Aplikace matematiky

Ivan Hlaváček

Dual finite element analysis for unilateral boundary value problems

Aplikace matematiky, Vol. 22 (1977), No. 1, 14–51

Persistent URL: `http://dml.cz/dmlcz/103675`

DUAL FINITE ELEMENT ANALYSIS FOR UNILATERAL BOUNDARY VALUE PROBLEMS

IVAN HLAVÁČEK

(Received February 25, 1976)

INTRODUCTION

In the analysis of the classical-bilateral-boundary value problems the dual variational approach can be used, yielding (i) approximate values of the solution and of its cogradient, (ii) a posteriori error bounds and two-sided estimates of the energy (cf. [1], [2], [3] and the literature therein).

It is the aim of the present paper to extend the dual approach to scalar second order elliptic equations with unilateral boundary conditions of the Signorini's type (cf. e.g. [4]). We consider coercive cases only to avoid the use of normalizing subspaces.

The variational formulations in terms of the scalar variable and of the corresponding cogradient vector variable, respectively (minimum of potential and of complementary energy) are established and justified on the basis of the duality theory and the saddle point theorem (cf. [5]).

Restricting the admissible functions to finite elements with piecewise linear polynomials on triangulations of the given domain, we are led to minimization problems over finite-dimensional convex sets. Some procedures of the quadratic programming are proposed for the numerical solution of these problems. Making use of one-sided approximations (cf. [7]) on the boundary by piecewise linear spline functions, we prove asymptotic orders of convergence, provided the solution is sufficiently smooth.

Finally, a posteriori error estimates and two-sided estimates of the energy are given.

1. THE SIGNORINI PROBLEM AND THE DUAL VARIATIONAL FORMULATIONS FOR SECOND ORDER ELLIPTIC EQUATIONS

With regard to the dual analysis we shall distinguish two classes of elliptic equations, namely (i) those with a strictly positive "absolute" term and (ii) those without an absolute term.

For the class (i) we choose the following simple model problem[1])

Problem $\mathscr{P}_1$:

$$-\Delta u + u = f \quad \text{on} \quad \Omega \subset R^n, \tag{1.1}$$

$$u \geqq 0, \quad \frac{\partial u}{\partial \nu} \geqq 0, \quad u \frac{\partial u}{\partial \nu} = 0 \quad \text{on} \quad \partial\Omega = \Gamma, \tag{1.2}$$

where $\partial u/\partial \nu$ denotes the normal derivative with respect to the outward normal ν.

In the class (ii) we restrict ourselves to coercive cases and therefore consider the following model problem.

Problem $\mathscr{P}_2$:

$$-\Delta u = f \qquad \text{on} \quad \Omega \subset R^n, \tag{1.3}$$

$$u = 0 \quad \text{on} \quad \Gamma_u \subset \Gamma, \tag{1.4}$$

$$u \geqq 0, \quad \partial u/\partial \nu \geqq 0, \quad u\, \partial u/\partial \nu = 0 \quad \text{on} \quad \Gamma_a = \Gamma \dot{-} \Gamma_u, \tag{1.5}$$

where Γ_u and Γ_a are nonempty sets which contain subsets open in Γ.

Let Ω be a bounded domain with Lipschitz boundary (see [6] for the definition of such domain). We shall use the Sobolev spaces $H^k(\Omega)$ of functions, the derivatives of which up to the order k exist (in the sense of distributions) and are square-integrable in Ω. The usual norm of u in $H^k(\Omega)$ will be denoted by $\|u\|_k$, $H^0(\Omega) = L_2(\Omega)$,

$$(f, g)_0 = \int_\Omega fg \, d\mathbf{x} .$$

Let the right hand sides of (1.1), (1.3) $f \in L_2(\Omega)$.

It is well-known that the Problem $\mathscr{P}_1$ can be recast as follows. Introduce the convex set

$$K_1 = \{v \mid v \in H^1(\Omega), \quad \gamma v \geqq 0\},$$

where γv denotes the trace of v on the boundary Γ, and the functional (potential energy)

$$\mathscr{L}_1(v) = \tfrac{1}{2}\|v\|_1^2 - (f, v)_0 .$$

---

[1]) All the results could be easily extended to equations

$$-\sum_{i,j=1}^{n} \frac{\partial}{\partial x_i}\left(a_{ij}(\mathbf{x}) \frac{\partial u}{\partial x_j}\right) + a_0(\mathbf{x})\, u = f,$$

where a_{ij}, a_0 are bounded measurable functions, if the matrix $[a_{ij}(\mathbf{x})]$ is symmetric, uniformly positive definite on Ω and $a_0(x) \geqq c > 0$ for the class (i) or $a_0(\mathbf{x}) = 0$ for the class (ii).

Then the problem to find $u \in K_1$ such that

(1.6) $$\mathscr{L}_1(u) \leqq \mathscr{L}_1(v) \; \forall v \in K_1$$

represents a variational formulation of the Problem $\mathscr{P}_1$.

For the Problem $\mathscr{P}_2$ we introduce the convex set

$$K_2 = \{v \mid v \in H^1(\Omega),\ \gamma v|_{\Gamma_u} = 0,\ \gamma v|_{\Gamma_a} \geqq 0\}$$

and the functional (potential energy)

$$\mathscr{L}_2(v) = \tfrac{1}{2}|v|_1^2 - (f, v)_0 \,,$$

where

$$|v|_1^2 = \int_\Omega |\text{grad } v|^2 \, d\mathbf{x} \,.$$

Then the problem to find $u \in K_2$ such that

(1.7) $$\mathscr{L}_2(u) \leqq \mathscr{L}_2(v) \; \forall v \in K_2$$

is a variational formulation of the Problem $\mathscr{P}_2$.

The minimization problems (1.6) and (1.7) will be called *primary.*

Both the primary problems can be reformulated in terms of the gradient-vector (cf. [1], [2], [3]). To this end, let us introduce the set

(1.8) $$Q = \{\lambda \mid \lambda \in [L_2(\Omega)]^n,\ \text{div } \lambda \in L_2(\Omega)\} \,,$$

where the differential operator

$$\text{div } \lambda = \sum_{i=1}^{n} \partial\lambda_i / \partial x_i$$

is defined in the distribution sense only:

(1.9) $$\int_\Omega \lambda \,.\, \text{grad } \varphi \, d\mathbf{x} = -\int_\Omega \varphi \, \text{div } \lambda \, d\mathbf{x} \; \forall \varphi \in C_0^\infty(\Omega) \,.$$

For $\lambda \in Q$, the functional $\lambda \,.\, \nu \in H^{-1/2}(\Gamma)$ can be defined by means of the relation

(1.10) $$\langle \lambda \,.\, \nu, \gamma v \rangle = \int_\Omega (\lambda \,.\, \text{grad } v + v \, \text{div } \lambda) \, d\mathbf{x} \; \forall v \in H^1(\Omega).$$ [1])

[1]) Note that any function $w \in H^{1/2}(\Gamma)$ can be identified with a trace γv of a function $v \in H^1(\Omega)$, the mapping $H^{1/2}(\Gamma) \to H^1(\Omega)$ being linear and continuous (cf. [6]).

We write $s|_\Gamma \geqq 0$ for an $s \in H^{-1/2}(\Gamma)$ if

$$\langle s, \gamma v \rangle \geqq 0 \quad \forall v \in K_1 .$$

Next introduce the set

$$\mathscr{U}_1 = \{\lambda \mid \lambda \in [L_2(\Omega)]^{n+1},\ \lambda = [\lambda', \lambda_{n+1}],\ \lambda' \in Q,$$
$$\lambda_{n+1} = f + \operatorname{div} \lambda',\ \lambda' . \nu|_\Gamma \geqq 0\}$$

and the functional (complementary energy)

$$\mathscr{S}_1(\lambda) = \tfrac{1}{2} \sum_{i=1}^{n+1} \|\lambda_i\|_0^2 .$$

The problem to find $\lambda^0 \in \mathscr{U}_1$ such that

$$\mathscr{S}_1(\lambda^0) \leqq \mathscr{S}_1(\lambda)\ \forall \lambda \in \mathscr{U}_1 \tag{1.11}$$

will be called *dual to the primary problem* (1.6).

For the problem $\mathscr{P}_2$, let us introduce the set

$$\mathscr{U}_2 = \{\lambda \mid \lambda \in Q,\ \operatorname{div} \lambda + f = 0,\ \lambda . \nu|_{\Gamma_a} \geqq 0\} ,$$

where $s|_{\Gamma_a} \geqq 0$ is defined as follows

$$\langle s, \gamma v \rangle \geqq 0 , \quad \forall v \in K_2 ,$$

and the functional (complementary energy)

$$\mathscr{S}_2(\lambda) = \tfrac{1}{2} \sum_{i=1}^{n} \|\lambda_i\|_0^2 .$$

The problem to find $\lambda^0 \in \mathscr{U}_2$ such that

$$\mathscr{S}_2(\lambda^0) \leqq \mathscr{S}_2(\lambda) \quad \forall \lambda \in \mathscr{U}_2 \tag{1.12}$$

will be called *dual to the primary problem* (1.7).

It is easy to prove that both the primary and the dual problems possess unique solutions. Moreover, there exists an interpretation of the solutions of the dual problems in terms of the solutions to the primary problems.

Theorem 1.1. 1. *Let u be the solution to the primary problem* (1.6) *and λ^0 of the dual problem* (1.11). *Then*

$$\lambda_i^0 = \partial u / \partial x_i , \quad i = 1, \ldots, n , \quad \lambda_{n+1}^0 = u .$$

2. *Let u be the solution to the primary problem* (1.7) *and* λ^0 *to the dual problem* 1.12. *Then*

$$\lambda^0 = \operatorname{grad} u .$$

Let us prove part 1 of the theorem. We shall need the following

Lemma 1.1. (*Saddle point*). *There exists* $w \in H^{1/2}(\Gamma)$, $w \geqq 0$ *such that*

$$\mathscr{S}_1(\lambda^0) - \langle \lambda^0 . v, \mu \rangle \leqq \mathscr{S}_1(\lambda^0) - \langle \lambda^0 . v, w \rangle \leqq \mathscr{S}_1(\lambda) - \langle \lambda . v, w \rangle \quad (1.13)$$
$$\forall \mu \in H^{1/2}(\Gamma), \quad \mu \geqq 0, \quad \forall \lambda \in Q_f ,$$

where

$$Q_f = \{\lambda \mid \lambda \in [L_2(\Omega)]^{n+1} \ \lambda = [\lambda', \lambda_{n+1}], \ \lambda' \in Q, \ \lambda_{n+1} = f + \operatorname{div} \lambda'\} .$$

The proof is based on the following Corollary of the Hahn-Banach theorem.

Let V be a normed vector-space, S and T two convex subsets of V such that S contains at least one interior point and T does not contain any interior point of S.

Then a linear bounded functional $F \in V'$, $F \neq 0$ and a number $\alpha \in R^1$ exist such that

$$\langle F, s \rangle \geqq \alpha \geqq \langle F, t \rangle \quad \forall s \in S \quad \forall t \in T . \quad (1.14)$$

Let us choose $V = R^1 \times H^{-1/2}(\Gamma)$. Define S as the set of all pairs $\{\mathscr{S}_1(\lambda) - \mathscr{S}_1(\lambda^0) + s_0; -\lambda . v + s\}$, where $\lambda \in Q_f$, $s_0 \in R^1$, $s_0 \geqq 0$, $s \in H^{-1/2}(\Gamma)$, $s|_\Gamma \geqq 0$. Let T be the set of pairs $\{-t_0. -t\}$, where $t_0 \in R^1$, $t_0 > 0$, $t \in H^{-1/2}(\Gamma)$, $t|_\Gamma \geqq 0$.

S and T are disjoint. In fact, assuming that

$$\mathscr{S}_1(\lambda) - \mathscr{S}_1(\lambda^0) + s_0 = -t_0 , \quad -\lambda v + s = -t ,$$

we deduce

$$\mathscr{S}_1(\lambda) < \mathscr{S}_1(\lambda^0) , \quad \lambda . v|_\Gamma \geqq 0 , \quad \lambda \in Q_f ,$$

consequently $\lambda \in \mathscr{U}_1$ and we obtain a contradiction with (1.11). Obviously, S and T are convex. In order to prove that S contains an interior point, let us consider the ball

$$B_\varepsilon = \{\psi \in H^{-1/2}(\Gamma) , \ \|\psi\|_{H^{-1/2}(\Gamma)} < \varepsilon\}$$

and the Neumann problems

$$-\Delta u + u = f \quad \text{in} \quad \Omega , \quad \partial u / \partial v = - \psi \quad \text{on} \quad \partial \Omega , \quad (1.15)$$

where $\psi \in B_\varepsilon$. Denote u_ψ the solution of (1.15) and

$$\lambda_i(u_\psi) = \partial u_\psi / \partial x_i , \quad i = 1, \ldots, n , \quad \lambda_{n+1}(u) = u_\psi . \quad (1.15')$$

Then $\lambda(u_\psi) \in Q_f$, $-\lambda(u_\psi) \, . \, v = \psi$. Let u_0 be the solution of (1.15) for $\psi = 0$ and denote $\lambda(u_0)$ the corresponding vector. Then the point $\{\mathscr{S}_1(\lambda(u_0)) - \mathscr{S}_1(\lambda^0) + \varepsilon, 0\}$ is an interior point of S, because

$$(\mathscr{S}_1(\lambda(u_0)) - \mathscr{S}_1(\lambda^0), \mathscr{S}_1(\lambda(u_0)) - \mathscr{S}_1(\lambda^0) + 2\varepsilon) \times B_\varepsilon \subset S \, .$$

In fact, the above interval can be obtained for $\lambda = \lambda(u_0)$, $0 < s_0 < 2\varepsilon$ and the ball B_ε for $s = 0$ and $\lambda = \lambda(u_\psi)$, $\psi \in B_\varepsilon$.

Consequently, there exist $\alpha_0 \in R^1$, $\varphi \in H^{1/2}(\Gamma)$ and $\alpha \in R^1$ such that

$$|\alpha| + \|\varphi\|_{H^{1/2}(\Gamma)} > 0 \, , \tag{1.16}$$

$$\alpha_0(\mathscr{S}_1(\lambda) - \mathscr{S}_1(\lambda^0) + s_0) + \langle -\lambda \, . \, v + s, \varphi \rangle \geqq \alpha \geqq -\alpha_0 t_0 + \langle -t, \varphi \rangle$$

$$\forall \lambda \in Q_f \, , \quad s_0 \in R^1 \, , \quad s_0 \geqq 0 \, , \quad s \in H^{-1/2}(\Gamma) \, , \quad s|_\Gamma \geqq 0 \, , \quad t_0 \in R^1 \, ,$$

$$t_0 > 0 \, , \quad t \in H^{-1/2}(\Gamma) \, , \quad t|_\Gamma \geqq 0 \, .$$

Hence $\alpha_0 \geqq 0$, $\varphi \geqq 0$ and

$$\alpha_0(\mathscr{S}_1(\lambda) - \mathscr{S}_1(\lambda^0)) - \langle \lambda \, . \, v, \varphi \rangle \geqq 0 \quad \forall \lambda \in Q_f \tag{1.17}$$

follows.

Suppose that $\alpha_0 = 0$. Then

$$\langle \lambda \, . \, v, \varphi \rangle \leqq 0 \quad \forall \lambda \in Q_f \, . \tag{1.18}$$

But we may write

$$\lambda \in Q_f \Rightarrow \lambda = q^0 + q^f \, , \quad q^f \in Q_f \, , \quad q^0 \in Q_0 \, ,$$

where q^f is a fixed chosen vector. Consequently,

$$\langle q^0 \, . \, v, \varphi \rangle \leqq - \langle q^f \, . \, v, \varphi \rangle = \text{const} \quad \forall q^0 \in Q_0 \, ,$$

where Q_0 (i.e. Q_f for $f = 0$) is a linear subspace of $[L_2(\Omega)]^{n+1}$. The operator β : $: \chi \to \chi \, . \, v$ maps Q_0 onto $H^{-1/2}(\Gamma)$, therefore (1.18) yields

$$\langle F, \varphi \rangle \leqq \text{const} \quad \forall F \in H^{-1/2}(\Gamma) \, ,$$

which implies $\varphi = 0$. Thus we arrive at a contradiction with (1.16) and α_0 must be positive.

Denoting $\varphi/\alpha_0 = w$, from (1.17) it follows

$$\mathscr{S}_1(\lambda^0) \leqq \mathscr{S}_1(\lambda) - \langle \lambda \, . \, v, w \rangle \quad \forall \lambda \in Q_f \, , \tag{1.19}$$

where $w \in H^{1/2}(\Gamma)$, $w \geqq 0$.

Inserting $\lambda = \lambda^0$ and using the definition of $\mathscr{U}_1$, we obtain

$$\langle \lambda^0 \, . \, v, w \rangle = 0 \, . \tag{1.20}$$

Hence

$$\langle \lambda^0 \, . \, v, \mu \rangle \geqq 0 = \langle \lambda^0 . \, v, w \rangle \quad \forall \mu \in H^{1/2}(\Gamma) \, , \quad \mu \geqq 0 \, . \tag{1.21}$$

The assertion (1.13) follows from (1.19) and (1.21). Q.E.D.

The proof of Part 1 of Theorem 1.1. Let us define

$$H = [L_2(\Omega)]^{n+1} \, , \quad (\lambda, \mu)_H = \sum_{i=1}^{n+1} (\lambda_i, \mu_i)_0 \, ,$$

$$H_1 = \{\lambda \in H \mid \exists v \in H_0^1(\Omega), \lambda = \lambda(v) \text{ (see } (1.15'))\}^1)$$

$$H_2 = \{\lambda \in H \mid \int_\Omega (\lambda' . \operatorname{grad} v + v\lambda_{n+1}) \, dx = 0 \;\; \forall v \in H_0^1(\Omega)\} \, .$$

Let $\tilde{u}$ be the solution of the Dirichlet problem

$$-\Delta \tilde{u} + \tilde{u} = f \quad \text{in} \quad \Omega \, , \quad \gamma \tilde{u} = w \quad \text{on} \quad \Gamma \, , \tag{1.22}$$

where $w \in H^{1/2}(\Gamma)$, $w \geqq 0$ is the function from Lemma 1.1. There exists $u_0 \in H_0^1(\Omega)$ such that $\gamma u_0 = w$ on Γ and we may write $\tilde{u} = u_0 + v_0$, $v_0 \in H_0^1(\Omega)$.

Let $\lambda \in Q_f$. Then $\lambda - \lambda(\tilde{u}) \in H_2$, because $\lambda(\tilde{u}) \in Q_f$. H_1 and H_2 are orthogonal subsets of H, as follows from their definition.

For $\lambda \in Q_f$ we may write

$$\mathscr{J}(\lambda) = \|\lambda - \lambda(u_0)\|_H^2 = \|\lambda - \lambda(\tilde{u}) + \lambda(v_0)\|_H^2 = \|\lambda - \lambda(\tilde{u})\|_H^2 + \|\lambda(v_0)\|_H^2 \, .$$

Hence $\mathscr{J}(q) \leqq \mathscr{J}(\lambda)$ for any $\lambda \in Q_f$ if and only if $q = \lambda(\tilde{u})$. Using the definitions of Q_f and (1.10) we obtain

$$(\lambda, \lambda(u_0))_H = \int_\Omega (\lambda' \, . \operatorname{grad} u_0 + u_0(\operatorname{div} \lambda' + f)] \, dx = \langle \lambda \, . \, v, w \rangle + (f, u_0)_0 \, ,$$

$$\tfrac{1}{2}\mathscr{J}(\lambda) = \mathscr{S}_1(\lambda) - \langle \lambda \, . \, v, w \rangle - (f, u_0)_0 + \tfrac{1}{2}\|\lambda(u_0)\|_H^2 \, .$$

Consequently, the right-hand inequality of (1.13), i.e.

$$\mathscr{S}_1(\lambda^0) - \langle \lambda^0 \, . \, v, w \rangle \leqq \mathscr{S}_1(\lambda) - \langle \lambda \, . \, v, w \rangle \;\; \forall \lambda \in Q_f \, ,$$

holds if and only if $\lambda^0 = q = \lambda(\tilde{u})$.

---

1) $H_0^1(\Omega)$ is the closure of $C_0^\infty(\Omega)$ in $H^1(\Omega)$.

Moreover, we have $\gamma\tilde{u} = w \geqq 0$ on Γ,

$$\lambda^0 . \nu|_\Gamma = \partial\tilde{u}/\partial\nu|_\Gamma \geqq 0 , \quad \partial\tilde{u}/\partial\nu \in H^{-1/2}(\Gamma) , \tag{1.23}$$

$$\langle \partial\tilde{u}/\partial\nu, \gamma\tilde{u} \rangle = \langle \lambda^0 . \nu, w \rangle = 0 \tag{1.24}$$

according to (1.20). Hence $\tilde{u}$ is a solution of the Problem $\mathscr{P}_1$ if the Signorini's conditions (1.2) are taken in the functional sense. Finally, we show that $\tilde{u}$ solves the variational problem (1.6) and from the uniqueness if follows that $u = \tilde{u}$.

In fact, since K_1 is a closed convex cone with the vertex $\{0\}$, u is the solution of (1.6) if and only if

$$(u, v)_1 \geqq (f, v)_0 , \quad \forall v \in K_1 , \tag{1.25}$$

$$(u, u)_1 = (f, u)_0 . \tag{1.26}$$

Let us verify (1.25) for $\tilde{u}$. From (1.22) and (1.10) it follows

$$0 \leqq \langle \partial\tilde{u}/\partial\nu, \gamma v \rangle = \int_\Omega (\text{grad } \tilde{u} . \text{grad } v + v \text{ divgrad } \tilde{u}) \, \mathrm{d}\mathbf{x} , \quad \forall v \in K_1 .$$

On the other hand, (1.22) yields that

$$\text{div grad } \tilde{u} = \tilde{u} - f .$$

Consequently, we have for any $v \in K_1$

$$0 \leqq \int_\Omega (\text{grad } \tilde{u} . \text{grad } v + \tilde{u}v) \, \mathrm{d}\mathbf{x} - \int_\Omega fv \, \mathrm{d}\mathbf{x} = (\tilde{u}, v)_1 - (f, v)_0 .$$

Similarly, from (1.24) we deduce

$$0 = \langle \partial\tilde{u}/\partial\nu, \gamma\tilde{u} \rangle = (\tilde{u}, \tilde{u})_1 - (f, \tilde{u})_0 . \qquad \text{Q.E.D.}$$

The proof of Part 2 of the Theorem 1.1 is analogous.

2. FINITE ELEMENT APPROXIMATIONS TO THE PRIMARY PROBLEMS

To propose a consistent dual finite element analysis, we shall consider straight triangular elements only and therefore study problems on polygonal domains. In fact, there would be difficulties with curved equilibrium elements for solving dual problems of class (ii). For class (i), it seems that the application of curved finite elements in the analysis is possible, but the construction of one-sided approximations (see Section 3) becomes then more complicated.

For simplicity, we restrict ourselves to plane polygonal domains, which leads, on the other hand, to some excessive regularity assumptions imposed upon the solution. We believe, however, that the results on the rate of convergence could hardly be reached without such regularity requirements.

Thus let $\Omega \subset R^2$ be a polygonal bounded domain. We carve it into triangles T generating triangulation $\mathscr{T}_h$. Denote h the maximal side of all triangles in $\mathscr{T}_h$. Let V_h be the space of continuous piecewise linear polynomials on the triangulation $\mathscr{T}_h$.

We say that the family of triangulations $\{\mathscr{T}_h\}$ with $0 < h \leqq 1$ is α-β-regular, if there exist positive α and β, independent of h and such that (i) the minimal angle of all triangles in $\mathscr{T}_h$ is not less than α for any h and (ii) the ratio between any two sides of $\mathscr{T}_h$ is less than β.

Define

$$K_{1h} = \{v \mid v \in V_h, v|_{\Gamma} \geqq 0\} = V_h \cap K_1 ,$$
$$K_{2h} = \{v \mid v \in V_h, v|_{\Gamma_u} = 0,\ v|_{\Gamma_a} \geqq 0\} = V_h \cap K_2 .$$

We say that $u_h \in K_{ih}$ $(i = 1, 2)$ is a *finite element approximation* to the primary problem $\mathscr{P}_i$, if

$$\mathscr{L}_i(u_h) \leqq \mathscr{L}_i(v) , \quad \forall v \in K_{ih} . \tag{2.1}$$

Since K_{ih} are closed convex subsets of $H^1(\Omega)$, it is easy to see that there exist unique solutions of the problems (2.1). To find them we can apply e.g. the *algorithm of Gauss-Seidel with constraints*, as follows.

Any $v \in V_h$ can be written in the form

$$v(x) = \sum_{j=1}^{M} v_j \varphi_j(x) ,$$

where φ_j represent the basis functions. Then it holds

$$v \in K_{1h} \Leftrightarrow \{v_{j_k} \geqq 0,\ k = 1, \ldots, p,\ p < M\} ,$$

where the indeces j_k correspond with the boundary nodes P_{j_k} of $\mathscr{T}_h$. Similarly, we have

$$v \in K_{2h} \Leftrightarrow \{v_{j_q} = 0 \quad \text{if} \quad P_{j_q} \in \bar{\Gamma}_u,\ v_{j_r} \geqq 0 \quad \text{if} \quad P_{j_r} \in \Gamma_a\} .$$

The functionals $\mathscr{L}_i(v)$ on K_{ih} may be written in the form

$$\mathscr{L}_i(v) = \mathbf{v}^{\mathsf{T}} \mathbf{A} \mathbf{v} - \mathbf{f}^{\mathsf{T}} \mathbf{v} ,$$

where

$$\mathbf{v} = (v_1, \ldots, v_M)^{\mathsf{T}} , \quad f_j = (f, \varphi_j)_0$$

$$A_{ij} = (\varphi_i, \varphi_j)_1 \quad \text{or} \quad A_{ij} = \int_{\Omega} \operatorname{grad} \varphi_i \,.\, \operatorname{grad} \varphi_j \, \mathrm{d}x ,$$

respectively, $(i, j = 1, \ldots, M)$.

We choose an initial vector $v^0 \in K_{ih}$. Then vectors $\mathbf{v}^{1,j}, \mathbf{v}^{2,j}, \ldots$ $(j = 1, 2, \ldots, M)$ are calculated step by step, where

$$\mathbf{v}^{m,j} = (v_1^m, \ldots, v_j^m, v_{j+1}^{m-1}, \ldots, v_M^{m-1}), \quad m = 1, 2, \ldots.$$

$$v_{j_k}^m = \mathrm{Proj}_{<0,\infty)} \tilde{v}_{j_k}^m \quad \text{for} \quad k = 1, \ldots, p \quad \text{and} \quad v_j^m = \tilde{v}_j^m \quad \text{otherwise, or}$$

$$v_{j_q}^m = 0 \quad \text{if} \quad P_{j_q} \in \bar{\Gamma}_u,$$

$$v_{j_r}^m = \mathrm{Proj}_{<0,\infty)} \tilde{v}_{j_r}^m \quad \text{if} \quad P_{j_r} \in \Gamma_a,$$

$$\tilde{v}_j^m = (-\sum_{i<j} A_{ji} v_i^m - \sum_{i>j} A_{ji} v_i^{m-1} + f_j)/A_{jj}, \quad (j = 1, \ldots, M),$$

$$\mathrm{Proj}_{<0,\infty)} w = \max\{0, w\}.$$

It is well known (see e.g. [5] Chapt. 4 § 1.4) that

$$\lim_{m \to \infty} \|\mathbf{v}^m - \omega\|_{R^M} = 0, \tag{2.2}$$

where

$$u_h = \sum_{j=1}^{M} \omega_j \varphi_j.$$

Next let us estimate the distance between the solutions u and u_h of the problems (1.6) or (1.7) and (2.1). To this end, we shall need the following

Lemma 2.1. (cf. [7]). *Let $\mathscr{J}(v)$ be the functional defined on a closed convex subset M of a Banach reflexive space B. Suppose that $\mathscr{J}$ is twice differentiable in B (in the sense of Gâteaux) and the second differential is positive definite and continuous, i.e., such positive constants α_0 and c exist that*

$$\alpha_0 \|z\|^2 \leqq \mathscr{J}''(u; z, z) \leqq c\|z\|^2, \quad \forall u \in M, \quad \forall z \in B.$$

Let $M_h \subset M$ be a closed convex set. Denote the minimizing elements of $\mathscr{J}(v)$ over M and M_h by u and u_h, respectively. Suppose that there exists $w_h \in M_h$ such that $2u - w_h \in M$. Then it holds

$$\|u - u_h\| \leqq \left(\frac{c}{\alpha_0}\right)^{1/2} \|u - w_h\|. \tag{2.3}$$

Proof. From the Taylor's theorem it follows that such ϑ exists that $0 < \vartheta < 1$ and

$$\mathscr{J}(u_h) = \mathscr{J}(u) + \mathscr{J}'(u, u_h - u) + \mathscr{J}''(u + \vartheta(u_h - u); u_h - u, u_h - u) \geqq \tag{2.4}$$

$$\geqq \mathscr{J}(u) + \alpha_0 \|u_h - u\|^2,$$

because

$$\mathscr{J}'(u, u_h - u) \geqq 0.$$

On the other hand, for any $v \in M_h$ we have

$$\mathscr{J}(v) = \mathscr{J}(u) + \mathscr{J}'(u, v - u) + \mathscr{J}''(u + \vartheta_1(v - u);\ v - u, v - u) \geqq \mathscr{J}(u_h).$$

Inserting $v = w_h$ and $v = 2u - w_h$ into the condition $\mathscr{J}'(u, v - u) \geqq 0$, we obtain

$$\mathscr{J}'(u, w_h - u) = 0$$

and consequently,

$$(2.5)\qquad \mathscr{J}(u_h) \leqq \mathscr{J}(w_h) = \mathscr{J}(u) + \mathscr{J}''(u + \vartheta_2(w_h - u);\ w_h - u, w_h - u) \leqq$$
$$\leqq \mathscr{J}(u) + c\|w_h - u\|^2 .$$

Finally, (2.4) and (2.5) result in (2.3). Q.E.D.

Applying Lemma 2.1 to problems (1.6), (1.7) and (2.1), we may set $\mathscr{J} = \mathscr{L}_i$, $M = K_i$, $M_h = K_{ih}$ $(i = 1, 2)$, $B = H^1(\Omega)$ for (1.6) and $B = \{v \in H^1(\Omega),\ \gamma v|_{\Gamma_u} = 0\}$ for (1.7). Then it is readily seen that $\alpha_0 = c = 1$ can be taken for $\mathscr{L}_1$ and $c = 1$, $\alpha_0 > 0$ for $\mathscr{L}_2$.

Thus if we find $w_h \in K_{ih}$ such that $2u - w_h \in K_i$ $(i = 1, 2)$ and w_h sufficiently close to u, then u_h is of the same order of accuracy as w_h. Fortunately, we can prove the following

Theorem 2.1. *Assume that $u \in H^2(\Omega)$ and $u \in H^2(\Gamma_m)$, $m = 1, \ldots, G$, where Γ_m denotes any side of the polygonal boundary Γ.*

Then there exists $w_h \in V_h$ such that

$$(2.6)\qquad 0 \leqq w_h \leqq u \quad \text{on} \quad \Gamma$$

and, if the triangulations are α-β-regular, it achieves the optimal order of approximation:

$$(2.7)\qquad \|u - w_h\|_1 \leqq Ch(\|u\|_2 + \sum_{m=1}^{G} \|u\|_{H_2(\Gamma_m)})$$

with C independent of h and u.

Proof is based on two auxiliary lemmas.

Lemma 2.2. (*One-sided approximation of u on the boundary*). *Let u satisfy the assumptions of Theorem 2.1. Then there exists a linear spline function $\psi_h \in C(\Gamma)$ with nodes given by the triangulation $\mathscr{T}_h$, such that*

$$(2.8)\qquad 0 \leqq \psi_h \leqq u \quad \text{on} \quad \Gamma ,$$

$$(2.9)\qquad \|u_I - \psi_h\|^2_{C(\Gamma)} \leqq h^3 \sum_{m=1}^{G} |u|^2_{2,\Gamma_m} ,$$

where u_I is the linear interpolate of u on Γ (with the same nodes)

$$\|\varphi\|_{C(\Gamma)} = \max_{s\in\Gamma} |\varphi(s)| ,$$

$$|\varphi|^2_{2,\Gamma_m} = \int_{\Gamma_m} [d^2\varphi/ds^2]^2 ds .$$

Proof. Consider the nodes of $\mathcal{T}_h$ on any closed polygonal part $\partial\Omega_k$[1]) of the boundary Γ and denote the corresponding arc parameters by $0 = s_1 < s_2 < \ldots < s_n$. Let φ_j, $(j = 1, 2, \ldots, n)$ be the basis linear spline functions on $\partial\Omega_k(\varphi_j(s_i) = \delta_{ji})$ and define

$$S_h = \{\boldsymbol{a} \in R_n \mid 0 \leqq \sum_{j=1}^n a_j \varphi_j(s) \leqq u(s) , \quad \forall s \in \partial\Omega_k\} .$$

We say that $\boldsymbol{a}^0 \in S_h$ is the maximal element of S_h if

$$\oint_{\partial\Omega_k} \sum_{j=1}^n a_j^0 \varphi_j \, ds \geqq \oint_{\partial\Omega_k} \sum_{j=1}^n a_j \varphi_j \, ds , \quad \forall \boldsymbol{a} \in S_h .$$

The maximal element of S_h exists. In fact, it is readily seen that S_h is bounded, since

$$0 \leqq a_j \leqq u(s_j) \leqq \|u\|_{C(\Gamma)} \leqq C\|u\|_2 ,$$

and closed in R^n. The integral to be maximized is a continuous function of $\boldsymbol{a}$, consequently, the maximum is attained in the compact set S_h.

Let us denote

$$\psi_h = \sum_{j=1}^n a_j^0 \varphi_j .$$

Then for any $j = 1, \ldots, n$ at least one of the following two conditions is satisfied:

(P 1) $$\psi_h(s_j) = u(s_j) ,$$

(P 2) $$\exists \sigma_j \in \langle s_{j-1}, s_j) \cup (s_j, s_{j+1}\rangle ,$$

(where we define $s_0 = s_n$, $s_{n+1} = s_1$) such that

$$\psi_h(\sigma_j) = u(\sigma_j) , \quad (d\psi_h/ds)(\sigma_j) = (du/ds)(\sigma_j) .$$[2])

[1]) Ω is a multiply connected domain, in general.
[2]) From $u \in H^2(\Gamma_m)$ it follows that $du/ds \in C(\Gamma_m)$.

In fact, let for some j_0 neither (P 1) nor (P 2) hold. Then obviously a positive ε exists such that

$$\psi^{\varepsilon} = \psi_h + \varepsilon\varphi_{j_0} \leqq u\,, \quad \forall s \in \partial\Omega_k\,.$$

But then

$$\oint_{\partial\Omega_k} \psi_h \,\mathrm{d}s < \oint_{\partial\Omega_k} \psi^{\varepsilon} \,\mathrm{d}s\,,$$

which contradicts the fact that $\boldsymbol{a}^0$ is the maximal element. In case of (P 2) we may write

$$u(s_j) - \psi_h(s_j) = \int_{\sigma_j}^{s_j} \frac{\mathrm{d}^2 u}{\mathrm{d}s^2}(z)\,(s_j - z)\,\mathrm{d}z\,,$$

$$|u(s_j) - \psi_h(s_j)|^2 \leqq h^3 \oint_{\partial\Omega_k} \left(\frac{\mathrm{d}^2 u}{\mathrm{d}s^2}\right)^2 \mathrm{d}z\,,$$

and (2.9) follows. Q.E.D.

Lemma 2.3. *Let $\varphi \in C(\Gamma)$ be a linear spline-function with the nodes determined by the α-β-regular triangulation $\mathscr{T}_h$.*

Then there exists $v_h \in V_h$ such that $v_h = \varphi$ on Γ and

$$\|v_h\|_1 \leqq Ch^{-1/2}\|\varphi\|_{C(\Gamma)}\,. \tag{2.10}$$

Proof. Denote $T \in \mathscr{T}_h$ the closed triangles and let Ω_h be the union of all $T \in \mathscr{T}_h$ such that $T \cap \Gamma \neq \varnothing$. Thus Ω_h is a "boundary strip" of Ω. Let $v_h \in V_h$, $v_h = 0$ in $\Omega - \Omega_h$ and $v_h = \varphi$ on Γ. Obviously, we have

$$\max_{x \in \Omega} |v_h(x)| = \|\varphi\|_{C(\Gamma)}\,.$$

In any triangle $T \in \Omega_h$ with the sides $a \leqq b \leqq c$, it holds

$$|\partial v_h/\partial x_k| = |n_k/n_3|\,, \quad (k = 1, 2)\,,$$

where $\boldsymbol{n}$ is the vector normal to the plane graph of $x_3 = v_h(x_1, x_2)$. Since

$$|n_3| \geqq a^2 \sin\alpha\,, \quad |n_k| \leqq 2c \max_{x \in T} |v_h| \leqq 2c\|\varphi\|_{C(\Gamma)}\,,$$

we obtain, making use of the α-β-regularity of $\mathscr{T}_h$,

$$|\partial v_h/\partial x_k| \leqq 2ca^{-2}(\sin\alpha)^{-1}\,\|\varphi\|_{C(\Gamma)} \leqq Ch^{-1}\|\varphi\|_{C(\Gamma)}\,, \quad \forall T \in \mathscr{T}_h\,,$$

where C is independent of h, φ and T, because

$$ca^{-2} \leqq c^{-1}(\sin\alpha)^{-2} \leqq (\min_{T \in \mathscr{T}_h} c)^{-1}(\sin\alpha)^{-2} \leqq \beta h^{-1}(\sin\alpha)^{-2}\,.$$

Finally, we have (cf. also "inverse inequalities" for V_h spaces)

$$\|v_h\|_1^2 = \int_{\Omega} \left(v_h^2 + \sum_{k=1}^{2} \left(\frac{\partial v_h}{\partial x_k} \right)^2 \right) dx \leqq C \text{ mes } \Omega_h \|\varphi\|_{C(\Gamma)}^2 (1 + h^{-2}) .$$

Since mes $\Omega_h < Ch$, (2.10) follows. Q.E.D.

Proof of Theorem 2.1. Let ψ_h be the one-side approximation of u from Lemma 2.2. Introducing

(2.11) $$\varphi = u_I - \psi_h ,$$

we construct the function $v_h \in V_h$ according to Lemma 2.3. Let u_I be the linear interpolate of u on the triangulation $\mathscr{T}_h$. Then the function

$$w_h = u_I - v_h \in V_h$$

satisfies (2.6) and (2.7). In fact, on Γ it holds

$$w_h = u_I - \varphi = \psi_h$$

and (2.8) implies (2.6).

Furthermore, it is well known that

(2.12) $$\|u - u_I\|_1 \leqq Ch\|u\|_2 .$$

From (2.9) and (2.10), (2.11) we obtain

(2.13) $$\|v_h\|_1 \leqq Ch^{-1/2}\|u_I - \psi_h\|_{C(\Gamma)} \leqq Ch\left(\sum_{m=1}^{G} |u|_{2,\Gamma_m}^2\right)^{1/2} .$$

Since

$$u - w_h = u - u_I + u_I - w_h = u - u_I + v_h ,$$

from (2.12) and (2.13) we deduce (2.7). Q.E.D.

Corollary 2.1. *Let the assumptions of Theorem 2.1 be satisfied. Let u_h be the finite element approximation to the primary problem $\mathscr{P}_1$. Then*

(2.14) $$\|u - u_h\|_1 = O(h) .$$

The proof follows from Theorem 2.1 and Lemma 2.1, because

$$2u - w_h \geqq u - w_h \geqq 0 \quad \text{on} \quad \Gamma ,$$

consequently, $2u - w_h \in K_1$ and $w_h \in K_{1h}$. Q.E.D.

As the *finite element approximations to the primary problem* $\mathscr{P}_2$ are concerned, one can prove an analogue of Theorem 2.1, where $u = 0$ on Γ_u and $u \in H^2(\Gamma_m \cap \Gamma_a)$ is assumed. Then the rate of convergence (2.14) is also valid.

3. FINITE ELEMENT APPROXIMATIONS FOR THE FIRST DUAL PROBLEM

In case of the dual problems, we have to distinguish strictly the problems (1.11) and (1.12), because of the different construction of the finite-dimensional subsets of $\mathscr{U}_1$ and $\mathscr{U}_2$.

Let us start with approximate solutions of the problem (1.11). Using the definition of $\mathscr{U}_1$, we transform (1.11) into an equivalent problem: to find $q^0 \in \mathscr{U}_0$ such that

$$I(q^0) \leqq I(q) \quad \forall q \in \mathscr{U}_0 , \tag{3.1}$$

where

$$\mathscr{U}_0 = \{q | q \in Q, \; q \cdot \nu|_\Gamma \geqq 0\} ,$$

$$I(q) = \tfrac{1}{2}\left(\sum_{i=1}^{n} \|q_i\|_0^2 + \|\text{div } q\|_0^2\right) + (f, \text{div } q)_0 . \tag{3.1'}$$

Then

$$\lambda_i^0 = q_i^0 , \quad (i = 1, \ldots, n) , \quad \lambda_{n+1}^0 = f + \text{div } q^0 .$$

Consider again the α-β-regular triangulations $\mathscr{T}_h$ of $\Omega \subset R^2$ and the spaces V_h of linear splines on $\mathscr{T}_h$. Introducing the subset

$$\mathscr{U}_{0h} = \mathscr{U}_0 \cap [V_h]^2 .$$

we may define:
a vector $q^h \in \mathscr{U}_{0h}$ will be called a finite element approximation to the dual problem (3.1), if

$$I(q^h) \leqq I(q) \quad \forall q \in \mathscr{U}_{0h} . \tag{3.2}$$

The linear space Q with the norm

$$\|q\|_Q = \left(\sum_{i=1}^{2} \|q_i\|_0^2 + \|\text{div } q\|_0^2\right)^{1/2}$$

is a Hilbert space, $\mathscr{U}_0$ is closed in Q and convex. Then $\mathscr{U}_{0h}$ is convex and closed in Q and the problem (3.2) possesses a unique solution.

Let $\{w^1, w^2, \ldots, w^N\}$ create a basis of V_h. Then

$$q \in \mathscr{U}_{0h} \Leftrightarrow \left\{q = \sum_{j=1}^{N} y_j w^j , \quad \mathbf{B}\mathbf{y} \geqq 0\right\} ,$$

where $\mathbf{B}$ is a $(p \times N)$ matrix and the rank of $\mathbf{B}$ equals p, $p < N$.

The conditions $\mathbf{B y} \geqq 0$ are generated by the boundary condition $q \cdot v|_\Gamma \geqq 0$.

The problem (3.2) can be rewritten as follows:

$$F(\mathbf{y}) = \tfrac{1}{2}\mathbf{y}^{\mathrm{T}}\mathbf{A}\mathbf{y} - \mathbf{y}^{\mathrm{T}}\mathbf{b} = \min \tag{3.3}$$

over the set

$$Y = \{\mathbf{y} \in R^N,\ \mathbf{B}\mathbf{y} \geqq 0\}\,, \tag{3.4}$$

where $\mathbf{A}$ is the corresponding Gramm matrix and $\mathbf{b}$ is a vector.

Since the matrix $\mathbf{B}$ is not diagonal, we do not employ the algorithm of Gauss-Seidel with constraints, but use e.g. the *algorithm of Uzawa* (cf. [5]), which transforms the side conditions into a simpler form.

Denote

$$R_+^p = \{\mathbf{z} \mid \mathbf{z} \in R^p,\ \mathbf{z} \geqq 0\}$$

and choose an arbitrary $\mathbf{z}^1 \in R_+^p$. Then we solve the problem

$$\mathbf{A}\mathbf{y}^1 = \mathbf{b} + \mathbf{B}^{\mathrm{T}}\mathbf{z}^1\,.$$

Having $\mathbf{y}^m$ and $\mathbf{z}^m$ $(m = 1, 2, \dots)$, we calculate

$$\mathbf{z}^{m+1} = P_R(\mathbf{z}^m - \varrho\mathbf{B}\mathbf{y}^m)\,,$$

$$\mathbf{A}\mathbf{y}^{m+1} = \mathbf{b} + \mathbf{B}^{\mathrm{T}}\mathbf{z}^{m+1}\,,$$

where P_R is the projection of R^p onto R_+^p, i.e.

$$z_j^{m+1} = \max\{0;\ z_j^m - \varrho(B\mathbf{y}^m)_j\}\,, \quad j = 1, \dots, p$$

and ϱ is a sufficiently small positive parameter.

It is well known (cf. [5] chpt. 4, § 5.1) that

$$\lim \|\mathbf{y}^m - \mathbf{y}^0\|_{R^N} = 0 \quad \text{for} \quad m \to \infty\,.$$

where $\mathbf{y}^0$ is the solution of (3.3), (3.4).

To obtain an estimate of the distance between the solutions q^0 and q^h of the problem (3.1) and (3.2), respectively, we may again apply Lemma 2.1, where $\mathscr{J} = I$, $M = \mathscr{U}_0$, $M_h = \mathscr{U}_{0h}$, $B = Q$, $\alpha_0 = c = 1$. Consequently, for any vector $t^h \in \mathscr{U}_{0h}$ such that $2q^0 - t^h \in \mathscr{U}_0$, it holds

$$\|q^0 - q^h\|_Q \leqq \|q^0 - t^h\|_Q\,. \tag{3.5}$$

A suitable vector t^h can be found, as follows from

Theorem 3.1. *Assume that* $q^0 \in [H^2(\Omega)]^2$ *and* $q^0 . \nu \in H^2(\Gamma_m)$ $m = 1, \ldots, G$, *where* Γ_m *denotes any side of the polygonal boundary* Γ.

Then there exists $t^h \in \mathscr{U}_{0h}$ *such that*

$$0 \leqq t^h . \nu \leqq q^0 . \nu \quad \text{on} \quad \Gamma \tag{3.6}$$

(*almost everywhere*) *and, if the triangulations are* α-β-regular,

$$\|q^0 - t^h\|_Q \leqq Ch\Big(\sum_{j=1}^{2} \|q_j^0\|_2^2 + \sum_{m=1}^{G} \|q^0 . \nu\|^2_{H^2(\Gamma_m)}\Big)^{1/2} \tag{3.7}$$

with C *independent of* h *and* q^0.

For the proof we need two auxiliary Lemmas.

Lemma 3.1. (*One-sided approximation of the flux on the boundary*). *Let* $q^0 \in \mathscr{U}_0$ *satisfy the assumptions of Theorem* 3.1. *Then there exist linear spline functions* $\psi_h^m \in C(\bar{\Gamma}_m)$, *with the nodes determined by the triangulation* $\mathscr{T}_h$ *and such that*

$$0 \leqq \psi_h^m \leqq q^0 . \nu \quad \text{on} \quad \Gamma_m \ \forall m, \tag{3.8}$$

$$\|q_I^0 . \nu - \psi_h\|^2_{\Gamma,\infty} \leqq h^3 \sum_{m=1}^{G} |q^0 . \nu|^2_{2,\Gamma_m}, \tag{3.9}$$

where ψ_h *is defined on* $\bigcup_m \Gamma_m$ *in such a way that its restrictions on to* Γ_m *equal* ψ_h^m, *furthermore*

$$\|\varphi\|_{\Gamma,\infty} = \max_{1 \leqq m \leqq G} \big(\sup_{s \in \Gamma_m} |\varphi(s)|\big),$$

q_I^0 *denotes the linear interpolate of* q^0 *on* $\mathscr{T}_h$ *and the seminorm* $|.|_{2,\Gamma_m}$ *has been defined in Lemma* 2.2.

Proof. For brevity, denote $q^0 . \nu = t$ and let t_I be the linear interpolate of t on $\bar{\Gamma}_m$ with the nodes of $\mathscr{T}_h$. Note that

$$t_I = q_I^0 . \nu .$$

The assertions (3.8), (3.9) can be proved like (2.8), (2.9) in Lemma 2.2, replacing only Γ by $\bar{\Gamma}_m$ and u by t.

Lemma 3.2. Let $\varphi_m \in C(\bar{\Gamma}_m)$, $m = 1, \ldots, G$ be the linear spline-functions with the nodes determined by the α-β-regular triangulation $\mathscr{T}_h$.

Then there exists $w^h \in [V_h]^2$ such that $w^h . \nu = \varphi_m$ on Γ_m for all $m = 1, \ldots, G$ and

$$\Big(\sum_{j=1}^{2} \|w_j^h\|_1^2\Big)^{1/2} \leqq Ch^{-1/2} \|\varphi\|_{\Gamma,\infty}, \tag{3.10}$$

where φ is a function such that its restrictions onto Γ_m coincide with φ_m for all m.

Proof. Consider again the "boundary strip" Ω_h of Ω like in the proof of Lemma 2.3. We set $w_j^h(b_i) = 0$, $j = 1, 2$ at all vertices $b_i \in \bar{\Omega} - \Gamma$. Then it suffices to determine and estimate the values $w_j^h(a_i)$ at the vertices $a_i \in \Gamma$.

1. Let a_i be a vertex of the polygonal boundary. Denote $\varphi(a_i +)$, $\varphi(a_i -)$ the limits of φ from both sides of the vertex a_i and $\nu+$, $\nu-$ the corresponding unit outward normals to the boundary. By the conditions be

$$w^h(a_i) \,.\, \nu- = \varphi(a_i -)\,,$$

$$w^h(a_i) \,.\, \nu+ = \varphi(a_i +)$$

the values $w_j^h(a_i)$, $j = 1, 2$ are determined and it holds

$$|w_j^h(a_i)| \leqq (|\varphi(a_i -)| + |\varphi(a_i +)|)\, |\sin \alpha_i|^{-1}\,, \quad j = 1, 2\,, \tag{3.11}$$

where α_i is the interior angle of Γ at the vertex a_i.

2. Let $a_i \in \Gamma_m$ be a vertex of $\mathscr{T}_h$ but not a vertex of Γ. Denote $\nu = (\nu_1, \nu_2)$ the unit normal to Γ_m and let

$$|\nu_k| = \max\{|\nu_1|, |\nu_2|\}\,.$$

Since $2\nu_k^2 \geqq \nu_1^2 + \nu_2^2 = 1$, we have $|\nu_k| \geqq 1/\sqrt{2}$. We choose

$$w_k^h(a_i) = \nu_k^{-1}\, \varphi(a_i)$$

and the remaining component

$$w_p^h(a_i) = 0\,, \quad (p \neq k)\,.$$

Then obviously

$$|w_j^h(a_i)| \leqq \sqrt{(2)}\, |\varphi(a_i)|\,, \quad j = 1, 2\,. \tag{3.12}$$

From (3.11), (3.12) it follows that a constant C exists such that

$$|w_j^h(a_i)| \leqq C\|\varphi\|_{\Gamma,\infty}\, (j = 1, 2)\,, \quad \forall a_i \in \Gamma$$

and consequently,

$$\max_{j=1,2} \|w_j^h\|_{C(\bar{\Omega})} \leqq C\|\varphi\|_{\Gamma,\infty}\,.$$

Using the α-β-regularity of $\mathscr{T}_h$, we can derive the estimate

$$|\partial w_j^h / \partial x_k| \leqq Ch^{-1}\|\varphi\|_{\Gamma,\infty} \quad \forall T \in \mathscr{T}_h$$

(cf. the proof of Lemma 2.3) with C independent of h, φ and T. The rest of the proof is parallel to that of Lemma 2.3. Q.E.D.

Proof of Theorem 3.1. Let ψ_h be the function defined in Lemma 3.1. Introducing

$$\varphi = q_I^0 \cdot \nu - \psi_h ,$$

we construct the vector-function $w^h \in [V_h]^2$ according to Lemma 3.2. Then the function

$$t^h = q_I^0 - w^h \in [V_h]^2$$

satisfies (3.6) and (3.7). In fact, on every Γ_m it holds

$$t^h \cdot \nu = q_I^0 \cdot \nu - \varphi_m = \psi_h^m$$

and (3.8) yields (3.6). Besides, we have

$$\|(q^0 - q_I^0)_j\|_1 \leqq Ch\|q_j^0\|_2 \quad (j = 1, 2) , \tag{3.13}$$

$$\|q\|_Q \leqq C(\sum_{j=1}^{2} \|q_j\|_1^2)^{1/2} \quad \forall q \in Q . \tag{3.14}$$

From Lemma 3.1 and 3.2 it follows that

$$(\sum_{j=1}^{2} \|w_j^h\|_1^2)^{1/2} \leqq Ch^{-1/2}\| q_I^0 \cdot \nu - \psi_h\|_{\Gamma,\infty} \leqq Ch(\sum_{m=1}^{G} |q^0 \cdot \nu|_{2,\Gamma_m}^2)^{1/2} . \tag{3.15}$$

Since we may write

$$\|q^0 - t^h\|_Q = \|q^0 - q_I^0 + w^h\|_Q \leqq \|q^0 - q_I^0\|_Q + \|w^h\|_Q ,$$

(3.13), (3.14) and (3.15) result in (3.7). Q.E.D.

Corollary 3.1. *Let the assumptions of Theorem* 3.1 *be satisfied. Let* q^h *be the finite element approximation to the problem* (3.1). *Then*

$$\|q^0 - q^h\|_Q = O(h) .$$

The proof follows from Theorem 3.1 and Lemma 2.1. In fact,

$$(2q^0 - t^h) \cdot \nu|_\Gamma \geqq q^0 \cdot \nu|_\Gamma - t^h \cdot \nu|_\Gamma \geqq 0$$

by virtue of (3.6). Therefore $2q^0 - t^h \in \mathscr{U}_0$, $t^h \in \mathscr{U}_{0h}$ and we may use (3.5), (3.7). Q.E.D.

Remark 3.1. If q^h is a solution of (3.2), then

$$\lambda^h = \{q_1^h, q_2^h, f + \operatorname{div} q^h\} \in \mathscr{U}_1$$

is an approximation to the dual problem (1.11). By virtue of Theorem 1.1 and Corollary 3.1, it holds for $h \to 0$

$$\|q_i^h - \partial u/\partial x_i\|_0 \quad = O(h)\,, \quad (i = 1, 2)\,,$$

$$\|\operatorname{div} q^h + f - u\|_0 = O(h)\,.$$

In Section 6 we present also some a posteriori error estimates for q^h.

4. FINITE ELEMENT APPROXIMATIONS TO THE SECOND DUAL PROBLEM

Let us consider the dual problem (1.12). Assume that the boundary Γ consists of J mutually disjoint closed polygons $\partial\Omega_j$ i.e., let

(A 1) $$\Gamma = \bigcup_{j=1}^{J} \partial\Omega_j\,, \quad \partial\Omega_j \cap \partial\Omega_k = \emptyset \quad \text{for} \quad j \neq k$$

and let

(A 2) $$\operatorname{mes}(\partial\Omega_j \cap \Gamma_u) > 0\,, \quad j = 1, \ldots, J\,.$$

Consider again the α-β-regular triangulations $\mathscr{T}_h$ of Ω and assume that (A 3) the boundary points of Γ_a are vertices of $\mathscr{T}_h$.

Let us construct a fixed vector $\bar{\lambda} \in Q$ such that

$$\operatorname{div} \bar{\lambda} + f = 0\,.$$

(We may choose e.g. $\bar{\lambda} = \{\bar{\lambda}_1, 0\}$, where

$$\bar{\lambda}_1 = -\int_0^{x_1} f(t, x_2)\,\mathrm{d}t)\,.$$

Then it is readily seen that

(4.1) $$\lambda \in \mathscr{U}_2 \Leftrightarrow \lambda - \bar{\lambda} = q \in K\,,$$

where

$$K = \{\boldsymbol{q} \mid \boldsymbol{q} \in Q,\ \operatorname{div} \boldsymbol{q} = 0,\ \boldsymbol{q} \,.\, \nu|_{\Gamma_a} = -\bar{\lambda} \,.\, \nu|_{\Gamma_a}\}\,.$$

Instead of the finite element subspaces V_h we have to work with subspaces of "equilibrium elements" which satisfy the equation

$$\operatorname{div} \boldsymbol{q} = 0$$

at least in the sense of distributions (cf. (1.9)). To this end we construct piecewise linear vector-functions as follows (cf. [2], [3] and the references therein).

For any triangle $T \in \mathscr{T}_h$ we define the set

$$\mathscr{M}(T) = \{\boldsymbol{q} \mid \boldsymbol{q} \in [P_1(T)]^2 \, , \; \operatorname{div} \boldsymbol{q} = 0\} \, .$$

Henceforth $P_1(A)$ denotes the space of linear polynomials on the set A.

Furthermore, let

$$\mathscr{N}_h = \{\boldsymbol{q} \mid \boldsymbol{q} \in [L_2(\Omega)]^2, \;\; \boldsymbol{q}|_T \in \mathscr{M}(T) \;\; \forall T \in \mathscr{T}_h, \;\; \boldsymbol{q} \, . \, \nu|_T + \boldsymbol{q} \, . \, \nu|_{T'} = 0 \quad \forall x \in T \cap T'\} \, ,$$

where the last conditions means that the "flux" $\boldsymbol{q} \, . \, \nu$ is continuous when crossing any interelement boundary between any two adjacent triangles T and T'.

It is easy to see that $\mathscr{N}_h$ is a linear finite-dimensional manifold and that $\mathscr{N}_h \subset Q$, because $\operatorname{div} \boldsymbol{q} = 0$ for any $\boldsymbol{q} \in \mathscr{N}_h$ in the sense of distributions.

Let us define the mapping

$$\Pi_T \in \mathscr{L}([H^1(T)]^2 \; ; \;\; [P_1(T)]^2)$$

by the conditions that the $L_2(S_i)$-projection of the flux $\lambda \cdot \nu|_{S_i}$ is equal to the flux $(\Pi_T \lambda) \, . \, \nu|_{S_i}$ for each side S_i of the triangle T. Define also the set

$$\mathscr{R}(\Omega) = \{\lambda \mid \lambda \in [H^1(\Omega)]^2, \; \operatorname{div} \lambda = 0\}$$

and the mapping r_h of $\mathscr{R}(\Omega)$ by the conditions

$$(r_h \lambda)|_T = \Pi_T \lambda \quad \forall T \in \mathscr{T}_h \, .$$

It can be proved (cf. [3]) that

(4.2) $$r_h \in \mathscr{L}(\mathscr{R}(\Omega); \mathscr{N}_h)$$

and

(4.3) $$\|\lambda - r_h \lambda\|_{[L_2(\Omega)]^2} \leqq C h^2 |\lambda|_2 \quad \forall \lambda \in [H^2(\Omega)]^2 \, ,$$

where C is independent of h and λ,

$$|\lambda|_2 = (\sum_{i=1}^{2} \sum_{|\alpha|=2} \|D^\alpha \lambda_i\|_0^2)^{1/2} \, .$$

Assume moreover that

(A 4) $$\exists \, \boldsymbol{G} \in [H^2(\Omega)]^2 \, , \quad \operatorname{div} \boldsymbol{G} = 0 \, , \quad \boldsymbol{G} \, . \, \nu|_{\Gamma_a} = -\bar{\lambda} \, . \, \nu|_{\Gamma_a} \, .$$

Let us denote $-\bar{\lambda}.\nu|_\Gamma = g$ and construct the function $g_h \in L_2(\Gamma_a)$ such that the restrictions $g_h|_{S_i}$ are the $L_2(S_k)$-projections of g onto $P_1(S_k)$ for each side $S_k \subset \Gamma_a$ of the triangulation $\mathscr{T}_h$. Note that g_h is piecewice linear with discontinuites at every node, in general.

Using (4.1), we can derive that the problem (1.12) is equivalent with the following problem: to find $\boldsymbol{q}^0 \in K$ such that

$$(4.4) \qquad J(\boldsymbol{q}^0) \leqq J(\boldsymbol{q}) \quad \forall \boldsymbol{q} \in K,$$

where

$$J(\boldsymbol{q}) = \tfrac{1}{2}\|\boldsymbol{q}\|^2 + (\bar{\lambda}, \boldsymbol{q}),$$

$$\|\boldsymbol{q}\|^2 = \sum_{i=1}^{2} \|q_i\|_0^2, \quad (\bar{\lambda}, \boldsymbol{q}) = \sum_{i=1}^{2} (\bar{\lambda}_i, q_i)_0 .$$

We say that $\boldsymbol{q}^h \in K_h$ is a finite element approximation to the dual problem (4.4) if

$$(4.5) \qquad J(\boldsymbol{q}^h) \leqq J(\boldsymbol{q}) \quad \forall \boldsymbol{q} \in K_h,$$

where

$$K_h = \{\boldsymbol{q} \mid \boldsymbol{q} \in \mathcal{N}_h, \ \boldsymbol{q} \cdot \nu|_{\Gamma_a} \geqq g_h\} .$$

Since K_h is convex and closed in $[L_2(\Omega)]^2$, the problem (4.5) has a unique solution.

Let us describe the algorithm of solving (4.5) in detail. Any element $\boldsymbol{q}' \in \mathcal{N}_h$ is determined by a vector β of N" flux parameters" (see [3]), N being equal to six-times the number of all triangles in $\mathcal{T}_h$. In each triangle $T_e \in \mathcal{T}_h$ with vertices a_1, a_2, a_3 it holds

$$\boldsymbol{w}^e = \boldsymbol{C}_e \beta^e ,$$

where

$$\boldsymbol{w}^e = [q_1^e(a_1), q_2^e(a_1), q_1^e(a_2), q_2^e(a_2), q_1^e(a_3), q_2^e(a_3)]^{\mathrm{T}}, \text{[1]}$$

$$\beta^e = [\beta_1^e, \ldots, \beta_6^e]^{\mathrm{T}}, \text{[2]}$$

$\boldsymbol{q}^e$ denotes the restriction of $\boldsymbol{q}$ onto T_e and the (6×6) matrix $\boldsymbol{C}_e$ is regular, because the inverse matrix is

$$\boldsymbol{C}_e^{-1} = \begin{bmatrix} \nu_1^{(3)} & \nu_2^{(3)} & 0 & . & . & . \\ \nu_1^{(1)} & \nu_2^{(1)} & 0 & . & . & . \\ 0 & 0 & \nu_1^{(1)} & \nu_2^{(1)} & 0 & . \\ 0 & 0 & \nu_1^{(2)} & \nu_2^{(2)} & 0 & . \\ . & . & . & . & \nu_1^{(2)} & \nu_2^{(2)} \\ . & . & . & . & \nu_1^{(3)} & \nu_2^{(3)} \end{bmatrix}$$

with $\nu^{(k)}$ denoting the outward unit normal to the side $S_k = \overline{a_k a_{k+1}}$, ($k = 1, 2, 3$ and $a_4 \equiv a_1$).

---

[1]) The superscript "T" denotes the transposed matrix.

[2]) β_j^k are defined as limits of the flux $\boldsymbol{q} \cdot \nu$ at the vertices of a side of T_e. The components of $\boldsymbol{\beta}$, however, are not independent — see (4.9), (4.10).

For $\mathbf{x} \in T_e$ we may write $\boldsymbol{q}$ in terms of the restricted basis functions:

$$q_i^e(x) = \sum_{j=1}^{3} q_i^e(a_j)\, \varphi_j(\mathbf{x}), \quad (i = 1, 2).$$

If we introduce the vector-functions

$$\Phi_1 = [\varphi_1, 0, \varphi_2, 0, \varphi_3, 0]^{\mathsf{T}},$$
$$\Phi_2 = [0, \varphi_1, 0, \varphi_2, 0, \varphi_3]^{\mathsf{T}},$$

then obviously

$$q_i = (\mathbf{w}^e)^{\mathsf{T}} \Phi_i = \Phi_i^{\mathsf{T}} \mathbf{w}^e, \quad (i = 1, 2),$$

$$\int_{Te} (q_i^e)^2 \, \mathrm{d}\mathbf{x} = (\mathbf{w}^e)^{\mathsf{T}} \int_{Te} \Phi_i \Phi_i^{\mathsf{T}} \, \mathrm{d}\mathbf{x}\, \mathbf{w}^e .$$

Consequently, we have

$$(4.6) \quad \sum_{i=1}^{2} \|q_i\|_0^2 = \sum_{i=1}^{2} \sum_{T_e \in \mathscr{T}_h} \int_{T_e} (q_i^e)^2 \, \mathrm{d}\mathbf{x} = \sum_{T_e} (\mathbf{w}^e)^{\mathsf{T}} \int_{T_e} (\Phi_1 \Phi_1^{\mathsf{T}} + \Phi_2 \Phi_2^{\mathsf{T}}) \, \mathrm{d}\mathbf{x}\, \mathbf{w}^e =$$
$$= \sum_{T_e} (\beta^e)^{\mathsf{T}} \mathbf{A}_e \beta^e = \beta^{\mathsf{T}} \mathbf{A} \beta ,$$

where $\mathbf{A}$ is symmetric and positive definite. In fact,

$$\mathbf{A}_e = \mathbf{C}_e^{\mathsf{T}} \int_{T_e} (\Phi_1 \Phi_1^{\mathsf{T}} + \Phi_2 \Phi_2^{\mathsf{T}}) \, \mathrm{d}\mathbf{x}\, \mathbf{C}_e$$

are symmetric and positive definite 6×6 matrices, because

$$(\beta^e)^{\mathsf{T}} \mathbf{A}_e \beta^e = \sum_{i=1}^{2} \int_{T_e} (q_i^e)^2 \, \mathrm{d}\mathbf{x} = 0 \Rightarrow \mathbf{w}^e = 0 \Rightarrow \beta^e = 0 .$$

Likewise, we may write

$$\sum_{i=1}^{2} (\bar{\lambda}_i, q_i)_0 = \sum_{i=1}^{2} \sum_{T_e} \int_{T_e} \bar{\lambda}_i \Phi_i^{\mathsf{T}} \mathbf{w}^e \, \mathrm{d}\mathbf{x} = -\sum_{T_e} (\mathbf{b}^e)^{\mathsf{T}} \beta^e = -\mathbf{b}^{\mathsf{T}} \beta ,$$

where

$$(4.7) \qquad \mathbf{b}^e = -\sum_{i=1}^{2} \mathbf{C}_e \int_{T_e} \bar{\lambda}_i \Phi_i \, \mathrm{d}\mathbf{x} .$$

Consequently, (4.5) is equivalent with the problem to find $\hat{\beta} \in \mathscr{B}$ such that

$$(4.8) \qquad \mathscr{J}(\hat{\beta}) \leqq \mathscr{J}(\mathrm{b}) \quad \forall \beta \in \mathscr{B} ,$$

where

$$\mathscr{J}(\beta) = \tfrac{1}{2} \beta^{\mathsf{T}} \mathbf{A} \beta - \mathbf{b}^{\mathsf{T}} \beta ,$$
$$\mathscr{B} = \{\beta \in R^N, \ \mathbf{D}\beta = 0, \ \mathbf{E}\beta \geqq \mathbf{e}(\bar{\lambda})\} .$$

Here $\mathbf{D}$ is a $p_1 \times N$ matrix corresponding (i) to $N/6$ conditions of the vanishing divergence – of the form (cf. [3])

$$l_1(\beta_2^e + \beta_3^e) + l_2(\beta_4^e + \beta_5^e) + l_3(\beta_1^e + \beta_6^e) = 0\,, \tag{4.9}$$

(l_i denotes the length of the side S_i of T_e) and (ii) to conditions of continuity for the fluxes along the interelement boundaries – of the form

$$\beta_i + \beta_k = 0\,. \tag{4.10}$$

Thus $\mathbf{D}\beta = 0$ if and only if thre corresponding function $\mathbf{q}$ belongs to $\mathcal{N}_h$.

Finally $\mathbf{E}$ is a $(p_2 \times N)$ matrix such that $\mathbf{E}\beta \geqq e(\bar{\lambda})$ corresponds with the condition $q \cdot v|_{\Gamma_a} \geqq g_h$. The rows of $\mathbf{E}$ consist of zeros and one unit. One can easily verify that under the assumptions (A 1), (A 2), (A 3) we have $p = p_1 + p_2 < N$.

Let us denote

$$\mathbf{B} = \begin{bmatrix} \mathbf{D} \\ \mathbf{E} \end{bmatrix}.$$

To solve the problem (4.8) we can use again e.g. the algorithm of Uzawa (see, however, Remarks 6.2–6.7).

Introduce the convex set

$$\Lambda = \{\mathbf{z} \in R^p \mid z_j \geqq e_j(\bar{\lambda}) \text{ for } j < p_1\}$$

and choose an arbitrary $\mathbf{z}^1 \in \Lambda$. Then we solve the problem

$$\mathbf{A}\beta^1 = \mathbf{b} + \mathbf{B}^{\mathsf{T}}\mathbf{z}^1\,.$$

Having β^m and $\mathbf{z}^m$, $(m = 1, 2, \ldots)$, we calculate

$$\mathbf{z}^{m+1} = P_\Lambda(\mathbf{z}^m - \varrho \mathbf{B}\beta^m)$$

$$\mathbf{A}\beta^{m+1} = \mathbf{b} + \mathbf{B}^{\mathsf{T}}\mathbf{z}^{m+1}\,,$$

where P_Λ is the projection of R^p onto Λ, i.e.

$$\mathbf{y} = P_\Lambda \mathbf{t} \Leftrightarrow \begin{cases} y_j = t_j\,, & j \leqq p_1\,, \\ y_j = \max\{e_j(\bar{\lambda}),\ t_j\}\,, & j > p_1\,. \end{cases}$$

and ϱ is a sufficiently small positive parameter.

It can be shown (cf. e.g. [5]-chpt. 4, § 5.1) that

$$\lim_{m \to \infty} \|\beta^m - \hat{\beta}\|_{R^N} = 0\,. \tag{4.11}$$

5. CONVERGENCE OF THE FINITE ELEMENT APPROXIMATIONS TO THE SECOND DUAL PROBLEM

In this section we analyze the distance between the solutions $\boldsymbol{q}^0$ and $\boldsymbol{q}^h$ of the problems (4.4) and (4.5), respectively. Then setting $\lambda^h = \bar{\lambda} + \boldsymbol{q}^h$, we obtain the distance between the solution $\lambda^0 = \bar{\lambda} + \boldsymbol{q}^0$ of (1.12) and λ^h, as $\lambda^0 - \lambda^h = \boldsymbol{q}^0 - \boldsymbol{q}^h$.

Define the following convex cones

$$\mathscr{C} = \{\boldsymbol{q} \in Q \mid \operatorname{div} \boldsymbol{q} = 0,\ \boldsymbol{q} \cdot \nu|_{\Gamma_a} \geqq 0\},$$

$$\mathscr{C}_h = \mathscr{C} \cap \mathscr{N}_h = \{\boldsymbol{q} \in \mathscr{N}_h \mid \boldsymbol{q} \cdot \nu|_{\Gamma_a} \geqq 0\}.$$

Under the assumption (A 4) it holds

$$\boldsymbol{q}^0 - \boldsymbol{G} = \boldsymbol{U} \in \mathscr{C}. \tag{5.1}$$

If we construct the projection $r_h\boldsymbol{G}$, then $r_h\boldsymbol{G} \in \mathscr{N}_h$ according to (4.2). Note that the flux $(r_h\boldsymbol{G}) \cdot \nu|_{S_k}$ coincides with the $L_2(S_k)$-projection of $\boldsymbol{G} \cdot \nu = -\bar{\lambda} \cdot \nu = g$ onto $P_1(S_k)$ on each triangle side $S_k \subset \Gamma_a$, therefore

$$(r_h\boldsymbol{G}) \cdot \nu = g_h \quad \text{on} \quad \Gamma_a .$$

Thus the difference

$$\boldsymbol{q}^h - r_h\boldsymbol{G} \equiv \boldsymbol{U}_h \in \mathscr{C}_h \tag{5.2}$$

and we come to the equivalence

$$\boldsymbol{q} \in K_h \Leftrightarrow \boldsymbol{q} - r_h\boldsymbol{G} \equiv \boldsymbol{V}_h \in \mathscr{C}_h .$$

Lemma 5.1. *Let a* $\boldsymbol{W}_h \in \mathscr{C}_h$ *exist such that* $2\boldsymbol{U} - \boldsymbol{W}_h \in \mathscr{C}$. *Then it holds*

$$\|\boldsymbol{q}^0 - \boldsymbol{q}^h\| \leqq \|\boldsymbol{U} - \boldsymbol{W}_h\| + \|\boldsymbol{G} - r_h\boldsymbol{G}\| . \tag{5.3}$$

Proof. Set $\boldsymbol{q} = \boldsymbol{G} + \boldsymbol{W}_h$. Then $\boldsymbol{q} \in K$ and

$$2\boldsymbol{q}^0 - \boldsymbol{q} = 2(\boldsymbol{G} + \boldsymbol{U}) - (\boldsymbol{G} + \boldsymbol{W}_h) = \boldsymbol{G} + 2\boldsymbol{U} - \boldsymbol{W}_h \in K .$$

The solution $\boldsymbol{q}^0$ of (4.4) satisfies the inequalities

$$DJ(\boldsymbol{q}^0, \boldsymbol{q} - \boldsymbol{q}^0) \geqq 0 ,$$

$$DJ(q^0, 2\boldsymbol{q}^0 - \boldsymbol{q} - \boldsymbol{q}^0) = DJ(q^0, \boldsymbol{q}^0 - \boldsymbol{q}) \geqq 0 ,$$

where

$$DJ(\boldsymbol{q}, \boldsymbol{p}) = (\boldsymbol{q}, \boldsymbol{p}) + (\bar{\lambda}, \boldsymbol{p})$$

is the Gâteaux differential of J. Consequently,

$$(5.4) \qquad 0 = DJ(\boldsymbol{q}^0, \boldsymbol{q} - \boldsymbol{q}^0) = DJ(\boldsymbol{q}^0, \boldsymbol{W}_h - \boldsymbol{U}) = (\boldsymbol{q}^0, \boldsymbol{W}_h - \boldsymbol{U}) + (\bar{\lambda}, \boldsymbol{W}_h - \boldsymbol{U}),$$

Second, let us choose $\boldsymbol{q} = \boldsymbol{G} + \boldsymbol{U}_h \in K$. Since $\boldsymbol{q} - \boldsymbol{q}^0 = \boldsymbol{U}_h - \boldsymbol{U}$, we obtain

$$(5.5) \qquad 0 \leqq DJ(\boldsymbol{q}^0, \boldsymbol{q} - \boldsymbol{q}^0) = DJ(\boldsymbol{q}^0, \boldsymbol{U}_h - \boldsymbol{U}) = (\boldsymbol{q}^0, \boldsymbol{U}_h - \boldsymbol{U}) + (\bar{\lambda}, \boldsymbol{U}_h - \boldsymbol{U}).$$

Finally, choosing $\boldsymbol{q} = r_h\boldsymbol{G} + \boldsymbol{W}_h \in K_h$, we may write

$$(5.6) \quad 0 \leqq DJ(\boldsymbol{q}^h, \boldsymbol{q} - \boldsymbol{q}^h) = DJ(\boldsymbol{q}^h, \boldsymbol{W}_h - \boldsymbol{U}_h) = (\boldsymbol{q}^h, \boldsymbol{W}_h - \boldsymbol{U}_h) + (\bar{\lambda}, \boldsymbol{W}_h - \boldsymbol{U}_h).$$

Using (5.4), (5.5) and (5.6), we obtain

$$(5.7) \qquad (\boldsymbol{q}^0 - \boldsymbol{q}^h, \boldsymbol{U}_h - \boldsymbol{W}_h) = (\boldsymbol{q}^0, \boldsymbol{U} - \boldsymbol{W}_h - \boldsymbol{U} + \boldsymbol{U}_h) + (\boldsymbol{q}^h, \boldsymbol{W}_h - \boldsymbol{U}_h) \geqq$$
$$\geqq (\bar{\lambda}, \boldsymbol{W}_h - \boldsymbol{U}) + (\bar{\lambda}, \boldsymbol{U} - \boldsymbol{U}_h) + (\bar{\lambda}, \boldsymbol{U}_h - \boldsymbol{W}_h) = 0.$$

Since

$$\boldsymbol{q}^0 - \boldsymbol{q}^h = \boldsymbol{G} + \boldsymbol{U} - (r_h\boldsymbol{G} + \boldsymbol{U}_h) = \boldsymbol{G} - r_h\boldsymbol{G} + \boldsymbol{U} - \boldsymbol{U}_h,$$

making use of (5.7), we may write

$$\|\boldsymbol{q}^0 - \boldsymbol{q}^h\|^2 = (\boldsymbol{q}^0 - \boldsymbol{q}^h, \boldsymbol{G} - r_h\boldsymbol{G}) + (\boldsymbol{q}^0 - \boldsymbol{q}^h, \boldsymbol{U} - \boldsymbol{U}_h + (\boldsymbol{U}_h - \boldsymbol{W}_h)) =$$
$$= (\boldsymbol{q}^0 - \boldsymbol{q}^h, \boldsymbol{G} - r_h\boldsymbol{G} + \boldsymbol{U} - \boldsymbol{W}_h) \leqq \|\boldsymbol{q}^0 - \boldsymbol{q}^h\| \{\|\boldsymbol{G} - r_h\boldsymbol{G}\| + \|\boldsymbol{U} - \boldsymbol{W}_h\|\}$$

Q.E.D.

The next problem is to show that there exists $\boldsymbol{W}_h \in \mathscr{C}_h$ sufficiently close to $\boldsymbol{U}$. Fortunately, we can prove the following

Theorem 5.1. *Let* $\boldsymbol{U} = \boldsymbol{q}^0 - \boldsymbol{G} \in [H^2(\Omega)]^2$ *and* $\boldsymbol{U} . \nu \in H^2(\Gamma_m \cap \Gamma_a)$ *for any side* Γ_m *of the polygonal boundary* Γ.

Then there exists $\boldsymbol{W}_h \in \mathscr{C}_h$ *such that* $2\boldsymbol{U} - \boldsymbol{W}_h \in \mathscr{C}$ *and*

$$\|\boldsymbol{U} - \boldsymbol{W}_h\| \leqq C\{h^2|\boldsymbol{U}|_2 + h^{3/2} \sum_{m=1}^{M} |\boldsymbol{U} . \nu|_{H^2(\Gamma_m \cap \Gamma_a)}\}.$$

The proof is based on two auxiliary lemmas.

Lemma 5.2. (*One-sided approximation of the boundary flux*). *Let the assumptions of Theorem* 5.1 *hold. Then there exists a piecewise linear function* ψ_h *on* Γ *with the nodes determined by the triangulation* $\mathscr{T}_h$ (*discontinuous at the nodes, in general*) *and such that*

$$(5.8) \qquad \int_{\partial\Omega_j} \psi_h \, \mathrm{d}s = \int_{\partial\Omega_j} (r_h\boldsymbol{U}) . \nu \, \mathrm{d}s, \quad j = 1, 2, \ldots, J,$$
$$0 \leqq \psi_h \leqq \boldsymbol{U} . \nu \quad \text{on} \quad \Gamma_a,$$

$$(5.9) \qquad \|(r_h\boldsymbol{U}) . \nu - \psi_h\|_{L_2(\Gamma)} \leqq Ch^2 \sum_{m=1}^{M} |\boldsymbol{U} . \nu|_{H^2(\Gamma_m \cap \Gamma_a)},$$

where

$$|v|_{H^2(\Gamma_m\cap\Gamma_a)} = \|d^2v/ds^2\|_{L_2(\Gamma_m\cap\Gamma_a)} .$$

Proof. Denote $\mathbf{U} . v = t$ and $(r_h\mathbf{U}) . v = t_h$. Let s_i be the parameters of the nodes of $\mathcal{T}_h$ on Γ. Consider the interval-side- $S_i = (s_i, s_{i+1}) \subset \Gamma_a$. Let t_I be the linear function such that $t_I = t$ at the end-points s_i and s_{i+1}. First we construct a function on S_i.

1. If $t \geqq t_I$ for all $s \in (s_i, s_{i+1})$, we set $\psi_h^i = t_I$. Then obviously $\psi_h^i \geqq 0$ on S_i and we obtain for the $L_2(S_i)$ norms

$$\|\psi_h^i - t_h\| \leqq \|t - t_I\| + \|t - t_h\| \leqq Ch^2\|t''\| , \tag{5.10}$$

where $t'' = d^2t/ds^2$. In fact, t_h is the $L_2(S_i)$-projection of t onto the subspace $P_1(S_i)$ and we may use the Bramble-Hilbert Lemma (cf. [8]) to get the estimate for $t - t_h$. The same Lemma can be employed to estimate $t - t_I$.

2. Let there exist points $s \in S_i$ with $t < t_I$. Since $H^2(\Gamma_m \cap \Gamma_a) \subset C^1(\overline{\Gamma_m \cap \Gamma_a})$, we can find a point $\sigma \in \overline{S_i}$ such that the tangent to the graph of t at σ lies under the graph of t and, if ψ_h^i is the function, the graph of which coincides with the tangent, then $\psi_h^i \geqq 0$ on $\overline{S_i}$. We have – for the $L_2(S_i)$ norms –

$$\|\psi_h^i - t_h\| \leqq \|\psi_h^i - t\| + \|t - t_h\| .$$

As in the proof of Lemma 2.2 we obtain

$$\max_{s \in \overline{S}_i} |t(s) - \psi_h^i(s)| \leqq h^{3/2}\|t''\|_{L_2(S_i)} ,$$

$$\|\psi_h^i - t\|_{L_2(S_i)} \leqq h^2\|t''\|_{L_2(S_i)} .$$

Therefore

$$\|\psi_h^i - t_h\| \leqq Ch^2\|t''\| . \tag{5.11}$$

In this way we construct a piecewise linear function ψ_h on the whole Γ_a such that

$$\psi_h|_{S_i} = \psi_h^i \quad \forall S_i \subset \Gamma_a .$$

From (5.10), (5.11) it follows that

$$\|\psi_h - t_h\|^2_{L_2(\Gamma_a)} = \sum_{S_i \subset \Gamma_a} \|\psi_h^i - t_h\|^2_{L_2(S_i)} \leqq Ch^4 \sum_{m=1}^{M} \|t''\|^2_{L_2(\Gamma_m\cap\Gamma_a)} . \tag{5.12}$$

Define

$$\psi_h = t_h + \Delta_j \quad \text{on} \quad \partial\Omega_j - \Gamma_a , \quad j = 1, 2, \ldots, J ,$$

where

$$\Delta_j = [\text{mes}\,(\partial\Omega_j - \Gamma_a)]^{-1} \int_{\partial\Omega_j\cap\Gamma_a} (t_h - \psi_h)\, ds ,$$

$$\Delta_j = 0 \quad \text{if} \quad \partial\Omega_j \cap \Gamma_a = \varnothing .$$

Then it is easy to verify (5.8). Moreover, we have

$$\|\psi_h - t_h\|^2_{L_2(\partial\Omega_j - \Gamma_a)} = \Delta_j^2 \operatorname{mes}(\partial\Omega_j - \Gamma_a) = \|t_h - \psi_h\|^2_{L_2(\partial\Omega_j \cap \Gamma_a)} \leqq$$
$$\leqq Ch^4 \sum_{m=1}^{M} \|t''\|^2_{L_2(\Gamma_m \cup \Gamma_a)} .$$

Thus we are led to the estimate (5.9). Q.E.D.

Lemma 5.3. *Let a piecewise linear (discontinuous) function φ on Γ be given (with the nodes determined by $\mathscr{T}_h$) such that*

(5.13) $$\int_{\partial\Omega_j} \varphi \, ds = 0 , \quad j = 1, 2, \ldots, J .$$

Then there exists a vector-function $\mathbf{w}^h \in \mathscr{N}_h$ such that

$$\mathbf{w}^h . \nu = \varphi \quad \text{on} \quad \Gamma ,$$

(5.14) $$\|\mathbf{w}^h\| \leqq Ch^{-1/2} \|\varphi\|_{L_2(\Gamma)} .$$

Proof. Consider the "boundary strip" Ω_h of Ω like in the proof of Lemma 2.3. We may write

$$\Omega_h = \bigcup_{j=1}^{J} \Omega_h^j$$

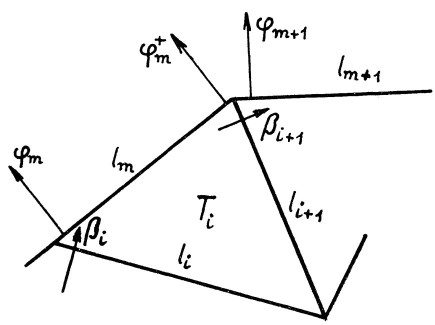

Fig. 1

where Ω_h^j is adjacent to the polygon $\partial\Omega_j$. We determine $\mathbf{w}^h \in \mathscr{N}_h$ by means of properly chosen flux parameters β on the sides of $\mathscr{T}_h$, such that

$$\operatorname{supp} \mathbf{w}^h \in \Omega_h .$$

To this end we consider the strip Ω_h^j. On $\partial\Omega_j$ we choose the flux parameters equal to the corresponding values of φ and equal to zero on $\partial\Omega_h^j - \partial\Omega_j$. As the sides connecting vertices of $\partial\Omega_j$ and $\partial\Omega_h^j - \partial\Omega_j$ are concerned, we set $\beta_k = 0$ at the "internal" vertices belonging to $\partial\Omega_h^j - \partial\Omega_j$, but the parameters β_i at the external vertices on $\partial\Omega_j$ remain to be determined. To each of these sides $l_1, l_2, \ldots, l_n$, a parameter β_i is attached, $i = 1, \ldots, n$ (see Fig. 1).

First assume that each $T_i \in \Omega_h^j$ has at most one side on $\partial\Omega_j$. The conditions of the form (4.9) and (4.10) generate a system of n equations

(5.15) $$\boldsymbol{a}\beta = \boldsymbol{b}\,,$$

where

$$a_{ii} = -l_i\,, \quad i = 1, 2, \ldots, n\,,$$

$$a_{i,i+1} = l_{i+1}\,, \quad i = 1, \ldots, n-1\,,$$

$$a_{n1} = l_1$$

and the remaining entries of the matrix $\boldsymbol{a}$ vanish. Furthermore,

$$b_i = -l_m(\varphi_m + \overset{+}{\varphi}_m)\,,$$

or $b_i = 0$ if $T_i \cap \partial\Omega_j$ is a vertex only. The assumption (5.13) implies that

$$\sum_{i=1}^{n} b_i = 0\,.$$

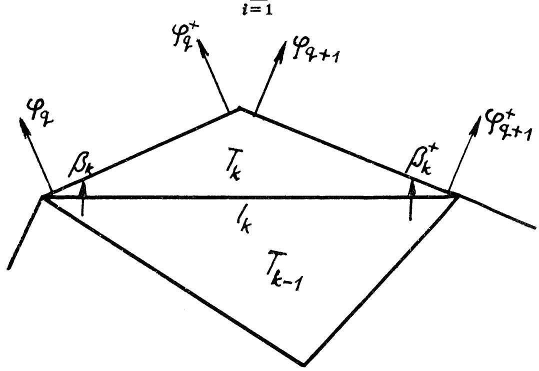

Fig. 2

Consequently, the last equation can be omitted. If we choose $\beta_1 = 0$, the remaining system has the following solution

(5.16) $$\beta_i = l_i^{-1}\sum_{p=1}^{i-1} b_p\,, \quad i = 2, 3, \ldots, n\,.$$

Second, let a triangle $T_k \in \Omega_h^j$ have two sides l_q and l_{q+1} on $\partial\Omega_j$ (see Fig. 2). We obtain the equation

$$(\beta_k^+ + \beta_k)\, l_k = -l_q(\varphi_q + \overset{+}{\varphi}_q) - l_{q+1}(\varphi_{q+1} + \overset{+}{\varphi}_{q+1})\,.$$

Then e.g. $\beta_k^+ = 0$ may be chosen and β_k calculated. The remaining system has again the form (5.15). It looks like the system for a "truncated" triangulation of $\partial\Omega_h^j - T_k$, with β_k and β_k^+ being considered as given outward flux parameters.

Using (5.16) and the α-β-regularity of $\mathscr{T}_h$, we obtain

$$(5.71) \qquad |\beta_i| = 2l_i^{-1}\left|\int_0^{S_i} \varphi \, ds\right| \leqq 2h_{\min}^{-1} \int_{\partial\Omega_j} |\varphi| \, ds \leqq Ch^{-1}\|\varphi\|_{L_2(\Gamma)},$$

$$(i = 1, \ldots, n\,; \quad j = 1, \ldots, M)\,.$$

The same estimate is true for β_k and the other parameters in case of Fig. 2. The estimate (5.17) is true also for the boundary parameters φ_m, φ_m^+. In fact, we have

$$\int_0^{l_m} \varphi^2 \, ds = (l_m/6)\left[\varphi_m^2 + (\varphi_m^+)^2 + (\varphi_m + \varphi_m^+)^2\right] \geqq (l_m/6)\left(|\varphi_m|^2 + |\varphi_m^+|^2\right),$$

consequently

$$|\varphi_m|^2 + |\varphi_m^+|^2 \leqq 6l_m^{-1}\|\varphi\|_{L_2(\Gamma)}^2 \leqq Ch^{-1}\|\varphi\|_{L_2(\Gamma)}^2\,.$$

Let $T \subset \Omega_h^j$ be an arbitrary triangle. From (5.17) we derive that

$$|w_k^h(Q)| \leqq (\sin\alpha)^{-1}\, Ch^{-1}\|\varphi\|_{L_2(\Gamma)}$$

holds for any vertex Q of T. Consequently, we obtain the same estimate for $|w_k^h(\mathbf{x})|$ at any $\mathbf{x} \in \Omega_h$ and therefore

$$\|\mathbf{w}^h\|^2 = \sum_{k=1}^{2} \int_{\Omega_h} (w_k^h)^2 \, d\mathbf{x} \leqq Ch^{-2}\|\varphi\|_{L_2(\Gamma)}^2 \operatorname{mes} \Omega_h \leqq Ch^{-1}\|\varphi\|_{L_2(\Gamma)}^2,$$

because $\operatorname{mes} \Omega_h \leqq Ch$. Q.E.D.

Proof of Theorem 5.1. Let ψ_h be the one-sided approximation from Lemma 5.2. We set

$$\varphi = (r_h U)\,.\,\nu - \psi_h$$

and consider the "extension" $\mathbf{w}^h \in \mathscr{N}_h$ from Lemma 5.3. Then the function $\mathbf{W}_h = r_h\mathbf{U} - \mathbf{w}^h$ satisfies the conditions of the Theorem. In fact, we have $\mathbf{W}_h \in \mathscr{N}_h$ and

$$\mathbf{W}_h\,.\,\nu = (r_h\mathbf{U})\,.\,\nu - \varphi = \psi_h$$

on Γ, therefore

$$0 \leqq \mathbf{W}_h\,.\,\nu \leqq \mathbf{U}\,.\,\nu$$

holds on Γ_a. Consequently, $\mathbf{W}_h \in \mathscr{C}_h$. Moreover,

$$\operatorname{div}(2\mathbf{U} - \mathbf{W}_h) = 0 \quad \text{in} \quad \Omega\,,$$

$$(2\mathbf{U} - \mathbf{W}_h)\,.\,\nu \geqq (\mathbf{U} - \mathbf{W}_h)\,.\,\nu \geqq 0 \quad \text{on} \quad \Gamma_a$$

and we conclude that $2\mathbf{U} - \mathbf{W}_h \in \mathscr{C}$. Using (4.3), Lemma 5.3 and 5.2, we obtain

$$\|\mathbf{U} - \mathbf{W}_h\| - \|\mathbf{U} - r_h\mathbf{U} + \mathbf{w}^h\| \leqq \|\mathbf{U} - r_h\mathbf{U}\| + \|\mathbf{w}^h\| \leqq$$
$$\leqq C(h^2|\mathbf{U}|_2 + h^{-1/2}\|\varphi\|_{L_2(\Gamma)}) \leqq$$
$$\leqq C\{h^2|\mathbf{U}|_2 + h^{3/2}\sum_{m=1}^{M}|\mathbf{U}\,.\,\nu|_{H^2(\Gamma_m \cap \Gamma_a)}\}\,. \qquad \text{Q.E.D}$$

Corollary 5.1. *Let the assumptions of Theorem* 5.1. *and* (A 1) – (A 4) *be satisfied. Then*

$$\|\lambda^0 - \lambda^h\| = \|\mathbf{q}^0 - \mathbf{q}^h\| = O(h^{3/2})\,.$$

The proof follows from Lemma 5.1, Theorem 5.1 and the estimate (4.3) applied to $\mathbf{G}$.

Corollary 5.2. *Let β^m be an iterative solution of* (4.8) *such that* (4.11) *holds. Define $\lambda^{h,m} = \bar{\lambda} + \mathbf{q}(\beta^m)$. Then*

(5.18) $$\lim \|\lambda^0 - \lambda^{h,m}\| = 0 \quad \text{for} \quad h \to 0,\ m \to \infty$$

if the assumptions of Theorem 5.1 *and* (A 1)–(A 4) *are satisfied.*

Proof. We may write $\mathbf{q}^h = \mathbf{q}(\hat{\beta})$ and using (4.6), (4.11),

(5.19) $$\|\mathbf{q}(\hat{\beta}) - \mathbf{q}(\beta^m)\|^2 = (\hat{\beta} - \beta^m)^\mathrm{T}\,\mathbf{A}(\hat{\beta} - \beta^m) \leqq \|\mathbf{A}\|\,\|\hat{\beta} - \beta^m\|^2 \to 0 \quad \text{for} \quad m \to \infty\,.$$

Since $\lambda^0 - \lambda^{h,m} = \mathbf{q}^0 - \mathbf{q}(\beta^m) = \mathbf{q}^0 - \mathbf{q}^h + \mathbf{q}(\hat{\beta}) - \mathbf{q}(\beta^m)$, by virtue of Corrolary 5.1 and (5.18)

$$\|\lambda^0 - \lambda^{h,m}\| \leqq \|\mathbf{q}^0 - \mathbf{q}^h\| + \|\mathbf{q}(\hat{\beta}) - \mathbf{q}(\beta^m)\| \to 0$$

provided $h \to 0$ and $m \to \infty$.

6. A POSTERIORI ERROR ESTIMATES AND TWO-SIDED BOUNDS OF ENERGY.

On the basis of the dual analysis described above we can deduce some a posteriori estimates of errors, like in the theory of classical-bilateral-boundary value problems.

Consider the problem $\mathscr{P}_1$ and its variational formulation (1.6). The solution u of (1.6) satisfies the variational inequality

(6.1) $$(u, v - u)_1 \geqq (f, v - u)_0 \quad \forall v \in K_1\,,$$

where $(\cdot, \cdot)_1$ denotes the scalar product associated with the norm $\|\cdot\|_1$. Let $v \in K_1$ be arbitrary. Then (6.1) implies

$$(6.2) \qquad 2[\mathscr{L}_1(v) - \mathscr{L}_1(u)] = \|v\|_1^2 - \|u\|_1^2 - 2(f, v - u)_0 \geqq$$
$$\geqq \|v\|_1^2 - \|u\|_1^2 - 2(u, v - u)_1 = \|v - u\|_1^2 .$$

The term $\mathscr{L}_1(u)$, however, is unknown in general and a lower bound for it is needed. From the duality theory it follows (cf. [5] chpt. 5, § 3) that

$$(6.3) \qquad \mathscr{L}_1(u) = \operatorname*{Max}_{\substack{\mu \in H^{-1/2}(\Gamma) \\ \mu \geqq 0}} \; \operatorname*{Min}_{v \in H^1(\Omega)} \{\mathscr{L}_1(v) - \langle \mu, \gamma v \rangle\} ,$$

consequently,

$$(6.4) \qquad \mathscr{L}_1(u) \geqq \operatorname*{Min}_{v \in H^1(\Omega)} \{\mathscr{L}_1(v) - \langle \mu, \gamma v \rangle\} \; \forall \mu \in H^{-1/2}(\Gamma) , \quad \mu \geqq 0 .$$

The minimum problem corresponds with the Neumann's problem

$$- \Delta u + u = f \quad \text{in} \quad \Omega , \quad \partial u / \partial \nu = \mu \quad \text{on} \quad \Gamma .$$

Denoting u_μ the solution of the latter problem, i.e. the minimizing element of $\mathscr{L}_1(v) - \langle \mu, \gamma v \rangle$ over $H^1(\Omega)$, we may again use the dual approach with the complementary energy $\mathscr{S}_1(\lambda)$ to get the lower bound for the potential energy (cf. [3]). Thus

$$(6.5) \qquad \mathscr{L}_1(u_\mu) - \langle \mu, \gamma u_\mu \rangle = \tfrac{1}{2}\|u_\mu\|_1^2 = -\mathscr{S}_1(\lambda(\mu)) ,$$

and introducing the set of admissible vector-functions

$$(6.6) \quad \Lambda_{f,\mu} = \{\lambda \mid \lambda \in [L_2(\Omega)]^{n+1}, \; (\lambda_1, \ldots, \lambda_n) \in Q, \; \lambda_{n+1} = f + \operatorname{div} \lambda, \; \lambda \,.\, \nu|_\Gamma = \mu\},$$

we can use the principle of minimum complementary energy

$$\mathscr{S}_1(\lambda(\mu)) \leqq \mathscr{S}_1(\lambda) \quad \forall \lambda \in \Lambda_{f,\mu}$$

and (6.4), (6.5) to obtain

$$(6.7) \qquad \mathscr{L}_1(u) \geqq -\mathscr{S}_1(\lambda) \quad \forall \lambda \in \Lambda_{f,\mu} .$$

From (6.2) and (6.7) we conclude that

$$(6.8) \qquad \tfrac{1}{2}\|v - u\|_1^2 = \mathscr{L}_1(v) + \mathscr{S}_1(\lambda)$$

holds for any $v \in K_1$ and $\lambda \in \Lambda_{f,\mu}$, where $\mu \in H^{-1/2}(\Gamma)$, $\mu \geqq 0$ is arbitrary.

The estimate (6.8) can be applied to the finite element approximations $u_h \in K_{1h} \subset K_1$. It remains, however, to choose μ and λ properly. To this end, we employ the finite element approximations to the dual problem.

Lemma 6.1. *Let $u_h \in K_{1h}$ and $q^h \in \mathscr{U}_{0h}$ be the finite element approximations to the primary problem (1.6) and to the dual problem (3.1), respectively. Then*

$$(6.9) \qquad \tfrac{1}{2}\|u_h - u\|_1^2 \leqq \mathscr{L}_1(u_h) + I(\boldsymbol{q}^h) + \tfrac{1}{2}\|f\|_0^2 ,$$

where u is the solution of (1.6) and I is the funvtional defined in (3.1').

Proof. From (6.3), (6.5) it follows

$$(6.10) \qquad \mathscr{L}_1(u) = \max_{\substack{\mu \in H^{-1/2}(\Gamma) \\ \mu \geqq 0}} (-\mathscr{S}_1(\lambda(\mu)) = -\min_{\mu} \mathscr{S}_1(\lambda(\mu)) =$$

$$= -\min_{\mu} \min_{\lambda \in \Lambda_{f,\mu}} \mathscr{S}_1(\lambda) = -\min_{\lambda \in \mathscr{U}_1} \mathscr{S}_1(\lambda) = -\mathscr{S}_1(\lambda^0) ,$$

where $\lambda^0 \in U_1$ is the solution of (1.11), because

$$\bigcup_{\substack{\mu \in H^{-1/2}(\Gamma) \\ \mu \geqq 0}} \Lambda_{f,\mu} = \mathscr{U}_1 .$$

Setting $\lambda^0 = [q_1^0, q_2^0, f + \operatorname{div} \boldsymbol{q}^0]$, we obtain

$$(6.11) \qquad \mathscr{S}_1(\lambda^0) = I(\boldsymbol{q}^0) + \tfrac{1}{2}\|f\|_0^2 \leqq I(\boldsymbol{q}^h) + \tfrac{1}{2}\|f\|_0^2 = \mathscr{S}_1(\lambda^h) .$$

Consequently, (6.2), (6.10) and (6.11) result in

$$(6.12) \qquad \tfrac{1}{2}\|u_h - u\|_1^2 \leqq \mathscr{L}_1(u_h) - \mathscr{L}_1(u) = \mathscr{L}_1(u_h) + \mathscr{S}_1(\lambda^0) \leqq$$

$$\leqq \mathscr{L}_1(u_h) + \mathscr{S}_1(\lambda^h) = \mathscr{L}_1(u_h) + I(\boldsymbol{q}^h) + \tfrac{1}{2}\|f\|_0^2 . \qquad \text{Q.E.D.}$$

Remark 6.1. The finite element approximation u_h is not known, in general. Moreover, a difference of great numbers may occur on the right-hand side of (6.9). Therefore we transform (6.9) as follows (cf. also Remark 6.2).

Theorem 6.1. *Let $\tilde{u}_h \in K_{1h}$ be any approximation to the primary problem (1.6) and $\boldsymbol{q}^h \in U_{0h}$ a finite element approximation to the dual problem (3.1). Then*

$$(6.13) \qquad \|\tilde{u}_h - u\|_1^2 \leqq \sum_{i=1}^{2} \|q_i - \partial \tilde{u}_h/\partial x_i\|_0^2 + \|f + \operatorname{div} \boldsymbol{q}^h - \tilde{u}_h\|_0^2 +$$

$$+ 2 \int_\Gamma \boldsymbol{q}^h . \nu \tilde{u}_h \, \mathrm{d}s = E(\boldsymbol{q}^h, \tilde{u}_h) .$$

Proof. For $\lambda^h = [q_1^h, q_2^h, f + \operatorname{div} \boldsymbol{q}^h]$ we deduce, using (6.12), that

$$\tfrac{1}{2}\|\tilde{u}_h - u\|_1^2 \leqq \mathscr{L}_1(\tilde{u}_h) + \mathscr{S}_1(\lambda^h) = \tfrac{1}{2}\|\tilde{u}_h\|_1^2 - (f, u_h)_0 + \sum_{i=1}^{3} \tfrac{1}{2}\|\lambda_i^h\|_0^2 =$$

$$= \tfrac{1}{2} \sum_{i=1}^{3} \{\|\lambda_i^h - \lambda_i(\tilde{u}_h)\|_0^2 + 2(\lambda_i^h, \lambda_i(\tilde{u}_h))_0\} - (f, \tilde{u}_h)_0 ,$$

where $\lambda_i(v) = \partial v/\partial x_i$, $(i = 1, 2)$ and $\lambda_3(v) = v$.

The definition (1.10) implies that

$$\sum_{i=1}^{3} (\lambda_i^h, \lambda_i(\tilde{u}_h)) - (f, u_h)_0 = \int_\Omega [\lambda^h \,.\, \text{grad}\, \tilde{u}_h + (\lambda_3^h - f)\, \tilde{u}_h] \, \mathrm{d}x =$$

$$= \langle \lambda^h \,.\, v, \gamma \tilde{u}_h \rangle = \int_\Gamma \boldsymbol{q}^h \,.\, v \tilde{u}_h \, \mathrm{d}s \geqq 0 \,. \qquad \text{Q.E.D.}$$

Remark 6.2. All terms of the upper bound in (6.13) are non-negative. For $\tilde{u}_h$ any approximate solution

$$\tilde{u}_h = \sum_{j=1}^{M} v_j^m \varphi_j$$

(see Section 2), obtained by means of the iterative Gauss-Seidel algorithm with constraints, can be substituted.

The "exact" q^h can be replaced by any

$$\boldsymbol{q}^{hm} = \sum_{j=1}^{N} y_j^m \mathbf{w}^j$$

if $\boldsymbol{q}^{hm} \in \mathscr{U}_{0h}$, i.e., if $\mathbf{y} \in Y$ (cf. (3.4)). The algorithm of Uzawa, however, fails to satisfy this requirement. Therefore, if the a posteriori error bounds are needed, we recommend to apply e.g. some of the procedures called "methods of feasible directions" by Zoutendijk [9].

Theorem 6.2. *Let* $\tilde{u}_h$ *and* $\boldsymbol{q}^h$ *be the same as in Theorem* 6.1. *Then it holds*

$$(6.14) \qquad -2\mathscr{L}_1(u_h) \leqq \|u\|_1^2 \leqq \sum_{i=1}^{2} \|q_i^h\|_0^2 + \|f + \text{div}\, \boldsymbol{q}^h\|_0^2 \equiv F(\boldsymbol{q}^h)\,,$$

$$-2\mathscr{L}_1(u_h) \leqq (f, u)_0 \leqq F(\boldsymbol{q}^h)\,.$$

Proof. Inserting $v = 0$ and $v = 2u$ into (6.1), we obtain

$$(6.15) \qquad \|u\|_1^2 = (f, u)_0\,.$$

Consequently, we have

$$(6.16) \qquad 2\mathscr{L}_1(u) = \|u\|_1^2 - 2(f, u)_0 = -\|u\|_1^2\,.$$

Therefore the left-hand inequality of (6.14) follows from

$$\mathscr{L}_1(\tilde{u}_h) \geqq \mathscr{L}_1(u_h) \geqq \mathscr{L}_1(u) \quad \forall u_h \in K_{1h}\,,$$

where u_h and u are the f.e. approximation and the solution of the primary problem (1.6), respectively.

Finally, (6.10), (6.11), (6.16) and (3.1′) lead to the inequality

$$\|u\|_1^2 = -2\mathscr{L}_1(u) = 2\mathscr{S}_1(\lambda^0) \leqq 2I(\boldsymbol{q}^h) + \|f\|_0^2 =$$

$$= \sum_{i=1}^{2} \|q_i^h\|_0^2 + \|\operatorname{div} \boldsymbol{q}^h\|_0^2 + 2(f, \operatorname{div} \boldsymbol{q}^h)_0 + \|f\|_0^2 = F(\boldsymbol{q}^h).$$

The second assertion follows from the first by virtue of (6.15).

Q.E.D.

Remark 6.3. Similar comments are valid for the practical use of (6.14) as for (6.13). The right-hand side $F(\boldsymbol{q}^{hm})$ can be calculated on the basis of a method of feasible directions.

Let us derive an a posteriori estimate of error for the finite element approximations to the dual problem (3.1). The solution λ^0 of the dual problem (1.11) satisfies the inequality

$$\sum_{i=1}^{3} (\lambda_i^0, \lambda_i - \lambda_i^0)_0 \geqq 0 \quad \forall \lambda \in \mathscr{U}_1 .$$

Therefore we may write

$$2[\mathscr{S}_1(\lambda) - \mathscr{S}_1(\lambda^0)] = \sum_{i=1}^{3} (\|\lambda_i\|_0^2 - \|\lambda_i^0\|_0^2) \geqq \tag{6.17}$$

$$\geqq \sum_{i=1}^{3} [\|\lambda_i\|_0^2 - (\lambda_i^0, \lambda_i)_0] = \sum_{i=1}^{3} [(\lambda_i, \lambda_i - \lambda_i^0)_0 -$$

$$- (\lambda_i^0, \lambda_i - \lambda_i^0)_0 + (\lambda_i^0, \lambda_i - \lambda_i^0)_0] \geqq \sum_{i=1}^{3} \|\lambda_i - \lambda_i^0\|_0^2 \quad \forall \lambda \in \mathscr{U}_1 .$$

From (6.10) we deduce

$$-\mathscr{S}_1(\lambda^0) = \mathscr{L}_1(u) \leqq \mathscr{L}_1(\tilde{u}_h) \quad \forall \tilde{u}_h \in K_{1h} .$$

Substituting the vector $\lambda^h = [q_1^h, q_2^h, f + \operatorname{div} \boldsymbol{q}^h]$ for $\lambda \in \mathscr{U}_1$, we obtain

$$\tfrac{1}{2} \sum_{i=1}^{3} \|\lambda_i - \lambda_i^0\|_0^2 \leqq \mathscr{L}_1(\tilde{u}_h) + \mathscr{S}_1(\lambda^h) = \tfrac{1}{2} E(\boldsymbol{q}^h, \tilde{u}_h)$$

like in the proof of Theorem 6.1. Thus we come to the following .

Remark 6.4. Let $\tilde{u}_h$, $\boldsymbol{q}^h$ and $E(\boldsymbol{q}^h, \tilde{u}_h)$ be the same as in Theorem 6.1. Then it holds

$$\sum_{i=1}^{3} \|\lambda_i^h - \lambda_i^0\|_0^2 \leqq D(\boldsymbol{q}^h, \tilde{u}_h) \tag{6.18}$$

for $\lambda^h = [q_1^h, q_2^h, f + \operatorname{div} \boldsymbol{q}^h]$.

The situation with the Problem $\mathscr{P}_2$ is slightly more complicated, as far as the duality is concerned.

First, an anlogue of (6.2) is true, i.e.,

$$2[\mathscr{L}_2(v) - \mathscr{L}_2(u)] \geqq |v - u|_1^2 \quad \forall v \in K_2 , \tag{6.19}$$

where u is the solution of (1.7).

To derive a corresponding saddle point theorem, we introduce the following subspace of $H^{1/2}(\Gamma_a)$:

$$H_0^{1/2}(\Gamma_a) = \{s \mid s \in H^{1/2}(\Gamma_a),\ \exists \sigma \in H^{1/2}(\Gamma),\ \sigma|_{\Gamma_u} = 0,\ \sigma|_{\Gamma_a} = s\} .$$

We define the linear functionals $\varphi \in [H_0^{1/2}(\Gamma_a)]'$ and say that $\varphi|_{\Gamma_a} \geqq 0$ if

$$\langle \varphi, s\rangle \geqq 0 \quad \forall s \in H_0^{1/2}(\Gamma_a) , \quad s \geqq 0 .$$

There exists a $\lambda \in [H_0^{1/2}(\Gamma_a)]'$, $\lambda|_{\Gamma_a} \geqq 0$ such that

$$\mathscr{L}_2(u) - \langle \mu, \gamma u\rangle \leqq \mathscr{L}_2(u) - \langle \lambda, \gamma u\rangle \leqq \mathscr{L}_2(v) - \langle \lambda, \gamma v\rangle \tag{6.20}$$

holds for all $\mu \in [H_0^{1/2}(\Gamma_a)]'$, $\mu|_{\Gamma_a} \geqq 0$ and all $v \in H^1(\Omega)$, $\gamma v|_{\Gamma_u} = 0$. The latter assertion can be proved on the basis of the Corollary of Hahn-Banach Theorem (see the proof of Lemma 1.1).

Denote $H_+^{-1/2}(\Gamma_a)$ the set of all admissible μ in (6.20) and V the set of admissible v. Then it holds

$$\mathscr{L}_2(u) = \operatorname*{Max}_{\mu \in H_+^{-1/2}(\Gamma_a)} \operatorname*{Min}_{v \in V} [\mathscr{L}_2(v) - \langle \mu, \gamma v\rangle] . \tag{6.21}$$

The minimum problem corresponds with the following mixed boundary value problem

$$-\Delta u = f \quad \text{in} \quad \Omega , \quad u|_{\Gamma_u} = 0 , \quad \partial u/\partial \nu|_{\Gamma_a} = \mu .$$

Denoting u_μ the solution of the latter problem, we may use the dual approach with the complementary energy $\mathscr{S}_2(\lambda)$ (cf.[3]) to obtain from (6.21) that

$$\begin{aligned}\mathscr{L}_2(u) &= \operatorname*{Max}_{\mu \in H_+^{-1/2}(\Gamma_a)} \mathscr{L}_2(u_\mu) = \operatorname*{Max}_{\mu} (-\mathscr{S}_2(\lambda(u_\mu))) = \\ &= -\operatorname*{Min}_{\mu} \operatorname*{Min}_{\lambda \in \Lambda_{f,\mu}} \mathscr{S}_2(\lambda) = -\operatorname*{Min}_{\lambda \in \mathscr{U}_2} \mathscr{S}_2(\lambda) = -\mathscr{S}_2(\lambda^0) ,\end{aligned} \tag{6.22}$$

where

$$\Lambda_{f,\mu} = \{\lambda \in Q \mid \operatorname{div} \lambda + f = 0,\ \lambda \cdot \nu|_{\Gamma_a} = \mu\} ,$$

so that

$$\bigcup_{\mu \in H_+^{-1/2}(\Gamma_a)} \Lambda_{f,\mu} = \mathscr{U}_2 .$$

Obviously, we have

$$\mathscr{S}_2(\lambda^0) \leqq \mathscr{S}_2(\lambda) \quad \forall \lambda \in \mathscr{U}_2 . \tag{6.23}$$

Unfortunately, the f.e. approximations $\boldsymbol{q}^h \in K_h$ (cf. (4.5)) are such that $\lambda^h = \bar{\lambda} + \boldsymbol{q}^h \notin \mathscr{U}_2$, in general. The condition $\bar{\lambda} + \boldsymbol{q}^h \in \mathscr{U}_2$ is satisfied, if $\bar{\lambda} . v \in P_1(S_k)$ for each side $S_k \subset \Gamma_a$ of $\mathscr{T}_h$. Then $g = -\bar{\lambda} . v = g_h$.

Remark 6.5. To obtain f.e. approximations $\lambda^h \in \mathscr{U}_2$ in general, we can attempt to find $\lambda^f \in Q$ such that

$$\text{div}\, \lambda^f + f = 0\,, \quad \lambda^f . v|_{\Gamma_a} = 0\,. \tag{6.24}$$

Then replacing $\bar{\lambda}$ by λ^f, we obtain $g = -\lambda^f . v = 0 = g_h$ on Γ_a, $e_j(\lambda^f) = 0$ for all $j > p_1$ (cf. (4.8) and the definition of $\mathscr{B}$) and $\lambda^h = \lambda^f + \boldsymbol{q}^h \in \mathscr{U}_2$.

There is an approach for the search of λ^f. To any $\omega \in H^2(\Omega)$ the vector $\lambda^\omega = \{-\partial\omega/\partial x_2, \partial\omega/\partial x_1\}$ satisfies the equation div $\lambda^\omega = 0$. Consequently, if we find $\omega \in H^2(\Omega)$ such that the trace of ω satisfies

$$\omega(s) = -\int_{s_0}^{s} (\bar{\lambda} . v)(t)\, \mathrm{d}t \quad \forall s \in \Gamma_a\,, \tag{6.25}$$

we obtain

$$\frac{\mathrm{d}\omega}{\mathrm{d}s} = -\frac{\partial\omega}{\partial x_2} v_1 + \frac{\partial\omega}{\partial x_1} v_2 = \lambda^\omega . v = -\bar{\lambda} . v \quad \text{on} \quad \Gamma_a\,.$$

Therefore $\lambda^f = \bar{\lambda} + \lambda^\omega$ satisfies (6.24).

Theorem 6.3. *Let $\tilde{u}_h \in K_{2h}$ be an approximation to the primary problem* (1.7) *and $\lambda^{hm} = \bar{\lambda} + \boldsymbol{q}^{hm} \in \mathscr{U}_2$ (or $\lambda^f + \boldsymbol{q}^{hm} \in \mathscr{U}_2$ – see Remark* 6.5) *an approximation to the dual problem* (1.12). *Then*

$$|u - \tilde{u}_h|_1^2 \leqq \sum_{i=1}^{2} \|\lambda_i^{hm} - \partial\tilde{u}_h/\partial x_i\|_0^2 + 2\int_{\Gamma_a} \lambda^{hm} . v\tilde{u}_h\, \mathrm{d}s = E(\lambda^{hm}, \tilde{u}_h)\,. \tag{6.26}$$

Proof is analogous to that of Theorem 6.1. It follows from (6.19), (6.22) and (6.23).

Remark 6.6. All terms in the right-hand side of (6.26) are non-negative. For $\boldsymbol{q}^{hm}$ we may substitute any iteration $\boldsymbol{q}(\beta^m)$ calculated by e.g. a method of feasible direction [9] but not by the Uzawa's algorithm, as the latter may fail to keep β^m in the set $\mathscr{B}$.

Theorem 6.4. *Let $\tilde{u}_h$ and λ^{hm} be the same as in Theorem* 6.3 *and Remark* 6.6. *Then it holds*

$$-2\mathscr{L}_2(\tilde{u}_h) \leqq |u|_1^2 \leqq 2\mathscr{S}_2(\lambda^{hm}) = \sum_{i=1}^{2} \|\lambda_i^{hm}\|_0^2\,,$$

$$-2\mathscr{L}_2(\tilde{u}_h) \leqq (f, u)_0 \leqq \sum_{i=1}^{2} \|\lambda_i^{hm}\|_0^2\,.$$

The proof is analogous to that of Theorem 6.2.

Remark 6.7. Let $\tilde{u}_h$, λ^{hm}, $E(\lambda^{hm}, \tilde{u}_h)$ be the same as in Theorem 6.3. Then it holds (cf. Remark 6.4)

$$\sum_{i=1}^{2} \|\lambda_i^{hm} - \lambda_i^0\|_0^2 \leqq E(\lambda^{hm}, \tilde{u}_h)\,.$$

Remark. Some results presented above can be employed in the dual finite element analysis of the unilateral problems with obstacles on the boundary [10].

Acknowledgement. The author is grateful to J. Haslinger, CSc. for his careful scrutiny of the manuscript and a number of valuable suggestions.

References

[1] *J. P. Aubin, H. G. Burchard:* Some aspects of the method of the hypercircle applied to elliptic variational problems. Num. sol. PDE-II, SYNSPADE (1970), 1—67.

[2] *I. Hlaváček:* Some equilibrium and mixed models in the finite element method. Proceedings of the St. Banach Internat. Math. Center, Warsaw, (1976).

[3] *J. Haslinger, I. Hlaváček:* Convergence of a finite element method based on the dual variational formulation. Apl. mat. 21 (1976), 43—65.

[4] *G. Fichera:* Boundary value problems of elasticity with unilateral constraints. Encyclopedia of Physics, ed. S. Flügge, Vol. VIa/2, Springer, Berlin 1972.

[5] *J. Céa:* Optimisation, théorie et algorithmes. Dunod, Paris 1971.

[6] *J. Nečas:* Les méthodes directes en théorie des équations elliptiques. Academia, Prague 1967.

[7] *U. Mosco, G. Strang:* One-sided approximations and variational inequalities. Bull. Am. Math. Soc. 80 (1974), 308—312.

[8] *J. H. Bramble, M. Zlámal:* Triangular elements in the finite element method. Math. Comp. 24 (1970), 809—820.

[9] *G. Zoutendijk:* Methods of feasible directions. Elsevier, Amsterdam 1960.

[10] *I. Hlaváček:* Dual finite element analysis for elliptic problems with obstacles on the boundary. Apl. mat. 22 (to appear).

Souhrn

DUÁLNÍ ANALÝZA JEDNOSTRANNÝCH OKRAJOVÝCH ÚLOH METODOU KONEČNÝCH PRVKŮ

IVAN HLAVÁČEK

Duální variační formulace okrajových úloh — principy minima potenciální resp. doplňkové energie — jsou rozšířeny na okrajové úlohy s nerovnostmi tzv. Signoriniho typu na hranici. Pomocí teorie sedlového bodu je dokázána souvislost obou variačních problémů. Jsou odvozeny algoritmy pro přibližné řešení obou úloh metodou konečných prvků na triangulaci dané oblasti s po částech lineárními polynomy, dále apriorní i aposteriorní odhady chyb a oboustranné odhady energie. K apriorním odhadům se používá tzv. jednostranných aproximací řešení resp. jeho toku na hranici, za předpokladu jisté regularity řešení.

Author's address: Ing. *Ivan Hlaváček*, CSc., Matematický ústav ČSAV, Žitná 25, 115 67 Praha 1.