Michal Křížek

On approximation of the Neumann problem by the penalty method

Applications of Mathematics, Vol. 38 (1993), No. 6, 459–469

Persistent URL: http://dml.cz/dmlcz/104568

ON APPROXIMATION OF THE NEUMANN PROBLEM BY THE PENALTY METHOD

MICHAL KŘÍŽEK, Praha

Summary. We prove that penalization of constraints occurring in the linear elliptic Neumann problem yields directly the exact solution for an arbitrary set of penalty parameters. In this case there is a continuum of Lagrange's multipliers. The proposed penalty method is applied to calculate the magnetic field in the window of a transformer.

Keywords: Neumann problem, penalty method, finite elements, magnetic field

AMS classification: 65N30, 35J50

1. INTRODUCTION

The necessity of solving a linear elliptic problem with the Neumann boundary conditions arises in many branches, e.g., in modelling transonic flows [5], in theories of shells [10] and elasticity [11], in electrical engineering [4]. For simplicity, consider first the following model Neumann problem:

$$\begin{aligned} -\Delta u &= f \quad \text{in } \Omega, \\ \frac{\partial u}{\partial \nu} &= 0 \quad \text{on } \partial\Omega, \end{aligned} \tag{1.1}$$

where $\Omega \subset R^d$ ($d \geqslant 1$) is a bounded domain with a Lipschitz boundary $\partial\Omega$, ν is the outer unit normal to $\partial\Omega$, and $f \in L^2(\Omega)$ is such that $\int_\Omega f \,\mathrm{d}x = 0$.

The associated bilinear form

$$a(v, w) = \int_\Omega \nabla v \cdot \nabla w \,\mathrm{d}x, \quad v, w \in H^1(\Omega), \tag{1.2}$$

is not $H^1(\Omega)$-elliptic, where $H^1(\Omega)$ is the Sobolev space of functions whose generalized derivatives belong to $L^2(\Omega)$. The nonellipticity causes the nonuniqueness of the true solution which often represents some troubles in numerical solution of (1.1). There are several ways how to handle the Neumann problem:

1. A natural approach is to use a dual variational formulation (see e.g. [11, p. 95]), since it has a unique solution. However, this formulation does not produce the

solution u but only its cogradient. Moreover, a particular solution of equilibrium equations has to be known a priori.

2. The standard variational formulation of (1.1) is usually given (see [3, p. 33]) in the quotient space $H^1(\Omega)/P_0(\Omega)$, in which $a(.,.)$ is elliptic. (Here $P_0(\Omega)$ stands for the space of constant functions.) But quotient spaces are unsuitable for finite element approximations.

3. The bilinear form (1.2) is elliptic also in the space

$$W = \left\{ v \in H^1(\Omega) \mid \int_\Omega v \, dx = 0 \right\}.$$

However, to construct finite element fields which would belong to W is practically difficult (especially in problems with more than one constraint [8]).

4. To fix the solution u in one point is also not advisable if $d > 1$, since u can have a singularity just at this point. Moreover, in [1, Theorem 4.1], Babuška demonstrates why we cannot fix functions from the Sobolev space $H^1(\Omega)$ in one point. (Note that $H^1(\Omega)$ is not contained in $C(\overline{\Omega})$ for $d > 1$.) The associated finite element schemes are then unstable

5. A direct finite element approximation of (1.1) leads to a system of linear algebraic equations with a singular matrix

$$A = (a(v^i, v^j))_{i,j=1}^n,$$

where $\{v^i\}_{i=1}^n \subset H^1(\Omega)$ are finite element basis functions such that $\sum_i c^i v^i \equiv 1$ in Ω for some coefficients $c^1, \ldots, c^n$. Since A is singular and positive semidefinite, we have to use some special solvers—see e.g. [12, p. 117] for a modified conjugate gradient method.

In the next section, we introduce another approach to solving the Neumann problem which will be based on the penalty method (or Lagrange's multipliers method). The associated finite element approximations then yield positive definite stiffness matrices.

2. Penalty method for a general Neumann problem

Throughout the paper we shall use the following assumptions:

(A1) Let V be a real or complex Banach space equipped with the norm $||.||$ and let $a(.,.)$ be a sesquilinear Hermitian (i.e., bilinear symmetric in the real case) continuous form on $V \times V$.

(A2) Let m be a positive integer, let $F_1, \ldots, F_m$ be linear continuous forms over V and let

$$W = \{v \in V \mid F_i(v) = 0,\ i = 1, \ldots, m\}.$$

(A3) Let $a(.,.)$ be W-elliptic, i.e., there exists a constant $C > 0$ such that

$$a(w, w) \geqslant C\|w\|^2\ \forall w \in W.$$

(A4) Let $\dim P = m$, where

$$P = \{p \in V \mid a(p, v) = 0\ \forall v \in V\}.$$

(A5) Let $b(.)$ be a linear continuous form over V such that

$$b(p) = 0 \quad \forall p \in P.$$

We shall deal with the problem: Find a function u satisfying

$$u \in W, \quad a(w, u) = b(w) \quad \forall w \in W. \tag{2.1}$$

By the Riesz theorem (or the Lax-Milgram lemma) this problem has a unique solution. Note that the standard weak formulation of a linear elliptic problem with the Neumann boundary conditions is typically of the form (2.1). However, due to the constraints occurring in the definition of W, a direct numerical solution of the problem (2.1) can be difficult. In Theorem 2.1 we will show that the penalty method enable us to solve (2.1) easily.

Define a functional $J : V \to R^1$ by the formula

$$J(v) = a(v, v) - b(v) - \overline{b(v)},$$

where the bar denotes the conjugate number. The problem (2.1) is then equivalent to the following one: Find $u \in W$ such that

$$J(u) = \min_{w \in W} J(w). \tag{2.2}$$

The *penalty method* for the problem (2.2) consists in finding $u_\lambda \in V$ such that

$$J_\lambda(u_\lambda) = \min_{v \in V} J_\lambda(v), \tag{2.3}$$

where

$$J_\lambda(v) = J(v) + \sum_{i=1}^{m} \lambda_i |F_i(v)|^2,$$

and where $\lambda = (\lambda_1, \ldots, \lambda_m)$, $\lambda_i > 0$ for $i = 1, \ldots, m$. The function u_λ is a solution of the problem (2.3) if and only if u_λ is a solution of the problem: Find $u_\lambda \in V$ such that

$$a(v, u_\lambda) + \sum_{i=1}^{m} \lambda_i F_i(v) \overline{F_i(u_\lambda)} = b(v) \quad \forall v \in V.$$

It is well-known (see e.g. [2, 12, 15]) that $u_\lambda \to u$ in V when all $\lambda_i \to \infty$. However, in our case the situation is much more simple. The next theorem shows that for any choice of λ with positive components we even have

$$u = u_\lambda.$$

Theorem 2.1. *Let* (A1)–(A5) *hold and let* $\lambda_1, \ldots, \lambda_m$ *be arbitrary positive numbers. Then there exists one and only one* $u \in V$ *such that*

$$a(v, u) + \sum_{i=1}^{m} \lambda_i F_i(v) \overline{F_i(u)} = b(v) \quad \forall v \in V. \tag{2.4}$$

Moreover, the forms $F_1, \ldots, F_m$ *are linearly independent and the solution* u *satisfies the conditions*

$$F_i(u) = 0, \quad i = 1, \ldots, m, \tag{2.5}$$

i.e., $u \in W$. *The penalty method thus gives the exact solution of the problem* (2.1) *for an arbitrary set of positive penalty parameters.*

Proof. First of all we show that each $v \in V$ can be decomposed so that $v = w + p$ for some $w \in W$ and $p \in P$. So let $v \in V$ be arbitrary. According to (A1) and (A3), $a(.,.)$ represents a scalar product on W and thus, by the Riesz theorem, there exists precisely one $w \in W$ such that

$$a(y, w) = a(y, v) \quad \forall y \in W, \tag{2.6}$$

because $y \mapsto a(y, v)$ is a linear continuous form. Let us denote by Q the mapping $v \mapsto w$. It is clearly a linear projection operator. Hence, by [14, Chap. 4.8], the space V can be expressed as the direct sum $V = \mathrm{Im}(Q) \oplus \mathrm{Ker}(Q) = W \oplus \mathrm{Ker}(Q)$ and

$$\dim \mathrm{Ker}(Q) \leqslant m, \tag{2.7}$$

(since for the time being we do not know whether the forms $F_1, \ldots, F_m$ are linearly independent over V).

Setting

$$p = v - w, \tag{2.8}$$

we find, in virtue of (2.6), that

$$p \in \operatorname{Ker}(Q). \tag{2.9}$$

For any $q \in P$ we have $a(y,q) = 0$ for all $y \in W$ and thus $Q(q) = 0$, i.e., $q \in \operatorname{Ker}(Q)$ and $P \subset \operatorname{Ker}(Q)$. However, from (2.7) and the assumption $\dim P = m$, we see that

$$P = \operatorname{Ker}(Q). \tag{2.10}$$

Consequently, due to (2.9), we obtain $p \in P$.

Further, we will follow some ideas of [7]. The left-hand side of (2.4) obviously represents a sesquilinear Hermitian continuous form on $V \times V$. We show now that it is V-elliptic. For any $v \in V$ we have by (2.8), (A4), (A2) and (A3) that

$$\begin{aligned}
a(v,v) &+ \sum_{i=1}^{m} \lambda_i |F_i(v)|^2 \\
&= a(w,w) + a(p,w) + \overline{a(p,w)} + a(p,p) + \sum_{i=1}^{m} \lambda_i |F_i(w) + F_i(p)|^2 \\
&= a(w,w) + \sum_{i=1}^{m} \lambda_i |F_i(p)|^2 \geqslant C\|w\|^2 + C_1\|p\|^2 \geqslant C_2\|w+p\|^2 = C_2\|v\|^2,
\end{aligned}$$

as all norms in the finite dimensional space P are equivalent. Thus the existence of a unique solution of (2.4) follows again from the Riesz theorem.

Next we check the relation (2.5). According to (2.10), $\dim \operatorname{Ker}(Q) = m$ and thus the forms $F_1, \ldots, F_m$ are linearly independent over V. Therefore, we may choose for any $i \in \{1, \ldots, m\}$ such an element v_i from the space

$$\{v \in V \mid F_1(v) = \ldots = F_{i-1}(v) = F_{i+1}(v) = \ldots = F_m(v) = 0\}$$

for which $F_i(v_i) = 1$. Hence,

$$F_i(v_j) = \delta_{ij}, \quad i,j = 1, \ldots, m.$$

Putting

$$p_i = v_i - w_i, \quad i = 1, \ldots, m,$$

where $w_i = Q(v_i)$, we see that

$$F_i(p_j) = \delta_{ij}, \quad i, j = 1, \ldots, m.$$

From here, (A4), (2.4) and (A5) we get for each $j \in \{1, \ldots, m\}$ that

$$\begin{aligned}\lambda_j \overline{F_j(u)} &= \sum_{i=1}^{m} \lambda_i F_i(p_j) \overline{F_i(u)} \\ &= a(p_j, u) + \sum_{i=1}^{m} \lambda_i F_i(p_j) \overline{F_i(u)} = b(p_j) = 0.\end{aligned}$$

Therefore, $F_j(u) = 0$ for $j = 1, \ldots, m$, i.e., $u \in W$ and by (2.4) we obtain

$$a(w, u) = b(w) \quad \forall w \in W.$$

□

Remark 2.2. Let V be a real Banach space. The subspace W from (A2) may also be defined as follows

$$W = \{w \in V \mid (F_i(w))^2 = 0, \ i = 1, \ldots, m\}.$$

The functionals $v \mapsto J(v) = a(v, v) - 2b(v)$ and $v \mapsto (F_i(v))^2$ are continuously Fréchet differentiable in V. Hence, from (2.2) and well-known results [13] it follows that if u is the solution of the problem (2.1) then there exist *Lagrange's multipliers* $\lambda_1, \ldots, \lambda_m$ so that

$$a(v, u) + \sum_{i=1}^{m} \lambda_i F_i(v) F_i(u) = b(v) \quad \forall v \in V.$$

Comparing this with (2.2), we observe an interesting fact that there is a continuum of Lagrange's multipliers in the considered case. One can a priori choose any set $\lambda_1, \ldots, \lambda_m$ of positive numbers.

3. Applications

Example 3.1. Let us look for a complex vector $\hat{x}$ which minimizes the quadratic functional

$$J(x) = x^H A x - b^H x - x^H b$$

over the set of linear constraints

$$Bx = 0,$$

where x^H stands for the conjugate transposed vector to x, A is an $n \times n$ Hermitian matrix (i.e., $A^H = A$) which is positive semidefinite, B is an $m \times n$ complex rectangular matrix such that $1 \leqslant \operatorname{rank}(B) = m < n$, and $x^H A x > 0$ for all $x \neq 0$ for which $Bx = 0$. Finally, let $b^H p = 0$ for all $p \in P$, where P is the space of those complex vectors p such that $p^H A x = 0$ for all complex vectors x. In view of Theorem 2.1, we see that $\hat{x}$ is a unique solution of the problem

$$(A + B^H B)x = b.$$

For exact penalty methods in a finite dimensional space see also [16].

Example 3.2. We describe how to calculate a reduced magnetic potential u in the window of an ideal transformer (see Fig. 1).

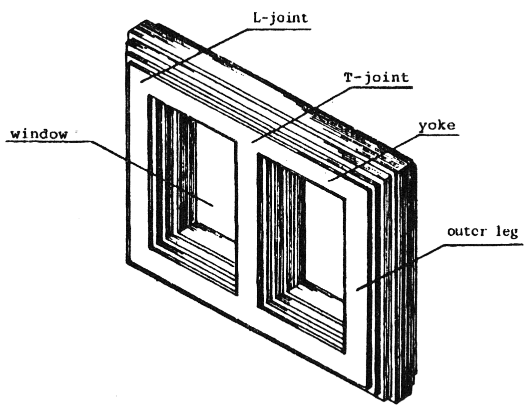

Fig. 1

From the stationary Maxwell equations in the standard cylindrical coordinates (r, θ, z), one can derive (under some simplifying assumptions—see [4]) that

$$-\frac{\partial}{\partial r}\Big(\frac{1}{r}\frac{\partial u}{\partial r}\Big) - \frac{\partial}{\partial z}\Big(\frac{1}{r}\frac{\partial u}{\partial z}\Big) = \mu_0 I \quad \text{in } \Omega, \tag{3.1}$$

where $\Omega = (r_1, r_2) \times (z_1, z_2)$, ($r_1 = 0.3$, $r_2 = 0.65$, $z_1 = 0$, $z_2 = 0.5$ [m]), is a rectangular domain which corresponds to the upper half of the window, $\mu_0 = 4\pi \cdot 10^{-7}$ [kg m/s^{2}A^2] is the vacuum permeability and I is the current density for which $\int_\Omega I \,\mathrm{d}r\,\mathrm{d}z = 0$. Since the permeability of the transformer magnetic core is much more greater than μ_0, we get the Neumann boundary condition

$$\frac{\partial u}{\partial \nu} = 0 \quad \text{on } \partial\Omega. \tag{3.2}$$

Define the associated forms as follows

$$a(v, w) = \int_\Omega \Big(\frac{\partial v}{\partial r}\frac{\partial w}{\partial r} + \frac{\partial v}{\partial z}\frac{\partial w}{\partial z}\Big)\frac{1}{r}\,\mathrm{d}r\,\mathrm{d}z,$$

$$b(v) = \mu_0 \int_\Omega I v \,\mathrm{d}r\,\mathrm{d}z$$

for $v, w \in V = H^1(\Omega)$, where

$$I(r, z) = \begin{cases} 10^6 & \text{in } \Omega_1, \\ -10^6 & \text{in } \Omega_2, \\ 0 & \text{in } \Omega \setminus (\Omega_1 \cup \Omega_2), \end{cases}$$

(the values of I are given in [A/m^2]). The position and shape of the domains Ω_1 and Ω_2 are sketched in Figure 2. They correspond to the primary and secondary windings.

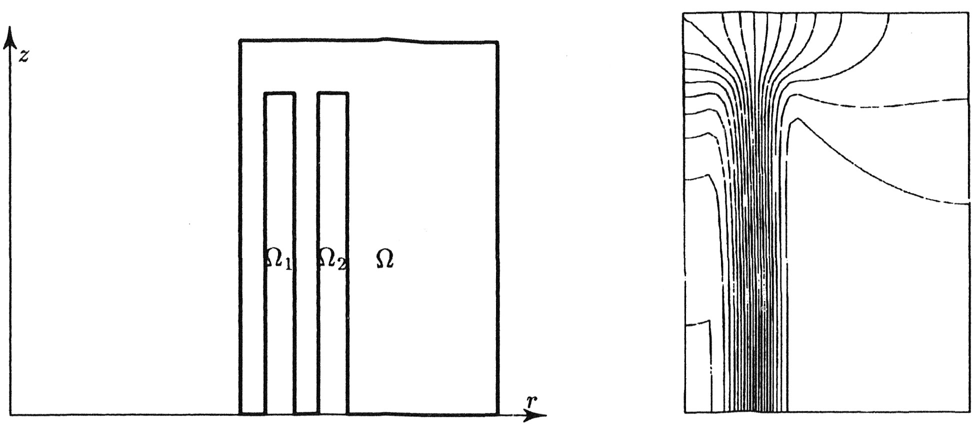

Fig. 2

Let us approximate V by bilinear finite elements, i.e., we set

$$V_h = \{v_h \in V \mid v_h|K \text{ is bilinear for all } K \in T_h\},$$

where the partition T_h consists of 35×50 rectangles. Consider now the problem: Find a function $u_h \in V_h$ such that

$$a(v_h, u_h) + \lambda \int_\Gamma v_h \mathrm{d}s \int_\Gamma u_h \mathrm{d}s = b(v_h) \quad \forall v_h \in V_h, \tag{3.3}$$

where $\lambda > 0$ is a given number and $\Gamma = [r_1, r_2] \times \{z_2\}$ is the upper side of $\overline{\Omega}$. Since the left-hand side of (3.3) represents a symmetric elliptic form, the associated stiffness matrix is symmetric positive definite. It is band if we number the usual basis functions of V_h row-wise. The choice of λ has no influence upon u_h whereas it has an influence on the condition number of the associated stiffness matrix, see [9]. The convergence of u_h to u in the $H^1(\Omega)$-norm follows now from the standard theory of finite elements (see [3]).

In Figure 2 we see the contour lines of u_h which seem to satisfy the Neumann boundary condition (3.2). Figure 3 illustrates the corresponding approximation of the radial and axial components of the magnetic field $(B_r, B_z) = \left(-\frac{1}{r}\frac{\partial u}{\partial z}, \frac{1}{r}\frac{\partial u}{\partial r}\right)$.

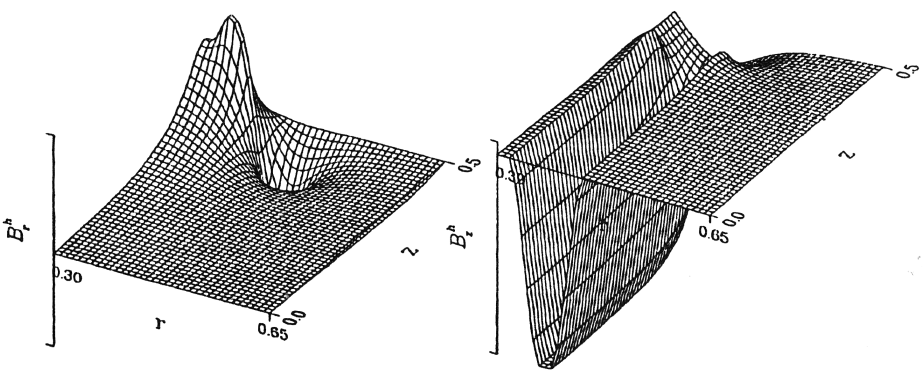

Fig. 3

R e m a r k 3.3. For further application to a linear elasticity problem we refer to [7], where the assumption (A3) is satisfied due to Korn's inequality [6]. In [8], periodic boundary conditions in linear elasticity are treated by the proposed method based upon the assumptions (A1)–(A5). The next theorem shows that the coerciveness of $a(.,.)$ (see (3.4) below) already implies that the space

$$P = \{p \in V \mid a(p, v) = 0 \quad \forall v \in V\}.$$

is finite dimensional. In other words, the assumption (A4) is fulfilled a priori.

Theorem 3.4. *Let $\Omega \subset R^d$, $d \in \{1, 2, \ldots\}$, be a bounded domain with a Lipschitz boundary and let*

$$V = \prod_{j=1}^{q} H^{k_j}(\Omega),$$

where $H^{k_j}(\Omega)$ are the Sobolev spaces with $k_j \geqslant 1$. Let $a(.,.)$ be a sesquilinear Hermitian form such that

$$a(v,v) + \|v\|_0^2 \geqslant c_0\|v\|^2 \quad \forall v \in V, \tag{3.4}$$

where $\|.\|_0$ is the $(L^2(\Omega))^q$-norm and $c_0 > 0$ is independent of v. Then $\dim P < \infty$.

Proof. Assume that P is not a finite-dimensional space. Then by the Riesz lemma (see e.g. [14, Theorem 3.12-E]) there exists a sequence $\{p_i\}_{i=1}^{\infty} \subset P$ such that $\|p_i\| = 1$ and

$$\|p_i - p\| > \tfrac{1}{2} \quad \forall p \in \operatorname{span}\{p_1, \ldots, p_{i-1}\}. \tag{3.5}$$

Using the compactness of the imbedding $V \hookrightarrow (L^2(\Omega))^q$, we find that there exists an $(L^2(\Omega))^q$-convergent subsequence of $\{p_i\}$ which will be still denoted by $\{p_i\}$. Hence, there exists an integer i_0 such that

$$\|p_i - p_j\|_0 < \tfrac{1}{2}\sqrt{c_0}$$

for all $i > j \geqslant i_0$. However, this contradicts (3.4) for $v = p_i - p_j$, since $a(p_i - p_j, p_i - p_j) = 0$ and $\|p_i - p_j\| > \frac{1}{2}$ by (3.5). □

Acknowledgement. The author wishes to thank D. Stará, K. Horák and J. Toivanen for their help in carrying out the figures. The work was supported by grant No. 11919 of the Grant Agency of the Czechoslovak Academy of Sciences. This support is gratefully acknowledged.

References

[1] *I. Babuška*: Uncertainties in engineering design: mathematical theory and numerical experience, In the Optimal Shape, (J. Bennet and M. M. Botkin eds.), Plenum Press (1986), also in Technical Note BN-1044, Univ. of Maryland (1985), 1–35.
[2] *J. Céa*: Optimization, théorie et algorithmes, Dunod, Paris, 1971.
[3] *P. G. Ciarlet*: The finite element method for elliptic problems, North-Holland, Amsterdam, 1978.
[4] *I. Doležel*: Numerical calculation of the leakage field in the window of a transformer with magnetic shielding, Acta Tech. ČSAV (1981), 563–588.
[5] *M. Feistauer*: Mathematical methods in fluid dynamics, Longman Scientific & Technical, Harlow, 1993.

[6] *I. Hlaváček, J. Nečas*: On inequalities of Korn's type, Arch. Rational Mech. Anal. *36* (1970), 305–334.
[7] *M. Křížek, W. G. Litvinov*: On the methods of penalty functions and Lagrange's multipliers in the abstract Neumann problem, Z. Angew. Math. Mech. (1993).
[8] *M. Křížek, Z. Milka*: On a nonconventional variational method for solving the problem of linear elasticity with Neumann or periodic boundary conditions, Banach Center Publ. (1993).
[9] *M. Křížek, P. Neittaanmäki, M. Vondrák*: A nontraditional approach for solving the Neumann problem by the finite element method, Mat. Apl. Comput. *11* (1992), 31–40.
[10] *W. G. Litvinov*: Optimization in elliptic boundary value problems with applications to mechanics, (in Russian), Nauka, Moscow, 1987.
[11] *J. Nečas, I. Hlaváček*: Mathematical theory of elastic and elasto-plastic bodies: an introduction, Elsevier, Amsterdam, 1981.
[12] *B. N. Pšeničnyj, Ju. M. Danilin*: Numerical methods in extremum problems, (Russian), Nauka, Moscow, 1975.
[13] *L. Schwartz*: Analyse mathématique, Vol. 1, Hermann, Paris, 1967.
[14] *A. E. Taylor*: Introduction to functional analysis, John Wiley & Sons, New York, 1958.
[15] *R. Temam*: Navier-Stokes equations, North-Holland, Amsterdam, 3rd revised edn, 1984.
[16] *D. E. Ward*: Exact penalties and sufficient conditions for optimality in nonsmooth optimization, Optim. Theory Appl. *57* (1988), 485–499.

Souhrn

O APROXIMACI NEUMANNOVA PROBLÉMU METODOU PENALIZACE

MICHAL KŘÍŽEK

V článku je dokázáno, že penalizací ohraničení, která se vyskytují v lineárním eliptickém Neumannově problému, dostaneme přímo přesné řešení pro libovolnou množinu penalizačních parametrů. V tomto případě také existuje kontinuum Lagrangeových multiplikátorů. Navržená metoda penalizace je použita k výpočtu magnetického pole v okně transformátoru.

Author's address: *Michal Křížek*, Mathematical Institute of the Czech Academy of Sciences, Žitná 25, 115 67 Prague 1.