulting matrix has maximal rank (cf. Velte [46]).

Clearly, as a consequence of approximately solving on the one hand the minimization problem using Ritz method and on the other hand the maximization problem using Trefftz method in finite-dimensional manifolds, we obtain the relation

$$\min_{\mathbf{u}_h \in \mathcal{V}_h + \mathbf{u}_0} \Pi(\mathbf{u}_h) \geq \min_{\mathbf{u} \in \mathcal{V} + \mathbf{u}_0} \Pi(\mathbf{u}) = \Pi(\hat{\mathbf{u}}) = \overset{*}{\Pi}(\hat{\mathbf{u}}) = \max_{\mathbf{w} \in \mathcal{V}^\perp + \mathbf{w}_0} \overset{*}{\Pi}(\mathbf{w}) \geq \max_{\mathbf{w}_h \in \mathcal{V}_h^\perp + \mathbf{w}_0} \overset{*}{\Pi}(\mathbf{w}_h). \tag{19}$$

### 2.4. The method of orthogonal projection

We now focus our attention on two further energy functionals given by

$$\Pi_1(\mathbf{u}) = \|\mathbf{u} - \mathbf{w}_0\|_E^2 \quad \text{and} \quad \Pi_2(\mathbf{w}) = \|\mathbf{w} - \mathbf{u}_0\|_E^2 . \tag{20}$$

The minimization problems associated with (3) now consist of finding the minima

$$\Pi_1(\hat{\mathbf{u}}) = \min_{\mathbf{u} \in \mathcal{V} + \mathbf{u}_0} \Pi_1(\mathbf{u}) \quad \text{and} \quad \Pi_2(\hat{\mathbf{w}}) = \min_{\mathbf{w} \in \mathcal{V}^\perp + \mathbf{w}_0} \Pi_2(\mathbf{w}). \tag{21}$$

It is readily seen that this is equivalent to finding the unique solutions $\hat{\mathbf{u}} \in \mathcal{V} + \mathbf{u}_0$ and $\hat{\mathbf{w}} \in \mathcal{V}^\perp + \mathbf{w}_0$ of the variational problems

$$a(\hat{\mathbf{u}} - \mathbf{w}_0, \mathbf{v}) = 0 \quad \forall \mathbf{v} \in \mathcal{V} \quad \text{and} \quad a(\hat{\mathbf{w}} - \mathbf{u}_0, \mathbf{v}) = 0 \quad \forall \mathbf{v} \in \mathcal{V}^\perp, \tag{22}$$

respectively. Hence, we conclude from the construction of the space $\mathcal{V}^\perp$ in (8) that $\hat{\mathbf{u}} \in \mathcal{V}^\perp + \mathbf{w}_0$ and $\hat{\mathbf{w}} \in \mathcal{V} + \mathbf{u}_0$ according to which it follows directly that both minimization problems (21) have the same unique solution $\hat{\mathbf{u}} = \hat{\mathbf{w}}$. A point to be emphasized is that the solution $\hat{\mathbf{u}} = \hat{\mathbf{w}}$ is on the one hand the orthogonal projection of $\mathbf{u}_0$ on $\mathcal{V}^\perp + \mathbf{w}_0$ and on the other hand the orthogonal projection of $\mathbf{w}_0$ on $\mathcal{V} + \mathbf{u}_0$. Therefore, this approach is called the *method of orthogonal projection* as introduced by Zaremba [49, 50] and Weyl [48]. A graphical interpretation of this method in $\mathbb{R}^2$ is depicted in Fig. 1. We end up with introducing finite-dimensional subspaces $\mathcal{V}_h \subset \mathcal{V}$ and $\mathcal{V}_h^\perp \subset \mathcal{V}^\perp$ in order to solve the minimization problems (21) numerically by means of a uniquely solvable linear system of equations.

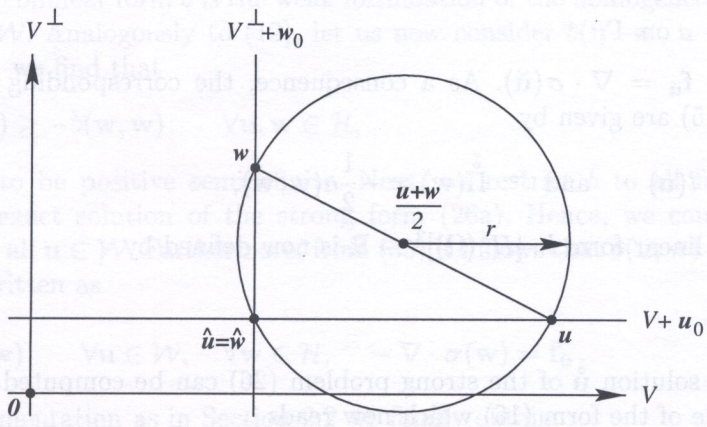


**Fig. 1.** The hypercircle method

### 2.5. The hypercircle method after Prager and Synge

To complete the picture, we show next how to compute the solution $\hat{\mathbf{u}} = \hat{\mathbf{w}}$ of the minimization problems (21) numerically using the *hypercircle method* as developed by Prager and Synge [31] as an alternative approach to the method of orthogonal projection. Since $\mathbf{u} \in \mathcal{V} + \mathbf{u}_0$ and $\mathbf{w} \in \mathcal{V}^\perp + \mathbf{w}_0$, we get $\mathbf{u} - \hat{\mathbf{u}} \in \mathcal{V}$ and $\mathbf{w} - \hat{\mathbf{u}} \in \mathcal{V}^\perp$. Consequently, from the definition of $\mathcal{V}^\perp$ in (8), we obtain the important result $a(\mathbf{u}-\hat{\mathbf{u}}, \mathbf{w}-\hat{\mathbf{u}}) = 0$. With this notation at hand, it follows the identity

$$\|(\mathbf{u} - \hat{\mathbf{u}}) + (\mathbf{w} - \hat{\mathbf{u}})\|_E^2 = \|(\mathbf{u} - \hat{\mathbf{u}}) - (\mathbf{w} - \hat{\mathbf{u}})\|_E^2 \tag{23}$$

which can be rewritten as

$$\left\| \frac{\mathbf{u}+\mathbf{w}}{2} - \hat{\mathbf{u}} \right\|_E^2 = \frac{1}{4}\|\mathbf{u} - \mathbf{w}\|_E^2 . \tag{24}$$

Obviously, this result is the equation of a circle, the so-called hypercircle, with radius $r = \frac{1}{2}\|\mathbf{u}-\mathbf{w}\|_E$ and midpoint $\frac{1}{2}(\mathbf{u}+\mathbf{w})$ as shown graphically in Fig. 1 for the case of $\mathbf{u}, \mathbf{w} \in \mathbb{R}^2$. Thus, the solution $\hat{\mathbf{u}} = \hat{\mathbf{w}}$ of the minimization problems (21) is lying on the hypercircle described by (24). Again, in the context of a numerically computed approximated solution, we introduce finite-dimensional subspaces $\mathcal{V}_h \subset \mathcal{V}$ and $\mathcal{V}_h^\perp \subset \mathcal{V}^\perp$ in order to obtain a uniquely solvable linear system of equations which results from the minimization of $r^2$, that is the right-hand side of (24). For further elaborations and details we refer to Synge [42–44], Rieder [36] and Velte [46].

## 3. Solution of Neumann problems with the generalized Trefftz method

### 3.1. A Neumann problem in linear elasticity

In this section we apply Trefftz method in conjunction with the a posteriori error analysis of finite element approximations (cf. Section 4) to a Neumann problem. Let us begin with the strong form of the following Neumann problem, derived from the linear elasticity problem (7), where we omitted the body forces $\bar{\mathbf{f}}$: find a solution $\hat{\mathbf{u}} \in [C^2(\Omega)]^3 \cap [C^1(\bar{\Omega})]^3$ such that

$$-\nabla \cdot \boldsymbol{\sigma}(\hat{\mathbf{u}}) = \mathbf{0} \qquad \text{in } \Omega, \tag{25a}$$

$$\boldsymbol{\sigma}(\hat{\mathbf{u}}) \cdot \mathbf{n} = \bar{\mathbf{t}} \qquad \text{on } \Gamma = \Gamma_N . \tag{25b}$$

Further, we additively decompose the solution $\hat{\mathbf{u}} = \breve{\mathbf{u}} + \mathring{\mathbf{u}}$ in such a fashion that $\breve{\mathbf{u}}$ and $\mathring{\mathbf{u}}$ fulfill the boundary conditions $\boldsymbol{\sigma}(\breve{\mathbf{u}}) \cdot \mathbf{n} = \bar{\mathbf{t}}$ on $\Gamma$ and $\boldsymbol{\sigma}(\mathring{\mathbf{u}}) \cdot \mathbf{n} = \mathbf{0}$ on $\Gamma$, respectively. This leads to the transformed boundary value problem

$$-\nabla \cdot \boldsymbol{\sigma}(\mathring{\mathbf{u}}) = \mathbf{f}_{\breve{\mathbf{u}}} \qquad \text{in } \Omega, \tag{26a}$$

$$\boldsymbol{\sigma}(\mathring{\mathbf{u}}) \cdot \mathbf{n} = \mathbf{0} \qquad \text{on } \Gamma, \tag{26b}$$

with the body forces $\mathbf{f}_{\breve{\mathbf{u}}} = \nabla \cdot \boldsymbol{\sigma}(\breve{\mathbf{u}})$. As a consequence, the corresponding transformed energy functionals (1) and (15) are given by

$$\mathring{\Pi}(\mathbf{u}) = \frac{1}{2}a(\mathbf{u}, \mathbf{u}) - l(\mathbf{u}) \qquad \text{and} \qquad \overset{*}{\mathring{\Pi}}(\mathbf{w}) = -\frac{1}{2}a(\mathbf{w}, \mathbf{w}), \tag{27}$$

where the continuous linear form $l : [H^1(\Omega)]^3 \to \mathbb{R}$ is now defined by

$$l(\mathbf{u}) = \int_\Omega \mathbf{f}_{\breve{\mathbf{u}}} \cdot \mathbf{u} \, \mathrm{d}V. \tag{28}$$

Similarly to (16), the solution $\mathring{\mathbf{u}}$ of the strong problem (26) can be computed by a minimization–maximization principle of the form (16) which now reads

$$\min_{\mathbf{u} \in \mathcal{V}} \mathring{\Pi}(\mathbf{u}) = \mathring{\Pi}(\mathring{\mathbf{u}}) = \overset{*}{\mathring{\Pi}}(\mathring{\mathbf{u}}) = \max_{\mathbf{w} \in \mathcal{V}^\perp + \mathbf{w}_0} \overset{*}{\mathring{\Pi}}(\mathbf{w}). \tag{29}$$

### 3.2. The generalized Trefftz method

It is elementary to verify that a solution of the Neumann problem (26), computed by the maximization of the corresponding complementary energy functional in the sense of Trefftz, has to fulfill the PDE (26a) as well as the Neumann boundary condition (26b) at least in a weak sense which can cause difficulties in finding functions fulfilling both, the field equations and the boundary conditions. In order to cope with this problem, we will now construct a complementary energy functional defined for functions satisfying only the PDE (26a) in the domain. In other words, the functions are not subjected to any boundary condition, see e.g. Mikhlin [24]. For this purpose, let us first introduce the quotient space $\mathcal{H} = [H^2(\Omega)]^3/\mathcal{Z}$, where

$$\mathcal{Z} = \{\mathbf{v} \in \mathcal{V}\,; \quad a(\mathbf{w}, \mathbf{v}) = 0 \quad \forall \mathbf{w} \in \mathcal{V}\} \tag{30}$$

denotes the space of rigid body motions on which we define an energy functional in the form of a symmetric, positive semidefinite bilinear form $b : \mathcal{H} \times \mathcal{H} \to \mathbb{R}$, as suggested by Birman [5], given by

$$b(\mathbf{u}, \mathbf{w}) = a(\mathbf{u}, \mathbf{w}) - \alpha \int_\Gamma \mathbf{u} \cdot \mathbf{w}\, \mathrm{d}S + \frac{1}{\alpha} \int_\Gamma (\boldsymbol{\sigma}(\mathbf{u}) \cdot \mathbf{n} - \alpha \mathbf{u}) \cdot (\boldsymbol{\sigma}(\mathbf{w}) \cdot \mathbf{n} - \alpha \mathbf{w})\, \mathrm{d}S \qquad \forall \mathbf{u}, \mathbf{w} \in \mathcal{H} \tag{31}$$

with constant $\alpha \in \mathbb{R}^+$ fulfilling the condition

$$\alpha \int_\Gamma \mathbf{u} \cdot \mathbf{u}\, \mathrm{d}S \leq a(\mathbf{u}, \mathbf{u}), \tag{32}$$

since $b$ is assumed to be positive semidefinite. Particularly, in the special case of $\mathbf{u} \in \mathcal{W} \subset \mathcal{H}$, where the subspace $\mathcal{W}$ is defined by

$$\mathcal{W} = \{\mathbf{v} \in \mathcal{H}\,; \quad ; \boldsymbol{\sigma}(\mathbf{v}) \cdot \mathbf{n} = \mathbf{0} \quad \text{on } \Gamma\}, \tag{33}$$

the bilinear form $b$ simplifies to

$$b(\mathbf{u}, \mathbf{w}) = a(\mathbf{u}, \mathbf{w}) - \int_\Gamma \mathbf{u} \cdot (\boldsymbol{\sigma}(\mathbf{w}) \cdot \mathbf{n})\, \mathrm{d}S \qquad \forall \mathbf{u} \in \mathcal{W}, \quad \forall \mathbf{w} \in \mathcal{H}. \tag{34}$$

Subsequently, after integration by parts, we obtain the result

$$b(\mathbf{u}, \mathbf{w}) = \int_\Omega \mathbf{u} \cdot (-\nabla \cdot \boldsymbol{\sigma}(\mathbf{w}))\, \mathrm{d}V \qquad \forall \mathbf{u} \in \mathcal{W}, \quad \forall \mathbf{w} \in \mathcal{H}, \tag{35}$$

which states that the bilinear form $b$ is the weak formulation of the homogeneous PDE, $-\nabla \cdot \boldsymbol{\sigma}(\mathbf{w}) = \mathbf{0}$ in $\Omega$, for all $\mathbf{u} \in \mathcal{W}$. Analogously to (12), let us now consider $b(\mathbf{u}-\mathbf{w}, \mathbf{u}-\mathbf{w})$ for all $\mathbf{u}, \mathbf{w} \in \mathcal{H}$. Following Velte [46], we find that

$$b(\mathbf{u}, \mathbf{u}) - 2b(\mathbf{u}, \mathbf{w}) \geq -b(\mathbf{w}, \mathbf{w}) \qquad \forall \mathbf{u}, \mathbf{w} \in \mathcal{H}, \tag{36}$$

since $b$ is assumed to be positive semidefinite. Next, we restrict $b$ to all functions $\mathbf{u} \in \mathcal{W}$ and $\mathbf{w} \in \mathcal{H}$ beeing an exact solution of the strong form (26a). Hence, we conclude from (34) that $b(\mathbf{u}, \mathbf{u}) = a(\mathbf{u}, \mathbf{u})$ for all $\mathbf{u} \in \mathcal{W}$. Furthermore, from (35) it follows that $b(\mathbf{u}, \mathbf{w}) = l(\mathbf{u})$. Consequently, Eq. (36) can be rewritten as

$$\mathring{\Pi}(\mathbf{u}) \geq -\frac{1}{2} b(\mathbf{w}, \mathbf{w}) \qquad \forall \mathbf{u} \in \mathcal{W}, \quad \forall \mathbf{w} \in \mathcal{H}, \quad -\nabla \cdot \boldsymbol{\sigma}(\mathbf{w}) = \mathbf{f}_{\ddot{u}}\,. \tag{37}$$

With the same argumentation as in Section 2.2 we finally obtain

$$\min_{\mathbf{u} \in \mathcal{V}} \mathring{\Pi}(\mathbf{u}) = \mathring{\Pi}(\mathring{\mathbf{u}}) = -\frac{1}{2} a(\mathring{\mathbf{u}}, \mathring{\mathbf{u}}) = -l(\mathring{\mathbf{u}}) = -\frac{1}{2} b(\mathring{\mathbf{u}}, \mathring{\mathbf{u}}) = \max_{\substack{\mathbf{w} \in \mathcal{H} \\ -\nabla \cdot \boldsymbol{\sigma}(\mathbf{w}) = \mathbf{f}_{\ddot{u}}}} -\frac{1}{2} b(\mathbf{w}, \mathbf{w}). \tag{38}$$

Here, the energy bilinear form $b$ is a generalized Trefftz functional. Clearly, for the corresponding discrete problem of finding a solution $\mathring{\mathbf{u}}_h \in \mathcal{V}_h \subset \mathcal{V}$ we get

$$\min_{\mathbf{u}_h \in \mathcal{V}_h} \mathring{\Pi}(\mathbf{u}_h) \geq \min_{\mathbf{u} \in \mathcal{V}} \mathring{\Pi}(\mathbf{u}) = \mathring{\Pi}(\mathring{\mathbf{u}}) = -\frac{1}{2} a(\mathring{\mathbf{u}}, \mathring{\mathbf{u}}) = -l(\mathring{\mathbf{u}})$$

$$= -\frac{1}{2} b(\mathring{\mathbf{u}}, \mathring{\mathbf{u}}) = \max_{\substack{\mathbf{w} \in \mathcal{H} \\ -\nabla \cdot \boldsymbol{\sigma}(\mathbf{w}) = \mathbf{f}_{\mathring{\mathbf{u}}}}} -\frac{1}{2} b(\mathbf{w}, \mathbf{w}) \geq \max_{\substack{\mathbf{w}_h \in \mathcal{H} \\ -\nabla \cdot \boldsymbol{\sigma}(\mathbf{w}_h) = \mathbf{f}_{\mathring{\mathbf{u}}}}} -\frac{1}{2} b(\mathbf{w}_h, \mathbf{w}_h), \quad (39)$$

with the finite-dimensional subspace $\mathcal{H}_h \subset \mathcal{H}$. Note that the functions $\mathbf{w}$ and $\mathbf{w}_h$ have to fulfill the PDE (26a) exactly in the domain $\Omega$, whereas the homogeneous boundary condition (26b) is not supposed to be fulfilled.

## 4. Finite element approximation and a posteriori error estimation

### 4.1. The $Q_k$-element

In this section we outline the basic theory of the finite element method as an alternative numerical approach to compute an approximated solution of the variational formulation (6) in finite-dimensional subspaces $\mathcal{V}_h \subset \mathcal{V}$, provided that we have homogeneous Dirichlet boundary conditions. Henceforth, we suppose for convenience that the elastic body $\bar{\Omega}$ is a polyhedral domain. In contrast to Ritz method, cf. Section 2.3, we introduce a partition $\mathcal{P}$ of $\bar{\Omega}$ given by a finite number of closed subdomains (hexahedrons) $\bar{\Omega}_e = \Omega_e \cup \partial\Omega_e$, where $\partial\Omega_e$ is the boundary of the element $\bar{\Omega}_e$, such that $\bar{\Omega} = \bigcup_{\bar{\Omega}_e \in \mathcal{P}} \bar{\Omega}_e$. Further, we suppose that the intersection of two elements $\bar{\Omega}_e$ and $\bar{\Omega}_f$ is either a vertex, an edge or a face if $\bar{\Omega}_e \cap \bar{\Omega}_f \neq \emptyset$, and that a family of partitions $\{\mathcal{P}\}$ is locally quasi-uniform, see Ainsworth and Oden [2].

Following standard notation (cf. Ciarlet [11]), we let $Q_k(\bar{\Omega}_e)$ denote the space of polynomials which are of degree $\leq k$ with respect to each variable and has the dimension $\dim Q_k(\bar{\Omega}_e) = (k+1)^3$. To complete the picture, we define by $\Sigma$ the set of degrees of freedom given by the values at the nodes $\mathbf{x}_n \in \bar{\Omega}_e$ with $1 \leq n \leq \dim Q_k(\bar{\Omega}_e)$.

As usual, let $\hat{\Omega} = [-1, 1]^3$ be the reference element. Moreover, we introduce a bijective mapping $\mathbf{F} : \hat{\Omega} \to \bar{\Omega}_e$ such that functions $\mathbf{v} \in [Q_k(\bar{\Omega}_e)]^3$ are given by $\mathbf{v} = \hat{\mathbf{v}} \circ \mathbf{F}^{-1} : \bar{\Omega}_e \to \mathbb{R}^3$ with $\hat{\mathbf{v}} \in [Q_k(\hat{\Omega})]^3$. We assume further that the mapping is in the same space and call the finite element to be of isoparametric type as presented by Irons and Zienkiewicz [16] (see also Johnson [20] and Brenner and Scott [6]).

With this notation at hand, we may now construct the finite element space

$$\mathcal{V}_h = \left\{ \mathbf{v}_h \in \left[C^0(\bar{\Omega})\right]^3 \; ; \quad \mathbf{v}_h|_{\bar{\Omega}_e} = \hat{\mathbf{v}}_h \circ \mathbf{F}^{-1} \; ; \quad \hat{\mathbf{v}}_h \in \left[Q_k(\hat{\Omega})\right]^3 \; ; \quad \forall \bar{\Omega}_e \in \mathcal{P} \right\} \quad (40)$$

using finite elements of the so-called $Q_k$-family. The discrete problem of finding the unique solution of the variational formulation (6) now reads: find an approximated solution $\hat{\mathbf{u}}_h \in \mathcal{V}_h$ such that

$$a(\hat{\mathbf{u}}_h, \mathbf{v}_h) = l(\mathbf{v}_h) \qquad \forall \mathbf{v}_h \in \mathcal{V}_h. \quad (41)$$

Clearly, due to the lemma of Lax–Milgram, a unique solution $\hat{\mathbf{u}}$ of this variational problem exists, since $\mathcal{V}_h \subset \mathcal{V}$.

### 4.2. An a posteriori error estimator based on local Neumann problems

Since it is obviously not possible to evaluate the discretization error $\mathbf{e} = \hat{\mathbf{u}} - \hat{\mathbf{u}}_h$ exactly, we show next how to derive an a posteriori error estimator of the error $\mathbf{e}$ by using an implicit error estimator rather than an explicit estimator. Therefore, we follow Bank and Weiser [4], Ainsworth and Oden [1, 2],

Stein and Ohnimus [40, 41] and Brink and Stein [8] to solve a posteriori a Neumann problem on each element which has the virtue that the global error can be bounded above without any multiplicative constant. Precisely, the local Neumann problem consists of finding a solution $\hat{\mathbf{u}}_e \in \mathcal{V}_e$ on each element $\bar{\Omega}_e$ such that

$$a_e(\hat{\mathbf{u}}_e\,,\,\mathbf{v}) = \int_{\partial\Omega_e} \tilde{\mathbf{t}}_e \cdot \mathbf{v}\,\mathrm{d}S - a_e(\hat{\mathbf{u}}_h\,,\,\mathbf{v}) \qquad \forall \mathbf{v} \in \mathcal{V}_e\,, \tag{42}$$

with the local space $\mathcal{V}_e = \{\mathbf{v}|_{\bar{\Omega}_e}\ ;\ \mathbf{v} \in \mathcal{V}\}$, the equilibrated tractions $\tilde{\mathbf{t}}_e \in [L_2(\partial\Omega_e)]^3$ and the bilinear form $a_e : \mathcal{V}_e \times \mathcal{V}_e \to \mathbb{R}$, restricted to an element, such that

$$a(\mathbf{u},\mathbf{v}) = \sum_{\bar{\Omega}_e \in \mathcal{P}} a_e\left(\mathbf{u}|_{\bar{\Omega}_e}\,,\,\mathbf{v}|_{\bar{\Omega}_e}\right) \qquad \forall \mathbf{u},\mathbf{v} \in \mathcal{V}. \tag{43}$$

Furthermore, we assume that the bilinear form $a_e$ is $\mathcal{V}_e/\mathcal{Z}_e$-elliptic, where

$$\mathcal{Z}_e = \{\mathbf{v} \in \mathcal{V}_e\ ;\quad a_e(\mathbf{w},\mathbf{v}) = 0 \quad \forall \mathbf{w} \in \mathcal{V}_e\} \tag{44}$$

is the kernel of the bilinear form $a_e$, i.e. the space of rigid body motions. Note that for simplicity the body forces $\bar{\mathbf{f}}$ are omitted in this formulation. A necessary condition for the solvability of the local Neumann problem (42) is that the equilibrated tractions $\tilde{\mathbf{t}}_e$ satisfy the Neumann boundary condition $\bar{\mathbf{t}}$ on $\partial\Omega_e \cap \Gamma_N$, the lemma of Cauchy $\tilde{\mathbf{t}}_e = -\tilde{\mathbf{t}}_f$ on $\bar{\Omega}_e \cap \bar{\Omega}_f$ and the *equilibration condition*

$$\int_{\partial\Omega_e} \tilde{\mathbf{t}}_e \cdot \mathbf{v} - a_e(\hat{\mathbf{u}}_h\,,\,\mathbf{v}) = 0 \qquad \forall \mathbf{v} \in \mathcal{Z}_e\,. \tag{45}$$

The basic idea of using equilibrated tractions to obtain an improved solution of finite element approximations was already introduced in the seventies by Bufler and Stein [10] and later by Stein and Ahmad [39]. In conjunction with the a posteriori error analysis, Ladevèze and Leguillon [23] first used equilibrated tractions based on averaged tractions on the element boundaries $\partial\Omega_e$. Stein and Ohnimus [40, 41] used a different access to this method by using energy equivalent nodal forces from the previous finite element solution. Moreover, they obtain the equilibrated tractions $\tilde{\mathbf{t}}_e$ by making use of admissible test functions of the space $\mathcal{V}_{e,h} = \{\mathbf{v}_h|_{\bar{\Omega}}\ ;\ \mathbf{v}_h \in \mathcal{V}_h\}$, rather than $\mathcal{Z}_e$, which guarantees exact equilibrium. Hence, it is possible to derive an anisotropic error estimator. For further details see Ohnimus [27]. Another equilibration method is shown by Ainsworth and Oden [1] using the idea of weighted tractions instead of averaged tractions. Since $\tilde{\mathbf{t}}_e \in [L_2(\partial\Omega_e)]^3$ and since we have no $C^0$-continuity requirements for the equilibrated tractions in tangential direction of element faces, due to the lemma of Cauchy, the equilibrated tractions can be computed on an element patch.

Now that we have at our disposal the solutions $\hat{\mathbf{u}}_e$ of the local Neumann problems (42) on each element, we are able to estimate the global discretization error $\mathbf{e}$ measured in the energy norm by

$$\|\mathbf{e}\|_E \le \left\{ \sum_{\bar{\Omega}_e \in \mathcal{P}} \|\hat{\mathbf{u}}_e\|_E^2 \right\}^{\frac{1}{2}} \tag{46}$$

which was proven by Ainsworth and Oden [1] and Brink and Stein [8]. Moreover, the latter extended the approach also to nonlinear, nearly incompressible problems. It was shown by many others, e.g. Rüter and Stein [38] in the context of anisotropic, nearly incompressible materials at finite strains, Ohnimus et al. [28] in the context of estimating the error locally and Brink and Stephan [9] in the context of finite elements coupled with boundary elements, that implicit error estimators based on the a posteriori solution of local Neumann problems are a powerful tool of estimating the error. One of the reasons for the success of this method is obviously that an upper bound of the error is guaranteed without any multiplicative constant. Furthermore, using test functions of sufficiently high-order yields a sharper bound of the error than obtained by explicit residual error estimators.

## 5. Relations of the generalized Trefftz method to error estimation

### 5.1. An upper bound of the error based on a generalized Trefftz functional

Considering the error estimator (46), the first point to bear in mind is that the upper bound property of the error is only fulfilled for functions $\hat{\mathbf{u}}_e$ in the infinite-dimensional space $\mathcal{V}_e$. Obviously, in a practical computation we exchange the infinite-dimensional space for a finite-dimensional subspace. As a consequence, the guaranteed upper bound property may get lost. In order to get around this problem, we show next how to use a generalized Trefftz functional to derive a guaranteed upper bound of the error $\mathbf{e}$ which is computable in a finite-dimensional subspace. Therefore, let us first introduce an improved solution of the discrete variational problem (41) given by $\tilde{\mathbf{u}}_e = \hat{\mathbf{u}}_e + \hat{\mathbf{u}}_h|_{\bar{\Omega}_e}$ such that the variational form of the Neumann problem (42) may then be formulated as follows

$$a_e(\tilde{\mathbf{u}}_e, \mathbf{v}) = \int_{\partial\Omega_e} \tilde{\mathbf{t}}_e \cdot \mathbf{v} \, \mathrm{d}S \qquad \forall \mathbf{v} \in \mathcal{V}_e . \tag{47}$$

Note that the corresponding strong form of equilibrium

$$-\nabla \cdot \boldsymbol{\sigma}(\tilde{\mathbf{u}}_e) = \mathbf{0} \qquad \text{in } \Omega_e , \tag{48a}$$

$$\boldsymbol{\sigma}(\tilde{\mathbf{u}}_e) \cdot \mathbf{n} = \tilde{\mathbf{t}}_e \qquad \text{on } \partial\Omega_e , \tag{48b}$$

is a slight modification of the strong form (25) with improved displacements $\tilde{\mathbf{u}}_e$ and improved surface tractions $\tilde{\mathbf{t}}_e$. In analogy to Section 3.1 we may additively decompose the solution $\tilde{\mathbf{u}}_e = \breve{\mathbf{u}} + \mathring{\mathbf{u}}$ on each element in two parts with the properties $\boldsymbol{\sigma}(\breve{\mathbf{u}}) \cdot \mathbf{n} = \tilde{\mathbf{t}}_e$ and $\boldsymbol{\sigma}(\mathring{\mathbf{u}}) \cdot \mathbf{n} = \mathbf{0}$ on the element boundary $\partial\Omega_e$. Doing this, we are now able to use the same generalized Trefftz functional $b(\mathbf{u}, \mathbf{w})$ for all $\mathbf{u}, \mathbf{w} \in \mathcal{H}$ as introduced in Section 3.2, restricted to each element $\bar{\Omega}_e \in \mathcal{P}$, such that

$$b(\mathbf{u}, \mathbf{w}) = \sum_{\bar{\Omega}_e \in \mathcal{P}} b_e \left( \mathbf{u}|_{\bar{\Omega}_e} , \mathbf{w}|_{\bar{\Omega}_e} \right) \qquad \forall \mathbf{u}, \mathbf{w} \in \mathcal{H}. \tag{49}$$

It is readily seen in (39) that $b_e$ achieves its maximum for $\mathbf{u}|_{\bar{\Omega}_e} = \mathbf{w}|_{\bar{\Omega}_e} = \mathring{\mathbf{u}}$. In this case, a necessary condition is that the first variation of $b_e$,

$$a_e(\mathring{\mathbf{u}}, \mathbf{v}) - \int_{\Omega_e} \mathbf{f}_{\breve{\mathbf{u}}} \cdot \mathbf{v} \, \mathrm{d}V - \int_{\partial\Omega_e} (\boldsymbol{\sigma}(\mathring{\mathbf{u}}) \cdot \mathbf{n}) \cdot \mathbf{v} \, \mathrm{d}S = 0 \qquad \forall \mathbf{v} \in \mathcal{H}_e , \tag{50}$$

vanishes for arbitrary functions $\mathbf{v} \in \mathcal{H}_e = \{\mathbf{v}|_{\bar{\Omega}_e} \, ; \, \mathbf{v} \in \mathcal{H}\}$. Now, we may insert this result into the generalized Trefftz functional (31) to obtain

$$-\frac{1}{2} b_e(\mathring{\mathbf{u}}, \mathring{\mathbf{u}}) = -\frac{1}{2} a_e(\mathring{\mathbf{u}}, \mathring{\mathbf{u}}) + \frac{1}{2} \int_{\partial\Omega_e} (\boldsymbol{\sigma}(\mathring{\mathbf{u}}) \cdot \mathbf{n}) \cdot \mathring{\mathbf{u}} \, \mathrm{d}S \tag{51}$$

after some algebraic manipulations. Note that (51) holds only for the solution $\mathring{\mathbf{u}}$. Next, we introduce an approximated solution $\mathring{\mathbf{u}}_h \in \mathcal{H}_{e,h} \subset \mathcal{H}_e$ of the variational problem (50) in a finite-dimensional subspace. A straightforward application of the estimate (39) then yields

$$\begin{aligned} b_e(\mathring{\mathbf{u}}_h, \mathring{\mathbf{u}}_h) &= a_e(\mathring{\mathbf{u}}_h, \mathring{\mathbf{u}}_h) - \int_{\partial\Omega_e} (\boldsymbol{\sigma}(\mathring{\mathbf{u}}_h) \cdot \mathbf{n}) \cdot \mathring{\mathbf{u}}_h \, \mathrm{d}S \\ &\geq a_e(\mathring{\mathbf{u}}, \mathring{\mathbf{u}}). \end{aligned} \tag{52}$$

In what follows, it proves convenient to add on both sides of the estimate (52) the bilinear forms $a_e(\breve{\mathbf{u}}, \breve{\mathbf{u}}) + 2a_e(\mathring{\mathbf{u}}, \breve{\mathbf{u}})$ such that this estimate takes the form

$$\begin{aligned} c_e[\mathring{\mathbf{u}}_h, \breve{\mathbf{u}}] &:= a_e(\mathring{\mathbf{u}}_h, \mathring{\mathbf{u}}_h) - a_e(\breve{\mathbf{u}}, \breve{\mathbf{u}}) + 2 \int_{\partial\Omega_e} \tilde{\mathbf{t}}_e \cdot \breve{\mathbf{u}} \, \mathrm{d}S - \int_{\partial\Omega_e} (\boldsymbol{\sigma}(\mathring{\mathbf{u}}_h) \cdot \mathbf{n}) \cdot \mathring{\mathbf{u}}_h \, \mathrm{d}S \\ &\geq a_e(\tilde{\mathbf{u}}_e, \tilde{\mathbf{u}}_e), \end{aligned} \tag{53}$$

where we made use of the relation

$$a_e(\mathring{\mathbf{u}}, \breve{\mathbf{u}}) = \int_{\partial\Omega_e} (\boldsymbol{\sigma}(\mathring{\mathbf{u}}) \cdot \mathbf{n}) \cdot \breve{\mathbf{u}} \, \mathrm{d}S - \int_{\Omega_e} (\nabla \cdot \boldsymbol{\sigma}(\mathring{\mathbf{u}})) \cdot \breve{\mathbf{u}} \, \mathrm{d}V \tag{54a}$$

$$= \int_{\Omega_e} (\nabla \cdot \boldsymbol{\sigma}(\breve{\mathbf{u}})) \cdot \breve{\mathbf{u}} \, \mathrm{d}V \tag{54b}$$

$$= \int_{\partial\Omega_e} \tilde{\mathbf{t}}_e \cdot \breve{\mathbf{u}} \, \mathrm{d}S - a_e(\breve{\mathbf{u}}, \breve{\mathbf{u}}). \tag{54c}$$

Furthermore, we introduced the "form" $c_e$ which is obviously not bilinear as indicated by the square brackets. Since the approximated solution $\mathring{\mathbf{u}}_h$ has to fulfill the governing PDE (26a) in the element domain $\Omega_e$, we may now simply choose $\mathring{\mathbf{u}}_h = -\breve{\mathbf{u}}$ which clearly fulfills the PDE (26a). We thus arrive at

$$c_e[-\breve{\mathbf{u}}, \breve{\mathbf{u}}] = 3 \int_{\partial\Omega_e} \tilde{\mathbf{t}}_e \cdot \breve{\mathbf{u}} \, \mathrm{d}S \geq a_e(\tilde{\mathbf{u}}_e \,, \tilde{\mathbf{u}}_e). \tag{55}$$

Finally, upon summing up the solutions of each local Neumann problem over all elements $\bar{\Omega}_e \in \mathcal{P}$, the error estimator (46) now reads

$$\|\mathbf{e}\|_E \leq \left\{ \sum_{\bar{\Omega}_e \in \mathcal{P}} \|\hat{\mathbf{u}}_e\|_E^2 \right\}^{\frac{1}{2}} \leq \left\{ \sum_{\bar{\Omega}_e \in \mathcal{P}} c_e[-\breve{\mathbf{u}}, \breve{\mathbf{u}}] - a_e(\hat{\mathbf{u}}_h \,, \hat{\mathbf{u}}_h) \right\}^{\frac{1}{2}}. \tag{56}$$

We remark that the computation of the upper bound can be rather cumbersome, since one has to compute a solution $\breve{\mathbf{u}}$ fulfilling the Neumann boundary condition $\boldsymbol{\sigma}(\breve{\mathbf{u}}) \cdot \mathbf{n} = \tilde{\mathbf{t}}_e$ on $\partial\Omega_e$. The main problem which arises is that the equilibrated tractions $\tilde{\mathbf{t}}_e$ lead to unsymmetric stress tensors $\boldsymbol{\sigma}$. To cope with this problem it is possible to use higher order tractions, such that symmetric stress tensors are obtained. However, in this case the $C^0$-continuity condition between neighbouring elements is not fulfilled, and the question remains if the upper bound is still guaranteed.

We show next that $c_e$ is asymptotically exact, i.e.

$$\lim_{\mathring{\mathbf{u}}_h \to \mathring{\mathbf{u}}} c_e[\mathring{\mathbf{u}}_h \,, \breve{\mathbf{u}}] = a(\tilde{\mathbf{u}}_e \,, \tilde{\mathbf{u}}_e) \tag{57}$$

holds. Therefore, we substitute the approximate solution $\mathring{\mathbf{u}}_h$ in (53) by the exact solution $\mathring{\mathbf{u}}$. Hence, we obtain

$$c_e[\mathring{\mathbf{u}}, \breve{\mathbf{u}}] = a_e(\mathring{\mathbf{u}}, \mathring{\mathbf{u}}) - a_e(\breve{\mathbf{u}}, \breve{\mathbf{u}}) + 2 \int_{\partial\Omega_e} \tilde{\mathbf{t}}_e \cdot \breve{\mathbf{u}} \, \mathrm{d}S \tag{58}$$

and, subsequently, with (54)

$$c_e[\mathring{\mathbf{u}}, \breve{\mathbf{u}}] = a_e(\mathring{\mathbf{u}}, \mathring{\mathbf{u}}) + 2a_e(\mathring{\mathbf{u}}, \breve{\mathbf{u}}) + a_e(\breve{\mathbf{u}}, \breve{\mathbf{u}}) \tag{59a}$$

$$= a_e(\mathring{\mathbf{u}} + \breve{\mathbf{u}}, \mathring{\mathbf{u}} + \breve{\mathbf{u}}). \tag{59b}$$

The assertion is thus proved that $c_e$ is asymptotically exact.

## 6. Conclusions

In this paper we dealt with a generalized Trefftz functional to obtain an upper bound of the discretization error of finite element approximations. Therefore we first introduced the well-known minimization principle of the total potential energy of an elastic body. A numerical solution of the minimization problem led directly to Ritz method. Conversely, the linear elasticity problem can be solved by maximizing the complementary energy functional which was constructed by using the

projection theorem. The corresponding numerical solution of the maximization problem thus led to Trefftz method. Moreover, in this context we presented two further numerical methods, namely the method of orthogonal projection and the related hypercircle method, to compute approximated solutions of minimization principles in finite-dimensional subspaces. Next, we introduced the Neumann problem in linear elasticity. It proved convenient to solve the Neumann problem by introducing a generalized Trefftz method, rather than the original Trefftz method, which has the virtue that the solution of the maximization problem is searched amongst all functions satisfying the governing PDE in the domain, but no boundary conditions. Upon introducing the finite element method and an implicit a posteriori error estimator based on equilibrated local Neumann problems, we transferred the upper bound property of the generalized Trefftz functional to the solution of the Neumann problems on each element. We were thus able to compute a guaranteed upper bound of the error in a finite-dimensional subspace without any multiplicative constant. Finally, we showed that this upper bound is asymptotically exact.

It remains to show that numerical examples confirm our presented theoretical results and yield a computable guaranteed upper bound of the discretization error. A computer implementation seems to be very attractive and is scheduled for the future.

## REFERENCES


[1] M. Ainsworth, J.T. Oden. A unified approach to a posteriori error estimation based on element residual methods. *Numer. Math.*, **65**: 23–50, 1993.

[2] M. Ainsworth, J.T. Oden. A posteriori error estimation in finite element analysis. *Comput. Methods Appl. Mech. Engrg.*, **142**: 1–88, 1997.

[3] A.M. Arthurs. *Complementary Variational Principles.* Clarendon Press, Oxford, 1970.

[4] R.E. Bank, A. Weiser. Some a posteriori error estimators for elliptic partial differential equations. *Math. Comp.*, **44**(170): 283–301, 1985.

[5] M.Sh. Birman. Variational methods of solution of boundary problems analogous to the method of Trefftz (in Russian). *Vestnik Leningrad. Univ./ Serija Matematiki, Mechaniki i Astronomii*, **13**: 69–89, 1956.

[6] S.C. Brenner, L.R. Scott. *The Mathematical Theory of Finite Element Methods.* Springer, New York, 1994.

[7] U. Brink, M. Kreienmeyer, K. Peters, E. Stein. Aspects of Trefftz' method in BEM and FEM and their coupling. *Comput. Assist. Mech. Eng. Sci.*, **4**: 327–344, 1997.

[8] U. Brink, E. Stein. A posteriori error estimation in large-strain elasticity using equilibrated local Neumann problems. *Comput. Methods Appl. Mech. Engrg.*, **161**: 77–101, 1998.

[9] U. Brink, E.P. Stephan. Implicit residual error estimators for the coupling of finite elements and boundary elements. *Math. Meth. Appl. Sci.*, **22**: 923–936, 1999.

[10] H. Bufler, E. Stein. Zur Plattenberechnung mittels finiter Elemente. *Ing. Archiv*, **39**: 248–260, 1970.

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

[12] P. Cooperman. An extension of the method of Trefftz for finding local bounds on the solutions of boundary value problems, and on their derivatives. *Quart. Appl. Math.*, **10**: 359–373, 1953.

[13] J.A.T. Freitas, J.P.M. Almeida, E.M.B.R. Pereira. Non-conventional formulations for the finite element method. *Struct. Eng. Mech.*, **4**: 655–678, 1996.

[14] W. Han. A posteriori error analysis for linearization of nonlinear elliptic problems and their discretizations. *Math. Meth. Appl. Sci.*, **17**: 487–508, 1994.

[15] I. Herrera. Trefftz method. In: C.A. Brebbia, ed., *Topics in Boundary Element Research — Basic Principles and Applications*, 225–253. Spinger, Berlin, 1984.

[16] B.M. Irons, O.C. Zienkiewicz. The isoparametric finite element system — a new concept in finite element analysis. In: *Proc. Conf. Recent Advances in Stress Analysis.* Royal Aero Soc., 1968.

[17] W.G. Jin, Y.K. Cheung, O.C. Zienkiewicz. Application of the Trefftz method in plane elasticity problems. *Int. J. Numer. Methods Engrg.*, **30**: 1147–1161, 1990.

[18] J. Jirousek. Basis for development of large finite elements locally satisfying all field equations. *Comput. Methods Appl. Mech. Engrg.*, **14**: 65–92, 1978.

[19] J. Jirousek, A. Venkatesh. A simple stress error estimator for hybrid-Trefftz $p$-version elements. *Int. J. Numer. Methods Engrg.*, **28**: 211–236, 1989.

[20] C. Johnson. *Numerical Solution of Partial Differential Equations by the Finite Element Method.* Cambridge University Press, Cambridge, 1987.

[21] D.W. Kelly. The self-equilibration of residuals and complementary a posteriori error estimates in the finite element method. *Int. J. Numer. Methods Engrg.*, **20**: 1491–1506, 1984.
[22] E. Kita, N. Kamiya. Trefftz method: An overwiew. *Advance Engineering Software*, **24**: 3–12, 1995.
[23] P. Ladevèze, D. Leguillon. Error estimate procedure in the finite element method and applications. *SIAM J. Numer. Anal.*, **20**(3): 485–509, 1983.
[24] S.G. Mikhlin. *Variational Methods in Mathematical Physics.* Pergamon Press, Oxford, 1964.
[25] J.L. Nowinski. *Applications of Functional Analysis in Engineering.* Plenum Press, New York, 1981.
[26] J.T. Oden, L. Demkowicz, W. Rachowicz, T.A. Westermann. Toward a universal *h-p* adaptive finite element strategy. Part 2, A posteriori error estimation. *Comput. Methods Appl. Mech. Engrg.*, **77**: 113–180, 1989.
[27] S. Ohnimus. General computation of equilibrated boundary tractions in adaptive FEM. Submitted to *Engrg. Comput.*, 1999.
[28] S. Ohnimus, E. Stein, E. Walhorn. Local error estimates of FEM for displacements and stresses in linear elasticity by solving local Neumann problems. *Int. J. Numer. Methods Engrg.*, 2001 (in print).
[29] K. Peters, E. Stein, W. Wagner. A new boundary-type finite element for 2d- and 3d-elastic structures. *Int. J. Numer. Methods Engrg.*, **37**: 1009–1025, 1994.
[30] T.H.H. Pian. Derivation of element stiffness matrices by assumed stress distributions. *AIAA J.*, **2**: 1333–1336, 1964.
[31] W. Prager, J.L. Synge. Approximations in elasticity based on the concept of function space. *Quart. Appl. Math.*, **5**: 241–269, 1947.
[32] Q.H. Qin. Hybrid-Trefftz finite element method for Reissner plates on an elastic foundation. *Comput. Methods Appl. Mech. Engrg.*, **122**: 379–392, 1995.
[33] J.N. Reddy. *Energy and Variational Methods in Applied Mechanics.* John Wiley & Sons, New York, 1984.
[34] S.I. Repin, L.S. Xanthis. A posteriori error estimation for elasto-plastic problems based on duality theory. *Comput. Methods Appl. Mech. Engrg.*, **138**: 317–339, 1996.
[35] S.I. Repin, L.S. Xanthis. A posteriori error estimation for nonlinear variational problems. *C. R. Acad. Sci. Paris Sér. I Math.*, **324**: 1169–1174, 1997.
[36] G. Rieder. Iterationsverfahren und Operatorgleichungen in der Elastizitätstheorie. *Abh. Braunschweig. Wiss. Gesell.*, **14**: 109–343, 1962.
[37] W. Ritz. über eine neue Methode zur Lösung gewisser Variationsprobleme der mathematischen Physik. *J. Reine Angew. Math.*, **135**: 1–61, 1908.
[38] M. Rüter, E. Stein. Analysis, finite element computation and error estimation in transversely isotropic nearly incompressible finite elasticity. *Comput. Methods Appl. Mech. Engrg.*, **190**: 519–541, 2000.
[39] E. Stein, R. Ahmad. An equilibrium method for stress calculation using finite element methods in solid- and structural-mechanics. *Comput. Methods Appl. Mech. Engrg.*, **10**: 175–198, 1977.
[40] E. Stein, S. Ohnimus. Coupled model- and solution-adaptivity in the finite-element method. *Comput. Methods Appl. Mech. Engrg.*, **150**: 327–350, 1997.
[41] E. Stein, S. Ohnimus. Anisotropic discretization- and model-error estimation in solid mechanics by local Neumann problems. *Comput. Methods Appl. Mech. Engrg.*, **176**: 363–385, 1999.
[42] J.L. Synge. The method of the hypercircle in function-space for boundary-value problems. *Proc. Roy. Soc.*, **A 191**: 447–467, 1947.
[43] J.L. Synge. The method of the hypercircle in elasticity when body forces are present. *Quart. Appl. Math.*, **6**: 15–19, 1948.
[44] J.L. Synge. *The Hypercircle in Mathematical Physics — A Method for the Approximate Solution of Boundary Value Problems.* University Press, Cambridge, 1957.
[45] E. Trefftz. Ein Gegenstück zum Ritzschen Verfahren. In: *Proceedings of the 2nd International Congress for Applied Mechanics*, 131–138. Zürich, 1926.
[46] W. Velte. *Direkte Methoden der Variationsrechnung.* B.G. Teubner, Stuttgart, 1976.
[47] W. Velte. Complementary Variational Principles. In: E. Stein, W.L. Wendland, eds., *Finite Element and Boundary Element Techniques from Mathematical and Engineering Point of View*, 1–32. Springer, Wien, 1988.
[48] H. Weyl. The method of orthogonal projection in potential theory. *Duke Math. J.*, **7**: 411–444, 1940.
[49] S. Zaremba. Sur le principe de minimum. *Bull. Int. Acad. Sci. Cracovie, Cl. Sci. Math. Nat.*, **7**: 197–264, 1909.
[50] S. Zaremba. Sur un problème toujours possible comprenant, à titre de cas particuliers, le problème de Dirichlet et celui de Neumann. *J. Math. Pures et Appl.*, **6**: 127–163, 1927.
[51] A.P. Zieliński, O.C. Zienkiewicz. Generalized finite element analysis with T-complete boundary solution functions. *Int. J. Numer. Methods Engrg.*, **21**: 509–528, 1985.