Київський національний університет імені Тараса Шевченка

Журнал обчислювальної та прикладної

2012 # М А Т Е М А Т И К И 3 (109)

Заснований у 1965 році

Адреса редакції: 03022 Київ, пр. Глушкова, 2, корп. 6
Київський національний університет імені Тараса Шевченка,
факультет кібернетики, кафедра обчислювальної математики,
тел.: (044) 259-04-36, E-mail: opmjournal@gmail.com
http://www.opmj.univ.kiev.ua

Адреса редакції серії: 79000 Львів, вул. Університетська, 1
Львівський національний університет імені Івана Франка,
Кафедра обчислювальної математики,
тел.: (032) 239-43-91, E-mail: kom@franko.lviv.ua
http://blues.franko.lviv.ua/ami/kom/jnam/jnam_link_ua.htm

Затверджено Вченою радою факультету кібернетики
від 17 вересня 2012 р. (протокол № 2)

Свідоцтво про державну реєстрацію КВ 4246 від 26.05.2000

Підписано до друку 17 вересня 2012 р.

Журнал обчислювальної та прикладної математики | **2012, № 3 (109)** с. 1–10 | Journal of Numerical & Applied Mathematics

UDC 519.6

ON THE COMBINATION OF SINGULAR AND HYPERSINGULAR BOUNDARY INTEGRAL EQUATIONS FOR THE NEUMANN BOUNDARY VALUE PROBLEM FOR AN ELLIPTIC EQUATION WITH VARIABLE COEFFICIENTS

Christina Babenko, Roman Chapko

Резюме. Для чисельного розв'язування внутрішньої задачі Неймана для еліптичного рівняння зі змінними коефіцієнтами запропоновано підхід, який приводить до системи граничних інтегральних рівнянь з сингулярними і гіперсингулярними ядрами. Дискретизацію інтегральних рівнянь здійснено методом квадратур із використанням тригонометричних квадратурних формул інтерполяційного типу. Приведено приклади чисельних експериментів.

Abstract. We consider the interior Neumann boundary value problem for an elliptic equation with variable coefficients. For the numerical solution of this problem we develop an approach, which leads to a system of boundary integral equations with strong- and hypersingular kernels. The full discretization is realized by the quadrature method with use of quadrature rules based on trigonometrical interpolation. The results of numerical experiments are presented.

1. Introduction

The boundary integral equation method is an effective tool for theoretical investigations and numerical solution of various boundary value problems. For the use of direct or indirect integral equation approach it is extremely important to know the fundamental solution for the considered differential equation. This is not a big problem for the large number of equations with constant coefficients. But in the case of variable coefficients the fundamental solution is very difficult to find and therefore the integral equation method is not used very often for such kind of problems. However, it is possible to involve the parametrix which describes the main part of the fundamental solution and doesn't satisfy the equation. Note that in the case of elliptic equation of the second order the parametrix is also known as Levi's function [7, 8]. As a result, a given boundary value problem can be reduced to a boundary-domain integral equation. This approach doesn't contain the main advantage of integral equation method related to the decrease of the dimension of the differential problem. Therefore we investigate another approach which does not have this disadvantage. This approach has been applied in [2] for the case of the Dirichlet boundary value condition. Its idea consists in the following: we introduce a set of closed nonintersecting curves in the solution domain and consider the differential equation

†*Key words.* Elliptic equation with variable coefficients; Levi's functions; System of boundary integral equations; Strong and hyper-singularities; Quadrature method.

on these curves. Next we construct potentials with the Levi's function and reduce the given boundary value problem to boundary integral equations with various singularities in the kernels.

In this paper we extend the described approach to the case of the Neumann boundary value condition with the use of strong and hypersingular integral equations.

Let $D \subset \mathbb{R}^2$ be a bounded simply connected domain with the boundary $\Gamma_0 \in C^3$. We search for the function $u : D \to \mathbb{R}$ which satisfies the elliptic equation

$$Lu(x) = \operatorname{div}(\sigma(x)\,\operatorname{grad} u(x)) = 0, \quad x \in D \tag{1}$$

and the Neumann boundary value condition

$$\sigma(x)\frac{\partial u}{\partial \nu}(x) = f(x), \quad x \in \Gamma_0. \tag{2}$$

Here ν is the outward unit normal on Γ_0, $\sigma \in L^\infty(\bar{D})$, $\sigma > 0$ and $f \in H^{-1/2}(\Gamma_0)$ are given functions and

$$\int_{\Gamma_0} f(y)ds(y) = 0.$$

It is known [9] that the solution $u \in H^1(D)$ of the problem (1), (2) can be determined uniquely up to an additive constant. Therefore we assume that the coordinate origin belongs to the domain D and add the condition $u(0) = 0$.

2. Modified problem and boundary integral equations

Definition 1. The function $P(x, y)$, $x, y \in D$ is called the parametrix (or Levi's function) of a differential operator L if

$$L_x P(x, y) = \delta(x - y) + R(x, y),$$

where δ is the Dirac function and the function R has weak singularity for $x = y$.

It is easy to make sure that for the operator in (1) the Levi's function has the form

$$P(x, y) = \frac{\ln|x - y|}{2\pi\sigma(y)}, \quad x, y \in \mathbb{R}^2, \; x \neq y$$

and the remainder function is

$$R(x, y) = \frac{(x - y)\cdot \operatorname{grad}\sigma(x)}{2\pi\sigma(y)|x - y|^2}, \quad x, y \in \mathbb{R}^2 \; x \neq y.$$

Now we introduce the set of smooth closed disjoint curves $\Gamma = \bigcup_{k=1}^N \Gamma_k$ in the domain D. Assume that all curves have following parametric representations

$$\Gamma_k = \{x_k(t) = (x_{1,k}(t), x_{2,k}(t)), t \in [0, 2\pi]\}, \quad k = 0, \dots, N,$$

where $x_k : \mathbb{R} \to \mathbb{R}^2$ are C^3 and 2π–periodic with $|x_k'(t)| > 0$ for all t.

We modify the problem (1), (2) as follows: find the function $\tilde{u} : \Gamma \to \mathbb{R}$, which satisfies the differential equation (1) on Γ and the boundary value condition (2).

Lets introduce the single layer potential

$$w(x) = \sum_{i=1}^{N} \int_{\Gamma_i} \varphi_i(y) P(x,y) ds(y), \quad x \in D \tag{3}$$

with unknown densities $\varphi_i \in L^2(\Gamma_i)$. Then from the equation (1) considered on Γ and the definition of the Levi's function we receive the system of integral equation

$$\varphi_k(x) + \sum_{i=1}^{N} \int_{\Gamma_i} \varphi_i(y) R(x,y) ds(y) = 0, x \in \Gamma_k,\ k = 1, \ldots, N.$$

The boundary value condition (2) also needs to be satisfied. In order to achieve this we will combine the representation (3) with a potential over the boundary Γ_0.

We can present the solution of the modified problem in the form

$$\tilde{u}(x) = \int_{\Gamma_0} \varphi_0(y) \sigma(y) \frac{\partial P(x,y)}{\partial \nu(y)} ds(y) + w(x), \quad x \in \Gamma. \tag{4}$$

Then from the definition of the Levi's function and properties of a logarithmic double layer potential the modified problem can be reduced to the system of boundary integral equations

$$\begin{cases} \varphi_k(x) + \displaystyle\sum_{i=1}^{N} \int_{\Gamma_i} \varphi_i(y) R(x,y) ds(y) + \\ \qquad + \displaystyle\int_{\Gamma_0} \varphi_0(y) \sigma(y) \frac{\partial R(x,y)}{\partial \nu(y)} ds(y) = 0, \quad x \in \Gamma_k, \\ \displaystyle\sum_{i=1}^{N} \int_{\Gamma_i} \varphi_i(y) \frac{\partial P(x,y)}{\partial \nu(x)} ds(y) + \\ \qquad + \displaystyle\frac{\partial}{\partial \nu(x)} \int_{\Gamma_0} \varphi_0(y) \sigma(y) \frac{\partial P(x,y)}{\partial \nu(y)} ds(y) = \frac{f(x)}{\sigma(x)}, \quad x \in \Gamma_0 \end{cases} \tag{5}$$

for $k = 1, \ldots, N$. Note here that in the case of $\sigma = 1$ the system (5) will be simplified to the integral equation

$$\frac{1}{2\pi} \frac{\partial}{\partial \nu(x)} \int_{\Gamma_0} \varphi(y) \frac{\partial \ln |x-y|}{\partial \nu(y)} ds(y) = f(x), x \in \Gamma_0. \tag{6}$$

It is known [1, 5], that the integral operator in this equation is not invertible. Therefore we replace the equation (6) by the following modification

$$\frac{1}{2\pi} \frac{\partial}{\partial \nu(x)} \int_{\Gamma_0} \varphi(y) \frac{\partial \ln |x-y|}{\partial \nu(y)} ds(y) + \alpha = f(x), x \in \Gamma_0, \int_{\Gamma_0} \varphi(y) ds(y) = 0. \tag{7}$$

Here $\varphi \in H^{1/2}(\Gamma_0)$ and $\alpha \in \mathbb{R}$ are unknown. Now the integral operator in (7) is invertible in corresponding Sobolev spaces [1, 5].

Thus we consider the following final system of integral equations related to (5)

$$\begin{cases} \varphi_k(x) + \sum_{i=1}^{N} \int_{\Gamma_i} \varphi_i(y) R(x,y) ds(y) + \\ \qquad + \int_{\Gamma_0} \varphi_0(y)\sigma(y) \frac{\partial R(x,y)}{\partial \nu(y)} ds(y) = 0, \quad x \in \Gamma_k, \\ \\ \sum_{i=1}^{N} \int_{\Gamma_i} \varphi_i(y) \frac{\partial P(x,y)}{\partial \nu(x)} ds(y) + \\ \qquad + \frac{\partial}{\partial \nu(x)} \int_{\Gamma_0} \varphi_0(y)\sigma(y) \frac{\partial P(x,y)}{\partial \nu(y)} ds(y) + \alpha = \frac{f(x)}{\sigma(x)}, \quad x \in \Gamma_0, \\ \\ \int_{\Gamma_0} \varphi_0(y) ds(y) = 0. \end{cases}$$

Taking into account the form of Levi's and remainder functions we can rewrite this system in the following parametric form

$$\begin{cases} (1-\delta_{k0})\mu_k(t) + \frac{1}{2\pi} \sum_{i=0}^{N} \int_0^{2\pi} \mu_i(\tau) H_{k,i}(t,\tau) d\tau + \\ \qquad + \delta_{k0}\alpha = g_k(t), \ k = 0, \dots, N, \\ \\ \int_0^{2\pi} \mu_0(\tau) d\tau = 0 \end{cases} \tag{8}$$

with unknown densities $\mu_k(t) = \varphi_k(x_k(t))$, $k = 0, \dots, N$ and an unknown constant α and with right hand sides

$$g_k(t) = \begin{cases} 0 & k = 1, \dots, N, \\ \\ \frac{f(x_0(t))}{\sigma(x_0(t))} & k = 0, \end{cases}$$

and 2π-periodic kernels $H_{k,i}(t,\tau) = 2\pi R(x_k(t), x_i(\tau))|x_i'(\tau)|$, $k = 1, \dots, N$, $i = 1, \dots, N$,

$$H_{k,0}(t,\tau) = \frac{2(x_{k,1}(t) - x_{0,1}(\tau))(x_{k,2}(t) - x_{0,2}(\tau))}{|x_k(t) - x_0(\tau)|^4} \times$$

$$\times (x_2'(\tau)\sigma_{x_2}'(x_k(t)) - x_1'(\tau)\sigma_{x_1}'(x_k(t))) + \frac{(x_{k,1}(t) - x_{0,1}(\tau))^2}{|x_k(t) - x_0(\tau)|^4} -$$

$$- \frac{(x_{k,2}(t) - x_{0,2}(\tau))^2 (x_1'(\tau)\sigma_{x_2}'(x_k(t)) + x_2'(\tau)\sigma_{x_1}'(x_k(t)))}{|x_k(t) - x_0(\tau)|^4} -$$

$$- \frac{(x_k(t) - x_0(\tau)) \cdot \operatorname{grad} \sigma(x_k(t)) \nu(x_0(\tau)) \cdot \operatorname{grad} \sigma(x_0(\tau))}{|x_k(t) - x_0(\tau)|^2 \sigma(x_0(\tau))},$$

$$H_{0,k}(t,\tau) = |x_k'(\tau)| \frac{(x_0(t) - x_k(\tau)) \cdot \nu(x_0(t))}{|x_0(t) - x_k(\tau)|^2 \sigma(x_k(\tau))}$$

for $k = 1, \dots, N$ and

$$H_{0,0}(t,\tau) = \Bigg\{ -\frac{\nu(x_0(t)) \cdot \nu(x_0(\tau))}{|x_0(t) - x_0(\tau)|^2} +$$

$$+ \frac{2\nu(x_0(t)) \cdot (x_0(t) - x_0(\tau)) \nu(x_0(\tau)) \cdot (x_0(t) - x_0(\tau))}{|x_0(t) - x_0(\tau)|^4} -$$

$$- \frac{\nu(x_0(t)) \cdot (x_0(t) - x_0(\tau)) \nu(x_0(\tau)) \cdot \operatorname{grad} \sigma(x_0(\tau))}{|x_0(t) - x_0(\tau)|^2 \sigma^2(x_0(\tau))} \Bigg\} |x_0'(\tau)|.$$

As we see some kernels in (8) have various singularities. We split the strong singularity in $H_{\ell,\ell}$, $\ell = 1, \dots, N$ in the following form

$$H_{\ell,\ell}(t,\tau) = H_{\ell,\ell}^{(1)}(t,\tau) \cot \frac{\tau - t}{2}$$

with smooth kernels

$$H_{\ell,\ell}^{(1)}(t,\tau) = \begin{cases} \tan \dfrac{\tau - t}{2} H_{\ell,\ell}(t,\tau) & \text{for} \quad t \neq \tau, \\[2ex] \dfrac{x_0'(t) \cdot \operatorname{grad} \sigma(x_0(t))}{2\sigma(x_0(t)) |x_0'(t)|} & \text{for} \quad t = \tau. \end{cases}$$

To handle the hypersingularity in the kernel $H_{0,0}$ we rewrite it as

$$H_{0,0}(t,\tau) = -\frac{1}{4|x_0'(t)| \sin^2 \dfrac{t - \tau}{2}} + \tilde{H}_{0,0}(t,\tau),$$

where

$$\tilde{H}_{0,0}(t,\tau) = H_{0,0}(t,\tau) + \frac{1}{4|x_0'(t)| \sin^2 \dfrac{t - \tau}{2}}$$

with the diagonal term

$$\tilde{H}_{0,0}(t,t) = -\frac{\nu(x_0(t)) \cdot x_0''(t)}{2|x_0'(t)|^4} + \frac{\nu(x_0(t)) \cdot x_0''(t) \; \nu(x_0(t)) \cdot \operatorname{grad} \sigma(x_0(t))}{\sigma(x_0(t)) |x_0'(t)|} +$$

$$+ \frac{|x_0'(t)|^4 - 2|x_0'(t)|^2 x_0'(t) \cdot x_0'''(t) + 3(x_0'(t) \cdot x_0''(t))^2}{12|x_0'(t)|^5} -$$

$$- \frac{3(x_{0,1}'(t) x_{0,2}''(t) - x_{0,2}'(t) x_{0,1}''(t))^2}{12|x_0'(t)|^5}.$$

Based on the uniqueness results of the boundary value problem (1)–(2) and the Riesz-Schauder theory for compact operators [6] we have the following result about well-posedness for the system of 2π-periodic integral equations (8).

Theorem 1. *Let $p > 1/2$. For every $f \in H^p[0, 2\pi]$ the system (8) posses an unique solution $\mu_0 \in H^{p+1}[0, 2\pi]$ and $\mu_k \in H^p[0, 2\pi]$, $k = 1, \dots, N$.*

3. Quadrature method

We begin by describing the appropriate quadrature rules. For this we consider trigonometric interpolation with $2n$ equidistant nodal points

$$t_j^{(n)} = \frac{j\pi}{n}, \quad j = 0, \dots, 2n-1$$

with respect to the $2n$-dimensional space of trigonometric polynomials, and use the following quadrature rules

$$\frac{1}{2\pi}\int_0^{2\pi} f(\tau)\,d\tau \approx \frac{1}{2n}\sum_{k=0}^{2n-1} f(t_k^{(n)}), \tag{9}$$

$$\frac{1}{2\pi}\int_0^{2\pi} f(\tau)\ln\left(4\sin^2\frac{t-\tau}{2}\right)d\tau \approx \sum_{k=0}^{2n-1} \tilde{R}_k^{(n)}(t)\,f(t_k^{(n)}), \tag{10}$$

$$\frac{1}{2\pi}\int_0^{2\pi} f(\tau)\cot\frac{\tau-t}{2}d\tau \approx \sum_{k=0}^{2n-1} \tilde{T}_k^{(n)}(t)\,f(t_k^{(n)}), \tag{11}$$

$$\frac{1}{2\pi}\int_0^{2\pi} f'(\tau)\cot\frac{\tau-t}{2}d\tau \approx \sum_{k=0}^{2n-1} T_k^{(n)}(t)\,f(t_k^{(n)}). \tag{12}$$

The weight functions are given by

$$\tilde{R}_k^{(n)}(t) = -\frac{1}{n}\sum_{m=1}^{n-1}\frac{1}{m}\cos m(t-t_k^{(n)}) - \frac{1}{2n^2}\cos n(t-t_k^{(n)}),$$

$$\tilde{T}_k^{(n)}(t) = -\frac{1}{n}\sum_{m=1}^{n-1}\sin m(t-t_k^{(n)}) - \frac{1}{2n}\sin n(t-t_k^{(n)}),$$

$$T_k^{(n)}(t) = -\frac{1}{n}\sum_{m=1}^{n-1} m\cos m(t-t_k^{(n)}) - \frac{1}{2}\cos n(t-t_k^{(n)}).$$

These quadratures are obtained by replacing f with its trigonometric interpolation polynomial and then integrating exactly [3, 6]. Note that some of given quadratures coincide with quadrature formulas used in the method of discrete charges [4].

Thus we use quadrature rules (9),(11) and (12) to approximate three types of integrals in the system of integral equations (8) and collocate the approximate equations to obtain the linear system

$$\mathbf{A}\tilde{\mu} = \mathbf{b}$$

with matrix coefficients

$$A_{k,0}^{(ij)} = \frac{1}{2n}H_{k,0}(t_i^{(n)}, t_j^{(n)}), \quad k = 1, \dots, N,$$

$$A_{0,k}^{(ij)} = \begin{cases} \dfrac{1}{2n} H_{0,k}(t_i^{(n)}, t_j^{(n)}), & k = 1, \dots, N, \\ -\dfrac{1}{2|x_0(t_i^{(n)})|} T_j(t_i^{(n)}) + \dfrac{1}{2M} \tilde{H}_{0,0}(t_i^{(n)}, t_j^{(n)}), & k = 0, \end{cases}$$

$$A_{k,\ell}^{(ij)} = \begin{cases} \dfrac{1}{2n} H_{k,\ell}(t_i^{(n)}, t_j^{(n)}), & k \neq \ell, \\ \tilde{T}_j(t_i^{(n)}) H_{k,\ell}(t_i^{(n)}, t_j^{(n)}), & k = \ell \end{cases}$$

and $A_{0,0}^{(2n,j)} = 1$, $A_{0,0}^{(i,2n)} = 1$ and with the right hand side $b_k^{(i)} = g_k(t_i^{(n)})$, $k = 0, \dots, N$, $i = 0, \dots, 2n-1$ and $b_0^{(2n)} = 0$.

To find the numerical solution of the modified problem we parametrize the reprezentation (4)

$$\tilde{u}(x_k(t)) = \frac{1}{2\pi} \sum_{\ell=0}^{N} \int_0^{2\pi} \mu_\ell(\tau) L_{k,\ell}(t, \tau) d\tau, \tag{13}$$

where $L_{k,\ell}(t,\tau) = \frac{\pi}{n} |x'_\ell(\tau)| P(x_k(t), x_\ell(\tau))$ for $\ell, k = 1, \dots, N$ and

$$L_{k,0}(t,\tau) = -\frac{(x_k(t) - x_0(\tau)) \cdot \nu(x_0(\tau))}{\sigma(x_0(\tau)) |x_k(t) - x_0(\tau)|^2} -$$

$$- \frac{\operatorname{grad} \sigma(x_0(\tau)) \cdot \nu(x_0(\tau))}{\sigma^2(x_0(\tau))} \ln |x_k(t) - x_0(\tau)|.$$

As we see the kernels $L_{\ell,\ell}$ have logarithmic singularity and we split it in the following form

$$L_{\ell,\ell}(t,\tau) = L_{\ell,\ell}^1(t,\tau) \ln \left(4 \sin \frac{t-\tau}{2} \right) + L_{\ell,\ell}^2(t,\tau)$$

with

$$L_{\ell,\ell}^1(t,\tau) = \frac{|x'_\ell(\tau)|}{2\sigma(x_\ell(\tau))}$$

and

$$L_{\ell,\ell}^2(t,\tau) = \begin{cases} L_{\ell,\ell}(t,\tau) - L_{\ell,\ell}^1(t,\tau) \ln \left(4 \sin \dfrac{t-\tau}{2} \right) & \text{for} \quad t \neq \tau, \\ \dfrac{|x'_\ell(t)|}{\sigma(x_\ell(t))} \ln |x'_\ell(t)| & \text{for} \quad t = \tau. \end{cases}$$

Now according to (13) and using quadratures (9) and (10) we have the following formula for the numerical solution of the modified problem

$$\tilde{u}_n(x_k(t)) = \sum_{\ell=1}^{N} \sum_{i=0}^{2n-1} \tilde{\mu}_\ell^{(i)} \tilde{L}_{k,\ell}(t, t_i^{(n)}),$$

where

$$\tilde{L}^2_{k,\ell}(t,t_i^{(n)}) = \begin{cases} \dfrac{1}{2n} L_{k,\ell}(t,t_i^{(n)}) & \text{for} \quad \ell \neq k, \\[2ex] L^1_{\ell,\ell}(t,t_i^{(n)}) \tilde{R}_i^{(n)}(t) + \dfrac{1}{2n} L^2_{\ell,\ell}(t,t_i^{(n)}) & \text{for} \quad \ell = k. \end{cases}$$

4. Numerical examples

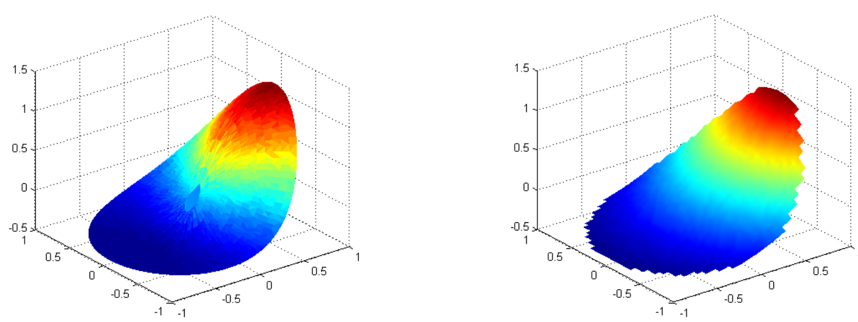
Example 1. We consider the domain D bounded by a circle Γ_0 with the radius $R = 1$. The given function σ and f are given as

$$\sigma(x) = 1 + x_1^2 + x_2^2, \ x \in D \quad \text{and} \quad f(x) = x_1 e^{x_1} \cos x_2 - x_2 e^{x_1} \sin x_2, \ x \in \Gamma_0.$$

The numerical solution of the boundary value problem (1),(2) received by proposed method is presented in the Fig. 1a. Here we used the following discretization parameters $n = 64$ and $N = 13$ and the set of curves

$$\Gamma_k = \left\{ x_k(t) = (1 - \frac{k}{N+1})(\cos t, \sin t), 0 \le t \le 2\pi \right\}, \ k = 0, \ldots, N.$$

The numerical result obtained by FEM method by PDE Toolbox in Matlab is illustrated in Fig. 1b. As we see both results are sufficiently close.



a) Numerical solution by BIEM b) Numerical solution by FEM

Fig. 1. Results of numerical experiments for the example 1

Example 2. Assume that the boundary curve Γ_0 and the set of curves Γ (see Fig. 2a) have the parametric representation

$$\Gamma_k = \left\{ x_k(t) = r(t)(1 - \frac{k}{N+1})(\cos t, \sin t), 0 \le t \le 2\pi \right\}, \ k = 0, \ldots, N$$

with the radial function

$$r(t) = \left(\left(\frac{1}{2} \cos t \right)^{10} + \left(\frac{2}{3} \sin t \right)^{10} \right)^{-0.1}.$$

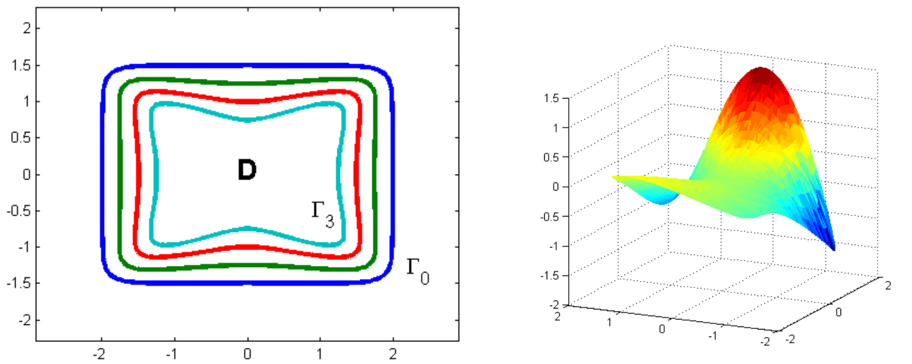
Let

$$\sigma(x) = 1 + e^{0.3(x_1^2 + x2)}, \ x \in D$$

and

$$f(x) = e^{x_1}\left(\cos x_2 \nu_1(x) - \sin x_2 \nu_2(x)\right), \ x \in \Gamma_0.$$

The numerical solution obtained via proposed method is given in Fig. 2b.



a) Set of curves Γ b) Numerical solution by BIEM

FIG. 2. Results of numerical experiments for the example 2

We considered the numerical solution of the interior planar Neumann boundary value problem for an elliptic differential equation of second order with variable coefficients. The proposed method is based on boundary integral equations. First we approximated the given problem by a modified problem on the introduced set of closed curves in the solution domain. Then the potentials with Levi's function are used for the modified problem. As result the system of boundary integral equations with singular and hypersingular kernels is received. The full discretization is realized by trigonometric quadrature method. The presented numerical examples confirmed the applicability of the proposed method.

BIBLIOGRAPHY

1. Atkinson K. The Numerical Solution of Integral Equations of the Second Kind / K. Atkinson.– Cambridge: Cambridge University Press, 1997.
2. Babenko Ch. On the numerical solution of a direct EIT problem by boundary integral equation method / Ch. Babenko, R. Chapko // Visnyk Lviv. Universytetu. Seriya prykladna matematyka ta informatyka.– 2012.– Vol. 18.– P. 22-30 (in Ukrainian).
3. Chapko R. On a quadrature method for a logarithmic integral equation of the first kind / R. Chapko, R. Kress // In Agarwal, ed.: World Scientific Series in Applicable Analysis.– Vol. 2. Contributions in Numerical Mathematics.– 1993.– P. 127-140.
4. Gandel Yu. V. Mathematical Questions of Discrete Charges Method / Yu. V. Gandel, T. S. Polanskaya.– Kharkiv: Kharkiv. University, 1991 (in Russian).
5. Kieser R. On a full discretization scheme for a hypersingular boundary integral equation over smooth curves / R. Kieser, B. Kleemann, R. Ratsfeld // Zeitschrieft für Analysis und Ihre Anwendungen.– 1992.– Vol. 11.– P. 385-396.
6. Kress R. Linear Integral Equations / R. Kress. 2nd. ed. Berlin: Springer-Verlag, 1999.
7. Miranda C. Partial Differential Equations of Elliptic Type / C. Miranda.– New York: Springer-Verlag, 1970.

8. Pomp A. The boundary-domain integral method for elliptic systems. With applications in shells / A. Pomp // Lecture Notes in Mathematics.– Vol. 1683, Berlin: Springer-Verlag, 1998.
9. McLean W. Strongly Elliptic Systems and Boundary Integral Operators / W. McLean.– Cambridge: Cambridge University Press, 2000.
10. Mikhailov S. E. Analysis of extended boundary-domain integral and integro-differential equations of some variable-coefficient BVP / S. E. Mikhailov // Advances in Boundary Integral Methods. Proceedings of the 5th UK Conference on Boundary Integral Methods (Edited by Ke Chen), University of Liverpool Publ., UK.– 2005.– P. 106-125.

Christina Babenko,
Ukrainian Engineering Pedagogics Academy, Kharkiv, 61003, Ukraine;
Roman Chapko,
Ivan Franko National University of Lviv,
1, Universytets'ka Str., Lviv, 79000, Ukraine

Received 24.03.2012

Журнал обчислювальної та прикладної математики | 2012, № 3 (109) с. 11–22 | Journal of Numerical & Applied Mathematics

UDC 519.6

RIGOROUS VANISHING SOLUTIONS OF A NONLINEAR HAMMERSTEIN INTEGRAL EQUATION RELATED TO PROBLEMS WITH FREE PHASE

Olena Bulatsyk, Ihor Tupychak, Yuriy Topolyuk

Резюме. Розглядається нелінійне інтегральне рівняння Гаммерштейна, яке виникає в задачах з вільною фазою. Досліджується новий клас дійсних та комплексних розв'язків цього рівняння. Розв'язки подаються у явному вигляді зі скінченим числом невідомих комплексних параметрів, що є нулями спеціально побудованого полінома, і скінченим числом дійсних параметрів – нулів цих розв'язків у їх області визначення. Для знаходження цих параметрів строго отримана нова коректна система трансцендентних рівнянь. Розв'язки цієї системи досліджуються чисельно. Аналізуються точки галуження цих розв'язків відносно дійсного параметра задачі.

Abstract. A nonlinear Hammerstein integral equation that arises in problems with free phase is considered. A new class of real and complex solutions of this equation is investigated. Solutions are represented in an explicit form with a finite number of unknown complex parameters being zeros of a specially built polynomial, and a finite number of real parameters – zeros of these solutions in their domain of definition. A new correctly determined form of earlier obtained transcendental equations system is found. The solutions of this system are numerically investigated. Their branching points are analyzed with respect to a real parameter of the problem.

1. Introduction

Problems with free phase, covering, in particular, the phase problem, attracted the attention of researchers for a long time [1]– [4]. A wide class of these problems complete the phase optimization problems, main idea of which consists in the mean square approximation of a given non-negative function by modulus of the functions being the result of action of a bounded operator on compactly supported complex functions [5]–[9]. These variational problems are reduced in a usual way to nonlinear integral equations of Hammerstein type as the Lagrangue-Euler equations for respective functionals. As a nonlinearity, these equations involve the phase factor (argument of unknown complex function) in the integrant.

One of the ways of solving such type of equations was an approach suggested and developed in [10], [11], and described in details in [12]. In this approach the solutions are represented in an explicit form with a finite number of unknown complex parameters. A system of transcendental equations for calculation of

[†] *Key words.* Nonlinear integral equation of Hammerstein type, finite-parametric solutions, branching of solutions, phase problem, vanishing solution.

these parameters was obtained. The approach was extended to a general class of Hammerstein equations of the considered type.

As it was noted in [12], above approach does not cover all solutions of the equations, it considers only nonvanishing solutions. The solutions having zeros in their definition domain were particularly considered in [13], [14]. Such problems arise, in particular, when the desired function to be approximated has zeros in its support domain. Among the physical problems of such type we should, in particular, mention the antenna synthesis problem by the given multi-lobe amplitude directivity pattern [15].

Similar problems were investigated in [13]. There were announced the ideas of analytical presentation of the solution for solving the problems. The main theoretical results are given in [14]. Partial numerical results are also considered there. Numerical results of investigation of real solutions of partial system of equations obtained in [14] are conducted in [16].
It should be also mentional the work [17], [18] on the approximation of functions defined on the real axis by the classes of entire functions and more universal approach in works [19], [20] which are close to the ideology of problems in our article.

In the article we consider more general one-dimensional case. Some results given here were announced in [21]. Real solutions, having zeros in the domain of finiteness of given non-negative function were considered in [16]. Some results concerning the solution branching with respect to the real parameter which is included in the kernel of nonlinear equation of Hammerstein type were described. It turns out that the sets of real and complex solutions are not separated. There are such values of the real kernel parameters, at which the complex solution branches into the real one. Branching of the initial complex solutions (this class of solutions will be described later) into other complex solutions is numerically investigated in this article. The starting point of the article is a system of complex transcendental equations, correctness of which is proved theoretically. Numerical results for several particular cases are presented and analyzed.

2. Problem formulation

Consider the Hammerstein integral equation

$$f(\xi) = \int_a^b K(\xi, \xi')F(\xi')e^{i \arg f(\xi')} d\xi'. \tag{1}$$

with the kernel

$$K(\xi, \xi') = \frac{s(\xi)q(\xi') - s(\xi')q(\xi)}{\tau(\xi) - \tau(\xi')}, \tag{2}$$

where $s(\xi)$, $q(\xi)$, $\tau(\xi)$ are real continuous functions, such that the systems of functions, $\{\tau^n(\xi)s(\xi)\}$, $\{\tau^n(\xi)q(\xi)\}$, $(n = 0, 1, \dots)$ are linearly independent, $F(\xi) \in L_2(a, b)$ is a given non-negative function. It is assumed that the solutions of the equation may have real zeros in the interval $[a, b]$.

3. The theoretical results

Let us represent the solution of equation (1) in the form

$$f(\xi)=\gamma\hat{f}(\xi)P_N(\tau)\prod_{j=1}^{M}(\xi-p_j)\,, \tag{3}$$

where $\hat{f}(\xi)$ is a real positive function in $\xi\in[a,b]$; γ is a complex constant with $|\gamma|=1$; $a\le p_j\le b$ are real zeros of function $f(\xi):f(p_j)=0$, M is a positive integer number;

$$P_N(\tau)=\prod_{k=1}^{N}(1-\eta_{Nk}\tau)$$

is polynomial of the degree N with complex, pairwise nonconjugated zeros η_{Nk}^{-1} :

$$\eta_{Nk}-\bar{\eta}_{Nm}\neq 0,\;\; k,m=1,2,...,N. \tag{4}$$

Without loss of generality, we can set $\gamma=1$. From (3) we obtain

$$e^{i\arg f(\xi)}=\frac{P_N(\tau)}{|P_N(\tau)|}\prod_{j=1}^{M}\operatorname{sgn}(\xi-p_j)\,. \tag{5}$$

The function $\hat{f}(\xi)$ can be uniquelly defined from the equality

$$\hat{f}(\xi)\left|\prod_{j=1}^{M}(\xi-p_j)\right|=$$
$$=\frac{1}{|P_N(\tau)|}\left|\int_a^b K(\xi,\xi')F(\xi')\frac{P_N(\tau')}{|P_N(\tau')|}\prod_{j=1}^{M}\operatorname{sgn}\left(\xi'-p_j\right)d\xi'\right| \tag{6}$$

which follows from (1).

Theorem 1. *Function $f(\xi)$ of the form (3) is a solution of equation (1) if and only if the real parameters p_j, $j=1,...,M$ and complex η_{Nk}, $k=1,...,N$, with the condition (4) satisfy the system of transcendental equations:*

$$T_{Nn}(p_j,\eta_{N1},\eta_{N2},...,\eta_{NN})=0,\;\; n=1,2,...,N,\;\; j=1,2,...,M, \tag{7}$$

$$\Phi_{Nn}(p_j,\eta_{N1},\eta_{N2},...,\eta_{NN})=0,\;\; n=1,2,...,N,\;\; j=1,2,...,M, \tag{8}$$

$$\Psi_{Nn}(p_j,\eta_{N1},\eta_{N2},\ldots,\eta_{NN})=0,\; n=1,2,...,N,\;\; j=1,2,...,M, \tag{9}$$

where

$$T_{Nn}=\int_a^b K(p_j,\xi')F(\xi')\frac{\operatorname{Re}\left[\bar{P}_N(\tau(p_j))P_N(\tau')\right]}{|P_N(\tau')|}\prod_{j=1}^{M}\operatorname{sgn}\left(\xi'-p_j\right)d\xi',$$

$$\Phi_{Nn}=\int_a^b \tau^{n-1}s(\xi)\frac{F(\xi)}{|P_N(\tau)|}\prod_{j=1}^{M}\operatorname{sgn}(\xi-p_j)\,d\xi, \tag{10}$$

$$\Psi_{Nn} = \int_a^b \tau^{n-1} q(\xi) \frac{F(\xi)}{|P_N(\tau)|} \prod_{j=1}^{M} \operatorname{sgn}(\xi - p_j)\, d\xi. \tag{11}$$

Proof. Necessity. Let function $f(\xi)$ represented as (3) with $\gamma = 1$ be the solution of equation (1). Substituting expressions (3), (5) into (1) and multiplying both sides by $\bar{P}_N(\xi)$ result in

$$\begin{gathered} \hat{f}(\xi)\, |P_N(\tau)|^2 \prod_{j=1}^{M} (\xi - p_j) = \\ = \bar{P}_N(\tau) \int_a^b K(\xi, \xi') F(\xi') \frac{P_N(\tau')}{|P_N(\tau')|} \prod_{j=1}^{M} \operatorname{sgn}\left(\xi' - p_j\right) d\xi'. \end{gathered} \tag{12}$$

After extracting the imaginary part from (12), we obtain

$$\int_a^b \left(\tau - \tau'\right) K(\xi, \xi') \frac{F(\xi')}{|P_N(\tau')|} \prod_{j=1}^{M} \operatorname{sgn}\left(\xi' - p_j\right) R_{N-1}(\tau, \tau') d\xi' = 0, \tag{13}$$

where

$$R_{N-1}(\tau, \tau') = \frac{2i\, \operatorname{Im}[P_N(\tau')\bar{P}_N(\tau)]}{\tau - \tau'} = \sum_{k,m=1}^{N} a_{km} \tau^{k-1} (\tau')^{m-1} \tag{14}$$

is a polynomial of two variables with matrix coefficients $A = \{a_{km}\}$. Substitution (14), (2) into (13) and interchanging the variables ξ and ξ', give

$$\begin{gathered} \sum_{k,m=1}^{N} a_{km} \left[q(\xi') \int_a^b \tau^{k-1} s(\xi) \frac{F(\xi)}{|P_N(\tau)|} \prod_{j=1}^{M} \operatorname{sgn}(\xi - p_j)\, d\xi - \right. \\ \left. - s(\xi') \int_a^b \tau^{k-1} q(\xi) \frac{F(\xi)}{|P_N(\tau)|} \prod_{j=1}^{M} \operatorname{sgn}(\xi - p_j)\, d\xi \right] \left(\tau'\right)^{m-1} \equiv 0. \end{gathered} \tag{15}$$

Since the functions $\{\tau^k s(\xi)\}$, $\{\tau^k q(\xi)\}$, $k = 0, 1, \ldots, N-1$, are linearly independent, expression (15) results in the following systems:

$$\sum_{k=1}^{N} a_{km} \Phi_{Nn} = 0, \;\; n = 1, 2, ..., N, \tag{16}$$

$$\sum_{k=1}^{N} a_{km} \Psi_{Nn} = 0, \;\; n = 1, 2, ..., N, \tag{17}$$

where Φ_{Nn}, Ψ_{Nn} are defined in (10), (11). They can be considered as independent systems of linear algebraic equations for the unknown Φ_{Nn}, Ψ_{Nn}, with the same matrix of coefficients A.

Determinant of matrix A has been found in [11]:

$$\det A = (-1)^{[N/2]} \prod_{k,m=1}^{N} (\bar{\eta}_{Nm} - \eta_{Nk}),$$

where the square brackets mean the integer part of the number. From (4) we get $\det A \neq 0$, and the equation systems (16), (17) have only zero solutions. This means that transcendental equations (8), (9) are satisfied.

Let the solution of equation (1) satisfy the condition $f(p_j) = 0$, $j = 1, ..., M$. Then, according to (1),

$$\int_a^b K(p_j, \xi') F(\xi') e^{i \arg f(\xi')} d\xi' = 0. \tag{18}$$

Multiplying both sides of (18) by $\bar{P}_N(\xi)$ and using (5) result in

$$\int_a^b K(p_j, \xi') F(\xi') \frac{\bar{P}_N(\tau(p_j)) P_N(\tau')}{|P_N(\tau')|} \prod_{j=1}^{M} \operatorname{sgn}\left(\xi' - p_j\right) d\xi' = 0. \tag{19}$$

After extracting the real part from (19), we obtain the system of equations (7). Imaginary part (19) gives the following system:

$$\begin{aligned} \int_a^b \left(\tau(p_j) - \tau'\right) K(p_j, \xi') \frac{F(\xi')}{|P_N(\tau')|} \times \\ \times \prod_{j=1}^{M} \operatorname{sgn}\left(\xi' - p_j\right) R_{N-1}(\tau(p_j), \tau') d\xi' = 0. \end{aligned} \tag{20}$$

System of equations (20) coincides with the system (13) in case of $\xi = p_j$ and $\tau = \tau(p_j)$.

Sufficiency. Let the system of transcendental equations (7), (8), (9) be satisfied for some integer N, complex numbers η_{Nk}, $k = 1, 2, ..., N$, which satisfy the condition (4) and real numbers p_j, $j = 1, ..., M$. We show that the function of the form (3) is a solution of equation (1) and p_j, $j = 1, ..., M$, are real zeros of the solution.

After reducing system (8), (9) to equalities (16), (17) and substituting (10), (11) into them we get the equality (15). Then, using (13), we have

$$\operatorname{Im}\left[\bar{P}_N(\tau) \int_a^b K(\xi, \xi') \frac{F(\xi')}{|P_N(\tau')|} \prod_{j=1}^{M} \operatorname{sgn}\left(\xi' - p_j\right) P_N(\tau') d\xi'\right] = 0.$$

Add the real function $\hat{f}(\xi) \left|P_N(\tau)\right|^2 \prod_{j=1}^{M} (\xi - p_j)$ under the symbol of imaginary part:

$$\operatorname{Im}\left[\hat{f}(\xi)\left|P_N(\tau)\right|^2\prod_{j=1}^{M}(\xi-p_j)+\right.$$
$$\left.+\bar{P}_N(\tau)\int\limits_a^b K(\xi,\xi')\frac{F(\xi')}{|P_N(\tau')|}\prod_{j=1}^{M}\operatorname{sgn}\left(\xi'-p_j\right)P_N(\tau')d\xi'\right]=0. \tag{21}$$

Dividing both sides of equality (21) by the positive function $|P_N(\tau)|$ and taking into account the equality

$$\hat{f}(\xi)\left|P_N(\tau)\right|=\frac{|f(\xi)|}{\left|\prod\limits_{j=1}^{M}(\xi-p_j)\right|}, \tag{22}$$

we get

$$\operatorname{Im}\left[|f(\xi)|\prod_{j=1}^{M}\operatorname{sgn}(\xi-p_j)+\right.$$
$$\left.+\frac{\bar{P}_N(\tau)}{|P_N(\tau)|}\int\limits_a^b K(\xi,\xi')\frac{F(\xi')}{|P_N(\tau')|}\prod_{j=1}^{M}\operatorname{sgn}\left(\xi'-p_j\right)P_N(\tau')d\xi'\right]=0. \tag{23}$$

On the other hand, (6) gives

$$\operatorname{Re}\left[|f(\xi)|\prod_{j=1}^{M}\operatorname{sgn}(\xi-p_j)+\right.$$
$$\left.+\frac{\bar{P}_N(\tau)}{|P_N(\tau)|}\int\limits_a^b K(\xi,\xi')\frac{F(\xi')}{|P_N(\tau')|}\prod_{j=1}^{M}\operatorname{sgn}\left(\xi'-p_j\right)P_N(\tau')d\xi'\right]=0. \tag{24}$$

Equalities (23) and (24) mean that the expression in brackets equals to zero and, according to (5), function (3) is a solution of integral equation (1). □

Let then the equation system (7) be satisfied. Using (24), (22), (3) with $\xi = p_j$, we get

$$\operatorname{Re}\left[f\left(p_j\right)\right]=0.$$

Similarly, (23), (13), (22), (3) give

$$\operatorname{Im}\left[f\left(p_j\right)\right]=0.$$

Thus, the parameters p_j, $j=1,...,M$ are zeros of the function f of form (3). The theorem is proved.

4. The numerical results

The Newton method was used for solving system (7), (8), (9). Integrals from the left part of the equation system were calculated by the Simpson method. The integration interval was divided into segments by points so, that the integrand was smooth there. Iterative process of Newton method can be described by the following formula:

$$\vec{x}^{(m+1)} = \vec{x}^{(m)} - \left(\Im'\left(\vec{x}^{(m)};c\right)\right)^{-1}\Im\left(\vec{x}^{(m)};c\right), \quad m=0,1,$$

where $\vec{x}^{(m)} = \left\{\eta_n'^{(m)}, \eta_n''^{(m)}, p_l^{(m)}\right\}$ is an approximation of zeros of the system (7), (8), (9) on the m-th step, $\eta_n^{(m)} = \eta_n'^{(m)} + i\eta_n''^{(m)}$, i is the imaginary unit, $\Im$ is a vector of the left parts of the equation system (7), (8), (9), $\Im'\left(\vec{x}^{(m)};c\right)$ is the Jacobian matrix of this system in the point $\vec{x}^{(m)}$; c is a real positive parameter; m is an iteration number. The end-point condition of the iterative process in the Newton method is

$$\max_{n=\overline{1,N}}\left|\eta_n'^{(m+1)} - \eta_n'^{(m)}\right| + \max_{n=\overline{1,N}}\left|\eta_n''^{(m+1)} - \eta_n''^{(m)}\right| +$$
$$+\max_{l=\overline{1,M}}\left|p_l^{(m+1)} - p_l^{(m)}\right| < \varepsilon.$$

The structure of the Jacobian matrix in general case is

$$\Im'\left(\eta_n', \eta_n'', p_l\right) = \begin{pmatrix} \left\{\frac{\partial\Phi_{Nj}}{\partial\eta_{Nk}'}\right\}_{j,k=1}^{N} & \left\{\frac{\partial\Phi_{Nj}}{\partial\eta_{Nk}''}\right\}_{j,k=1}^{N} & \left\{\frac{\partial\Phi_{Nj}}{\partial p_k}\right\}_{j,k=1}^{N,M} \\ \left\{\frac{\partial\Psi_{Nj}}{\partial\eta_{Nk}'}\right\}_{j,k=1}^{N} & \left\{\frac{\partial\Psi_{Nj}}{\partial\eta_{Nk}''}\right\}_{j,k=1}^{N} & \left\{\frac{\partial\Psi_{Nj}}{\partial p_k}\right\}_{j,k=1}^{N,M} \\ \left\{\frac{\partial T_{Nj}}{\partial\eta_{Nk}'}\right\}_{j,k=1}^{M,N} & \left\{\frac{\partial T_{Nj}}{\partial\eta_{Nk}''}\right\}_{j,k=1}^{M,N} & \left\{\frac{\partial T_{Nj}}{\partial p_k}\right\}_{j,k=1}^{M,M} \end{pmatrix}.$$

Equation system (7), (8), (9) was investigated numerically for the case $s(\xi) = \sin c\xi$, $q(\xi) = \cos c\xi$, $\tau = \xi$, $a = 1$, $b = -1$, $c > 0$. In this case equation (1) has the form

$$f(\xi) = \int_{-1}^{1} \frac{\sin c(\xi - \xi')}{\xi - \xi'} F(\xi') \exp(i \arg f(\xi')) d\xi'. \tag{25}$$

We consider two types of given non-negative functions $F(\xi) : F_1(\xi) = |\xi - t|$ and $F_2(\xi) = \sin(\pi \cdot |\xi - t|/(1+|t|))$, $t \in (-1,1)$, having one zero in the integration domain.

Real solutions. We consider first the real solutions that correspond to $N = 0$. According to (3), each real solution of equation (1) is represented in the form:

$$f\left(\xi'\right)=\int\limits_{-1}^{1}F\left(\xi\right)\frac{\sin c\left(\xi-\xi'\right)}{\xi-\xi'}\prod_{j=1}^{M}\operatorname{sgn}\left(\xi-p_j\right)d\xi. \tag{26}$$

It follows from (7), that the real parameters p_j of are found from the equation system

$$\int\limits_{-1}^{1}F\left(\xi\right)\frac{\sin c\left(\xi-p_l\right)}{\xi-p_l}\prod_{k=1}^{M}\operatorname{sgn}\left(\xi-p_k\right)d\xi=0,\quad l=1,...,M. \tag{27}$$

The cases $M=1$ and $M=2$ are investigated numerically. In these cases the real solution of equation (1) has one or two zeros in the interval [-1,1], respectively.

In case $M=1$, system (27) becomes the transcendental equation with respect to $p=p_1$:

$$\int\limits_{-1}^{1}F\left(\xi\right)\frac{\sin c(\xi-p)}{|\xi-p|}d\xi=0. \tag{28}$$

We solve this equation by the chord method. Integrals in the left hand sides of the equation are calculated by Simpson method. To apply this method, the integration interval [-1;1] is divided into parts by zeros of the functions F, f so that the integrand is smooth in each of these parts. Then, in particular, unknown parameter p occurs in the integration limits.

Solutions of equation (28) are shown in Fig. 1. Different solutions depending on parameter c are marked by p_{1j}, where j means the solution number.

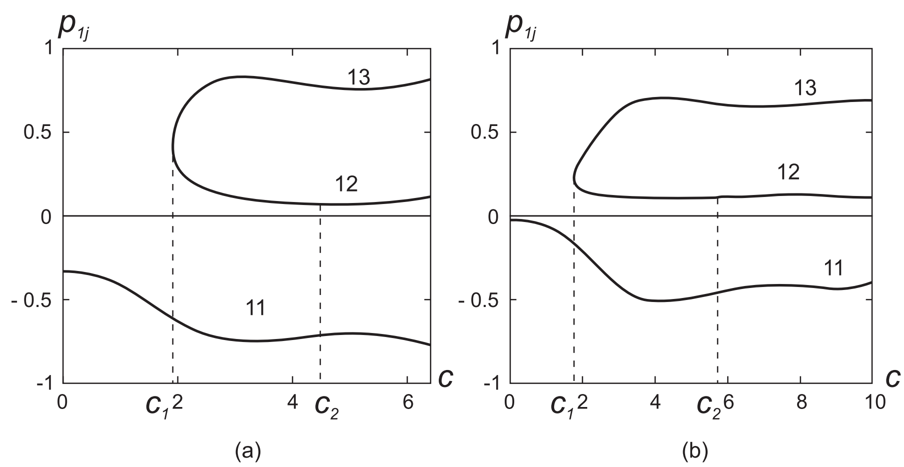


FIG. 1. Solutions of equation (28): (a) $F_1(\xi)=|\xi-t|$, (b) $F_2(\xi)=\sin(\pi\cdot|\xi-t|/(1+|t|))$; $t=0.1$

As follows from Fig. 1, number of solutions varies depending on the real parameter. For $c<c_1$ only one solution p_{11} exists. At the point $c=c_1$ two

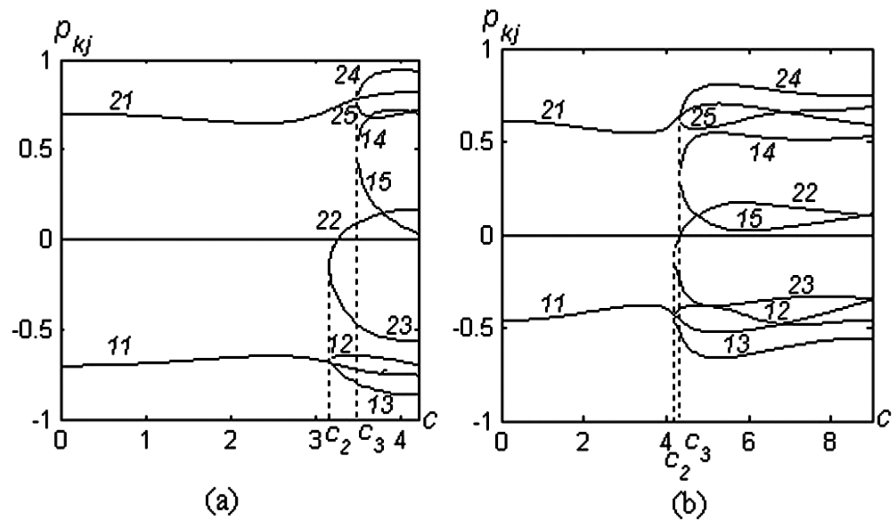


FIG. 2. Solutions of equation system (27) at $M = 2$: (a) $F_1(\xi) = |\xi - t|$, (b) $F_2(\xi) = \sin(\pi \cdot |\xi - t|/(1 + |t|))$; $t = 0.1$

more solutions (p_{12} and p_{13}) appear. These results show that the point c_1 is an isolated bifurcation point, namely a point of appearance of new solutions. For $c>c_1$ we already have three solutions of equation (28).

The more complicated situation arises in the case when two zeros of real solution of equation (1) exist in the interval $[-1; 1]$ ($M = 2$). In this case the equation system (27) is solved by the Newton method. As before, the interval $[-1; 1]$ is divided into parts so that the integrand is smooth in each of these parts.

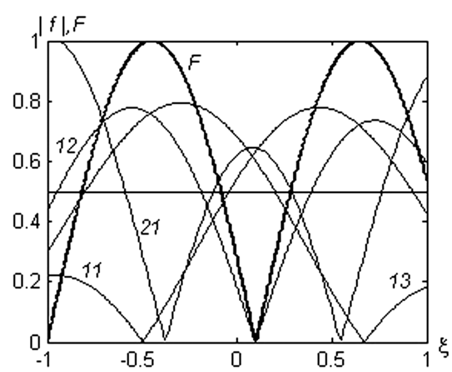


FIG. 3. Different solutions to (25) for given function $F_2(\xi) = \sin(\pi \cdot |\xi - t|/(1 + |t|))$ at $c = 3.5$, $t = 0.1$

The solutions with two zeros p_1, p_2 exist for arbitrary value of $c>0$. The curves corresponding to these solutions are marked with symbols $1j$ and $2j$, respectively. Similarly to the previous case, there is a pair of new solutions with parameters $\{p_{12}, p_{22}\}$ and $\{p_{13}, p_{23}\}$ at the point $c = c_2$. One more pair of solutions $\{p_{14}, p_{24}\}$ and $\{p_{15}, p_{25}\}$ appears at the point $c = c_3$. Thus, when $c>c_3$, there are 5 different solutions of system (27) each of which has two real zeros in the interval: $\{p_{11}, p_{21}\}$, $\{p_{12}, p_{22}\}$, $\{p_{13}, p_{23}\}$, $\{p_{14}, p_{24}\}$, $\{p_{15}, p_{25}\}$.

Moduli $|f_i(\xi)|$ of all found solutions (26) with $M = 1$ and $M = 2$ of the initial equation (25) in case $F_2(\xi) = \sin(\pi \cdot |\xi - t|/(1 + |t|))$, $c = 3.5$ are shown in Fig. 3. It turns out that the closest in modulus to the given function $F_2(\xi)$, is f_{12} having one real zero at the point $\xi = p_{12}$.

Complex solutions. Complex solutions of the system of transcendental equations (7), (8), (9) in the considered particular case were numerically investigated for $N = 1$, $M = 1$. These solutions have one real and one complex parameter. The equation system for this case is of the form

$$\int_{-1}^{1} F(\xi)\frac{\sin c\,(\xi - p_1)}{\xi - p_1}\frac{\operatorname{Re}\left[\bar{P}_1(p_1)P_1(\xi)\right]}{|P_1(\xi)|}\operatorname{sgn}(\xi - p_1)\,d\xi = 0, \tag{29}$$

$$\int_{-1}^{1} \sin c\xi\frac{F(\xi)}{|P_1(\xi)|}\operatorname{sgn}(\xi - p_1)\,d\xi = 0, \quad \int_{-1}^{1} \cos c\xi\frac{F(\xi)}{|P_1(\xi)|}\operatorname{sgn}(\xi - p_1)\,d\xi = 0,$$

where $P_1(\xi) = 1 - \eta_1\xi$. This system was solved by the Newton method.

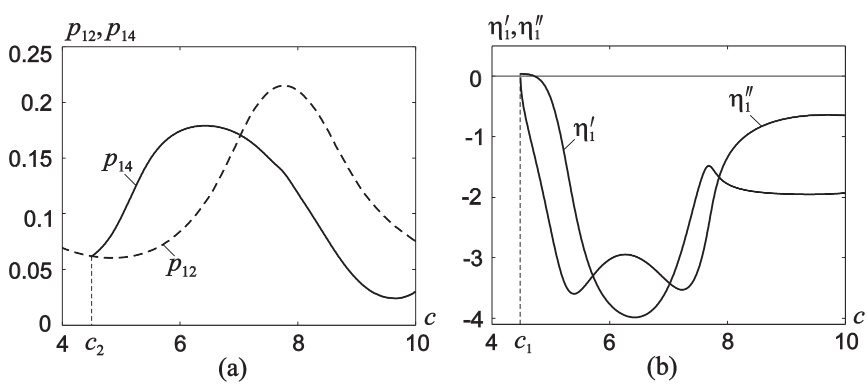


FIG. 4. Parameters of complex solution with $M = 1$, $N = 1$ to equation (29) for $F_1(\xi) = |\xi - 0.1|$

We investigate the real solution f_{12} (with one parameter p_{12}) of initial equation (25). This solution appears at the point $c = c_1$ (see Fig. 1, curve 12). At the point $c = c_2$ two complex conjugated solutions are branched off from f_{12}. At $c > c_2$ they have one real parameter p_{14} and one of two complex parameters $\eta_{1,2} = \eta_1' \pm i\eta_1''$.

Numerical results are shown in Fig. 4 and Fig. 5 for the given functions $F_1(\xi)$ and $F_2(\xi)$, respectively. The results demonstrate that sets of real and complex solutions of equation (1) are not isolated and real solutions branch into the

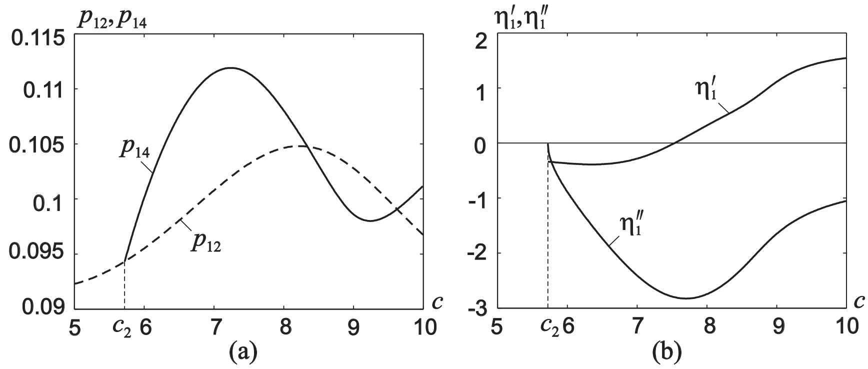


FIG. 5. Parameters of complex solution with $M = 1$, $N = 1$ to equation (29) for $F_2(\xi) = \sin(\pi \cdot |\xi - 0.1|/1.1)$

complex ones (the example of transforming the complex nonvanishing solutions into the real vanishing ones for this type of given function $F(\xi)$ see in [21]).

5. Conclusions

A nonlinear Hammerstein integral equation arisen in problems with free phase has been considered. A new class of real and complex solutions of this equation has been investigated. Solutions have been represented in an explicit form with a finite number of unknown complex parameters being zeros of a complex polynomial, and a finite number of real parameters – zeros of these solutions in their domain of definition. A new correctly determined form of earlier obtained transcendental equations system has been found. The solutions of this system have been numerically investigated for a particular case. The branching points of these solutions with the respect to a real parameter of the problem have been analyzed.

Bibliography

1. Goncharsky A. V. Mathematical models and the phase problem in inverse problem of X-ray diffraction / A. V. Goncharsky, A. A. Stepanov // Mathem. Model.– 1994.– Vol. 6.– P. 117-127 (in Russian).
2. Yezhov P. V. Correlation method for pattern recognition / P.V. Yezhov, A.V. Kuzmenko, V.A. Komarov // Semiconductor Physics, Quantum Electronics & Optoelectronics.– 2002.– Vol. 5.– P. 331-433.
3. Hwi Kim Iterative Fourier transform algorithm withregularization for the optimal design of diffractive optical elements / Kim Hwi, Yang Byungchoon, Lee Byoungho // J. Opt. Soc. Am. A.– 2004.– Vol. 21.– P. 2353-2365.
4. Kuzmenko A. V. Weighting method of the Fourier-kinoform synthesis / A. V. Kuzmenko // Semiconductor Physics, Quantum Electronics & Optoelectronics.– 2008.– Vol. 11.– P. 303-306.
5. Ferwerda H. A. Problem of the wave front phase reconstruction according to amplitude distribution and coherent functions / H.A. Ferwerda In: Inverse Scattering Problems in Optics. Ed. H.P. Baltes.– Berlin: Springer, 1978.
6. Kuznetsova T. I. On the phase problem in optics / T. I. Kuznetsova // Sov. Phys. Usp.– 1988.– Vol. 31.– P. 364-371.

7. Boikova A. T. Solution of phase retrieval problem by a maximum entropy method / A. T. Boikova // Radiophysics and Quantum Electronics.– 1996.– Vol. 39.– P. 321-327.
8. Pol'skikh S. D. The phase problem: analysis of local extrema and image reconstruction algorithms / S. D. Pol'skikh // Journal of Communications Technology and Electronics.– 2008.– Vol. 53.– P. 223-237.
9. Samoilenko M. V. Reconstruction of the amplitude–phase distribution of the field of the received signal in the aperture of a phased antenna array from the measured power / M. V. Samoilenko // Journal of Communications Technology and Electronics.– 2009.– Vol. 54.– P. 1058-1063.
10. Voitovich N. N. Mean square approximation of compactly supported functions with free phase by functions with bounded spectrum / N. N. Voitovich, Yu. P. Topolyuk, O. M. Gis // Dopovidi NAN Ukrainy.– 1999.– Vol. 10.– P. 3-7.
11. Voitovich N. N. Approximation of compactly supported functions with free phase by functions with bounded spectrum / N. N. Voitovich, Yu. P. Topolyuk, O. O. Reshnyak // Fields Institute Communications. AMS.– 2000.– Vol. 25.– P. 531-541.
12. Bulatsyk O. O. Phase Optimization Problems. Applications in Wave Field Theory / O. O. Bulatsyk, B. Z. Katsenelenbaum, Yu. P. Topolyuk, N. N. Voitovich.– Weinheim: WILEY-VCH, 2010.
13. Gis O. M. Closed vanishing solution of the nonlinear problems of the antenna synthesis / O. M. Gis, Yu. P. Topolyuk // Direct and Inverse Problems of Electromagnetic and Acoustic Wave Theory (DIPED-99).– Proc. Int. Seminar / Workshop, Lviv.– 1999.– P. 75-79.
14. Gis O. M. The mean square approximation of nonnegative finite functions by the modulus of functions with the finite spectrum (the case of vanishing solutions of approximation on support the finite function) / O. M. Gis, Yu. P. Topolyuk // Mathem. meth. fiz.-mesh. pola.– 2001.– Vol. 44.– P. 48-54.
15. Andriychuk M. I. Antenna Synthesis according to the Amplitude Radiation Pattern. Numerical Methods and Algorithms / M. I. Andriychuk, N. N. Voitovich, P. A. Savenko, V. P. Tkachuk.– Kiev: Naukova Dumka, 1993 (in Russian).
16. Tupychak I. V. Real vanishing solutions to nonlinear equation related to modified phase problem / I. V. Tupychak, N. N. Voitovich // Direct and Inverse Problems of Electromagnetic and Acoustic Wave Theory (DIPED-2006).– Proc. Int. Seminar / Workshop, Tbilisi.– 2001.– P. 169-173.
17. Stepanets A. I. Classes of functions defined on the real axis, and their approximations by entire functions. I / A. I. Stepanets // Ukr. Math. J.– 1990.– Vol. 42.– P. 102-112 (in Russian).
18. Stepanets A. I. Classes of functions defined on the real axis, and their approximations by entire functions. II / A. I. Stepanets // Ukr. Math. J.– 1990.– Vol. 42. – P. 210-222 (in Russian).
19. Stepanets A. I. Extremal problems of approximation theory in linear spaces / A. I. Stepanets // Ukr. Math. J.– 2003.– Vol. 55.– P. 1662-1698 (in Russian).
20. Stepanets A. I. Extremal problems for integrals of nonnegative functions / A. I. Stepanets // Izvestya RAN, Ser. Math.– 2010.– Vol. 74.– P. 169-224 (Russian).
21. Bulatsyk O. O. Complex vanishing solutions to nonlinear equation related to modified phase problem / O. O. Bulatsyk, I. V. Tupychak, Yu. P. Topolyuk // Direct and Inverse Problems of Electromagnetic and Acoustic Wave Theory (DIPED-2010).– Proc. Int. Seminar / Workshop, Tbilisi.– 2010.– P. 183-186.

Olena Bulatsyk, Ihor Tupychak, Yuriy Topolyuk,
Pidstrygach Institute for Applide Problems of Mechanics and Mathematics, 3b Naukova str., Lviv, 79061, Ukraine

Received 20.04.2012

Журнал обчислювальної та прикладної математики | **2012, № 3 (109)** с. 23–33 | Journal of Numerical & Applied Mathematics

UDC 517.958; 519.63

NUMERICAL INVESTIGATION OF A PLAIN STRAIN STATE FOR A BODY WITH THIN COVER USING DOMAIN DECOMPOSITION

Ivan Dyyak, Yarema Savula, Andriy Styahar

Резюме. Розглядається модель, яка описує напружено деформований стан двовимірного гетерогенного тіла з тонким покриттям. Спочатку доведено збіжність ітеративного алгоритму, побудованого на основі поєднання методу скінченних елементів (МСЕ) та методу граничних елементів (МГЕ) з використанням декомпозиції областей. Після цього алгоритм проілюстровано на прикладі двовимірної задачі для тіла з покриттям.

Abstract. We consider a model, that describes the plain stress state of the $2D$ heterogeneous elastic body with the thin cover. First we prove the convergence of the iterative algorithm based on finite element method/boundary element method (FEM/BEM) coupling using domain decomposition. Further we illustrate this algorithm with an example of $2D$ problem for the body with a cover.

1. Introduction

A lot of structures, both natural and artificial, contain thin covers or thin inclusions. Therefore, the problem of analyzing the stress-strain state of such bodies is of great importance. Typically they consist of two or more homogeneous parts that have a big differences in physical dimensions and properties between them. A lot of aspects of the problems, related to this subject, were analyzed (see for example [2, 4, 5, 7, 8]). In this paper we use the combined model, where the parts of the body with comparable physical dimensions are described by the linear elasticity equations, whereas the sress state of the thin cover is described by Tymoshenko shell theory equations [5]. These parts are connected using the appropriate coupling conditions on the common boundaries.

In order to perform numerical analysis of our model we solve the corresponding problems in thin shells by finite element method (FEM) with bubble basis functions, and the other parts of the body are solved numerically using boundary element method (BEM) with linear basis functions; the iterative domain decomposition algorithm is then used to connect the solutions in both domains.

In this paper we also prove the properties of our model and prove the convergence of the algorithm.

[†] *Key words.* Elasticity theory, boundary element method, finite element method, domain decomposition.

2. Problem statement

Let us consider a problem of plane strain of cylindrical body Ω_1 with the cover Ω_2.

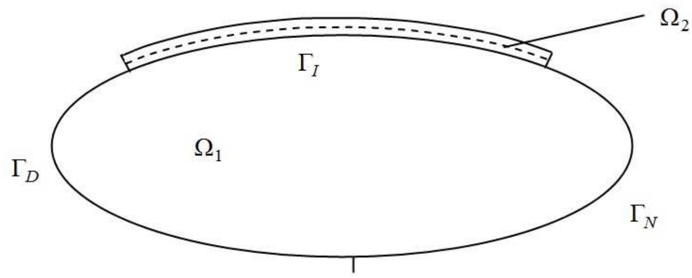


Fig. 1. Body with cover

The plane strain stress of the body in Ω_1 can be described by [1]

$$\frac{\partial \sigma_{11}}{\partial x_1} + \frac{\partial \sigma_{12}}{\partial x_2} = f_1, \qquad \frac{\partial \sigma_{21}}{\partial x_1} + \frac{\partial \sigma_{22}}{\partial x_2} = f_2 \tag{1}$$

that holds for $x \in \Omega_1$, $x = x_1, x_2$. Here $f = f_1, f_2$ denotes the volume forces that act on the body in Ω_1. From the Hook's law it follows that the components of the stress tensor can be written as

$$\sigma_{ij} = \frac{1}{2} E_1 \left(\frac{\partial u_i}{\partial x_j} + \frac{\partial u_j}{\partial x_i} \right), \qquad i, j = 1, 2,$$

where $u(x) = u_1(x), u_2(x)$ is the displacement vector with u_i being the displacements in the directions x_i for $i = 1, 2$; E_1 is the Young's modulus of the body in Ω_1. In the following we assume that no volume forces act on the body in Ω_1.

Let us denote by n the outer normal vector to Ω_1, and by τ – the tangent vector. Equations (1) are considered together with the boundary conditions

$$u_v = 0, \quad u_\tau = 0, \quad x \in \Gamma_D$$

and

$$\sigma_{vv} = 0, \quad \sigma_{v\tau} = 0, \quad x \in \Gamma_N,$$

where u_v and u_τ are the components of the stress tensor in the coordinate system n, τ. Similarly, σ_{vv} and $\sigma_{v\tau}$ are the components of the stress tensor in the n, τ coordinate system.

For the description of the cover in Ω_2 we use the equations of Timoshenko shell theory for the cylindrical shell of the form [5]

$$
\begin{gathered}
-\frac{1}{A_1}\frac{dT_{11}}{d\xi_1}-k_1T_{13}=p_1,\\
-\frac{1}{A_1}\frac{dT_{13}}{d\xi_1}+k_1T_{11}=p_3,\\
-\frac{1}{A_1}\frac{dM_{11}}{d\xi_1}+T_{13}=m_1,\quad -1\le\xi_1\le 0,
\end{gathered}
\tag{2}
$$

where v_1, w, γ_1 are the displacements and angle of revolution in the shell; T_{11}, T_{13}, M_{11} are the forces and moments in the shell; $A_1 = A_1(\xi_1)$, $k_1 = k_1(\xi_1)$ correspond to Lame parameter and median surface curvature parameter; p_1, p_3, m_1 are given functions; it holds

$$
T_{11}=\frac{E_2h}{1-v_2^2}\varepsilon_{11},\quad T_{13}=k'G'h\varepsilon_{13},\quad M_{11}=\frac{E_2h^3}{12\left(1-v_2^2\right)}\chi_{11},
\tag{3}
$$

$$
\varepsilon_{11}=\frac{1}{A_1}\frac{d\mathrm{v}_1}{d\xi_1}+k_1w,\quad \varepsilon_{13}=\frac{1}{A_1}\frac{dw}{d\xi_1}+\gamma_1-k_1\mathrm{v}_1,\quad \chi_{11}=\frac{1}{A_1}\frac{d\gamma_1}{d\xi_1},
\tag{4}
$$

$$
\begin{gathered}
p_1=\left(1+k_1\frac{h}{2}\right)\sigma_{13}^{+}-\left(1-k_1\frac{h}{2}\right)\sigma_{13}^{-},\\
p_3=\left(1+k_1\frac{h}{2}\right)\sigma_{33}^{+}-\left(1-k_1\frac{h}{2}\right)\sigma_{33}^{-},\\
m_1=\frac{h}{2}\left(\left(1+k_1\frac{h}{2}\right)\sigma_{13}^{+}-\left(1-k_1\frac{h}{2}\right)\sigma_{13}^{-}\right).
\end{gathered}
\tag{5}
$$

Here E_2 is the Young's modulus for the shell, v_2 is the Poisson's ratio; g_1, g_3 are the components of the volume forces vector, that act on the shell; $\sigma_{ij}^{+}, \sigma_{ij}^{-}$, $i, j = 1, 3$ are the components of the stress tensor on the outer ($\xi_3 = \frac{h}{2}$) and inner ($\xi_3 = -\frac{h}{2}$) surfaces of the shell. It is known, that in the case of isotropic bodies we have $k' = \frac{5}{6}$, $G' = \frac{E_2}{2(1+v_2)}$.

At each end of the thin cover we impose boundary conditions either on the displacements v_1, w and γ_1 or on the forces T_{11}, T_{13} and moment M_{11} in the shell (if the end is subjected to load or free). At the outer surface of the shell we prescribe to σ_{13}^{+} and σ_{33}^{+} some given stresses.

Remark 1. *The choice of 2D curvilinear coordinate system for the shell as ξ_1, ξ_3 (instead of ξ_1, ξ_2) is based on the fact, that 2D problem is obtained from the 3D case by assuming the cylinder being infinite in the direction of ξ_2.*

On the boundary Γ_I, common to both Ω_1 and Ω_2 we prescribe the following coupling conditions [5]*:*

$$u_v = w, \quad u_\tau = \mathrm{v}_1 - \frac{h}{2}\gamma_1,$$
$$\sigma_{vv} = \sigma_{33}^{-}, \quad \sigma_{v\tau} = \sigma_{13}^{-}. \tag{6}$$

Let us rewrite the coupling conditions (6) on Γ_I as follows:

$$u_v = w, \quad u_\tau = \mathrm{v}_1 - \frac{h}{2}\gamma_1,$$
$$A_1\left(1 - k_1\frac{h}{2}\right)\sigma_{vv} - A_1\left(1 - k_1\frac{h}{2}\right)\sigma_{33}^{-} = 0, \tag{7}$$
$$A_1\left(1 - k_1\frac{h}{2}\right)\sigma_{v\tau} - A_1\left(1 - k_1\frac{h}{2}\right)\sigma_{13}^{-} = 0.$$

3. The properties of the Steklov-Poincare operators and convergence of the domain decomposition iterative algorithm

Let us suppose that on the inferface Γ_I the displacement is equal to $\varphi = \varphi_1, \varphi_2$, $\varphi_i \in H^1(\Gamma_I)$, $i = 1, 2$. In the following we consider the Steklov-Poincare operator S for our problem as well as local Steklov-Poincare operators S_i, that correspond to Ω_i, $i = 1, 2$. Therefore, we have from (7)

$$\langle S\varphi, \psi\rangle_{\Gamma_I} = \langle S_1\varphi, \psi\rangle_{\Gamma_I} + \langle S_2\varphi, \psi\rangle_{\Gamma_I}, \qquad \forall\varphi, \psi \in H^1(\Gamma_I) \times H^1(\Gamma_I)$$

$$\langle S_1\varphi, \psi\rangle_{\Gamma_I} = \left\langle A_1\left(1 - k_1\frac{h}{2}\right)G_I\sigma_{vv}(\varphi), \psi_1\right\rangle_{\Gamma_I} +$$
$$+\left\langle A_1\left(1 - k_1\frac{h}{2}\right)G_I\sigma_{v\tau}(\varphi), \psi_2\right\rangle_{\Gamma_I}, \tag{8}$$
$$\langle S_2\varphi, \psi\rangle_{\Gamma_I} = \left\langle -A_1\left(1 - k_1\frac{h}{2}\right)\sigma_{33}^{-}(\varphi), \psi_1\right\rangle_{\Gamma_I} +$$
$$+\left\langle A_1\left(1 - k_1\frac{h}{2}\right)\sigma_{13}^{-}(\varphi), \psi_2\right\rangle_{\Gamma_I},$$

where $G_I\sigma$ is the trace of σ on Γ_I; $\langle u, v\rangle_{\Gamma_I}$ denotes the bilinear form which formally can be written as

$$\langle u, v\rangle_{\Gamma_I} = \int_{\Gamma_I} uv d\Gamma_I.$$

First we prove that there exists a unique solution to the problem for Steklov-Poincare operators. For this purpose we will use the Lax-Milgram lemma.

Let Ω_2^* be a midline of Ω_2. Without loss of generality we assume that $g_1 = g_3 = \sigma_{13}^{+} = \sigma_{33}^{+} = 0$. Moreover, one notices that all the displacements defined in Ω_2 are continuous with respect to ξ_3, since both equations and boundary conditions are independent of ξ_3. Using the coupling conditions (7), one can rewrite (8) as

$$\begin{gathered}\langle S_2\varphi,\psi\rangle_{\Gamma_I}=\left\langle -A_1\left(1-k_1\frac{h}{2}\right)\sigma_{33}^{-}\left(\varphi\right),\tilde{w}\right\rangle_{\Gamma_I}+\\ +\left\langle -A_1\left(1-k_1\frac{h}{2}\right)\sigma_{13}^{-}\left(\varphi\right),\left(\tilde{\mathrm{v}}_1-\frac{h}{2}\tilde{\gamma}_1\right)\right\rangle_{\Gamma_I}=\\ =\left(-A_1\left(1-k_1\frac{h}{2}\right)\sigma_{33}^{-},\tilde{w}\right)_{\Omega_2^*}+\left(-A_1\left(1-k_1\frac{h}{2}\right)\sigma_{13}^{-},\tilde{\mathrm{v}}_1\right)_{\Omega_2^*}+\\ +\left(A_1\frac{h}{2}\left(1-k_1\frac{h}{2}\right)\sigma_{13}^{-},\tilde{\gamma}_1\right)_{\Omega_2^*},\end{gathered}\tag{9}$$

where

$$(u,v)_{\Omega_2^*}=\int\limits_{\Omega_2^*}uv\,d\Omega_2^*.$$

Let us substitute into (9) the corresponding left sides of the system of equations (2)-(5):

$$\begin{gathered}\langle S_2\varphi,\psi\rangle_{\Gamma_I}=\left(-\frac{dT_{13}}{d\xi_1}+k_1A_1T_{11},\tilde{w}\right)_{\Omega_2^*}+\\ +\left(-\frac{dT_{11}}{d\xi_1}-k_1A_1T_{13},\tilde{\mathrm{v}}_1\right)_{\Omega_2^*}+\left(-\frac{dM_{11}}{d\xi_1}+A_1T_{13},\tilde{\gamma}_1\right)_{\Omega_2^*}.\end{gathered}$$

After integrating by parts one can easily notice that the coerciveness and symmetry of the Steklov-Poincare operator S_2 follows from the properties of the corresponding operator defined on the midline Ω_2^* which has been proven in [2]. Therefore, one obtains

$$\begin{gathered}\langle S_2\varphi,\varphi\rangle_{\Gamma_I}\geq c^2\int\limits_{-1}^{0}\left(\left(\frac{d\mathrm{v}_1}{d\xi_1}\right)^2+\left(\frac{dw}{d\xi_1}\right)^2+\left(\frac{d\gamma_1}{d\xi_1}\right)^2\right)d\Omega_2^*+\\ +c^2\int\limits_{-1}^{0}\left(\mathrm{v}_1^2+w^2+\gamma_1^2\right)d\Omega_2^*,\quad c\neq 0.\end{gathered}$$

Further,

$$\langle S_2\varphi,\varphi\rangle_{\Gamma_I}\geq c_1^2\int\limits_{-1}^{0}\left(\left(\frac{dw}{d\xi_1}\right)^2+\left(\frac{d\mathrm{v}_1}{d\xi_1}-\frac{h}{2}\frac{d\gamma_1}{d\xi_1}\right)^2\right)d\Omega_2^*+$$

$$+c_1^2\int_{-1}^{0}\left(w^2+\left(\mathrm{v}_1-\frac{h}{2}\gamma_1\right)^2\right)d\Omega_2^*,\quad c_1\neq 0.$$

Thus, S_2 is coercive. The linearity of S_2 follows directly from the linearity of the corresponding operator in Ω_2^*.

Let us now prove the continuity of S_2. For this purpose, firstly one proves the continuity of the following operator in Ω_2^*

$$(Ay,\tilde{y})_{\Omega_2^*}=\left(-\frac{dT_{13}}{d\xi_1}+k_1A_1T_{11},\tilde{w}\right)_{\Omega_2^*}+$$
$$+\left(-\frac{dT_{11}}{d\xi_1}-k_1A_1T_{13},\tilde{\mathrm{v}}_1\right)_{\Omega_2^*}+\left(-\frac{dM_{11}}{d\xi_1}+A_1T_{13},\tilde{\gamma}_1\right)_{\Omega_2^*},$$

where $y=\mathrm{v}_1,w,\gamma_1,\quad \tilde{y}=\tilde{\mathrm{v}}_1,\tilde{w},\tilde{\gamma}_1$. Using Cauchy-Schwarz inequality, one obtains for $y,\tilde{y}\in H^1(\Gamma_I)\times H^1(\Gamma_I)\times H^1(\Gamma_I)$

$$(Ay,\tilde{y})_{\Omega_2^*}=\int_{-1}^{0}\left(T_{13}\frac{d\tilde{w}}{d\xi_1}+k_1A_1T_{11}\tilde{w}\right)d\xi_1+$$
$$+\int_{-1}^{0}\left(T_{11}\frac{d\tilde{\mathrm{v}}_1}{d\xi_1}-k_1A_1T_{13}\tilde{\mathrm{v}}_1\right)d\xi_1+\int_{-1}^{0}\left(M_{11}\frac{d\tilde{\gamma}_1}{d\xi_1}+A_1T_{13}\tilde{\gamma}_1\right)d\xi_1=$$
$$=\int_{-1}^{0}\left(k'G'h\left(\frac{1}{A_1}\frac{dw}{d\xi_1}+\gamma_1-k_1\mathrm{v}_1\right)\frac{d\tilde{w}}{d\xi_1}+\right.$$
$$\left.+k_1A_1\frac{E_2h}{1-v_2^2}\left(\frac{1}{A_1}\frac{d\mathrm{v}_1}{d\xi_1}+k_1w\right)\tilde{w}\right)d\xi_1+$$
$$+\int_{-1}^{0}\left(\frac{E_2h}{1-v_2^2}\left(\frac{1}{A_1}\frac{d\mathrm{v}_1}{d\xi_1}+k_1w\right)\frac{d\tilde{\mathrm{v}}_1}{d\xi_1}-\right.$$
$$\left.-k_1A_1k'G'h\left(\frac{1}{A_1}\frac{dw}{d\xi_1}+\gamma_1-k_1\mathrm{v}_1\right)\tilde{\mathrm{v}}_1\right)d\xi_1+$$
$$+\int_{-1}^{0}\left(\frac{E_2h^3}{12\left(1-v_2^2\right)}\frac{1}{A_1}\frac{d\gamma_1}{d\xi_1}\frac{d\tilde{\gamma}_1}{d\xi_1}+A_1k'G'h\left(\frac{1}{A_1}\frac{dw}{d\xi_1}+\gamma_1-k_1\mathrm{v}_1\right)\tilde{\gamma}_1\right)d\xi_1\leq$$
$$\leq k'G'h\frac{1}{A_1^m}\left[\int_{-1}^{0}\left(\frac{dw}{d\xi_1}\right)^2d\xi_1\right]^{\frac{1}{2}}\left[\int_{-1}^{0}\left(\frac{d\tilde{w}}{d\xi_1}\right)^2d\xi_1\right]^{\frac{1}{2}}+$$

$$+k'G'h\left[\int_{-1}^{0}(\gamma_1)^2\,d\xi_1\right]^{\frac{1}{2}}\left[\int_{-1}^{0}\left(\frac{d\tilde{w}}{d\xi_1}\right)^2 d\xi_1\right]^{\frac{1}{2}}+$$

$$+k'G'h\left|k_1^M\right|\left[\int_{-1}^{0}(\mathrm{v}_1)^2\,d\xi_1\right]^{\frac{1}{2}}\left[\int_{-1}^{0}\left(\frac{d\tilde{w}}{d\xi_1}\right)^2 d\xi_1\right]^{\frac{1}{2}}+$$

$$+\frac{E_2h}{1-v_2^2}\left|k_1^M\right|\left[\int_{-1}^{0}\left(\frac{d\mathrm{v}_1}{d\xi_1}\right)^2 d\xi_1\right]^{\frac{1}{2}}\left[\int_{-1}^{0}(\tilde{w})^2\,d\xi_1\right]^{\frac{1}{2}}+$$

$$+\left(A_1\left|k_1\right|\right)^M\frac{E_2h}{1-v_2^2}\left[\int_{-1}^{0}(w)^2\,d\xi_1\right]^{\frac{1}{2}}\left[\int_{-1}^{0}(\tilde{w})^2\,d\xi_1\right]^{\frac{1}{2}}+$$

$$+\frac{E_2h}{1-v_2^2}\frac{1}{A_1^m}\left[\int_{-1}^{0}\left(\frac{d\mathrm{v}_1}{d\xi_1}\right)^2 d\xi_1\right]^{\frac{1}{2}}\left[\int_{-1}^{0}\left(\frac{d\tilde{\mathrm{v}}_1}{d\xi_1}\right)^2 d\xi_1\right]^{\frac{1}{2}}+$$

$$+\frac{E_2h}{1-v_2^2}\left|k_1^M\right|\left[\int_{-1}^{0}(w)^2\,d\xi_1\right]^{\frac{1}{2}}\left[\int_{-1}^{0}\left(\frac{d\tilde{\mathrm{v}}_1}{d\xi_1}\right)^2 d\xi_1\right]^{\frac{1}{2}}+$$

$$+\frac{E_2h}{1-v_2^2}\left|k_1^M\right|\left[\int_{-1}^{0}\left(\frac{dw}{d\xi_1}\right)^2 d\xi_1\right]^{\frac{1}{2}}\left[\int_{-1}^{0}(\tilde{\mathrm{v}}_1)^2\,d\xi_1\right]^{\frac{1}{2}}+$$

$$+k'G'h\left(A_1\left|k_1\right|\right)^M\left[\int_{-1}^{0}(\gamma_1)^2\,d\xi_1\right]^{\frac{1}{2}}\left[\int_{-1}^{0}(\tilde{\mathrm{v}}_1)^2\,d\xi_1\right]^{\frac{1}{2}}+$$

$$+k'G'h\left(A_1k_1^2\right)^M\left[\int_{-1}^{0}(\mathrm{v}_1)^2\,d\xi_1\right]^{\frac{1}{2}}\left[\int_{-1}^{0}(\tilde{\mathrm{v}}_1)^2\,d\xi_1\right]^{\frac{1}{2}}+$$

$$+\frac{E_2h^3}{12\left(1-v_2^2\right)}\frac{1}{A_1^m}\left[\int_{-1}^{0}\left(\frac{d\gamma_1}{d\xi_1}\right)^2 d\xi_1\right]^{\frac{1}{2}}\left[\int_{-1}^{0}\left(\frac{d\tilde{\gamma}_1}{d\xi_1}\right)^2 d\xi_1\right]^{\frac{1}{2}}+$$

$$+k'G'h\left[\int_{-1}^{0}\left(\frac{dw}{d\xi_1}\right)^2 d\xi_1\right]^{\frac{1}{2}}\left[\int_{-1}^{0}(\tilde{\gamma}_1)^2\,d\xi_1\right]^{\frac{1}{2}}+$$

$$+k'G'hA_1^M\left[\int_{-1}^{0}(\gamma_1)^2\,d\xi_1\right]^{\frac{1}{2}}\left[\int_{-1}^{0}(\tilde{\gamma}_1)^2\,d\xi_1\right]^{\frac{1}{2}}+$$

$$+k'G'h\left(A_1\left|k_1\right|\right)^M\left[\int_{-1}^{0}(\mathrm{v}_1)^2\,d\xi_1\right]^{\frac{1}{2}}\left[\int_{-1}^{0}(\tilde{\gamma}_1)^2\,d\xi_1\right]^{\frac{1}{2}}\leq$$

$$\leq C^2\left\|y\right\|_{H^1(\Omega_2^*)}\left\|\tilde{y}\right\|_{H^1(\Omega_2^*)},\quad C\neq 0.$$

In the above $f^M=\sup\limits_{\Omega_2^*}f$, $f^m=\inf\limits_{\Omega_2^*}f$. As a result, the continuity of the operator A is proven. Taking into account the continuity of the operator A, we can conclude

$$\langle S_2\varphi,\psi\rangle_{\Gamma_I}\leq$$

$$\leq C^2\left[\int_{-1}^{0}\left(\left(\frac{d\mathrm{v}_1}{d\xi_1}\right)^2+\left(\frac{dw}{d\xi_1}\right)^2+\left(\frac{d\gamma_1}{d\xi_1}\right)^2+\mathrm{v}_1^2+w^2+\gamma_1^2\right)d\Omega_2^*\right]^{\frac{1}{2}}\times$$

$$\times\left[\int_{-1}^{0}\left(\left(\frac{d\tilde{\mathrm{v}}_1}{d\xi_1}\right)^2+\left(\frac{d\tilde{w}}{d\xi_1}\right)^2+\left(\frac{d\tilde{\gamma}_1}{d\xi_1}\right)^2+\tilde{\mathrm{v}}_1^2+\tilde{w}^2+\tilde{\gamma}_1^2\right)d\Omega_2^*\right]^{\frac{1}{2}},\quad C\neq 0.$$

Thus, one obtains

$$\langle S_2\varphi,\psi\rangle_{\Gamma_I}\leq$$

$$\leq C_1^2\left[\int_{-1}^{0}\left(\left(\frac{dw}{d\xi_1}\right)^2+\left(\frac{d\mathrm{v}_1}{d\xi_1}-\frac{h}{2}\frac{d\gamma_1}{d\xi_1}\right)^2+w^2+\left(\mathrm{v}_1-\frac{h}{2}\gamma_1\right)^2\right)d\Omega_2^*\right]^{\frac{1}{2}}\times$$

$$\times\left[\int_{-1}^{0}\left(\left(\frac{d\tilde{w}}{d\xi_1}\right)^2\left(\frac{d\tilde{\mathrm{v}}_1}{d\xi_1}-\frac{h}{2}\frac{d\tilde{\gamma}_1}{d\xi_1}\right)^2+\tilde{w}^2+\left(\tilde{\mathrm{v}}_1-\frac{h}{2}\tilde{\gamma}_1\right)^2\right)d\Omega_2^*\right]^{\frac{1}{2}},\quad C_1\neq 0.$$

Let us consider now the local Steklov-Poincare operator S_1.

$$\langle S_1\varphi, \psi\rangle_{\Gamma_I} = \left\langle A_1\left(1 - k_1\frac{h}{2}\right) G_I\sigma_{vv}\left(\varphi\right), \psi_1\right\rangle_{\Gamma_I} +$$

$$+\left\langle A_1\left(1 - k_1\frac{h}{2}\right) G_I\sigma_{v\tau}\left(\varphi\right), \psi_2\right\rangle_{\Gamma_I}.$$

It can be shown similarly to the case of linear elasticity that the operator S_1 is coercive, symmetric, linear and continuous on $H^{1/2}\left(\Gamma_I\right)$ [3, 6]. From the equivalence of the $H^{1/2}\left(\Gamma_I\right)$ and $L_2\left(\Gamma_I\right)$ norms with the use of Friedrichs' inequality, we obtain, that the operator S_1 is linear, continuous, symmetric and coercive on $H^1\left(\Gamma_I\right)$.

To conclude, the Steklov-Poincare operator S is linear, continuous, symmetric and coercive on $H^1\left(\Gamma_I\right)$ as the sum of the operators having such properties. By the Lax-Milgram lemma, our problem for the Steklov-Poincare operator has a unique solution on $H^1\left(\Gamma_I\right)$.

We remark that for the case of nonzero volume forces as well as nonzero boundary conditions, the proof can be carried out in a similar way.

Let Q, Q_1 and Q_2 be the corresponding preconditioners in the domain decomposition algorithm [6]. It is known, that in the case of Dirichlet-Neumann iterations these preconditioners can be expressed through S_1 and S_2 as [6]

$$\begin{aligned} &Q = Q_1 + Q_2, \\ &\langle Q_1\varphi, \psi\rangle_{\Gamma_I} = \langle S_1\varphi, \psi\rangle_{\Gamma_I}, \\ &\langle Q_2\varphi, \psi\rangle_{\Gamma_I} = \langle S_2\varphi, \psi\rangle_{\Gamma_I} \end{aligned} \tag{10}$$

Since the Steklov-Poincare operators S_1 and S_2 are linear, continuous, symmetric and coercive on $H^1\left(\Gamma_I\right)$, we conclude that the operators Q, Q_1 and Q_2 also possess these properties.

Therefore, by the convergence of the Dirichlet-Neumann iterations, the following method is convergent for $0 < \theta < \theta_{max}$:

$$\varphi^{k+1} = \varphi^k + \theta Q_2^{-1}\left(G - Q\varphi^k\right), \quad k = 0, 1, 2, ...$$

where G is the right-hand side of the equation $Q\varphi = G$.

It is worth mentioning that all the properties of the continuous operators can be transferred to the corresponding discrete operators, and in the case of quasi-uniform mesh, these properties also hold for the discrete operators [6].

4. Numerical example

In this section we consider a rectangular object lying in Ω that consists of a concrete main part in Ω_1 with a thin steel cover Ω_2 attached to its top. The physical dimensions are as follows: $x_1^b = 0.05$, $x_2^b = 0.05$, $x_1^e = 1.05$, $x_2^e = 0.55$, $h = 0.02$. The physical parameters for the main part are $\nu = 0.33$, $E = 25000MPa$, for the shell – $\nu = 0.33$, $E = 200000MPa$. The body is kept fixed on both sides and subjected to the load on the bottom of $p = 1MPa/m^2$ (see Fig. 2) with zero load on top.

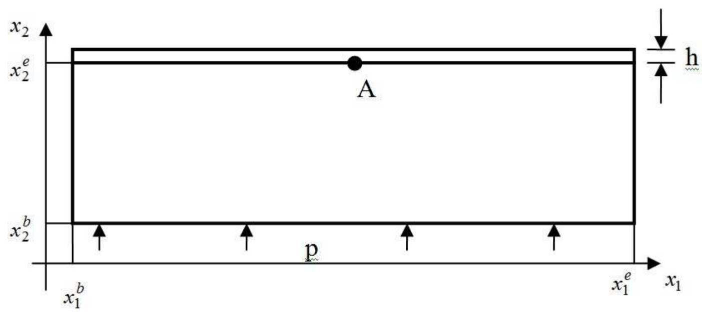


FIG. 2. Numerical Example

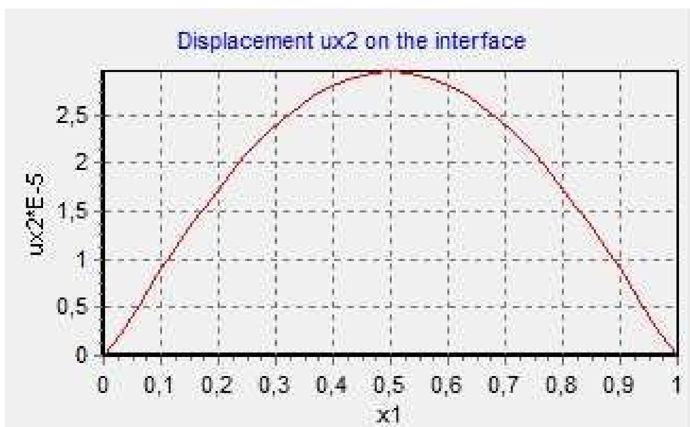


FIG. 3. Displacements in x_2 direction on the interface

The solution on each iteration in the main part is done by BEM with linear basis functions with the Galerkin method applied to integral representation formula [1]

$$\frac{1}{2}u_i = \int_\Gamma \left(F_{ij}\left(x,y\right)t_j\left(y\right)\right)d\Gamma + \int_\Gamma \left(G_{ij}\left(x,y\right)u_j\left(y\right)\right)d\Gamma, \quad i = 1,2,$$

where F_{ij} and G_{ij} are the Green's function and the co-normal derivative of Green's function respectively; $t_i = \sigma_{ij}n_j$ are the tractions.

The solution in Ω_2 is seeked as the linear combination of bubble basis functions which are defined on each element by

$$\Phi_0\left(\xi\right) = \frac{1-\xi}{2}, \quad \Phi_1\left(\xi\right) = \frac{1+\xi}{2}$$

$$\Phi_j\left(\xi\right) = \sqrt{\frac{2j-1}{2}}\int_{-1}^{\xi} P_{j-1}\left(t\right)dt, \quad j = 2,3...,$$

where $P_j(\xi)$ are the Legendre polynomials. The solution in both domains is then combined using the iterative algorithm (10).

For our example we choose 96 equally spaced boundary elements. The relaxation parameter θ is taken to be equal 0.00225

In Fig. 3 the displacement in x_2 direction along the interface is shown. The displacement achieves its maximum in the middle point A of the interface.

Bibliography

1. Brebbia C. A. Boundary Element Techniques / C. A. Brebbia, J. C. Telles, L. C. Wrobel.– Heidelberg: Springer-Verlag, 1984.
2. Dyyak I. *D*-adaptive mathematical model of solid body with thin coating / I. Dyyak, Ya. Savula // Mathematical studies.– 1997.– Vol. 7, № 1.– P. 103-109 (in Ukrainian).
3. Hsiao G. C. Boundary integral equations / G. C. Hsiao, W. L. Wendland.– Heidelberg: Springer-Verlag, 2008.
4. Krevs V. On the application of domain decomposition method to the problem of heat transfer for the bodies with thin layer / V. Krevs, Ya. Savula // Visnyk Lviv. Universytetu. Seriya mech.-math.– 1996.– № 44 (in Ukrainian).
5. Pelekh B. Generalized shell theory / B. Pelekh.– Lviv, 1978 (in Ukrainian).
6. Quarteroni A. Domain decomposition methods for partial differential equations / A. Quarteroni, A. Valli.– Oxford, 1999.
7. Sulym H. The foundations of the mathematical theory of thermoelastic equilibrium for deformable solid bodies with thin inclusions / H. Sulym.– Lviv, 2007 (in Ukrainian).
8. Vynnytska L. The stress-strain state of elastic body with thin inclusion / L. Vynnytska, Ya. Savula // *Ph*-math. modelling and information technologies.– 2008.– № 7.– P. 21-29 (in Ukrainian).

Ivan Dyyak, Yarema Savula, Andriy Styahar,
Ivan Franko National University of Lviv,
1, Universytets'ka Str., Lviv, 79000, Ukraine

Received 30.07.2012

Журнал обчислювальної та прикладної математики
2012, № 3 (109) с. 34–52
Journal of Numerical & Applied Mathematics

UDC 519.62

EXACT THREE-POINT DIFFERENCE SCHEME FOR SECOND ORDER NONLINEAR ORDINARY DIFFERENTIAL EQUATIONS WITH BOUNDARY CONDITIONS OF THE THIRD KIND

Lyubomyr Gnativ, Marta Król, Myroslav Kutniv

Резюме. Для нелінійних звичайних диференціальних рівнянь другого порядку з похідною в правій частині та крайовими умовами третього роду побудовано та обґрунтовано точну триточкову різницеву схему на нерівномірній сітці. Доведено існування та єдиність розв'язку цієї схеми, збіжність методу простої ітерації для її розв'язування.

Abstract. Exact three-point difference scheme on a nonuniform grid for the second-order nonlinear ordinary differential equations with derivative in the right-hand side and boundary conditions of the third kind is constructed and justified. The existence and uniqueness of solution of this scheme, the convergence of the method of simple iteration for its solution are proved.

1. Introduction

The exact three-point difference scheme (ETDS) and three-point difference schemes (TDS) of high order accuracy on a uniform grid for the second-order nonlinear ordinary differential equations with no derivative in the right-hand side and Dirichlet boundary conditions is constructed and justified in [10, 11]. These results on a nonuniform grid were generalized and developed in [9] and for monotone boundary value problems in [1, 7]. Difference boundary conditions of the third kind is constructed in [6, 8].

In this chapter for the nonlinear boundary value problem (BVP)

$$\frac{d}{dx}\left[k(x)\frac{du}{dx}\right] = -f\left(x, u, \frac{du}{dx}\right), \quad x \in (0,1), \tag{1}$$

$$k(0)\frac{du(0)}{dx} - \beta_1 u(0) = -\mu_1, \quad -k(1)\frac{du(1)}{dx} - \beta_2 u(1) = -\mu_2, \tag{2}$$

where $k(x), f(x,u,\xi)$ are given functions and $\beta_1, \beta_2, \mu_1, \mu_2$ are given numbers, exact three-point difference scheme is constructed. We prove the existence and the uniqueness of the solution of the ETDS and convergence of the method of simple iteration its solution for the operator of BVP (1), (2) with monotone conditions.

---

†*Key words.* Nonlinear boundary value problem, exact three-point difference schemes, method of simple iteration.

2. Existence and uniqueness of a solution

The function $u(x) \in W_2^1(0,1)$ is a week solution of problem (1), (2), if $\forall u(x), v(x) \in W_2^1(0,1)$ satisfies the relation

$$\int_0^1 k(x)\frac{du}{dx}\frac{dv}{dx}dx + (\beta_1 u(0) - \mu_1)v(0) + (\beta_2 u(1) - \mu_2)v(1) =$$

$$= \int_0^1 f\left(x, u, \frac{du}{dx}\right) v(x)dx.$$

Sufficient conditions for the existence and uniqueness of a weak solution of problem (1), (2) are given in the next theorem.

Theorem 1. *Let the following assumptions be satisfied*

$$0 < c_1 \le k(x) \le c_2 \quad \forall x \in [0,1]\ , \quad k(x) \in Q^1[0,1], \tag{3}$$

$$\begin{array}{l} f_{u\xi}(x) \equiv f\,(x,u,\xi) \in Q^0[0,1] \quad \forall u, \xi \in \mathbb{R}^1, \\ f_x\,(u,\xi) \equiv f\,(x,u,\xi) \in C(\mathbb{R}^2) \quad \forall x \in [0,1], \end{array} \tag{4}$$

$$|f\,(x,u,\xi) - f_0(x)| \le c(|u|)[g(x) + |\xi|] \quad \forall x \in [0,1],\ u, \xi \in \mathbb{R}^1, \tag{5}$$

$$[f\,(x,u,\xi) - f\,(x,v,\eta)]\,(u - v) \le 0 \quad \forall x \in [0,1],\ u, v, \xi, \eta \in \mathbb{R}^1, \tag{6}$$

$$\beta_1 > 0, \quad \beta_2 > 0, \tag{7}$$

then, the BVP (1), (2) has a unique solution $u(x) \in W_2^1(0,1)$*, with* $u(x), k(x)\dfrac{du}{dx} \in C[0,1]$.

Here $c(t)$ *is a continuous function,* $f_0(x) \in L_2(0,1)$*,* $g(x) \in L_1(0,1)$*,* c_1, c_2, c_3 *are constants,* $Q^p[0,1]$ *is the class of functions having p piece-wise continuous derivatives and a finite number of discontinuity points of first kind.*

Proof. Due to (4) and (5) the function $f\,(x,u,\xi)$ satisfies the Caratheodory conditions [3, p.63] and belongs to the class $L_1(0,1)$ (see e.g.[3, c.113]), we can define the operator $A\,(x,u)$ the identity

$$(A\,(x,u)\,,v) = \int_0^1 k(x)\frac{du}{dx}\frac{dv}{dx}dx - \int_0^1 \tilde{f}\left(x, u, \frac{du}{dx}\right) v(x)dx +$$

$$+ (\beta_1 u(0) - \mu_1)v(0) + (\beta_2 u(1) - \mu_2)v(1) \quad \forall u(x), v(x) \in W_2^1(0,1),$$

where

$$\tilde{f}(x,u,\xi) = f(x,u,\xi) - f_0(x).$$

Note that the function $u(x) \in W_2^1(0,1)$ is absolutely continuous on $[0,1]$, and its generalized derivative $\dfrac{du}{dx}$ is equal to the classical derivative almost everywhere on $[0,1]$ (see e.g. [3, c.74]). Thus, $u(x) \in C[0,1]$, $\dfrac{du}{dx} \in L_2(0,1)$.

Let us show that the operator $A\,(x,u)$ is bounded. Actually, taking into account the Cauchy-Bunyakovsky-Schwarz inequality, the conditions (3) and (5), the inequality $c(|u|) \le C_2$ for all $x \in [0,1]$ as well as $\|v\|_{C[0,1]} \le C_1 \|v\|_{1,2,(0,1)}$ for all $v(x) \in W_2^1(0,1)$ (see e.g.[3, c.112]) we obtain

$$|(A(x,u),v)| \le \left\{ \int_0^1 \left[k(x)\frac{du}{dx} \right]^2 dx \right\}^{1/2} \left\{ \int_0^1 \left[\frac{dv}{dx} \right]^2 dx \right\}^{1/2} +$$

$$+ \|v\|_{C[0,1]} \left[\int_0^1 \left| \tilde{f}\left(x,u,\frac{du}{dx}\right) \right| dx + (\beta_1+\beta_2)\|u\|_{C[0,1]} + |\mu_1+\mu_2| \right] \le$$

$$\le \Big[\left(c_2 + C_1^2(\beta_1+\beta_2)\right) \|u\|_{1,2,(0,1)} +$$

$$+C_1 \left\| \tilde{f} \right\|_{0,1,(0,1)} + C_1|\mu_1+\mu_2| \Big] \|v\|_{1,2,(0,1)} \le$$

$$\le \Big[\left(c_2 + C_1(C_2 + C_1(\beta_1+\beta_2))\right) \|u\|_{1,2,(0,1)} + C_1 C_2 \|g\|_{0,1,(0,1)} +$$

$$+C_1|\mu_1+\mu_2| \Big] \|v\|_{1,2,(0,1)},$$

where

$$\|u\|_{C[0,1]} = \max_{x\in[0,1]} |u(x)|, \quad \|u\|_{0,1,(0,1)} = \int_0^1 |u(x)|dx,$$

$$\|u\|_{0,2,(0,1)} = \left[\int_0^1 (u(x))^2 \, dx \right]^{1/2},$$

$$\|u\|_{1,2,(0,1)} = \left[\int_0^1 (u(x))^2 \, dx + \int_0^1 \left(\frac{du}{dx}\right)^2 dx \right]^{1/2}.$$

If $u_n \to u_0$ in $W_2^1(0,1)$, then $\tilde{f}\left(x,u_n,\frac{du_n}{dx}\right) \to \tilde{f}\left(x,u_0,\frac{du_0}{dx}\right)$, $k(x)\frac{du_n}{dx} \to k(x)\frac{du_0}{dx}$ in $L_1(0,1)$ (see e.g.[3, c.113]). Thus, for $\forall v(x) \in W_2^1(0,1)$ we have

$$\lim_{n\to\infty} (A(x,u_n),v) = \lim_{n\to\infty} \left[\int_0^1 k(x)\frac{du_n}{dx}\frac{dv}{dx}dx - \int_0^1 \tilde{f}\left(x,u_n,\frac{du_n}{dx}\right) v(x)dx + \right.$$

$$\left. +(\beta_1 u_n(0) - \mu_1)v(0) + (\beta_2 u_n(1) - \mu_2)v(1) \right] =$$

$$= \int_0^1 k(x)\frac{du_0}{dx}\frac{dv}{dx}dx - \int_0^1 \tilde{f}\left(x,u_0,\frac{du_0}{dx}\right) v(x)dx +$$

$$+ (\beta_1 u(0) - \mu_1)v(0) + (\beta_2 u(1) - \mu_2)v(1) = (A(x,u_0),v),$$

i.e. the operator $A(x,u)$ is demicontinuous.

Let us show that the operator $A(x,u)$ is strongly monotone. Due to the conditions (3), (6) and (7), taking into account the Friedrichs inequality (see e.g. [2, c.187])

$$\int_0^1 u^2(x)dx \le \frac{16}{\pi^2}\int_0^1 \left(\frac{du}{dx}\right)^2 dx + \frac{4}{\pi}\left[u^2(0)+u^2(1)\right],$$

we obtain

$$(A(x,u) - A(x,v), u-v) = \int_0^1 k(x)\left[\frac{du}{dx} - \frac{dv}{dx}\right]^2 dx -$$

$$-\int_0^1 \left[f\left(x,u(x),\frac{du}{dx}\right) - f\left(x,v(x),\frac{dv}{dx}\right)\right][u(x)-v(x)]\,dx +$$

$$+\beta_1(u(0)-v(0))^2 + \beta_2(u(1)-v(1))^2 \ge c_1\left\|\frac{du}{dx} - \frac{dv}{dx}\right\|^2_{0,2,(0,1)} +$$

$$+\beta_1(u(0)-v(0))^2 + \beta_2(u(1)-v(1))^2 \ge$$

$$\ge \min\left\{\frac{\pi^2 c_1}{16}, \frac{\pi\beta_1}{4}, \frac{\pi\beta_2}{4}\right\}\|u-v\|^2_{0,2,(0,1)}.$$

From the strong monotonicity follows the coerciveness of $A(x,u)$.

Thus, the Browder-Minty theorem (see [3, c.204]) guaranties the existence of a unique solution $u \in W_2^1(0,1)$ of problem (1), (2). ☐

Since

$$k(x)\frac{du}{dx} = \int_0^x f\left(t,u,\frac{du}{dt}\right)dt + C$$

almost everywhere on [0,1] (see e.g. [3, c.134]), i.e. the flux $k(x)\dfrac{du}{dx}$ is the undefined Lebesgue integral, this function is absolutely continuous on [0,1], and the claim $k(x)\dfrac{du}{dx} \in C[0,1]$ is shown.

3. Existence of an exact three-point difference scheme

On the closed $(0,1)$ we introduce an nonuniform grid

$$\hat{\omega}_h = \left\{x_j \in (0,1), \quad j=1,2,...,N-1, \quad h_j = x_j - x_{j-1} > 0, \quad \sum_{j=1}^N h_j = 1\right\}$$

such that the discontinuity points of functions $k(x), f(x,u,\xi)$ coincide with the nodes of the grid $\hat{\omega}_h$. Denote by ρ the set of all discontinuity points and assume that N is such that $\rho \subseteq \hat{\omega}_h$. At points of discontinuity we use the continuity conditions for BVP(1), (2)

$$u(x_i-0)=u(x_i+0),\quad k(x)\frac{du}{dx}\Big|_{x=x_i-0} = k(x)\frac{du}{dx}\Big|_{x=x_i+0} \quad \forall x_i\in\rho.$$

We will use the following notation

$$e_\alpha^j=(x_{j-2+\alpha},x_{j-1+\alpha}),\quad \bar e_\alpha^j=[x_{j-2+\alpha},x_{j-1+\alpha}].$$

Consider the boundary value problems

$$\begin{aligned}&\frac{d}{dx}\left(k(x)\frac{dY_2^0(x,u)}{dx}\right)=-f\left(x,Y_2^0(x,u),\frac{dY_2^0(x,u)}{dx}\right),\quad x\in e_2^0,\\ &k(x_0)\frac{dY_2^0(x_0,u)}{dx}-\beta_1Y_2^0(x_0,u)=-\mu_1,\quad Y_2^0(x_1,u)=u(x_1),\end{aligned}\tag{8}$$

$$\begin{aligned}&\frac{d}{dx}\left(k(x)\frac{dY_\alpha^j(x,u)}{dx}\right)=-f\left(x,Y_\alpha^j(x,u),\frac{dY_\alpha^j(x,u)}{dx}\right),\quad x\in e_\alpha^j,\\ &Y_\alpha^j(x_{j-2+\alpha},u)=u(x_{j-2+\alpha}),\quad Y_\alpha^j(x_{j-1+\alpha},u)=u(x_{j-1+\alpha}),\\ &j=3-\alpha,4-\alpha,...,N-\alpha,\quad \alpha=1,2,\end{aligned}\tag{9}$$

$$\begin{aligned}&\frac{d}{dx}\left(k(x)\frac{dY_1^N(x,u)}{dx}\right)=-f\left(x,Y_1^N(x,u),\frac{dY_1^N(x,u)}{dx}\right),\quad x\in e_1^N,\\ &Y_1^N(x_{N-1},u)=u(x_{N-1}),\\ &-k(x_N)\frac{dY_1^N(x_N,u)}{dx}-\beta_2Y_1^N(x_N,u)=-\mu_2.\end{aligned}\tag{10}$$

Lemma 1. *Let the assumptions of Theorem* 1 *be satisfied. Then each of the problems (8)-(10) has a unique solution* $Y_\alpha^j(x,u)\in W_2^1(e_\alpha^j)$, $j=2-\alpha,3-\alpha,...,N+1-\alpha,\ \alpha=1,2$, *with*

$$Y_\alpha^j(x,u),k(x)\frac{dY_\alpha^j(x,u)}{dx}\in C(\bar e_\alpha^j)$$

and for the solution BVP (1), (2) it holds

$$u(x)=Y_\alpha^j(x,u),\quad x\in\bar e_\alpha^j.\tag{11}$$

Proof. We introduce the nonlinear operators for problems (8)-(10) by the equations

$$\begin{aligned}(A_2^0(x,Y_2^0),v)=&\int_{x_0}^{x_1}k(x)\frac{dY_2^0(x,u)}{dx}\frac{dv(x)}{dx}dx-\\ &-\int_{x_0}^{x_1}\tilde f\left(x,Y_2^0(x,u),\frac{dY_2^0(x,u)}{dx}\right)v(x)dx+(\beta_1Y_2^0(0,u)-\mu_1)v(0),\end{aligned}$$

$$\left(A_\alpha^j\left(x,Y_\alpha^j\right),v\right)=$$
$$=\int\limits_{x_{j-2+\alpha}}^{x_{j-1+\alpha}} k(x)\frac{dY_\alpha^j(x,u)}{dx}\frac{dv(x)}{dx}dx-\int\limits_{x_{j-2+\alpha}}^{x_{j-1+\alpha}}\tilde{f}\left(x,Y_\alpha^j(x,u),\frac{dY_\alpha^j(x,u)}{dx}\right)v(x)dx,$$
$$(A_1^N(x,Y_1^N),v)=\int\limits_{x_{N-1}}^{x_N} k(x)\frac{dY_1^N(x,u)}{dx}\frac{dv(x)}{dx}dx-$$
$$-\int\limits_{x_{N-1}}^{x_N}\tilde{f}\left(x,Y_1^N(x,u),\frac{dY_1^N(x,u)}{dx}\right)v(x)dx+(\beta_2Y_1^N(1,u)-\mu_2)v(1),$$
$$\tilde{f}(x,u,\xi)=f(x,u,\xi)-f_0(x),$$
that true for $\forall Y_\alpha^j(x,u),v(x)\in W_2^1(e_\alpha^j)$.

Let us show that the operators $A_2^0\left(x,Y_2^0\right)$, $A_\alpha^j\left(x,Y_\alpha^j\right)$, $A_1^N\left(x,Y_1^N\right)$ are bounded. Taking into account the Cauchy-Bunyakovsky-Schwarz inequality, the conditions (3), (5) with $c(|Y_\alpha^j(x,u)|)\le C_2$, $\forall x\in\bar{e}_\alpha^j$ and inequality $\|v\|_{C(\bar{e}_\alpha^j)}\le C_1\|v\|_{1,2,e_\alpha^j}$, $\forall v(x)\in W_2^1(e_\alpha^j)$, we obtain

$$\left|\left(A_2^0\left(x,Y_2^0\right),v\right)\right|\le\left\{\int\limits_{x_0}^{x_1}\left[k(x)\frac{dY_2^0}{dx}\right]^2dx\right\}^{1/2}\left\{\int\limits_{x_0}^{x_1}\left[\frac{dv}{dx}\right]^2dx\right\}^{1/2}+$$
$$+\|v\|_{C(\bar{e}_2^0)}\left[\int\limits_{x_0}^{x_1}\left|\tilde{f}\left(x,Y_2^0,\frac{dY_2^0}{dx}\right)\right|dx+\beta_1\left\|Y_2^0\right\|_{C(\bar{e}_2^0)}+|\mu_1|\right]\le$$
$$\le\left[\left(c_2+C_1^2\beta_1\right)\left\|Y_2^0\right\|_{1,2,e_2^0}+C_1\left\|\tilde{f}\right\|_{0,1,e_2^0}+C_1\left|\mu_1\right|\right]\|v\|_{1,2,e_2^0}\le$$
$$\le\left[\left(c_2+C_1\left(C_2+C_1\beta_1\right)\right)\left\|Y_2^0\right\|_{1,2,e_2^0}+C_1C_2\left\|g\right\|_{0,1,e_2^0}+C_1\left|\mu_1\right|\right]\|v\|_{1,2,e_2^0},$$
$$\left|\left(A_\alpha^j\left(x,Y_\alpha^j\right),v\right)\right|\le\left\{\int\limits_{x_{j-2+\alpha}}^{x_{j-1+\alpha}}\left[k(x)\frac{dY_\alpha^j}{dx}\right]^2dx\right\}^{1/2}\left\{\int\limits_{x_{j-2+\alpha}}^{x_{j-1+\alpha}}\left[\frac{dv}{dx}\right]^2dx\right\}^{1/2}+$$
$$+\|v\|_{C(\bar{e}_\alpha^j)}\int\limits_{x_{j-2+\alpha}}^{x_{j-1+\alpha}}\left|\tilde{f}\left(x,Y_\alpha^j(x,u),\frac{dY_\alpha^j}{dx}\right)\right|dx\le$$
$$\le\left[c_2\left\|Y_\alpha^j\right\|_{1,2,e_\alpha^j}+C_1\left\|\tilde{f}\right\|_{0,1,e_\alpha^j}\right]\|v\|_{1,2,e_\alpha^j}\le$$
$$\le\left[\left(c_2+C_1C_2\right)\left\|Y_\alpha^j\right\|_{1,2,e_\alpha^j}+C_1C_2\left\|g\right\|_{0,1,e_\alpha^j}\right]\|v\|_{1,2,e_\alpha^j},$$
$$\left|\left(A_1^N\left(x,Y_1^N\right),v\right)\right|\le\left\{\int\limits_{x_{N-1}}^{x_N}\left[k(x)\frac{dY_1^N}{dx}\right]^2dx\right\}^{1/2}\left\{\int\limits_{x_{N-1}}^{x_N}\left[\frac{dv}{dx}\right]^2dx\right\}^{1/2}+$$

$$+\|v\|_{C(\bar{e}_1^N)}\left[\int\limits_{x_{N-1}}^{x_N}\left|\tilde{f}\left(x,Y_1^N,\frac{dY_1^N}{dx}\right)\right|dx+\beta_2\left\|Y_1^N\right\|_{C(\bar{e}_1^N)}+|\mu_2|\right]\le$$

$$\le\left[\left(c_2+C_1^2\beta_2\right)\left\|Y_1^N\right\|_{1,2,e_1^N}+C_1\left\|\tilde{f}\right\|_{0,1,e_1^N}+C_1|\mu_2|\right]\|v\|_{1,2,e_1^N}\le$$

$$\le\left[\left(c_2+C_1\left(C_2+C_1\beta_2\right)\right)\left\|Y_1^N\right\|_{1,2,e_1^N}+C_1C_2\|g\|_{0,1,e_1^N}+C_1|\mu_2|\right]\|v\|_{1,2,e_1^N}.$$

The demicontinuity of operators $A_2^0\left(x,Y_2^0\right), A_\alpha^j\left(x,Y_\alpha^j\right), A_1^N\left(x,Y_1^N\right)$ follows from the condition (5). Really (see[3, p.113]), if $Y_{\alpha n}^j(x,u)\to Y_{\alpha 0}^j(x,u)$ in $W_2^1(e_\alpha^j)$, then

$$\tilde{f}\left(x,Y_{\alpha n}^j(x,u),\frac{dY_{\alpha n}^j(x,u)}{dx}\right)\to\tilde{f}\left(x,Y_{\alpha 0}^j(x,u),\frac{dY_{\alpha 0}^j(x,u)}{dx}\right),$$

$$k(x)\frac{dY_{\alpha n}^j(x,u)}{dx}\to k(x)\frac{dY_{\alpha 0}^j(x,u)}{dx},$$

$$j=2-\alpha,3-\alpha,...,N+1-\alpha,\quad \alpha=1,2,$$

in space $L_1(e_\alpha^j)$. Thus, for $\forall v(x)\in W_2^1(e_\alpha^j)$

$$\lim_{n\to\infty}\left(A_2^0\left(x,Y_{2,n}^0\right),v\right)=\lim_{n\to\infty}\left[\int\limits_{x_0}^{x_1}k(x)\frac{dY_{2,n}^0}{dx}\frac{dv}{dx}dx-\right.$$

$$\left.-\int\limits_{x_0}^{x_1}\tilde{f}\left(x,Y_{2,n}^0,\frac{dY_{2,n}^0}{dx}\right)v(x)dx+\left(\beta_1Y_{2,n}^0(0,u)-\mu_1\right)v(0)\right]=$$

$$=\int\limits_{x_0}^{x_1}k(x)\frac{dY_{2,0}^0}{dx}\frac{dv}{dx}dx-\int\limits_{x_0}^{x_1}\tilde{f}\left(x,Y_{2,0}^0,\frac{dY_{2,0}^0}{dx}\right)v(x)dx+$$

$$+\left(\beta_1Y_{2,0}^0(0,u)-\mu_1\right)v(0)=\left(A_2^0\left(x,Y_{2,0}^0\right),v\right),$$

$$\lim_{n\to\infty}\left(A_\alpha^j\left(x,Y_{\alpha n}^j\right),v\right)$$

$$=\lim_{n\to\infty}\left\{\int\limits_{x_{j-2+\alpha}}^{x_{j-1+\alpha}}k(x)\frac{dY_{\alpha n}^j}{dx}\frac{dv(x)}{dx}dx-\int\limits_{x_{j-2+\alpha}}^{x_{j-1+\alpha}}\tilde{f}\left(x,Y_{\alpha n}^j,\frac{dY_{\alpha n}^j}{dx}\right)v(x)dx\right\}=$$

$$=\int\limits_{x_{j-2+\alpha}}^{x_{j-1+\alpha}}k(x)\frac{dY_{\alpha 0}^j}{dx}\frac{dv}{dx}dx-\int\limits_{x_{j-2+\alpha}}^{x_{j-1+\alpha}}\tilde{f}\left(x,Y_{\alpha 0}^j,\frac{dY_{\alpha 0}^j}{dx}\right)v(x)dx=$$

$$=\left(A_\alpha^j\left(x,Y_{\alpha 0}^j\right),v\right),$$

$$\lim_{n\to\infty}\left(A_1^N\left(x,Y_{1,n}^N\right),v\right)=$$

$$=\lim_{n\to\infty}\left[\int\limits_{x_{N-1}}^{x_N}k(x)\frac{dY_{1,n}^N}{dx}\frac{dv}{dx}dx-\right.$$

$$\left.-\int\limits_{x_{N-1}}^{x_N}\tilde{f}\left(x,Y_{1,n}^N,\frac{dY_{1,n}^N}{dx}\right)v(x)dx+\left(\beta_2Y_{1,n}^N(1,u)-\mu_2\right)v(1)\right]=$$

$$=\int\limits_{x_{N-1}}^{x_N}k(x)\frac{dY_{1,0}^N}{dx}\frac{dv}{dx}dx-\int\limits_{x_{N-1}}^{x_N}\tilde{f}\left(x,Y_{1,0}^N,\frac{dY_{1,0}^N}{dx}\right)v(x)dx+$$

$$+\left(\beta_2Y_{1,0}^N(1,u)-\mu_2\right)v(1)=\left(A_1^N\left(x,Y_{1,0}^N\right),v\right),$$

that operators $A_2^0\left(x,Y_2^0\right), A_\alpha^j\left(x,Y_\alpha^j\right), A_1^N\left(x,Y_1^N\right)$ are demicontinuous.

Let us show that the operators $A_2^0\left(x,Y_2^0\right)$, $A_\alpha^j\left(x,Y_\alpha^j\right)$, $A_1^N\left(x,Y_1^N\right)$ are strongly monotone. Due to the conditions (4), (7), taking into account the Friedrichs inequalities (see e.g. [2, c.187])

$$\int\limits_a^b u^2(x)dx\le\frac{16(b-a)^2}{\pi^2}\int\limits_a^b\left(\frac{du}{dx}\right)^2dx+\frac{\pi(b-a)}{4}u^2(a),$$

$$\int\limits_a^b u^2(x)dx\le\frac{16(b-a)^2}{\pi^2}\int\limits_a^b\left(\frac{du}{dx}\right)^2dx+\frac{\pi(b-a)}{4}u^2(b)$$

we obtain

$$\left(A_2^0(x,Y_2^0)-A_2^0(x,\tilde{Y}_2^0),Y_2^0-\tilde{Y}_2^0\right)=\int\limits_{x_0}^{x_1}k(x)\left(\frac{dY_2^0}{dx}-\frac{d\tilde{Y}_2^0}{dx}\right)^2dx-$$

$$-\int\limits_{x_0}^{x_1}\left[f\left(x,Y_2^0,\frac{dY_2^0}{dx}\right)-f\left(x,\tilde{Y}_2^0,\frac{d\tilde{Y}_2^0}{dx}\right)\right]\left[Y_2^0(x,u)-\tilde{Y}_2^0(x,u)\right]dx+$$

$$+\beta_1\left(Y_2^0(0,u)-\tilde{Y}_2^0(0,u)\right)^2\ge c_1\left\|\frac{dY_2^0}{dx}-\frac{d\tilde{Y}_2^0}{dx}\right\|_{0,2,e_2^0}^2+$$

$$+\beta_1\left(Y_2^0(0,u)-\tilde{Y}_2^0(0,u)\right)^2\ge\min\left\{\frac{\pi^2c_1}{16},\frac{\pi\beta_1}{4}\right\}\left\|Y_2^0-\tilde{Y}_2^0\right\|_{0,2,e_2^0}^2$$

$$\left(A_\alpha^j(x,Y_\alpha^j)-A_\alpha^j(x,\tilde{Y}_\alpha^j),Y_\alpha^j-\tilde{Y}_\alpha^j\right)=\int\limits_{x_{j-2+\alpha}}^{x_{j-1+\alpha}}k(x)\left(\frac{dY_\alpha^j}{dx}-\frac{d\tilde{Y}_\alpha^j}{dx}\right)^2dx-$$

$$-\int\limits_{x_{j-2+\alpha}}^{x_{j-1+\alpha}}\left[f\left(x, Y_\alpha^j, \frac{dY_\alpha^j}{dx}\right)-f\left(x, \tilde{Y}_\alpha^j, \frac{d\tilde{Y}_\alpha^j}{dx}\right)\right]\left[Y_\alpha^j(x,u)-\tilde{Y}_\alpha^j(x,u)\right]dx \geq$$

$$\geq c_1\left\|\frac{dY_\alpha^j}{dx}-\frac{d\tilde{Y}_\alpha^j}{dx}\right\|_{0,2,e_\alpha^j}^2,$$

$$\left(A_1^N(x,Y_1^N)-A_1^N(x,\tilde{Y}_1^N), Y_1^N-\tilde{Y}_1^N\right)=\int\limits_{x_{N-1}}^{x_N}k(x)\left(\frac{dY_1^N}{dx}-\frac{d\tilde{Y}_1^N}{dx}\right)^2dx-$$

$$-\int\limits_{x_{N-1}}^{x_N}\left[f\left(x, Y_1^N, \frac{dY_1^N}{dx}\right)-f\left(x, \tilde{Y}_1^N, \frac{d\tilde{Y}_1^N}{dx}\right)\right]\left[Y_1^N(x,u)-\tilde{Y}_1^N(x,u)\right]dx+$$

$$+\beta_2\left(Y_1^N(1,u)-\tilde{Y}_1^N(1,u)\right)^2 \geq c_1\left\|\frac{dY_1^N}{dx}-\frac{d\tilde{Y}_1^N}{dx}\right\|_{0,2,e_1^N}^2+$$

$$+\beta_2\left(Y_1^N(1,u)-\tilde{Y}_1^N(1,u)\right)^2 \geq \min\left\{\frac{\pi^2 c_1}{16}, \frac{\pi\beta_2}{4}\right\}\left\|Y_1^N-\tilde{Y}_1^N\right\|_{0,2,e_1^N}^2$$

From the strong monotonicity follows the coerciveness of operators $A_2^0\left(x, Y_2^0\right)$, $A_\alpha^j\left(x, Y_\alpha^j\right)$, $A_1^N\left(x, Y_1^N\right)$.

Thus, the Browder-Minty theorem (see e.g.[3, p.204]) guaranties the existence of a unique solutions of problems (8)-(10).

Since

$$k(x)\frac{dY_\alpha^j(x,u)}{dx}=\int\limits_{x_{j-2+\alpha}}^{x_{j-1+\alpha}}f\left(t, Y_\alpha^j(t,u), \frac{dY_\alpha^j(t,u)}{dt}\right)dt+C,$$

then the function is absolutely continuous on $\bar{e}_\alpha^j$, that $k(x)\dfrac{dY_\alpha^j(x,u)}{dx}\in C(\bar{e}_\alpha^j)$, $j=2-\alpha, 3-\alpha, ..., N+1-\alpha,\ \alpha=1,2.$

Since $Y_\alpha^j(x,u)$ is the solution of (8)-(10), this function is also the solution of BVP (1), (2) which is unique due to the assumptions of our lemma. □

Now we are at the position to prove the next statement

Theorem 2. *Let the assumptions of Theorem* 1 *be satisfied. Then there exists the following ETDS for problem (1), (2)*

$$(au_{\bar{x}})_{\hat{x}}=-\hat{T}^x\left(f\left(\xi, u(\xi), \frac{du(\xi)}{d\xi}\right)\right), \quad x\in\hat{\omega}_h, \tag{12}$$

$$a_1u_{x,0}-\beta_1u_0=-\mu_1-h_1\hat{T}^{x_0}\left(f\left(\xi, u(\xi), \frac{du(\xi)}{d\xi}\right)\right),$$

$$-a_Nu_{\bar{x},N}-\beta_2u_N=-\mu_2-h_N\hat{T}^{x_N}\left(f\left(\xi, u(\xi), \frac{du(\xi)}{d\xi}\right)\right).$$

This ETDS has a unique solution $u(x)$ $\forall x \in \hat{\omega}_h$ which coincides with the solution (1), (2) at the points of the grid $\hat{\omega}_h$, where

$$u_{\bar{x},j} = \frac{u_j - u_{j-1}}{h_j}, \quad u_{\hat{x},j} = \frac{u_{j+1} - u_j}{\hbar_j}, \quad \hbar_j = \frac{h_j + h_{j+1}}{2},$$

$$a(x_j) = \left[\frac{1}{h_j} V_1^j(x_j)\right]^{-1},$$

$$\hat{T}^{x_j}(w(\xi)) = [\hbar_j V_1^j(x_j)]^{-1} \int\limits_{x_{j-1}}^{x_j} V_1^j(\xi) w(\xi) d\xi + [\hbar_j V_2^j(x_j)]^{-1} \int\limits_{x_j}^{x_{j+1}} V_2^j(\xi) w(\xi) d\xi,$$

$$\hat{T}^{x_0}(w(\xi)) = [h_1 V_1^1(x_1)]^{-1} \int\limits_{x_0}^{x_1} V_2^0(\xi) w(\xi) d\xi, \tag{13}$$

$$\hat{T}^{x_N}(w(\xi)) = [h_N V_1^N(x_N)]^{-1} \int\limits_{x_{N-1}}^{x_N} V_1^N(\xi) w(\xi) d\xi,$$

$$V_1^j(x) = \int\limits_{x_{j-1}}^{x} \frac{dx}{k(x)}, \quad V_2^j(x) = \int\limits_{x}^{x_{j+1}} \frac{dx}{k(x)}.$$

The function $u(x)$ on the right-hand side of (12) is defined by (11) and depends only on $u(x_j)$, $j = 0, 1, ..., N$.

Proof. Applying the operator $\hat{T}^{x_j}$ to both sides of equation (1) we obtain

$$\hat{T}^{x_j}\left(\frac{d}{d\xi}\left(k(\xi)\frac{du(\xi)}{d\xi}\right)\right) = -\hat{T}^{x_j}\left(f\left(\xi, u(\xi), \frac{du(\xi)}{d\xi}\right)\right), \quad j = 0, 1, 2, ..., N,$$

where

$$\hat{T}^{x_0}\left(\frac{d}{d\xi}\left(k(\xi)\frac{du(\xi)}{d\xi}\right)\right) = \left[h_1 V_1^1(x_1)\right]^{-1} \int\limits_{x_0}^{x_1} V_2^0(\xi) \frac{d}{d\xi}\left[k(\xi)\frac{du(\xi)}{d\xi}\right] d\xi,$$

$$\hat{T}^{x_j}\left(\frac{d}{d\xi}\left(k(\xi)\frac{du(\xi)}{d\xi}\right)\right) = \left[\hbar_j V_1^j(x_j)\right]^{-1} \int\limits_{x_{j-1}}^{x_j} V_1^j(\xi) \frac{d}{d\xi}\left[k(\xi)\frac{du(\xi)}{d\xi}\right] d\xi +$$

$$+ \left[\hbar_j V_2^j(x_j)\right]^{-1} \int\limits_{x_j}^{x_{j+1}} V_2^j(\xi) \frac{d}{d\xi}\left[k(\xi)\frac{du(\xi)}{d\xi}\right] d\xi, \quad j = 1, 2, ..., N-1,$$

$$\hat{T}^{x_N}\left(\frac{d}{d\xi}\left(k(\xi)\frac{du(\xi)}{d\xi}\right)\right) = \left[h_N V_1^N(x_N)\right]^{-1} \int\limits_{x_{N-1}}^{x_N} V_1^N(\xi) \frac{d}{d\xi}\left[k(\xi)\frac{du(\xi)}{d\xi}\right] d\xi.$$

The integration by parts implies

$$\hat{T}^{x_0}\left(\frac{d}{d\xi}\left(k(\xi)\frac{du(\xi)}{d\xi}\right)\right)=\frac{1}{h_1}\left(a_1 u_{x,0}-\beta_1 u_0+\mu_1\right),$$
$$\hat{T}^{x_j}\left(\frac{d}{d\xi}\left(k(\xi)\frac{du(\xi)}{d\xi}\right)\right)=(au_{\bar{x}})_{\hat{x},j},\quad j=1,2,...,N-1,$$
$$\hat{T}^{x_N}\left(\frac{d}{d\xi}\left(k(\xi)\frac{du(\xi)}{d\xi}\right)\right)=\frac{1}{h_N}\left(-a_N u_{\bar{x},N}-\beta_2 u_N+\mu_2\right),$$

which together with (8)-(10) proves the existence of the ETDS (12), (13).

To prove the uniqueness of the ETDS (12), (13) we consider the operator

$$A_h\left(x_j,u\right)=$$
$$=\begin{cases}-\dfrac{2}{h_1}\left(a_1 u_{x,0}-\beta_1 u_0+\mu_1-h_1\hat{T}^{x_0}\left(f\left(\xi,u,\dfrac{du}{d\xi}\right)\right)\right), & j=0,\\ -(au_{\bar{x}})_{\hat{x},j}-\hat{T}^{x_j}\left(f\left(\xi,u,\dfrac{du}{d\xi}\right)\right), & j=1,2,...,N-1,\\ \dfrac{2}{h_N}\left(a_N u_{\bar{x},N}+\beta_2 u_N-\mu_2-h_N\hat{T}^{x_N}\left(f\left(\xi,u,\dfrac{du}{d\xi}\right)\right)\right), & j=N\end{cases}$$

which is defined in the finite-dimensional Hilbert space of grid functions $H(\hat{\bar{\omega}}_h)$, with the scalar products

$$(u,v)_{\hat{\bar{\omega}}_h}=\sum_{\xi\in\hat{\omega}_h}\hbar(\xi)u(\xi)v(\xi)+0,5h_1u_0v_0+0,5h_Nu_Nv_N$$
$$(u,v)_{\hat{\omega}_h^+}=\sum_{\xi\in\hat{\omega}_h^+}h(\xi)u(\xi)v(\xi),\quad \hat{\omega}_h^+=\hat{\omega}_h\cup x_N,$$

and the norms

$$\|u\|_{0,2,\hat{\bar{\omega}}_h}=(u,u)^{1/2}_{\hat{\bar{\omega}}_h},\quad \|u\|_{0,2,\hat{\omega}_h^+}=(u,u)^{1/2}_{\hat{\omega}_h^+},$$
$$\|u\|_{1,2,\hat{\bar{\omega}}_h}=\left(\|u\|^2_{0,2,\hat{\bar{\omega}}_h}+\|u_{\bar{x}}\|^2_{0,2,\hat{\omega}_h^+}\right)^{1/2}.$$

Due to condition (5) the operator $A_h\left(x,u\right)$ is continuous. Let as show that the operator $A_h\left(x,u\right)$ is strongly monotone. Actually, taking into account the equality

$$\sum_{\xi\in\hat{\bar{\omega}}_h}\hbar(\xi)\hat{T}^{\xi}(w(\eta))g(\xi)=\sum_{j=1}^{N}\int\limits_{x_{j-1}}^{x_j}\hat{g}(\eta)w(\eta)d\eta=\int\limits_0^1\hat{g}(\eta)w(\eta)d\eta,$$

$$\hat{g}(\eta)=g(x_j)\frac{V_1^j(\eta)}{V_1^j(x_j)}+g(x_{j-1})\frac{V_2^{j-1}(\eta)}{V_1^j(x_j)},\quad x_{j-1}\le\eta\le x_j,$$

and the first difference Green's formula (see. [5, p.234]), we have

$$
\begin{aligned}
&(A_h(x,u) - A_h(x,v), u-v)_{\hat{\omega}_h} = \left(a(u_{\bar{x}} - v_{\bar{x}})^2, 1\right)_{\hat{\omega}_h^+} + \\
&+ \beta_1(u_0 - v_0)^2 + \beta_2(u_N - v_N)^2 - \\
&- \sum_{\xi\in\hat{\omega}_h} \hbar(\xi)\hat{T}^{\xi}\left(f\left(\eta, u(\eta), \frac{du(\eta)}{d\eta}\right) - f\left(\eta, v(\eta), \frac{dv(\eta)}{d\eta}\right)\right)[u(\xi) - v(\xi)] = \\
&= \left(a(u_{\bar{x}} - v_{\bar{x}})^2, 1\right)_{\hat{\omega}_h^+} + \beta_1(u_0 - v_0)^2 + \beta_2(u_N - v_N)^2 - \\
&- \int_0^1 [\hat{u}(\eta) - \hat{v}(\eta)]\left[f\left(\eta, u(\eta), \frac{du(\eta)}{d\eta}\right) - f\left(\eta, v(\eta), \frac{dv(\eta)}{d\eta}\right)\right] d\eta
\end{aligned}
$$

where the functions $u(x)$and $v(x)$ are defined by (11). Then using (6), we have

$$
\begin{aligned}
&(A_h(x,u) - A_h(x,v), u-v)_{\hat{\omega}_h} = \left(a(u_{\bar{x}} - v_{\bar{x}})^2, 1\right)_{\hat{\omega}_h^+} + \\
&+ \beta_1(u_0 - v_0)^2 + \beta_2(u_N - v_N)^2 - \\
&- \int_0^1 [u(\eta) - v(\eta)]\left[f\left(\eta, u(\eta), \frac{du(\eta)}{d\eta}\right) - f\left(\eta, v(\eta), \frac{dv(\eta)}{d\eta}\right)\right] d\eta - \\
&- \int_0^1 [\hat{u}(\eta) - \hat{v}(\eta) - u(\eta) + v(\eta)] \times \\
&\qquad \times \frac{d}{d\eta}\left\{k(\eta)\frac{d}{d\eta}[\hat{u}(\eta) - \hat{v}(\eta) - u(\eta) + v(\eta)]\right\} d\eta \geq \\
&\geq \left(a(u_{\bar{x}} - v_{\bar{x}})^2, 1\right)_{\hat{\omega}_h^+} + \beta_1(u_0 - v_0)^2 + \beta_2(u_N - v_N)^2 + \\
&+ \int_0^1 k(\eta)\left\{\frac{d}{d\eta}[\hat{u}(\eta) - \hat{v}(\eta) - u(\eta) + v(\eta)]\right\}^2 d\eta.
\end{aligned}
$$

Since (see [5, p.244]) $\gamma_1 \|u\|^2_{0,2,\hat{\omega}_h} \leq \left(u^2_{\bar{x}}, 1\right)_{\hat{\omega}_h^+} + \beta_1 u_0^2 + \beta_2 u_N^2,\ \gamma_1 > 0$, then have

$$
\begin{aligned}
&(A_h(x,u) - A_h(x,v), u-v)_{\hat{\omega}_h} \geq \\
&\geq \left(a(u_{\bar{x}} - v_{\bar{x}})^2, 1\right)_{\hat{\omega}_h^+} + \beta_1(u_0 - v_0)^2 + \beta_2(u_N - v_N)^2 \geq \\
&\geq \max\{c_1, 1\}\left[\left((u_{\bar{x}} - v_{\bar{x}})^2, 1\right)_{\hat{\omega}_h^+} + \beta_1(u_0 - v_0)^2 + \beta_2(u_N - v_N)^2\right] \geq \\
&\geq \max\{c_1, 1\}\gamma_1 \|u - v\|^2_{0,2,\hat{\omega}_h},
\end{aligned} \tag{14}
$$

i. e. the operator $A_h(x,u)$ is strongly monotone. This yields (see e.g.[4, p.461]) the uniqueness of the solution of the equation $A_h(x,u) = 0$. □

Lemma 2. *Let the assumptions of Theorem* 1 *be satisfied and*

$$
|f(x,u,\xi) - f(x,v,\eta)| \leq L\{|u-v| + |\xi - \eta|\} \quad \forall x \in (0,1),\ u,v,\xi,\eta \in \mathbb{R}^1.
$$

Then the iteration method

$$B_h \frac{u^{(n)} - u^{(n-1)}}{\tau} + A_h\left(x, u^{(n-1)}\right) = 0, \quad x \in \hat{\bar{\omega}}_h,$$
$$u^{(0)}(x) = \frac{\mu_1 + \mu_2 + \mu_1\beta_2 V_1(1)}{\beta_1 + \beta_2 + \beta_1\beta_2 V_1(1)} \frac{V_2(x)}{V_1(1)} + \frac{\mu_1 + \mu_2 + \mu_2\beta_1 V_1(1)}{\beta_1 + \beta_2 + \beta_1\beta_2 V_1(1)} \frac{V_1(x)}{V_1(1)}, \tag{15}$$

$$B_h u = \begin{cases} -\dfrac{2}{h_1}\left(a_1 u_{x,0} - \beta_1 u_0\right), & j = 0, \\ -(a u_{\bar{x}})_{\hat{x},j}, & j = 1, 2, ..., N-1, \\ \dfrac{2}{h_N}\left(a_N u_{\bar{x},N} + \beta_2 u_N\right), & j = N, \end{cases}$$

$$A_h(x_j, u) =$$
$$= \begin{cases} -\dfrac{2}{h_1}\left(a_1 u_{x,0} - \beta_1 u_0 + \mu_1 - h_1 \hat{T}^{x_0}\left(f\left(\xi, u, \dfrac{du}{d\xi}\right)\right)\right), & j = 0, \\ -(a u_{\bar{x}})_{\hat{x},j} - \hat{T}^{x_j}\left(f\left(\xi, u, \dfrac{du}{d\xi}\right)\right), & j = 1, 2, ..., N-1, \\ \dfrac{2}{h_N}\left(a_N u_{\bar{x},N} + \beta_2 u_N - \mu_2 - h_N \hat{T}^{x_N}\left(f\left(\xi, u, \dfrac{du}{d\xi}\right)\right)\right), & j = N \end{cases}$$

with

$$\tau = \tau_0 = \left[\left(1 + 2L\left(\frac{2K_1K_2}{\gamma_1}\right)^{1/2}\right)\left(1 + \frac{\sqrt{2}(1+\pi^2)L}{c_1\pi^2}\right)\right]^{-2},$$

$$K_1 = \max\left\{\frac{1}{c_1}\left(\frac{4}{\gamma_1} + \frac{c_2}{c_1}\right), \frac{4}{\gamma_1}\right\}, \quad K_2 = \max\left\{\frac{1}{c_1}, 1\right\},$$

$$\gamma_1 = \frac{8\left(\beta_1 + \beta_2 + \beta_1\beta_2\right)^2}{\left(2 + \beta_1\right)\left(2 + \beta_2\right)\left(2\beta_1 + 2\beta_2 + \beta_1\beta_2\right)}$$

converges in the space H_{B_h} *and the error estimate*

$$\left\|u^{(n)} - u\right\|_{B_h} \le q^n \left\|u^{(0)} - u\right\|_{B_h}, \tag{16}$$

where

$$q = \sqrt{1 - \tau_0}, \quad \|u\|_{B_h} = (B_h u, u)^{1/2}_{\hat{\bar{\omega}}_h}.$$

Proof. The operator B_h is selfadjoint and of positive definite $B_h = B_h^* > 0$. From the first difference Green's formula implies that

$$(B_h u, u)_{\hat{\bar{\omega}}_h} = \left(a u_{\bar{x}}^2, 1\right)_{\hat{\omega}_h^+} + \beta_1 u_0^2 + \beta_2 u_N^2,$$

and from (14) we obtain

$$\left(A_h(x, u) - A_h(x, v), u - v\right)_{\hat{\bar{\omega}}_h} \ge \|u - v\|^2_{B_h}. \tag{17}$$

Using the Cauchy-Bunyakovski-Schwarz inequality we can now deduce

$$\left(A_h\left(x,u\right)-A_h\left(x,v\right),z\right)_{\hat{\hat{\omega}}_h}=(B_h(u-v),z)_{\hat{\hat{\omega}}_h}-$$
$$-\sum_{\xi\in\hat{\hat{\omega}}_h}\hbar(\xi)T^{\xi}\left(f\left(\eta,u(\eta),\frac{du}{d\eta}\right)-f\left(\eta,v(\eta),\frac{dv}{d\eta}\right)\right)z(\xi)=$$
$$=(B_h(u-v),z)_{\hat{\hat{\omega}}_h}-\int\limits_0^1\left[f\left(\eta,u(\eta),\frac{du}{d\eta}\right)-f\left(\eta,v(\eta),\frac{dv}{d\eta}\right)\right]\hat{z}(\eta)d\eta\le$$
$$\le\|u-v\|_{B_h}\|z\|_{B_h}+$$
$$+\left\{\int\limits_0^1\left[f\left(\eta,u(\eta),\frac{du}{d\eta}\right)-f\left(\eta,v(\eta),\frac{dv}{d\eta}\right)\right]^2d\eta\right\}^{1/2}\left\{\int\limits_0^1[\hat{z}(\eta)]^2d\eta\right\}^{1/2}\le$$
$$\le\|u-v\|_{B_h}\|z\|_{B_h}+\sqrt{2}L\|u-v\|_{1,2,(0,1)}\|\hat{z}\|_{0,2,(0,1)}.$$

Since $V_1^j(x)\le V_1^j(x_j),\ V_2^{j-1}(x)\le V_1^j(x_j)\ \forall x\in[x_{j-1},x_j]$, we have

$$\|\hat{z}\|^2_{0,2,(0,1)}=\sum_{j=1}^N\int\limits_{x_{j-1}}^{x_j}\left[z_j\frac{V_1^j(x)}{V_1^j(x_j)}+z_{j-1}\frac{V_2^{j-1}(x)}{V_1^j(x_j)}\right]^2dx\le$$
$$\le 2\sum_{j=1}^N\int\limits_{x_{j-1}}^{x_j}\left\{z_j^2\left[\frac{V_1^j(x)}{V_1^j(x_j)}\right]^2+z_{j-1}^2\left[\frac{V_2^{j-1}(x)}{V_1^j(x_j)}\right]^2\right\}dx\le 4\|z\|^2_{0,2,\hat{\hat{\omega}}_h},\tag{18}$$

$$\left\|\frac{d\hat{z}}{dx}\right\|^2_{0,2,(0,1)}\le\sum_{j=1}^N\int\limits_{x_{j-1}}^{x_j}\left[\frac{1}{k(x)}\frac{1}{V_1^j(x_j)}\left(z_j-z_{j-1}\right)\right]^2dx\le$$
$$\le\frac{c_2}{c_1^2}\sum_{j=1}^N h_ja_jz^2_{\bar{x},j}=\frac{c_2}{c_1^2}\left(az^2_{\bar{x}},1\right)_{\hat{\omega}_h^+}.\tag{19}$$

So,

$$\left(A_h\left(x,u\right)-A_h\left(x,v\right),z\right)_{\hat{\hat{\omega}}_h}\le\|u-v\|_{B_h}\|z\|_{B_h}+$$
$$+2\sqrt{2}L\|u-v\|_{1,2,(0,1)}\|z\|_{0,2,\hat{\hat{\omega}}_h}.\tag{20}$$

Let us now show that

$$\|u-v\|_{1,2,(0,1)}\le\sqrt{K_1}\left(1+\frac{\sqrt{2}(1+\pi^2)L}{c_1\pi^2}\right)\|u-v\|_{B_h}.\tag{21}$$

We write $u(x)=\tilde{u}(x)+\hat{u}(x)$, and reduce the problem

$$\frac{d}{dx}\left[k(x)\frac{du}{dx}\right]=-f\left(x,u,\frac{du}{dx}\right),\quad x\in e_1^j,$$
$$u(x_{j-1})=u_{j-1},\quad u(x_j)=u_j,\quad j=1,2,...,N$$

to

$$\frac{d}{dx}\left[k(x)\frac{d\tilde{u}}{dx}\right] = -f\left(x, \tilde{u}+\hat{u}, \frac{d\tilde{u}}{dx}+\frac{d\hat{u}}{dx}\right), \quad x \in e_1^j,$$
$$\tilde{u}(x_{j-1}) = 0, \quad \tilde{u}(x_j) = 0, \quad j = 1, 2, ..., N,$$

Considering (3), (6) and using a Lipschitz condition we get

$$\frac{\pi^2 c_1}{\pi^2+1}\left\|\tilde{u}-\tilde{v}\right\|^2_{0,2,e_1^j} \le c_1\left\|\frac{d\tilde{u}}{dx}-\frac{d\tilde{v}}{dx}\right\|^2_{1,2,e_1^j} \le \int\limits_{x_{j-1}}^{x_j} k(x)\left[\frac{d\tilde{u}}{dx}-\frac{d\tilde{v}}{dx}\right]^2 dx =$$

$$= \int\limits_{x_{j-1}}^{x_j}\left[f\left(x, \tilde{u}(x)+\hat{u}(x), \frac{d\tilde{u}}{dx}+\frac{d\hat{u}}{dx}\right) - \right.$$
$$\left. -f\left(x, \tilde{v}(x)+\hat{v}(x), \frac{d\tilde{v}}{dx}+\frac{d\hat{v}}{dx}\right)\right]\left[\tilde{u}(x)-\tilde{v}(x)\right]dx =$$

$$= \int\limits_{x_{j-1}}^{x_j}\left[f\left(x, \tilde{u}(x)+\hat{u}(x), \frac{d\tilde{u}}{dx}+\frac{d\hat{u}}{dx}\right) - \right.$$
$$\left. -f\left(x, \tilde{v}(x)+\hat{u}(x), \frac{d\tilde{v}}{dx}+\frac{d\hat{u}}{dx}\right)\right]\left[\tilde{u}(x)-\tilde{v}(x)\right]dx +$$
$$+ \int\limits_{x_{j-1}}^{x_j}\left[f\left(x, \tilde{v}(x)+\hat{u}(x), \frac{d\tilde{v}}{dx}+\frac{d\hat{u}}{dx}\right) - \right.$$
$$\left. -f\left(x, \tilde{v}(x)+\hat{v}(x), \frac{d\tilde{v}}{dx}+\frac{d\hat{v}}{dx}\right)\right]\left[\tilde{u}(x)-\tilde{v}(x)\right]dx$$

$$\le \left\{\int\limits_{x_{j-1}}^{x_j}\left[f\left(x, \tilde{v}(x)+\hat{u}(x), \frac{d\tilde{v}}{dx}+\frac{d\hat{u}}{dx}\right) - \right.\right.$$
$$\left.\left. -f\left(x, \tilde{v}(x)+\hat{v}(x), \frac{d\tilde{v}}{dx}+\frac{d\hat{v}}{dx}\right)\right]^2 dx\right\}^{1/2} \times$$
$$\times \left\{\int\limits_{x_{j-1}}^{x_j}\left[\tilde{u}(x)-\tilde{v}(x)\right]^2 dx\right\}^{1/2} \le$$
$$\le \sqrt{2}L\left\|\hat{u}-\hat{v}\right\|_{1,2,e_1^j}\left\|\tilde{u}-\tilde{v}\right\|_{1,2,e_1^j},$$

Hence we get

$$\left\|\tilde{u}-\tilde{v}\right\|_{1,2,e_1^j} \le \frac{\sqrt{2}(1+\pi^2)L}{\pi^2 c_1}\left\|\hat{u}-\hat{v}\right\|_{1,2,e_1^j}.$$

So taking into account inequalities (18), (19) and inequality (see [5, p.244])

$$\gamma_1\left\|u\right\|^2_{0,2,\hat{\bar{\omega}}_h} \le \left(u^2_{\bar{x}}, 1\right)_{\hat{\omega}_h^+} + \beta_1 u_0^2 + \beta_2 u_N^2,$$

we have

$$\|u-v\|_{1,2,(0,1)} \le \|\tilde{u}-\tilde{v}\|_{1,2,(0,1)} + \|\hat{u}-\hat{v}\|_{1,2,(0,1)} \le$$

$$\le \left(1+\frac{\sqrt{2}(1+\pi^2)L}{c_1\pi^2}\right)\|\hat{u}-\hat{v}\|_{1,2,(0,1)} \le$$

$$\le \left(1+\frac{\sqrt{2}(1+\pi^2)L}{c_1\pi^2}\right)\left(4\|u-v\|^2_{0,2,\hat{\omega}_h} + \frac{c_2}{c_1^2}\left(a\left(u_{\bar{x}}-v_{\bar{x}}\right)^2, 1\right)_{\hat{\omega}_h^+}\right)^{1/2} \le$$

$$\le \left(1+\frac{\sqrt{2}(1+\pi^2)L}{c_1\pi^2}\right)\left\{\frac{1}{c_1}\left(\frac{4}{\gamma_1}+\frac{c_2}{c_1}\right)\left(a\left(u_{\bar{x}}-v_{\bar{x}}\right)^2, 1\right)_{\hat{\omega}_h^+} +\right.$$

$$\left. + \frac{4}{\gamma_1}\left[\beta_1\left(u_0-v_0\right)^2+\beta_2\left(u_N-v_N\right)^2\right]\right\}^{1/2} \le$$

$$\le \sqrt{K_1}\left(1+\frac{\sqrt{2}(1+\pi^2)L}{c_1\pi^2}\right)\|u-v\|_{B_h}.$$

Based on the (18), (21) from inequality (20) we obtain

$$\left(A_h(x,u)-A_h(x,v), z\right)_{\hat{\hat{\omega}}_h} \le \|u-v\|_{B_h}\|z\|_{B_h} +$$

$$+ 2L\sqrt{2K_1}\left(1+\frac{\sqrt{2}(1+\pi^2)L}{c_1\pi^2}\right)\|u-v\|_{B_h}\|z\|_{0,2,\hat{\hat{\omega}}_h} \le$$

$$\le \|u-v\|_{B_h}\|z\|_{B_h} + 2L\left(\frac{2K_1}{\gamma_1}\right)^{1/2}\left(1+\frac{\sqrt{2}(1+\pi^2)L}{c_1\pi^2}\right)\|u-v\|_{B_h}\times$$

$$\times\left(\frac{1}{c_1}\left(az_{\bar{x}}^2, 1\right)_{\hat{\omega}_h^+}+\beta_1 z_0^2+\beta_2 z_N^2\right)^{1/2} \le$$

$$\le \left(1+2L\left(\frac{2K_1K_2}{\gamma_1}\right)^{1/2}\right)\left(1+\frac{\sqrt{2}(1+\pi^2)L}{c_1\pi^2}\right)\|u-v\|_{B_h}\|z\|_{B_h}.$$

Setting $z = B_h^{-1}\left(A_h(x,u)-A_h(x,v)\right)$, we obtain

$$\left\|B_h^{-1}\left(A_h(x,u)-A_h(x,v)\right)\right\|_{B_h} \le$$

$$\le \left(1+2L\left(\frac{2K_1K_2}{\gamma_1}\right)^{1/2}\right)\left(1+\frac{\sqrt{2}(1+\pi^2)L}{c_1\pi^2}\right)\|u-v\|_{B_h}. \qquad (22)$$

Implies from (22), (17)

$$\left(A_h(x,u)-A_h(x,v), B_h^{-1}\left(A_h(x,u)-A_h(x,v)\right)\right)_{\hat{\hat{\omega}}_h} \le$$

$$\le \left[\left(1+2L\left(\frac{2K_1K_2}{\gamma_1}\right)^{1/2}\right)\left(1+\frac{\sqrt{2}(1+\pi^2)L}{c_1\pi^2}\right)\right]^2\|u-v\|^2_{B_h} \le$$

$$\le \left[\left(1+2L\left(\frac{2K_1K_2}{\gamma_1}\right)^{1/2}\right)\left(1+\frac{\sqrt{2}(1+\pi^2)L}{c_1\pi^2}\right)\right]^2\times$$

$$\left(A_h(x,u)-A_h(x,v), u-v\right)_{\hat{\hat{\omega}}_h}.$$

Then (see [5, p.502]) the iteration method (15) converges in the space H_{B_h} as well as the estimate (16). □

Note that the space H_{B_h} coincides with the space $L_2(\hat{\bar{\omega}}_h)$ and the conditions of equivalence of norms are executed.

$$\gamma_1 \|u\|_{0,2,\hat{\bar{\omega}}_h} \leq \|u\|_{B_h} \leq \gamma_2 \|u\|_{0,2,\hat{\bar{\omega}}_h}.$$

Lemma 3. *Let the assumptions of Lemma 2 be satisfied. Then the method of simple iteration (15) and in addition to (16) the following estimate holds*

$$\left\| k\frac{du^{(n)}}{dx} - k\frac{du}{dx} \right\|_{0,2,\hat{\bar{\omega}}_h} \leq M \left\| u^{(n)} - u \right\|_{B_h} \leq Mq^n,$$

where

$$\|u\|_{0,2,\hat{\bar{\omega}}_h} = \left\{ \sum_{j=1}^{N-1} \hbar_j u_j^2 + \frac{1}{2} h_1 u_0^2 + \frac{1}{2} h_N u_N^2 \right\}^{1/2} = \left\{ \frac{1}{2} \sum_{j=1}^{N} h_j (u_j^2 + u_{j-1}^2) \right\}^{1/2}.$$

Proof. Taking into account equality

$$k\frac{du}{dx}\bigg|_{x=x_j} = a_j u_{\bar{x},j} + \frac{1}{V_1^j(x_j)} \int_{x_{j-1}}^{x_j} V_1^j(\xi) \frac{d}{d\xi}\left[k(\xi)\frac{du}{d\xi}\right] d\xi =$$

$$= a_j u_{\bar{x},j} - \frac{1}{V_1^j(x_j)} \int_{x_{j-1}}^{x_j} V_1^j(\xi) f\left(\xi, u(\xi), \frac{du}{d\xi}\right) d\xi,$$

$$k\frac{du}{dx}\bigg|_{x=x_{j-1}} = a_j u_{\bar{x},j} - \frac{1}{V_2^{j-1}(x_{j-1})} \int_{x_{j-1}}^{x_j} V_2^{j-1}(\xi) \frac{d}{d\xi}\left[k(\xi)\frac{du}{d\xi}\right] d\xi =$$

$$= a_j u_{\bar{x},j} + \frac{1}{V_1^j(x_j)} \int_{x_{j-1}}^{x_j} V_2^{j-1}(\xi) f\left(\xi, u(\xi), \frac{du}{d\xi}\right) d\xi,$$

the inequality $(a+b)^2 \leq 2(a^2+b^2)$ as well as the Cauchy-Bunyakovsky-Schwarz inequality and a Lipschitz condition we obtain

$$\left\| k\frac{du^{(n)}}{dx} - k\frac{du}{dx} \right\|_{0,2,\hat{\bar{\omega}}_h} = \left\{ \frac{1}{2} \sum_{j=1}^{N} h_j \left[a_j u_{\bar{x},j}^{(n)} - a_j u_{\bar{x},j} - \right.\right.$$

$$\left. - \frac{1}{V_1^j(x_j)} \int_{x_{j-1}}^{x_j} V_1^j(\xi) \left(f\left(\xi, u^{(n)}(\xi), \frac{du^{(n)}}{d\xi}\right) - f\left(\xi, u(\xi), \frac{du}{d\xi}\right)\right) d\xi \right]^2 +$$

$$+\frac{1}{2}\sum_{j=1}^{N} h_j \left[a_j u_{\bar{x},j}^{(n)} - a_j u_{\bar{x},j} + \frac{1}{V_1^j(x_j)}\times\right.$$

$$\left.\left.\times \int\limits_{x_{j-1}}^{x_j} V_2^{j-1}(\xi)\left(f\left(\xi, u^{(n)}(\xi), \frac{du^{(n)}}{d\xi}\right) - f\left(\xi, u(\xi), \frac{du}{d\xi}\right)\right) d\xi\right]^2\right\}^{1/2} \leq$$

$$\leq \left\{2\sum_{j=1}^{N} h_j \left[a_j u_{\bar{x},j}^{(n)} - a_j u_{\bar{x},j}\right]^2 + \sum_{j=1}^{N} h_j [V_1^j(x_j)]^{-2}\times\right.$$

$$\times \left[\int\limits_{x_{j-1}}^{x_j} V_1^j(\xi)\left(f\left(\xi, u^{(n)}(\xi), \frac{du^{(n)}}{d\xi}\right) - f\left(\xi, u(\xi), \frac{du}{d\xi}\right)\right) d\xi\right]^2 +$$

$$+\sum_{j=1}^{N} h_j [V_1^j(x_j)]^{-2}\times$$

$$\left.\times \left[\int\limits_{x_{j-1}}^{x_j} V_2^{j-1}(\xi)\left(f\left(\xi, u^{(n)}(\xi), \frac{du^{(n)}}{d\xi}\right) - f\left(\xi, u(\xi), \frac{du}{d\xi}\right)\right) d\xi\right]^2\right\}^{1/2} \leq$$

$$\leq \left\{2\sum_{j=1}^{N} h_j \left[a_j u_{\bar{x},j}^{(n)} - a_j u_{\bar{x},j}\right]^2 + \sum_{j=1}^{N} h_j \int\limits_{x_{j-1}}^{x_j} \frac{\left[V_1^j(\xi)\right]^2 + \left[V_2^{j-1}(\xi)\right]^2}{\left[V_1^j(\xi) + V_2^{j-1}(\xi)\right]^2} d\xi \times\right.$$

$$\left.\times \int\limits_{x_{j-1}}^{x_j} \left[f\left(\xi, u^{(n)}(\xi), \frac{du^{(n)}}{d\xi}\right) - f\left(\xi, u(\xi), \frac{du}{d\xi}\right)\right]^2 d\xi\right\}^{1/2} \leq$$

$$\leq \left\{2\sum_{j=1}^{N} h_j a_j^2 \left[u_{\bar{x},j}^{(n)} - u_{\bar{x},j}\right]^2 +\right.$$

$$\left.+L^2\sum_{j=1}^{N} h_j \int\limits_{x_{j-1}}^{x_j} \left[\left|u^{(n)}(\xi) - u(\xi)\right| + \left|\frac{du^{(n)}}{d\xi} - \frac{du}{d\xi}\right|\right]^2 d\xi\right\}^{1/2} \leq$$

$$\leq \sqrt{2c_2}\left(a\left(u_{\bar{x}}^{(n)} - u_{\bar{x}}\right)^2, 1\right)_{\hat{\omega}_h^+} + \sqrt{2}L\left\|u^{(n)} - u\right\|_{1,2,(0,1)}.$$

Then based on the inequality (21) and Lemma 2 we have

$$\left\|k\frac{du^{(n)}}{dx} - k\frac{du}{dx}\right\|_{0,2,\hat{\bar{\omega}}_h} \leq \sqrt{2}\left[\sqrt{c_2} + \sqrt{K}_1\left(1 + \frac{\sqrt{2}(1+\pi^2)L}{c_1\pi^2}\right)\right]\times$$

$$\times \left\| u^{(n)} - u \right\|_{B_h} = M_1 \left\| u^{(n)} - u \right\|_{B_h} \leq M q^n.$$

So, in this page ETDS is constructed and justified, which you can develop (see [6]) a three-point difference schemes of high order accuracy for the numerical solution of the BVP (1), (2). □

Bibliography

1. Gnativ L. B. Generalized three-point difference schemes of high order of accuracy for nonlinear ordinary differential equations of second order / L. B. Gnativ, M. V. Kutniv, A. I. Chukhrai // Journal of Mathematical Sciences.– 2010.– Vol. 167, № 1.– P. 62-75.
2. Rektorys K. Variational Methods in Mathematical Physics and Engineering / K. Rektorys.– Moscow: Mir, 1985 (in Russian).
3. Kufner A. Nonlinear Differential Equations / A. Kufner, S. Fucik.– Moscow: Nauka, 1988 (in Russian).
4. Trenogin V. A. Functional Analysis / V. A. Trenogin.– Moscow: Nauka, 1980 (in Russian).
5. Samarskii A. A. Numerical Methods for Grid Equations / A. A. Samarskii, E. S. Nikolaev.– Moscow: Nauka, 1978 (in Russian).
6. Gavrilyuk I. P. Exact and Truncated Difference Schemes for Boundary Value ODEs. (International Seris of Numerical Mathematics Vol. 159) / I. P. Gavrilyuk, M. Hermann, V. L. Makarov., M. V. Kutniv.– Basel: Springer AG, 2011.
7. Kutniv M. V. Accurate three-point difference schemes for second-order monotone ordinary differential equations and their implementation / M. V. Kutniv // Computational Mathematics and Mathematical Physics.– 2000.– Vol. 40, № 3.– P. 368-382.
8. Kutniv M. V. Tree-point difference schemes of high accuracy order for second order nonlinear ordinary differential equations with the boundary conditions of the third kind / M. V. Kutniv // Visnyk of Lviv University. Series Applied Mathematics and Computer Science.– 2002.– № 4.– P. 61-66 (in Ukrainian).
9. Kutniv M. V. Modified three-point difference schemes of high-accuracy order for second order nonlinear ordinary differential equations / M. V. Kutniv // Computational Methods in Applied Mathematics.– 2003.– Vol. 3, № 2.– P. 287-312.
10. Kutniv M. V. Accurate three-point difference schemes for second-order nonlinear ordinary differential equations and their implementation / M. V. Kutniv, V. L. Makarov, A. A. Samarskii // Computational Mathematics and Mathematical Physics.– 1999.– Vol. 39, № 1.– P. 45-60.
11. Makarov V. L. Exact three-point difference schemes for second-order nonlinear ordinary differential equations and their implementation / V. L. Makarov, A. A. Samarskii // Soviet Math. Dokl.– 1991.– Vol. 41, № 4.– P. 495-500.

Lyubomyr Gnativ, Myroslav Kutniv,
Lviv Polytechnic National University, 12 S. Bandery str., Lviv, 79013, Ukraine;
Marta Król,
Rzeszow University of Technology, 8 Powstancow Warszawy str., 35-959, Rzeszow, Polska

Received 04.05.2012

Журнал обчислювальної та прикладної математики | 2012, № 3 (109) с. 53–71 | Journal of Numerical & Applied Mathematics

UDC 517.958:534+519.6

FORMULATION AND WELL-POSEDNESS OF THE VARIATIONAL PROBLEM OF VISCOUS HEAT-CONDUCTING FLUID ACOUSTICS

Vitaliy Horlatch, Ira Klymenko, Georgiy Shynkarenko

Резюме. На підставі законів збереження сформульовано лінійну початково-крайову та відповідну їй варіаційну задачу у термінах невідомих вектора зміщень та температури, яка описує процес поширення акустичних хвиль у в'язкій теплопровідній рідині з урахуванням зв'язаності механічного та температурного полів. Окреслено клас регулярності вхідних даних варіаційної задачі, який гарантує єдність та неперервну залежність шуканого розв'язку в енергетичній нормі задачі. На додаток доведено існування розв'язку розглядуваної задачі як границі послідовності напівдискретних (за просторовими змінними) апроксимацій Гальоркіна.

Abstract. On the basis of conservation laws, we formulate linear initial-boundary value problem and corresponding variational problem in terms of displacement vector and temperature, which describes the process of spreading of acoustic waves in viscous heat-conducting fluid taking into account connectivity of mechanical and thermal fields. We determined input data regularity for the variational problem, which guarantee uniqueness and continuous dependence of the solution in the energy norm of the problem. In addition we prove the existence of the solution of the problem as a limit of a sequence of the semi-discrete spatial Galerkin approximations.

1. Introduction

In most applications, when considering acoustic vibrations, the viscosity of fluid is neglected, hence considering it to be 'ideal'[5, 3]. However, there is a considerable number of problems, which are first of all connected to spreading of the high-frequency vibrations and vibrations at frequencies close to resonance, for which neglecting medium viscosity (even for traditionally "ideal" water or air) leads to considerable inaccuracies in solutions [1, 2, 10]. Furthermore, analysis of dissipative loss of energy in such problems, as well as estimation of reciprocal influence of acoustic and thermal processes are impossible without introducing viscosity of the medium to the model. The general principles of building corresponding models of acoustics of viscous heat-conducting fluid ("dissipative acoustics" is a widely-used term) are studied in papers [11, 6, 7, 9, 10].

† *Key words.* Thermohydrodynamycs, dissipative acoustics, initial-boundary value problem, variation problem, balance equation, the semi-discrete Galerkin method, well-posedness of variation formulation.

In paper [9], for numerical analysis of problems of dissipative acoustics with additional assumption of vortex-free flow in fluid, it is proposed to use Raviart-Thomas finite element approximations, and time integration schemes for semi-discretized problem are built by means of Galerkin method. However, the authors [2] proved earlier the correctness of application of classical approximations of finite element method for solving problems of spreading acoustic waves in viscous fluids and fluid-structure systems in terms of unknown displacements without additional assumptions. It is proposed that a similar approach should be used for problems of thermal and hydro acoustics.

This paper is organized as follows. In section 2, with reference to conservation law, we state a fundamental system of non-linear differential equations and phenomenological relations, which describe the motion of viscous heat-conducting Newtonian fluid, and complement them with possible initial and boundary conditions. Although the obtained system of equations is open in relation to density, mass, velocity, temperature, entropy of the fluid, the hypotheses of acoustics and thermodynamics applied in sections 3 and 4 allowed us to formulate a linear initial-boundary value problem of acoustics with closed system of equations of motion and heat conductivity in terms of acoustic displacement vector and temperature. In section 5 we state variational formulation of this problem as the main object of our study and in section 6 we characterize the components of its equations with regard to continuity and ellipticity. Based on these, in section 7 we describe an important instrument for research of the variational problem – a concretized energy equations of dissipative acoustics. A priori estimates, constructed on this basis in sections 8 and 9, make it possible to determine (quite usable) conditions of regularity of input data of the problem, which guarantee uniqueness and stability of its solution. To prove existence of this solution, in section 10 we recourse to space semi-discretization Galerkin method [4], and in section 11 we show that approximations built in such a way converge to such displacement vector and temperature, which satisfy variation equations of the problem of dissipative acoustics.

2. Fundamental equations of thermohydrodynamics of Newtonian fluid

Below we will consider mathematical models which describe motion of a viscous fluid, which in each moment of time $t \in [0,T], 0 < T < +\infty$, occupies connected bounded domain Ω of points $x = (x_1, \dots, x_d)$ of Euclidian space $\mathbb{R}^d$ (in applications $d = 1,2$ or 3). We denote as Γ the domain boundary Ω, $\Gamma = \partial\Omega$, and assume that it is Lipschitz-continuous. The latter hypothesis guarantees that almost everywhere on Γ we can build a unit vector of outward normal

$$n = (n_1, \dots, n_d), \quad n_i := \cos(n, x_i).$$

It is well known that physical features of fluid are defined by coefficients of bulk viscosity η and shear viscosity $\mu = const > 0$, and its state can be characterized by means of *velocity vector* $v = \{v_i(x,t)\}_{i=1}^d$ of its particles, density of its mass

$$\rho = \rho(x,t) \geq 0$$

and scalar field of *hydrostatic pressure* $p = p(x,t)$. If the above-mentioned characteristics of the fluid are defined, then with the help of Cauchy relations we can find the components of *strain tensor*

$$e_{ij}(\upsilon) := \frac{1}{2}\left(\frac{\partial \upsilon_i}{\partial x_j} + \frac{\partial \upsilon_j}{\partial x_i}\right), \quad i,j = 1, \dots, d, \tag{1}$$

and components of *stress tensor*

$$\sigma_{ij}(\upsilon, p) := -p\delta_{ij} + \tau_{ij}(\upsilon), \quad i,j = 1, \dots, d, \tag{2}$$

where $\tau_{ij}(\upsilon)$ - components *of viscous stress tensor*,

$$\tau_{ij}(\upsilon) := 2\mu\, e_{ij}(\upsilon) + (\eta - \tfrac{2}{3}\mu)\delta_{ij}\nabla.\upsilon, \quad i,j = 1, \dots, d, \tag{3}$$

δ_{ij} -Kronecker's symbol,

$$\delta_{ij} := \begin{cases} 1, & i = j, \\ 0, & i \neq j. \end{cases}$$

Modeling of fluid flows reduces to initial-boundary value problems for the partial differential equation system, which are based on the *laws of mass conservation*, momentum, energy, etc. [10, 11]. So, for example, the *law of mass conservation* of continuous medium states that given the absence of sources for mass increase, the density $\rho = \rho(x,t)$ and the vector of fluid velocity $\upsilon = \{\upsilon_i(x,t)\}_{i=1}^d$ satisfy the so-called *equation of mass continuity*

$$D_t\rho + \rho\nabla.\upsilon = 0 \quad in \quad \Omega \times (0,T]. \tag{4}$$

At the same time, *laws of momentum conservation* can be presented as a system of Navier-Stokes equations

$$\rho D_t\upsilon_i - \frac{\partial}{\partial x_m}\sigma_{im}(\upsilon, p) = \rho f_i, \quad i = 1, \dots, d, \quad in \quad \Omega \times (0,T], \tag{5}$$

where vector $f = \{f_i(x,t)\}_{i=1}^d$ describes volume forces which act on the considered fluid.

Finally, the *law of energy conservation* leads to *equation of continuity of entropy* $s = s(x,t)$ formulated as

$$\rho\theta D_t s + \nabla.q(\theta) - \tau(\upsilon) : e(\upsilon) = \rho g \quad in \quad \Omega \times (0,T], \tag{6}$$

where $g = g(x,t)$ is intensity of distributed in the fluid volume sources of heat, $q = \{q_i(x,t)\}_{i=1}^d$ is *vector of heat flow*, which is connected in most important cases to the temperature $\theta = \theta(x,t)$ and coefficient of heat conductivity $\chi > 0$ of fluid through *phenomenological Fourier law*

$$q(\theta) = -\chi\nabla\theta \quad in \quad \Omega \times (0,T]. \tag{7}$$

Here and further on we shall use the summation convention from 1 to d with repeated indexes, eliminating the sign of summation itself; e.g. scalar product in space $\mathbb{R}^d$ is written as

$$a.b \equiv a_ib_i := \sum_{i=1}^{d} a_ib_i \qquad \forall a = \{a_i\}_{i=1}^d, \ b = \{b_i\}_{i=1}^d \ \in \mathbb{R}^d,$$

and

$$\sigma : e \equiv \sigma_{mi}e_{im} := \sum_{i,m=1}^{d} \sigma_{mi}e_{im} \qquad \forall \sigma = \{\sigma_{ij}\}_{i,j=1}^{d}\,,\ e = \{e_{ij}\}_{i,j=1}^{d} \in \mathbb{R}^{d\times d}.$$

Finally, in the equations (4)-(7) we utilize widely-used symbols for full and partial derivatives of a scalar function by time variable and its gradient by spatial variable.

$$D_t w := w' + v.\nabla w, \quad w' := \frac{\partial}{\partial t} w(x,t), \quad \nabla w := \left\{ \frac{\partial w}{\partial x_m} \right\}_{m=1}^{d}.$$

Let us complement the system (1)-(7) with appropriate initial and boundary conditions. If on the outer surface of fluid $\Gamma_\sigma \subset \Gamma$ is affected by the applied stress vector $\hat{\sigma} = \{\hat{\sigma}_i(x,t)\}_{i=1}^{d}$, then the law of momentum conservation leads to the following boundary condition for stress:

$$\sigma_{ij}(v,p)n_j = \hat{\sigma}_i \quad i = 1, \ldots, d, \qquad on \quad \Gamma_\sigma \times [0,T]. \tag{8}$$

Similarly, if a part of the boundary $\Gamma_q \subset \Gamma$ is affected by heat flow, the normal component of which is determined by the function $\hat{q} = \hat{q}(x,t)$, then according to the law of energy conservation, the boundary condition will be

$$n.q(\theta) = \hat{q} \qquad on \quad \Gamma_q \times [0,T]. \tag{9}$$

Finally, if, for example, particles of the remaining fluid surface $\Gamma_v := \Gamma \backslash \Gamma_\sigma$ move in compliance with the known rule at the speed $\hat{v} = \{\hat{v}_i(x,t)\}$, then the boundary condition on this part of the surface should be

$$v = \hat{v} \qquad on \quad \Gamma_v \times [0,T], \ \ \Gamma_v := \Gamma \backslash \Gamma_\sigma. \tag{10}$$

Similarly, if it is known that the part of the surface $\Gamma_\theta := \Gamma \backslash \Gamma_q$ is maintained at the defined temperature, $\hat{\theta} = \hat{\theta}(x,t)$, then the boundary condition assigned to it is

$$\theta = \hat{\theta} \qquad on \quad \Gamma_\theta \times [0,T], \ \ \Gamma_\theta := \Gamma \backslash \Gamma_q. \tag{11}$$

We have to mention that there might be boundary conditions for different classes of applications, as a rule, formulated as linear combinations of condition components (8), (9) and (10), (11) correspondigly.

Finally, considering the specifics of the structure of system relations and equations (1)-(7), namely, the absence of pressure derivatives by time variable in it, we come to a conclusion that during the study of viscous fluid motion it is sufficient to reduce it to studying the initial conditions and values of mass density, velocity vector and temperature

$$\rho|_{t=0} = \rho_0, \quad v|_{t=0} = v_0, \quad \theta|_{t=0} = \theta_0 \quad in \quad \Omega. \tag{12}$$

The obtained nonlinear problem of thermohydrodynamics (1)-(12) contains less equations $(d+2)$ than, the unknowns $(d+4)$, and must be complemented by additional equations based on phenomenological deductions. For this purpose we shall use the hypotheses of acoustic approximation, which will allow us not only to find a closed equation system, but also to linearize it.

3. Linear equation system of dissipative acoustics in terms of acoustic displacement and temperature

Below we assume that, for one reason or another, there are connections between the unknowns $\{\rho, p, s, \theta\}$, which are expressed as

$$p = p(\rho, \theta), \quad s = s(\rho, \theta).$$

It is known that pressure is related to density and temperature by the following thermodynamic connections [10]:

$$\left.\frac{\partial p}{\partial \rho}\right|_{\theta} = \frac{c^2}{\gamma}, \qquad \left.\frac{\partial p}{\partial \theta}\right|_{\rho} = \frac{c^2 \rho \alpha}{\gamma},$$

where c is velocity of sound, α coefficient of thermal expansion, $\gamma = c_p c_v^{-1}$, c_p and c_v specific heat of fluid at constant pressure and volume respectively. Then to accuracy of an additive constant

$$p = p_0 + c^2 \gamma^{-1} [\rho + \rho \alpha \theta].$$

In addition we can linearize the obtained rule in the following way:

$$p(x,t) \cong p_0 + c^2 \gamma^{-1} [\rho(x,t) + \rho_0 \alpha \theta(x,t)], \tag{13}$$

where ρ_0 is mass density distribution of fluid in the state undisturbed by acoustic factors. Here we implicitly assume that the mass density of fluid admits the following decomposition

$$\begin{cases} \rho(x,t) = \rho_0 + \rho_*(x,t) \qquad \forall x \in \Omega \ \ \forall t \in [0,T], \\ \rho_*|_{t=0} = 0 \qquad in \quad \Omega, \\ ||\rho_*|| \ll ||\rho_0||. \end{cases} \tag{14}$$

Now we shall convey the velocity of fluid motion as a sum formulated as

$$\begin{cases} \upsilon(x,t) = \upsilon_0(x) + \upsilon_*(x,t) \qquad \forall x \in \Omega \ \ \forall t \in [0,T], \\ \upsilon_*|_{t=0} = 0 \qquad in \quad \Omega, \qquad ||\upsilon_*|| \ll ||\upsilon_0|| \end{cases} \tag{15}$$

And turn to the continuity equation from (4). Bearing in mind the hypotheses (14) and (15), we shall linearize it in the following way

$$\begin{aligned} \rho' + \upsilon.\nabla\rho + \rho\nabla.\upsilon \cong \rho'_* + \rho_0 \nabla.\upsilon_* + \upsilon_0.\nabla\rho_* \cong \\ \cong \rho'_* + \rho_0 \nabla.\upsilon_* = 0 \qquad in \ \ \Omega \times \ (0,T]], \end{aligned}$$

And later integrate the obtained approximation into a time interval $(0,t)$, $0 < t \le T$. As a result, we find out that

$$\begin{aligned} \rho_*(x,t) = -\rho_0 \nabla. \textstyle\int_0^t \upsilon_*(x,\tau) d\tau = \\ = -\rho_0 \nabla.u(x,t) \quad \forall x \in \Omega \ \ \forall t \in [0,T], \end{aligned} \tag{16}$$

where $u = u(x,t)$ – vector of acoustic displacement of fluid particles

$$u(x,t) := u_0(x) + \int_0^t \upsilon_*(x,\tau) d\tau \qquad in \ \ \Omega \times (0,T].$$

Taking into account (13) and (16), we come to a final expression for the linear approximation of acoustic pressure in fluid

$$\begin{aligned} p(x,t) &\cong p_0 + c^2\gamma^{-1}[\rho_*(x,t) + \rho_0\alpha\theta(x,t)] \cong \\ &\cong p_0 + c^2\gamma^{-1}\rho_0[-\nabla.u(x,t) + \alpha\theta(x,t)] \equiv \\ &\equiv p_0 + \pi(u,\theta) \qquad\qquad \forall x \in \Omega \ \ \forall t \in [0,T]. \end{aligned} \tag{17}$$

Introducing the vector of acoustic displacements $u = u(x,t)$ also leads to change of notation and structure of stress tensor of fluid, such as

$$\begin{aligned} \sigma_{ij}(\upsilon,p) &= -p\delta_{ij} + \tau_{ij}(\upsilon) \cong \\ &\cong -p_0\delta_{ij} + \pi(u,\theta)\delta_{ij} + \tau_{ij}(u') = \\ &= -p_0\delta_{ij} + \bar{\sigma}_{ij}(u,\theta) \qquad \forall x \in \Omega \ \ \forall t \in [0,T]. \end{aligned}$$

In other words, taking into consideration the relation (17), pressure is excluded when determining the stress tensor, instead we include the dependence of its components from the temperature of fluid. Taking into account the hypotheses of acoustics and linearization of convective constituents, the motion equations (5) undergo some changes, such as

$$\begin{aligned} &\rho[\upsilon_i'(t) + \upsilon_m\frac{\partial \upsilon_i}{\partial x_m}] - \frac{\partial\sigma_{im}(\upsilon,p)}{\partial x_m} - \rho f_i \cong \\ &\cong \rho_0 u_i''(t) + \frac{\partial}{\partial x_i}p_0 - \frac{\partial\bar{\sigma}_{im}(u')}{\partial x_m} - \rho_0 f_i = 0. \end{aligned}$$

It follows that considering the hypotheses of acoustics and the linearization of motion equations, performed above, lead to excluding pressure and density from the unknown, and after this procedure the motion equations acquire the form

$$\begin{aligned} &\rho_0 u_i''(t) - \frac{\partial\bar{\sigma}_{im}(u')}{\partial x_m} = \rho_0 f_i - \frac{\partial}{\partial x_i}p_0, \\ &\bar{\sigma}_{ij}(u) := -\pi(u,\theta)\delta_{ij} + \tau_{im}(u'), \\ &\pi(u,\theta) := c^2\gamma^{-1}\rho_0[-\nabla.u + \alpha\theta], \\ &\tau_{ij}(u') := 2\mu\, e_{ij}(u') + (\eta - \tfrac{2}{3}\mu)\delta_{ij}\nabla\cdot u', \\ &e_{ij}(u) := \frac{1}{2}\left(\frac{\partial u_i}{\partial x_j} + \frac{\partial u_j}{\partial x_i}\right) \qquad in \quad \Omega\times(0,T]. \end{aligned}$$

Since entropy is related to density and temperature through thermodynamic links expressed as [10]

$$(\frac{\partial s}{\partial \rho})_\theta = -\frac{c^2\alpha}{\rho\gamma}, \quad (\frac{\partial s}{\partial \theta})_\rho = \frac{c_V}{\theta},$$

then

$$\begin{aligned} \frac{\partial s}{\partial t} &= -\frac{c^2\alpha}{\rho\gamma}\frac{\partial\rho}{\partial t} + \frac{c_V}{\theta}\frac{\partial\theta}{\partial t} \cong -\frac{c^2\alpha}{\rho_0\gamma}\frac{\partial\rho}{\partial t} + \frac{c_V}{\theta_0}\frac{\partial\theta}{\partial t} \cong \\ &\cong \frac{c^2\alpha}{\rho_0\gamma}\rho_0\nabla.u' + \frac{c_V}{\theta_0}\frac{\partial\theta}{\partial t} = \frac{c^2\alpha}{\gamma}\nabla.u' + \frac{c_V}{\theta_0}\frac{\partial\theta}{\partial t} \end{aligned}$$

and after substitution of this expression in the equation of conservation of entropy in (6) and its linearization, we will come to the equation of thermal conductivity of viscous fluid

$$\rho\theta D_t s + \nabla\cdot q + \tau(u') : e(u') - \rho g \cong$$

$$\cong \rho_0\theta_0\left[\frac{c^2\alpha}{\gamma}\nabla . u' + \frac{c_V}{\theta_0}\frac{\partial\theta}{\partial t}\right] - \nabla\cdot[\chi\nabla\theta] - \tau(u') : e(u') - \rho_0 g$$

or

$$\rho_0 c_V\frac{\partial\theta}{\partial t} - \nabla\cdot[\chi\nabla\theta] + c^2\gamma^{-1}\rho_0\theta_0\alpha\nabla . u' = \rho_0 g \ \ in \ \ \Omega\times(0,T].$$

4. Linearized initial-boundary value problem of dissipative acoustics

Summarizing the results of section 3, we come to the following linearized initial-boundary value problem of dissipative acoustics with a closed system of fundamental equations:

Find displacement $u = \{u_i(x,t)\}_{i=1}^d$ *and temperature* $\theta = \theta(x,t)$ *which satisfy the linearized system of equations of dissipative acoustics*

$$\begin{cases} \rho_0 c_V\theta_0^{-1}\dfrac{\partial\theta}{\partial t} - \theta_0^{-1}\nabla.[\chi\nabla\theta] + c^2\gamma^{-1}\rho_0\alpha\nabla . u' = \rho_0\theta_0^{-1}g, \\ \rho_0 u_i''(t) + \dfrac{\partial}{\partial x_i}\pi(u,\theta) - \dfrac{\partial\tau_{im}(u')}{\partial x_m} = \rho_0 f_i - \dfrac{\partial}{\partial x_i}p_0, \\ \pi(u,\theta) := c^2\gamma^{-1}\rho_0[-\nabla . u + \alpha\theta]\,, \\ \tau_{ij}(u') := 2\mu\, e_{ij}(u') + (\eta - \frac{2}{3}\mu)\delta_{ij}\nabla . u', \\ e_{ij}(u) := \dfrac{1}{2}\left(\dfrac{\partial u_i}{\partial x_j} + \dfrac{\partial u_j}{\partial x_i}\right) \qquad in \quad \Omega\times(0,T], \end{cases} \tag{18}$$

boundary conditions

$$\begin{cases} \sigma_{ij}n_j = \hat{\sigma}_i, & on \quad \Gamma_\sigma\times[0,T],\ \Gamma_\sigma\subset\Gamma, \\ u = \hat{u}, & on \quad \Gamma_v\times[0,T],\ \Gamma_v := \Gamma\backslash\Gamma_\sigma, \\ q.n = \hat{q}, & on \quad \Gamma_q\times[0,T],\ \Gamma_q\subset\Gamma, \\ \theta = \hat{\theta} & on \quad \Gamma_\theta\times[0,T],\ \Gamma_\theta := \Gamma\backslash\Gamma_q \end{cases} \tag{19}$$

and initial conditions

$$u|_{t=0} = u_0, \quad u'|_{t=0} = v_0, \quad \theta|_{t=0} = \theta_0 \quad in \quad \Omega. \tag{20}$$

5. Variational problem of dissipative acoustics

To build a variational formulation of the initial-boundary value problem (18)-(20), we first (taking into account Dirichlet boundary conditions) introduce the space of admissible displacement vectors

$$V := \{v = \{v_i\}_{i=1}^d \in [H^1(\Omega)]^d :\ v = 0 \ on\ \Gamma_u\}$$

and the space of admissible temperatures

$$G := \{\zeta\in H^1(\Omega) :\ \zeta = 0 \ \ on \ \ \Gamma_\theta\}$$

respectively.

Now we shall multiply the equation of heat conductivity of the problem (18)-(20) by arbitrary function $\zeta \in G$ and integrate the obtained result over the domain Ω using integration by parts

$$\begin{aligned}
&\textstyle\int_\Omega \rho_0\theta_0^{-1}g(t)\zeta dx = \\
&= \textstyle\int_\Omega \{\rho_0 c_V\theta_0^{-1}\theta'(t) - \theta_0^{-1}\nabla\cdot[\chi\nabla\theta(t)] + c^2\gamma^{-1}\rho_0\alpha\nabla.u'(t)\}\zeta dx\rho_0 = \\
&= \textstyle\int_\Omega \{\rho_0 c_V\theta_0^{-1}\theta'(t)\zeta + \theta_0^{-1}\nabla\zeta.[\chi\nabla\theta(t)] + c^2\gamma^{-1}\rho_0\alpha\nabla.u'(t)\}dx + \\
&+ \textstyle\int_{\Gamma_q} \theta_0^{-1}\zeta q_m(\theta)n_m d\gamma = \\
&= \textstyle\int_\Omega [\rho_0 c_V\theta_0^{-1}\theta'(t)\zeta + \theta_0^{-1}\nabla\zeta.[\chi\nabla\theta(t)] + c^2\gamma^{-1}\rho_0\alpha\nabla.u'(t)]dx + \\
&+ \textstyle\int_{\Gamma_q} \theta_0^{-1}\hat{q}(t)\zeta d\gamma \qquad\qquad \forall\zeta \in G.
\end{aligned}$$

Let us introduce bilinear and linear forms

$$\begin{cases}
\chi(\theta,\zeta) := \int_\Omega \theta_0^{-1}\chi\nabla\zeta.\nabla\theta dx \\
s(\theta,\zeta) := \int_\Omega \rho_0 c_V\theta_0^{-1}\theta\zeta dx \qquad \forall\theta,\zeta \in G, \\
b(v,\zeta) := \int_\Omega c^2\gamma^{-1}\rho_0\alpha\zeta(\nabla.v)dx \qquad \forall v \in V \;\; \forall\zeta \in G
\end{cases} \tag{21}$$

and

$$< z,\zeta >:= \int_\Omega \rho_0\theta_0^{-1}g\zeta dx - \int_{\Gamma_q} \theta_0^{-1}\hat{q}\zeta d\gamma \qquad \forall\zeta \in G$$

and re-write the equation obtained above as

$$s(\theta'(t),\zeta) + \chi(\theta(t),\zeta) + b(u'(t),\zeta) =< z(t),\zeta > \qquad \forall\zeta \in G.$$

Similarly, we shall multiply the equation of motion of the problem (18)-(20) by arbitrary vector $v \in V$ and integrate the obtained result over the domain Ω

$$\begin{aligned}
&\textstyle\int_\Omega \rho_0 f(t).vdx = \\
&= \textstyle\int_\Omega \left\{\rho_0 u_i''(t) + \dfrac{\partial}{\partial x_m}[\pi[u(t),\theta(t)]\delta_{im} - \tau_{im}(u'(t))]\right\} v_i dx = \\
&= \textstyle\int_\Omega \rho v.\, u''(t)dx + \int_\Omega c^2\gamma^{-1}\rho_0[\nabla.u(t)](\nabla.v)dx - \\
&\qquad - \textstyle\int_\Omega c^2\gamma^{-1}\rho_0\alpha\theta(t)\nabla.vdx + \\
&\quad + \textstyle\int_\Omega \tau(u'(t)) : e(v)dx - \int_{\Gamma_\sigma} v\cdot\hat{\sigma}(t)d\gamma \qquad \forall v \in V.
\end{aligned}$$

Taking the obtained equation into account, we introduce the forms

$$\begin{cases}
m(u,v) := \int_\Omega \rho_0 u.vdx, \\
a(u,v) := \int_\Omega \tau(u) : e(v)dx \equiv \\
\qquad \equiv \int_\Omega [2\mu\, e(u) : e(v) + (\eta - \frac{2}{3}\mu)(\nabla.u)(\nabla.v)]dx, \\
c(u,v) := \int_\Omega c^2\gamma^{-1}\rho_0(\nabla.u)(\nabla.v)dx, \qquad \forall u,v \in V,
\end{cases}$$

$$< l,v >:= m(f - \rho_0^{-1}\nabla p_0, v) + \int_{\Gamma_\sigma} v.\hat{\sigma}d\gamma \quad \forall v \in V \tag{22}$$

and finally write the variational formulation of the initial-boundary value problem of dissipative acoustics

$$\begin{cases} \textit{Find pair } \{u(t),\theta(t)\} \in V \times G \textit{ such that} \\ m(u''(t),v) + a(u'(t),v) + c(u(t),v) - \\ \qquad -b(v,\theta(t)) = < l(t), v >, \\ s(\theta'(t),\zeta) + \chi(\theta(t),\zeta) + b(u'(t),\zeta) = \\ \qquad = < z(t),\zeta > \quad \forall t \in (0,T], \\ m(u'(0)-v_0,v) = 0, \quad a(u(0)-u_0,v) = 0, \qquad \forall v \in V, \\ \qquad s(\theta(0)-\theta_0,\zeta) = 0 \qquad \forall \zeta \in G. \end{cases} \tag{23}$$

Let us remark that bilinear form $b(\ .\ ,\ .\ ) : G \times V \to \mathbb{R}$, we determined in (21), binds variational equations of the problem (23) into a system for determining thermal and mechanical fields of acoustic wave. On the other hand, as we shall see later, this bilinear form describes the mechanism of heat-to-work conversion, and, since it is present in both variational equations, a contraria.

6. Properties of components of variational problem of dissipative acoustics

To perform the analysis of properties of bilinear forms and linear functional which constitute the structure of variational problem (23), we shall first introduce the following notation for spaces of scalar and vector functions

$$H := L^2(\Omega), \quad \mathbf{H} := H^d, \quad H(div;\Omega) := \{v \in \mathbf{H} : \ \ \nabla.v \in H\}.$$

Taking into account the additive values of the problem data (22), it is easy to notice that continuous symmetric bilinear forms

$$\begin{aligned} m(u,v) &= \int_\Omega \rho_0 u.v dx \qquad \forall u,v \in \mathbf{H}, \\ s(\theta,\zeta) &= \int_\Omega \rho_0 c_v \theta_0^{-1} \theta\zeta dx \qquad \forall\ \theta,\zeta \in H \end{aligned} \tag{24}$$

are scalar products on spaces $\mathbf{H}$ and H and as consequence, form norms on them

$$\begin{aligned} |v||_{\mathbf{H}} &:= \sqrt{m(v,v)} \qquad \forall v \in \mathbf{H}, \\ ||\zeta||_H &:= \sqrt{s(\zeta,\zeta)} \qquad \forall\ \zeta \in H, \end{aligned}$$

Equivalent to the norms of spaces $[L^2(\Omega)]^d$ and $L^2(\Omega)$ respectfully.

Similarly, taking into consideration Korn inequality, continuous symmetric bilinear forms

$$\begin{aligned} a(u,v) &= \int_\Omega [2\mu e_{ij}(u)e_{ij}(v) + (\eta - \tfrac{2}{3}\mu)(\nabla.u)(\nabla.v)]dx \quad \forall u,v \in V, \\ \chi(\theta,\zeta) &= \int_\Omega \theta_0^{-1}(\chi\nabla\theta).(\nabla\zeta)dx \qquad \forall\ \theta,\zeta \in G \end{aligned} \tag{25}$$

are scalar products on spaces $\mathbf{V}$ and G respectively, and as consequence, form norms on them

$$||v||_V := \sqrt{a(v,v)} \qquad \forall v \in V \quad (\textit{equivalent} \ || \cdot ||_{[H^1(\Omega)]^d}),$$
$$||\zeta||_G := \sqrt{\chi(\zeta,\zeta)} \qquad \forall \ \zeta \in G \ \ (\textit{equivalent} \ || \cdot ||_{H^1(\Omega)} \ ).$$

The properties of bilinear forms of variational problem that we have mentioned here are well known for problems of elastodynamics and heat conductivity which, as a matter of fact, form the core structure of variational problem of dissipative acoustics.

One of the specific properties of the problem of dissipative acoustic is illustrated by a continuous symmetric bilinear form

$$c(u,v) = \int_\Omega c^2\rho_0\gamma^{-1}(\nabla.u)(\nabla.v)dx \quad \forall u,v \in V,$$

which is non-negative on the space of admissible displacements **V** and creates seminorm in space $H(div;\Omega)$. We shall denote the latter as follows:

$$|v|_V := \sqrt{c(v,v)} \qquad \forall v \in V.$$

And finally, the bilinear form

$$b(v,\zeta) := \int_\Omega c^2\gamma^{-1}\rho_0\alpha\zeta(\nabla.v)dx \qquad \forall v \in V \ \ \forall \zeta \in G,$$

which determines the interaction mechanism of thermal and mechanical fields in the process of spreading acoustic waves, is continuous on the space $V \times G$. Linear functionals also possess this property

$$< z,\zeta >= \int_\Omega \rho_0\theta_0^{-1}g\zeta dx - \int_{\Gamma_q}\theta_0^{-1}\hat{q}\zeta d\gamma \qquad \forall \zeta \in G, \tag{26}$$

$$< l,v >= m(f - \rho_0^{-1}\nabla p_0, v) + \int_{\Gamma_\sigma} v.\hat{\sigma}d\gamma \quad \forall v \in V \tag{27}$$

in case that external sources of mechanics and thermal energy of the problem possess the following properties of regularity

$$g \in H, \ \ \hat{q} \in L^2(\Gamma_q), \ p_0 \in H^1(\Omega),$$
$$f \in H, \ \ \hat{\sigma} \in [L^2(\Gamma_\sigma)]^d.$$

7. Energy equalities of dissipative acoustics

We shall accept for the problem equations (23) for admissible functions $v = u'(t)$ and $\zeta = \theta(t)$ and add the first pair of variational equations. As a result of elimination of summands with the value of bilinear form $b(u'(t),\theta(t))$ (which indicates energy conversion without losses!) and using norms from p.6, we shall obtain energy equations of this problem

$$\frac{1}{2}\frac{d}{dt}[||u'(t)||_H^2 + |u(t)|_V^2 + ||\theta(t)||_H^2] + ||u'(t)||_V^2 + ||\theta(t)||_G^2 =$$
$$=< l(t),u'(t) > + < z(t),\theta(t) > \qquad \forall t \in (0,T]$$

Or after integrating over arbitrary time interval $[0,t]$, $0 \le t \le T$,

$$
\begin{aligned}
&\frac{1}{2}[||u'(t)||_H^2 + |u(t)|_V^2 + ||\theta(t)||_H^2] + \textstyle\int_0^t [||u'(t)||_V^2 + ||\theta(\tau)||_G^2] d\tau = \\
&= \frac{1}{2}[||v(0)||_H^2 + |u(0)|_V^2 + ||\theta(0)||_H^2] + \\
&\qquad + \textstyle\int_0^t [< l(\tau), u'(\tau) > + < z(\tau), \theta(\tau) >] d\tau \quad \forall t \in [0, T].
\end{aligned} \tag{28}
$$

We shall write the last equation as

$$
\begin{aligned}
&K_S[u'(t)] + P_S[u(t)] + P_C[\theta(t)] + \textstyle\int_0^t \{D_S[u(\tau)] + D_C[\theta(\tau)]\} d\tau = \\
&= K_S[v_0] + P_S[u_0] + P_C[\theta_0] + Q_S[u'(t)] + Q_C[\theta(t)] \quad \forall t \in [0,\ T],
\end{aligned}
$$

where

$$
K_S[u'(t)] := \frac{1}{2}||u'(t)||_H^2, \quad P_S[u(t)] := \frac{1}{2}|u(t)|_V^2,
$$
$$
D_S[u'(t)] := ||u'(t)||_V^2
$$

are instantaneous values of kinetic and potential energy, and its dissipation caused by kinetic motion of fluid, in the function

$$
P_C[\theta(t)] := ||\theta(t)||_H^2, \quad D_C[\theta(t)] := ||\theta(t)||_G^2
$$

they are instantaneous values of energy and its losses, caused by the existence of heat flow pattern of fluid,

$$
Q_S[u'(t)] := \int_0^t < l(\tau), u'(\tau) > d\tau, \quad Q_C[\theta(t)] := \int_0^t < \mu(\tau), \theta(\tau) > d\tau.
$$

8. Data regularity of a problem of dissipative acoustics

Let us consider the conditions of data regularity for the variation problem (22), as functions of space and time variables, which can be determined on the basis of equality analysis (28). In particular, to allow the total energy of acoustic field of fluid

$$
E[u(t), \theta(t)] := \frac{1}{2}[||u'(t)||_H^2 + |u(t)|_V^2 + ||\theta(t)||_H^2]
$$

take finite values in each moment of time $t \in (0, T]$, it is necessary that the following conditions are held

$$
u' \in L^\infty(0, T; H), \ \ u \in L^\infty(0, T; H(div; \Omega)), \ \ \theta \in L^\infty(0, T; H).
$$

Similarly, to allow the the losses of acoustic field of fluid

$$
D[u(t), \theta(t)] := \int_0^t [||u'(t)||_V^2 + ||\theta(\tau)||_G^2] d\tau
$$

take finite values in each moment of time $(0, t] \subset (0, T]$, it is necessary that the following conditions are held

$$
u' \in L^2(0, T; V), \quad \theta \in L^2(0, T; G).
$$

Thus, appropriate solutions of the variational problem of dissipative acoustics should satisfy the following conditions

$$\begin{cases} u' \in L^\infty(0,T;H) \cap L^2(0,T;V), \\ u \in L^\infty(0,T;H(div;\Omega)), \\ \theta \in L^\infty(0,T;H) \cap \theta \in L^2(0,T;G). \end{cases}$$

Now based on the requirement

$$\left| \int_0^t [< l(\tau), u'(\tau) > + < z(\tau), \theta(\tau) >] d\tau \right| < +\infty \quad \forall t \in (0,T]$$

we find sufficient requirements of regularity for energy sources, such as,

$$l \in L^2(0,T;V'), \;\; z \in L^2(0,T;G')$$

Or in more detail, taking into consideration the structures (26) and (27) of these functionals

$$\begin{cases} f \in L^2(0,T;H), \quad \hat{\sigma} \in L^2(0,T;[L^2(\Gamma_\sigma)]^d), \\ g \in L^2(0,T;H), \quad \hat{q} \in L^2(0,T;L^2(\Gamma_q)). \end{cases}$$

The latter sum

$$E[u(0),\theta(0)] := \frac{1}{2}[||u'(0)||_H^2 + |u(0)|_V^2 + ||\theta(0)||_H^2]$$

of energy equality (28) shows that the total energy of the acoustic field at the initial moment of time $t = 0$ will have finite values, if the initial data of the problem of dissipative acoustics are selected according to the rules

$$v_0 \in H, \quad u_0 \in V, \quad \theta_0 \in H.$$

9. Uniqueness and stability of solution of the variational problem of dissipative acoustics

Now we are ready to prove the next theorem

Theorem 1. *Assume that the variational problem of dissipative acoustics* (23), *whose data satisfy the conditions of regularity*

$$v_0 \in H, \quad u_0 \in V, \;\; \theta_0 \in H \tag{29}$$

and

$$\begin{cases} f \in L^2(0,T;H), \quad \hat{\sigma} \in L^2(0,T;[L^2(\Gamma_\sigma)]^d), \\ g \in L^2(0,T;H), \quad \hat{q} \in L^2(0,T;L^2(\Gamma_q)), \end{cases} \tag{30}$$

has the solution $\psi(t) = \{u(t),\, \theta(t)\}$.

Then the pair $\psi(t) = \{u(t),\, \theta(t)\}$ *will be the unique solution to the problem* (23) *and*

$$\begin{cases} L^\infty(0,T;H(div;\Omega)), \quad u' \in L^\infty(0,T;H) \cap L^2(0,T;V), \\ \theta \in L^\infty(0,T;H) \cap L^2(0,T;G); \end{cases}$$

Moreover, the solution $\psi(t) = \{u(t), \theta(t)\}$ *is continuously dependent on the problem data* (23) *and under these conditions the following a priori estimate is correct*

$$\begin{aligned} &\frac{1}{2}[||u'(t)||_H^2 + |u(t)|_V^2 + ||\theta(t)||_H^2] + \int_0^t [||u'(t)||_V^2 + ||\theta(\tau)||_G^2] d\tau \leq \\ &\leq C\left\{[||v_0||_H^2 + |u_0|_V^2 + ||\theta_0||_H^2] + \int_0^t [||l(\tau)||_{V'}^2 + ||(\tau)||_{G'}^2] d\tau\right\}, \\ &\qquad \forall t \in [0, T]. \end{aligned} \tag{31}$$

with constant $C > 0$*, the value of which is independent of quantities under consideration.*

Proof. Bearing in mind the conditions (30)

$$l \in L^2(0,T;V'), \;\; z \in L^2(0,T;G'),$$

we conclude that the following estimates are correct

$$\begin{aligned} &|< l(\tau), u'(\tau) >| \leq ||l(\tau)||_{V'}||u'(\tau)||_V \leq \frac{1}{2}\left\|u'(\tau)\right\|_V^2 + \frac{1}{2}\left\|l(\tau)\right\|_{V'}^2, \\ &|< z(\tau), \theta(\tau) >| \leq \frac{1}{2}\left\|\theta(\tau)\right\|_G^2 + \frac{1}{2}\left\|z(\tau)\right\|_{G'}^2, \qquad \forall t \in [0, T]. \end{aligned} \tag{32}$$

From the initial condition of the problem (23)

$$m(u'(0) - v_0, v) = 0, \qquad \forall v \in H$$

After substituting $v = u'(0)$ and $v = v_0$ we obtain that

$$||u'(0)||_H^2 = m(u'(0), v_0) = m(v_0, u'(0)) = m(v_0, v_0) = ||v_0||_H^2. \tag{33}$$

Applying the same principle

$$|u(0)||_V = ||u_0||_V, \;\; ||\theta(0)||_H = ||\theta_0||_H. \tag{34}$$

Next, taking into account the results from p.6, we find $C = const > 0$, such that

$$|v|_V \leq C||v||_V \qquad \forall v \in V$$

and, in particular,

$$|u(0)|_V \leq C||u(0)||_V = C||u_0||_V. \tag{35}$$

Summarizing (32)-(34) and (35) in energy equality (28), we come to an estimate (31).

Based on the same estimate, by contradiction, we demonstrate the uniqueness of the problem solution (23). □

Corollary 1. *Let us assume that the hypotheses of theorem 1 are satisfied in relation to the variation problem of dissipative acoustics* (23).

Then the natural norm for its solution $\psi(t) = \{u(t),\, \theta(t)\}$ *is*

$$||\psi(t)\,||^2 := ||u'(t)||^2_H + |u(t)|^2_V + ||\theta(t)||^2_H +$$

$$+\int_0^t [||u'(\tau)||^2_V + ||\theta(\tau)||^2_G] d\tau \qquad \forall t \in [0,T].$$

10. Galerkin semi-discretization of variational problem of dissipative acoustics

Let us assume that $\{V_h\}$ та $\{G_h\}$ are sequences of finite-dimensional spaces, such that

$$\begin{cases} V_h \subset V,\;\; G_h \subset G \qquad \forall h > 0, \\ \dim V_h = N = N(h) \to \infty, \\ \dim G_h = M = M(h) \to \infty, \;\; if \;\; h \to 0, \\ \bigcup\limits_{h>0} V_h \; dense \;\; in \;\; V, \;\; \bigcup\limits_{h>0} G_h \, dense \;\; in \;\; G. \end{cases}$$

On this basis we determine the sequence of semi-discrete Galerkin approximations $\{\psi_h\}_{h>0} = \{(u_h, \theta_h)\}_{h>0}$ expressed as solutions of the following variational problems:

$$\begin{array}{l} given\, h > 0;\, find\, pair\; \psi_h(t) = (u_h(t),\, \theta_h(t)) \,\in\, V_h \times G_h \, such\, that \\ \begin{cases} m(u''_h(t), v) + a(u'_h(t), v) + c(u_h(t), v) - \\ \qquad -b(\theta_h(t), v) =< l(t), v >, \\ s(\theta'_h(t), \zeta) + k(\theta_h(t), \zeta) + b(\zeta, u'_h(t)) =< z(t), \zeta > \qquad \forall t \in (0,T], \\ m(u'_h(0) - v_0, v) = 0, \;\; a(u_h(0) - u_0, v) = 0 \qquad \forall v \in V_h, \\ \qquad s\,(\theta_h(0) - \theta_0, \zeta) = 0 \qquad \forall \zeta \in G_h. \end{cases} \end{array} \tag{36}$$

To concretize the structure of problems we have just formulated and the required approximations $(u_h, \theta_h) \in L^2(0, T; V_h \times G_h)$, let us select certain bases $\{\phi_k(x)\}_{k=1}^N$ and $\{\varphi_k(x)\}_{k=1}^M$ of spaces V_h and G_h respectively. First of all, this selection univalently determines the form of each sequence member of semi-discrete approximations as a linear combination

$$u_h(x,t) = \sum_{k=1}^N u_k(t)\phi_k(x)\;,$$

$$\theta_h(x,t) = \sum_{k=1}^M \vartheta_k(t)\varphi_k(x) \quad \forall (x,t) \in \Omega \times [0,T]$$

with unknown coefficients $U(t) = \{u_k(t)\}_{k=1}^N$ and $\Theta(t) = \{\vartheta_m(t)\}_{m=1}^M$, and secondly, after application of Galerkin procedure, allows obtaining Cauchy problem for finding the above-mentioned coefficients

$$\left\{\begin{array}{l} MU''(t)+AU'(t)+CU(t)-B\,\Theta(t) \;=L(t), \\ \quad S\,\Theta'(t)+K\Theta(t)+B^TU'(t)=Z(t) \quad \forall t\in(0,T], \\ MU'(0)=Y^0, \quad AU(0)=U^0, \\ \quad S\,\Theta(0)=\Theta^0. \end{array}\right. \tag{37}$$

Here the components of matrices and vectors of the right side of equation are calculated according to the rules

$$C=\{c(\phi_i,\phi_k)\}_{i,k=1}^N\,,\quad B=\{b(\varphi_i,\phi_k)\}_{i,k=1}^{M,N}\,,\quad K=\{k(\varphi_i,\varphi_k)\}_{i,k=1}^M\,,$$

$$L(t)=\{<l(t),\phi_i>\}_{i=1}^N\,,\quad Z(t)=\{<z(t),\varphi_i>\}_{i=1}^M \qquad \forall T\in(0,t],$$

and

$$Y_0=\{m(\upsilon_0,\phi_k)\}_{k=1}^N\,,\quad U_0=\{a(u_0,\phi_k)\}_{k=1}^N\,,\ \Theta_0=\{s(\theta_0,\varphi_i>\}_{i=1}^M.$$

Since the rest of the matrices

$$M=\{m(\phi_i,\phi_k)\}_{i,k=1}^N\,,\ A=\{a(\phi_i,\phi_k)\}_{i,k=1}^N\,,\quad S=\{s(\varphi_i,\varphi_k)\}_{i,k=1}^M$$

are the Gram matrices in systems of linearly independent functions $\{\phi_k(x)\}_{k=1}^N$ and $\{\varphi_k(x)\}_{k=1}^M$ (in relation to scalar products described in p. 6, see (24) and (25)), it follow that they are positively defined. This fact guarantees the possibility of unique solution of the system of ordinary differential equations of Cauchy problem (37) and also systems of linear algebraic equations of its initial conditions in relation to vectors $U(0)$, $U'(0)$ and $\Theta(0)$. From here it follows that for each constant $h>0$ the Cauchy problem (37) has a unique solution$\{U(t),\Theta(t)\}$, which allows finding univalently the semi-discrete Galerkin approximation $(u_h,\theta_h)\in L^2(0,T;V_h\times G_h)$ as (36).

Theorem 2. *Let us assume that the data of variational problem of dissipative acoustics* (23) *is characterized by the conditions of regularity*

$$\upsilon_0\in H,\quad u_0\in V,\ \ \theta_0\in H$$

and

$$\left\{\begin{array}{l} f\in L^2(0,T;H), \quad \hat{\sigma}\in L^2(0,T;[L^2(\Gamma_\sigma)]^d), \\ g\in L^2(0,T;H), \quad \hat{q}\in L^2(0,T;L^2(\Gamma_q)). \end{array}\right.$$

Then for each value of discretization parameter $h>0$ the following statements will be true:

(i) the semi discretized problem has a unique solution (36) $\psi_h=\{u_h,\,\theta_h\}$ *and*

$$\left\{\begin{array}{l} u_h\in L^\infty[0,T;H(div;\Omega)], \quad u_h'\in L^\infty(0,T;H)\cap L^2(0,T;V), \\ \theta_h\in L^\infty(0,T;H)\cap L^2(0,T;G); \end{array}\right.$$

(ii) semi-discrete approximation $\psi_h=\{u_h,\,\theta_h\}$ is continuously dependent on the problem data (23)*, more, the following a priori estimate is correct*

$$\frac{1}{2}[||u_h'(t)||_H^2 + |u_h(t)|_V^2 + ||\theta_h(t)||_H^2] + \int_0^t [||u_h'(t)||_V^2 + ||\theta_h(\tau)||_G^2] d\tau \leq$$
$$\leq C \left\{ [||v_0||_H^2 + |u_0|_V^2 + ||\theta_0||_H^2] + \int_0^t [||l(\tau)||_{V'}^2 + ||z(\tau)||_{G'}^2] d\tau \right\}$$
$$\forall t \in [0,T] \quad \forall h > 0.$$

with constant $C > 0$, the value of which is independent of quantities under consideration.

11. Existence of solution variation problem of dissipative acoustics

Theorem 3. *Let us assume that the data of problem of dissipative acoustics* (23) *are characterized by regularity conditions* (29) *and* (30). *Then the variational problem* (23) *has a unique solution $\psi = \{u, \theta\}$ and*

$$\begin{cases} u_h \in L^\infty[0,T; H(div;\Omega)], \quad u_h' \in L^\infty(0,T;H) \cap L^2(0,T;V), \\ \theta_h \in L^\infty(0,T;H) \cap L^2(0,T;G); \end{cases}$$

moreover

$$\frac{1}{2}[||u'(t)||_H^2 + |u(t)|_V^2 + ||\theta(t)||_H^2] + \int_0^t [||u'(t)||_V^2 + ||\theta(\tau)||_G^2] d\tau \leq$$
$$\leq C \left\{ [||v_0||_H^2 + |u_0|_V^2 + ||\theta_0||_H^2] + \int_0^t [||l(\tau)||_{V'}^2 + ||z(\tau)||_{G'}^2] d\tau \right\},$$
$$\forall t \in [0,T].$$

with constant $C > 0$, the value of which is independent of quantities under consideration.

Proof. Bearing in mind the theorem 1 we need to estimate the existence of solution (23).

As it follows from the theorem 10.1, the sequence of semi-discrete Galerkin approximations $\psi_h = \{u_h, \theta_h\}$ (and also $\{u_h'\}$) form at $h \to 0$ bounded sets in the space $L^\infty(0,T;V) \times [L^\infty(0,T;H) \cap L^2(0,T;G)]$ (respectively$L^\infty(0,T;H) \cap L^2(0,T;V)$).

Therefore, among them we can select convergent subsequence $\psi_\Delta = \{u_\Delta, \theta_\Delta\}$ and $\{u_\Delta'\}$such that

$$\begin{cases} \psi_\Delta = \{u_\Delta, \theta_\Delta\} \underset{\Delta \to 0}{\to} \psi = \{u, \theta\} \ \ in \ \ L^2(0,T; V \times G) \ \ \text{weakly}, \\ u_\Delta' \underset{\Delta \to 0}{\to} u' \ \ in \ \ L^2(0,T;V) \ \ \text{weakly}. \end{cases}$$

After that it remains for us to show that the limit $\psi = \{u, \theta\}$ obtained in this way from space $L^2(0,T; V \times G)$ is the solution of the problem (23); more precise, it is the matter of direct verification to prove that the pair $\psi = \{u, \theta\}$ satisfies the equation of this problem.

For this purpose we select the spaces $V_h \subset V$, $G_h \subset G$ and $W := \{g \in C^1([0,T]) \mid g(T) = 0\}$. Let us assume that as before $\{\phi_k(x)\}_{k=1}^N$ and $\{\varphi_k(x)\}_{k=1}^M$ are bases of the spaces V_h and G_h respectively and

$$v_h(t) = \sum_{i=1}^{n} q_i(t)\phi_i \in V_h \quad \forall q_i \in W, \qquad g_h(t) = \sum_{i=1}^{k} \eta_i(T)\varphi_i \in G_h \quad \forall \eta_i \in W.$$

Due to the problem (36) we have

$$\begin{cases} m(u''_\Delta(t), v_h(t)) + a(u'_\Delta(t), v_h(t)) + c(u_\Delta(t), v_h(t)) - \\ \qquad -b(\theta_\Delta(t), v_h(t)) = < l(t), v_h(t) >, \\ s(\theta'_\Delta(t), g_h(t)) + k(\theta_\Delta(t), g_h(t)) + b(g_h(t), u'_\Delta(t)) = \\ \qquad = < \mu(t), g_h(t) > \quad \forall t \in (0,T]. \end{cases}$$

After time integration over the interval $(0,T)$ when applying integration by parts and initial conditions from (36), we obtain

$$\begin{cases} \int_0^T \{-m(u'_\Delta, v'_h) + a(u'_\Delta, v_h) + c(u_\Delta, v_h) - b(\theta_\Delta, v_h) - < l, v_h >\} d\tau = \\ \qquad = -m(u'_\Delta(0), v_h(0)) = -m(v_0, v_h(0)), \\ \int_0^T -s(\theta_\Delta, g'_h) + k(\theta_\Delta, g_h) + b(g_h, u'_\Delta) - < \mu, g_h >\} d\tau = \\ \qquad = -s(\theta_\Delta(0), g_h(0)) = -s(\theta_0, g_h(0)). \end{cases}$$

In the derived equations we proceed to the limit with $\Delta \to 0$, and then again perform integration by parts, we obtain

$$\begin{cases} \int_0^T \{m(u'', v_h) + a(u', v_h) + c(u, v_h) - b(\theta, v_h) - < l, v_h >\} d\tau = \\ \qquad = m(u'(0) - v_0, v_h(0)) \qquad \forall v_h \in C^1([0,T]; V_h) \\ \int_0^T s(\theta', g_h) + k(\theta, g_h) + b(g_h, u') - < \mu, g_h >\} d\tau = \\ \qquad = s(\theta(0) - \theta_0, g_h(0)) \qquad \forall g_h \in C^1([0,T]; G_h). \end{cases}$$

Since V_h is dense in space V, and G_h is dense in space G, the final equations is true for each $v \in C^1([0,T]; V)$ and $g \in C^1([0,T]; G)$.

$$\begin{cases} m(u'', v) + a(u', v) + c(u, v) - b(\theta, v) = < l, v >, \\ \qquad s(\theta', g) + k(\theta, g) + b(g, u') = < \mu, g >, \\ m(u'(0) - v_0, v) = 0 \quad \forall v \in V, \; s(\theta(0) - \theta_0, g) = 0 \quad \forall g \in G. \end{cases}$$

Finally, from the initial conditions and considering (36)

$$a(u_0, v) = a(u_\Delta(0), v) \to a(u(0), v) \qquad \forall v \in V.$$

It follows that the pair $\psi = \{u, \theta\}$ is the solution of the variational problem (23). Moreover, for this solution the energy equation (28) and estimate (31) stay true. The uniqueness of solution of variational problem (23) results from (31) and proof by contradiction. □

12. Conclusions

On the basis of the conservation laws, we have formulated fundamental equations, phenomenological relations, initial and boundary conditions that describe the motion of Newtonian viscous heat-conducting fluid in terms of mass density, vector of velocity, pressure, entropy and temperature. By applying for this non-closed model of hydrodynamics the hypotheses of acoustic disturbances of fluid by linearization, we have found the initial boundary value problem and corresponding variational problem only in terms of vector of acoustic displacements and temperature, which describes the process of spreading acoustic waves with consideration of connectivity of mechanical and thermal fields. We have determined the regularity class of input data of variational problem, which guarantee uniqueness and continuous dependence of the required solution in the energy norm of the problem. In addition, the existence of solution of the considered problem has been presented as a limit of sequence of semi-discrete (by spatial variables) Galerkin approximations.

The obtained results form a fully-functional system for successful modeling and analysis of numeric schemes for solving problems of dissipative acoustics. In particular, one of such schemes can be obtained by direct application of the one-step recurrent scheme for time integration of semi-discretized variational problem (36) using classic approximation spaces of the finite element method [8]. The results of modeling and analysis of convergence of such schemes will be presented in the nearest future.

Bibliography

1. Blokhintsev D. I. Acoustics of a nonhomogeneous moving medium / D. I. Blokhintsev.– Moscow: Nauka, 1981.– 208 p. (in Russian).
2. Voytovych V. M. Displacement-based modelling of acoustic fluid-structure interaction problem / V. M. Voytovych, V .M. Horlatch, Y. V. Kondratyuk, H. A. Shynkarenko // Applied problems of mechanics and mathematics.– 2005. № 2.– P. 108-118 (in Ukrainian).
3. Grinchenko V. T. Fundamentals of Acoustics / V. T. Grinchenko, Sh. V. Vovk, V. T. Matsipura.– Kiev: Naukova Dumka, 2007.– 640 p. (in Ukrainian).
4. Duvaut G. Inequalities in mechanics and physics / G. Duvaut, J.-L. Lions.– Moscow: Nauka, 1980.– 384 p. (in Russian).
5. Isakovich M. A. General Acoustics / M. A. Isakovich.– Moscow: Nauka, 1973.– 502 p. (in Russian).
6. Krasilnikov V. A. Introduction to Physical Acoustics / V. A. Krasilnikov, V. Krylov.– Moscow: Nauka, 1984.– 400 p. (in Russian).
7. Landau L. D. Theoretical Physics: Hydrodynamics / L. D. Landau, E. M. Lifshitz.– Moscow: Nauka, 1986.– 736 p. (in Russian).
8. Shynkarenko H. A. Projection-mesh methods for solving initial boundary value problems / H. A. Shynkarenko.– KA: SMC IVO, 1991.– 88 p. (in Ukrainian).

9. Bermudez A. Two discretization schemes for a time-domain dissipative acoustics problem / A. Bermudez, R. Rodrigues, D. Santamarina // Mathematical Models and Methods in Applied Sciences.– 2006.– Vol. 16, Issue 10.– P. 1559-1598.
10. Bruneau M. Fundamentals of acoustics / M. Bruneau.– London: ISTE, 2006. – 636 p.
11. Pierce A. D. Acoustics: An introduction to its physical principles and applications / A. D. Pierce.– NY: ASA, 1991.– 678 p.

Vitaliy Horlatch, Ira Klymenko, Georgiy Shynkarenko,
Ivan Franko National University of Lviv,
1, Universytets'ka Str., Lviv, 79000, Ukraine
Georgiy Shynkarenko,
Opole University of Technology,
5, Luboszycka Str., Opole, 45043, Poland

Received 13.07.2012

Журнал обчислювальної та прикладної математики | **2012, № 3 (109)** с. 72–88 | Journal of Numerical & Applied Mathematics

UDC 519.642

BALANCING PRINCIPLE FOR ITERATED TIKHONOV METHOD OF SEVERELY ILL-POSED PROBLEMS

Ganna Myleiko, Sergei Solodky

Резюме. В даній статті розглядається проблема наближеного розв'язування жорстко некоректних задач зі збуреними вхідними даними. До регулювання таких задач було застосовано ітерований метод Тіхонова з правилом зупинки згідно принципу рівноваги. Для запропонованого підходу була знайдена порядкова оцінка похибки на класі задач, що досліджуються.

Abstract. Considered in this paper are the problem of approximate solving severely ill-posed problems with perturbed input data. In oder to regularize these problems the iterated Tikhonov method with balancing principle as stop rule was applied. For this suggesting approach an order of accuracy on the class of problems under investigation was found.

1. Introduction

In this paper we consider the problem of approximate solving severely ill-posed problems represented in the form of operator equation of the first kind

$$Ax = y, \tag{1}$$

where $A : X \to Y$ is linear compact injective operator between Hilbert spaces X and Y. Let us denote inner products in these spaces by $(\cdot, \cdot)$ and corresponding norms by $\| \cdot \|$. The symbol $\| \cdot \|$ stands also for standart operator norm. It will become clear from the context which exactly space or norm is under consideration. Suppose also that an available perturbation $y_\delta \in Y : \|y - y_\delta\| \leq \delta, \delta > 0$, is known instead of the right-hand side y and a perturbed operator $A_h : \|A - A_h\| \leq h, h > 0$, is known instead of A, where $A_h : X \to Y$ is also linear compact injective one.

Usually, equation (1) is referred to as a severely ill-posed problem if its solution $x_0 = A^{-1}y$ has a finite "smoothness" in some sense, but A is an infinitely smoothing operator.

A distinguishing characteristic of such kind of problems is the fact that x_0 belongs to some subspace V continuously embedded in X, the singular values of the canonical embedding operator J_V from V into X tend to zero with polynomical rate, while the singular values $\{\sigma_l\}_{l=1}^{\infty}$ of the operator A tend to zero exponentially.

Following [2], [7] suppose that x_0 belongs to the set

$$M_{p,\rho}^K(A) := \{x : x = (\underbrace{\ln \ldots \ln}_{K\text{-times}}(A^*A)^{-1})^{-p} v, \quad \|v\| \leq \rho\}, \tag{2}$$

†*Key words.* Severely ill-posed problem, balancing principle, iterated Tikhonov method.

when some unknown $0 < p \le p_1$, $K = 1,2,...$, and known $\rho > 0$, where the operator function $(\underbrace{\ln ... \ln}_{K\text{-times}}(A^*A)^{-1})^{-p}$ well defined by the spectral decomposition

$$A^*A = \sum_{l=1}^{\infty} \sigma_l^2 (\Psi_l, \cdot)\Psi_l$$

of the operator A^*A, i.e.

$$(\underbrace{\ln ... \ln}_{K\text{-times}}(A^*A)^{-1})^{-p} v = \sum_{l=1}^{\infty} (\underbrace{\ln ... \ln}_{K\text{-times}}(\sigma_l^{-2}))^{-p} (\Psi_l, v)\Psi_l.$$

Further, without loss of generality we assume that

$$\|A\| \le M_K,\ M_K = m_K^{1/2}, m_k = \begin{cases} e^{-1}, & k = 1, \\ e^{-\frac{1}{m_{k-1}}}, & k = 2, ..., K \end{cases},$$

i.e.

$$\sigma_l \le m_K, \quad l = 1, 2, ... \ .$$

Example 1. To illustrate severely ill-posed problems let us consider a problem from satellite gravity gradiometry. With the surfaces of the Earth and the satellite orbit assumed to be sphericals with radius $r_1 < r_2$, correspondently, $\Omega_{r_i} = \left\{u \in \mathbb{R}^3, |u| = r_i\right\}, i = 1, 2$, then one of the problems arising in this theory (see, e.g.,[4], [11]) could be formulated as an equation (1) with the operator

$$Ax(u) := \frac{1}{4\pi r_1} \int\limits_{\Omega_{r_1}} \frac{d^2}{dr_2^2} \left(\frac{r_2^2 - r_1^2}{|u - v|^3} \right) x(v) d\Omega_{r_1}(v), \quad u \in \Omega_{r_2}. \qquad (3)$$

In satellite gradiometry the exact solution of equation (1) with operator (3) is usually considered to be an element of the spherical Sobolev space

$$\mathcal{H}^s := \{ f \in L_2(\Omega_{r_1}) : \|f\|_s^2 = \sum_{l=0}^{\infty} \sum_{k=1}^{2l+1} \left(l + \frac{1}{2} \right)^{2s} |\langle Y_{l,k}^{(1)}, f \rangle|^2 < \infty \ \}$$

for some positive index s, where

$$Y_{l,k}^{(1)}(\omega) = \frac{1}{r_1} Y_{m,j}(\frac{\omega}{r_1}), \quad \omega \in \Omega_{r_1},$$

$$\langle Y_{l,k}^{(1)}, x \rangle = \int\limits_{\Omega_{r_1}} Y_{l,k}^{(1)}(v) x(v) d\Omega_{r_1}(v)$$

and $\{Y_{m,j}, m = 0, 1, ..., j = 1, 2, ..., 2m + 1\}$ is a set of spherical harmonics L_2-orthonormalized with respect to the unit sphere in $\mathbb{R}^3$.

As for the singular values σ_l of the operator (3) the following relation (see, e.g., [12])

$$\ln \sigma_l^{-2} \asymp l + \frac{1}{2}$$

is valid, then there are some constants $c_2 > c_1 > 0$ such that for any $f \in \mathcal{H}^s$ two-sided estimate

$$c_1\|f\|_s \leq \| \ln^s(A^*A)^{-1}f\| \leq c_2\|f\|_s$$

is valid. It, in particular, means that any element of $\mathcal{H}^s$ belongs to the set (2) with $K = 1$ and $p = s$.

Example 2. Let us consider a two-dimensional model of the scattering by a perfectly reflecting periodic structure. According to Bao [3], Hettlich and Kirsch [5], we can formulate the problem as follows. Let $f \in C^2(\mathbb{R})$ be 2π-periodic function with $f(x) > 0$ for all $x \in \mathbb{R}$. We set

$$\Omega_f = \{(x, y) : y > f(x), x \in \mathbb{R}\}.$$

Then by

$$\partial\Omega_f = \{(x, y) : y = f(x), x \in \mathbb{R}\}$$

we denote a periodic interface which should be determined from scattering data. For this end, we introduce an incident field $u^I(x, y; k)$ given by

$$u^I(x, y; k) = \exp\{ik(x \sin\theta - y \cos\theta)\}, \tag{4}$$

which is a time-harmonic electromagnetic plane wave. Here $i = \sqrt{-1}$ and the constant $k \in \mathbb{R}$ is the refraction index of the material occupying Ω_f, and is given by $k = \omega c_0^{-1}\sqrt{\varepsilon\mu}$, where ω is the angular frequency, c_0 is the speed of light, $\mu > 0$ is the magnetic permeability and ε is the dielectric coefficient. Moreover, in (4), θ is regarded as the angle of incidence.

We assume that

$$0 < |\theta| < \frac{\pi}{2}$$

and

$$0 < k < \frac{1}{2\pi}.$$

Then the resulting scattering field $u^S(x, y; k)$ satisfies the Helmholtz equation with the perfect reflection boundary condition

$$\Delta u^S + k^2 u^S = 0 \text{ in } \Omega_f, \tag{5}$$

$$u^S + u^I = 0 \text{ on } \partial\Omega_f, \tag{6}$$

u^S satisfies so-called outgoing wave condition:

$$u^S = \sum_{n\in\mathbb{Z}} u_n e^{i(\alpha_n x+\beta_n y)}, \text{ if } y > \|f\|_{C[0;2\pi]}. \tag{7}$$

In this example the function u^S under consideration is regarded as complex-valued. Here, we set

$$\alpha_n = n + k \sin\theta, \quad \beta_n = \sqrt{k^2 - (n + k \sin\theta)^2}, \quad 0 \leq \arg\beta_n < \pi. \tag{8}$$

Moreover, we impose the $(k \sin\theta)$-quasi-periodicity condition over u^S

$$u^S(x + 2\pi, y; k) = \exp(2\pi ik \sin\theta)u^S(x, y; k) \tag{9}$$

for all $(x, y) \in \mathbb{R}^2$ (see, e.g., [3]).

Now we can state our inverse problem.

Determine $y = f(x), \quad x \in \mathbb{R}$, from measurement $u^S(x, y; k)$, $x \in (0; 2\pi)$, where u^S satisfies (5)-(7) and (9).

By the $(k \sin\theta)$-quasi-periodicity, setting

$$u = u(x, y; k) = u^I(x, y; k) + u^S(x, y; k).$$

We can rewrite (5)-(7) and (9) in terms of the total field u:

$$\Delta u + k^2 u = 0 \text{ in } \Omega_f, \tag{10}$$

$$u = 0 \text{ on } \partial\Omega_f, \tag{11}$$

$$u(x + 2\pi, y; k) = \exp(2\pi i k \sin\theta) u(x, y; k), \tag{12}$$

$$u - u^I \text{ satisfies the outgoing wave condition.} \tag{13}$$

Since k is fixed such that (8) is true, we simply write $u(x, y)$ in place of $u(x, y; k)$. Then our inverse problem is equivalent to determine $y = f(x), \quad x \in \mathbb{R}$, from measurement

$$u(x, 0), x \in (0; 2\pi),$$

where u satisfies (10)-(13).

For fixed positive constants $M_0, \quad M, \quad k$ and $a_0, \quad a$ such that $0 < M \le a_0 \le a$ and $0 < k < 1$, we set

$$\mathcal{F} = \{f \in C^{3+k}(\mathbb{R}) : \|f\|_{C^{3+k}[0;2\pi]} \le M_0, \quad f \quad \text{is} \quad 2\pi\text{-periodic},$$
$$\frac{d^j f}{dx^j}(0) = \frac{d^j f}{dx^j}(2\pi), \quad j = 0, 1, 2, 3,$$
$$f(0) = f(2\pi) = -a_0, \quad -a \le f(x) \le -M,$$
$$0 \le x \le 2\pi\}$$

as an admissible set of unknown surfaces.

Denote

$$\|f\|_{C^{3+k}[0;2\pi]} = \sum_{j=0}^{3} \left\| \frac{d^j f}{dx^j} \right\|_{C[0;2\pi]} + \sup_{0<x,x' \le 2\pi, x \ne x'} \frac{|(\frac{d^3 f}{dx^3})(x) - (\frac{d^3 f}{dx^3})(x')|}{|x - x'|^k}.$$

Let us set

$$\Omega_f = \{(x, y) : y > f(x), x \in \mathbb{R}\} \text{ for } f \in \mathcal{F}.$$

For $f_j \in \mathcal{F}, j = 1, 2$, let us consider

$$\Delta u_j + k^2 u_j = 0 \text{ in } \Omega_{f_j},$$
$$u_j = 0 \text{ on } \partial\Omega_{f_j},$$
$$u_j \text{ is } (k \sin\theta)\text{-quasi-periodicity, i.e.}$$
$$u_j(x + 2\pi, y) = \exp(2\pi i k \sin\theta) u_j(x, y).$$

We further assume that $u_j - u^I$ satisfies the outgoing wave condition.

Theorem (2.1) [5] shows that in stated above conditions there exists a constant $C = C(k, \theta, \mathcal{F}) > 0$ such that

$$\|f_1 - f_2\|_{C[0;2\pi]} \le \frac{C}{|\ln|\ln \frac{1}{\|(u_1 - u_2)(\cdot, 0)\|_{H^1(0;2\pi)}}||}$$

provided that for all $f_1, f_2 \in \mathcal{F}$. Hence, solution of equation (5) belongs to the set (2) with $K = 2$ and $p = 1$.

As far as the history of studing severely ill-posed problems, we should notice, that these studies could be traced back to work [8], where the estimate of accuracy for the Tikhonov regularization were found for equations (1) with operators of both finite and infinite smoothness. Moreover, some regularization methods for severely ill-posed problems were considered in [6], where, in particular, a general class regularization methods (according to Bakushinskiy; see, e.g., [1]) were suggested for solving (1) in the case of perturbed operators and the right-hand sides; for choosing a regularization parameter was employed a modification from [10]. Further, severely ill-posed problems were considered, in particular, in works [7], [2], [12], [13]. In [12] the approch for solving ill-posed problems (1) with solutions from (2) for $K = 1$ was proposed. It suggests a combination of usual Tikhonov's regularization with Morozov's discrepancy principle. The indicated combination allows to achieve the order-optimal accuracy (in the logarithmic scale) $O(\ln^{-1}\frac{1}{\delta})$ of recovering solution from the set $M^1_{p,\rho}(A)$ for any $p > p_0 > 0$. In [13] for solving the same problem Tikhonov's method was employed again; however, for the stop rule was considered the balancing principle. This approach also allows to attain the order-optimal accuracy $O(\ln^{-1}\frac{1}{\delta})$ of recovering solutions from pointed set for all $0 < p \le 1$. Notice, that studies initiated in [12] were extended in [14] to the more wide class of ill-posed problems (1) with solutions (2) for any $K = 1, 2, ...$ and $p > p_0 > 0$. Herewith the order-optimal accuracy of recovering solutions $O((\underbrace{\ln ... \ln}_{K\text{-times}}\frac{1}{\delta})^{-p})$ was obtained.

Unlike the works described above, in the present paper for regularization of severely ill-posed problems (1) with solutions (2) for $K \ge 1$, and perturbed operators and the right-hand sides iterated Tikhonov's method will be applied, and a regularization parameter will be chosen in accordance with the balancing principle. Subsequently we will demonstrate that the suggested approach for solving (1)-(2), which consists in combination of iterative Tikhonov's method and balancing principle, provides accuracy $O((\underbrace{\ln ... \ln}_{K\text{-times}}\frac{1}{h+\delta})^{-p})$.

We recall that iterated Tikhonov's method consists in a choosing a natural m, initial approximation $x^{h,\delta}_{0,\alpha}$, and consistently computation of elements $x^{h,\delta}_{i,\alpha}$, $i = 1, 2, ..., m$, by the rule

$$x^{h,\delta}_{i,\alpha} = \alpha(A_h^*A_h + \alpha I)^{-1}x^{h,\delta}_{i-1,\alpha} + \alpha(A_h^*A_h + \alpha I)^{-1}A_h^*y_\delta, \tag{14}$$

where $m \ge p_1$ and as the approximate solution we take $x^{h,\delta}_{m,\alpha}$. If $x^{h,\delta}_{0,\alpha} = 0$ then the element $x^{h,\delta}_{m,\alpha}$ can be rewritten in the form of

$$x^{h,\delta}_{m,\alpha} = \sum_{i=1}^{m} \alpha^{i-1}(A_h^*A_h + \alpha I)^{-i}A_h^*y_\delta. \tag{15}$$

Obviously, any numerical realization of the Tikhonov method requires us to curry out all computations with a finite-dimensional approximation $A_{h,n}$ instead of A_h. Thus we assume finite-dimensional approximation $A_{h,n}$ with $\mathrm{rank}(A_{h,n}) = n$ to be chosen such that

$$\|A_h - A_{h,n}\| \leq \varepsilon, \text{ where } \varepsilon = \begin{cases} \delta\rho^{-1} & , \quad 0 < h \leq \delta, \\ h & , \quad h > \delta \end{cases}. \tag{16}$$

Further, along with (15) we will also consider auxiliary elements:

$$x_{m,\alpha} = \sum_{i=1}^{m} \alpha^{i-1}(A^*A + \alpha I)^{-i} A_h^* y, \tag{17}$$

$$x_{m,\alpha,n}^{h} = \sum_{i=1}^{m} \alpha^{i-1}(A_{h,n}^* A_{h,n} + \alpha I)^{-i} A_{h,n}^* y, \tag{18}$$

$$x_{m,\alpha,n}^{h,\delta} = \sum_{i=1}^{m} \alpha^{i-1}(A_{h,n}^* A_{h,n} + \alpha I)^{-i} A_{h,n}^* y_\delta. \tag{19}$$

Recall that generating function of the iterated Tikhonov method has the form (see [15, p.21])

$$g_{m,\alpha}(\lambda) := \sum_{i=1}^{m} \alpha^{i-1}(\alpha + \lambda)^{-i} = \frac{1}{\lambda}(1 - \frac{\alpha^m}{(\alpha+\lambda)^m}), \quad \lambda \neq 0,$$

and satisfies inequality (see [15, p.22])

$$\sup_{0<\lambda<\infty} \sqrt{\lambda} g_{m,\alpha}(\lambda) \leq \sqrt{\frac{m}{\alpha}}.$$

2. Auxiliary statements

We shall later need the following auxiliary results and facts.

Thus, for any linear operators $A, B \in \mathcal{L}(X, Y)$ and natural m the decomposition (see [15, p. 92])

$$A^m - B^m = \sum_{j=0}^{m-1} A^j (A - B) B^{m-j-1} \tag{20}$$

holds true.

Lemma 1. *(see* [15, p. 34]*) If g is bounded, Borel measurable function with respect to the* $[0; M_K]$,
$A \in \mathcal{L}(X, Y)$, $\|A\| \leq M_K$ *then*

$$A^* g(AA^*) = g(A^*A)A^*,$$

$$Ag(A^*A) = g(AA^*)A.$$

In addition, it is well-known that for any bounded linear operator B

$$B(\alpha I + B^*B)^{-1} = (\alpha I + BB^*)^{-1}B,$$
$$\|(\alpha I + B^*B)^{-1}\| \le \alpha^{-1}, \quad \|(\alpha I + B^*B)^{-1}B^*\| \le \frac{1}{2\sqrt{\alpha}}, \tag{21}$$
$$\|B(\alpha I + B^*B)^{-1}B^*\| \le 1$$

hold.

Before proceeding further we establish a nomber of auxiliary assortations which will be need later for analysis of approximating properties of suggesting approch.

Lemma 2. *Let*

$$\|A\| \le M_K, \; M_K = m_K^{1/2}, \; m_k = \begin{cases} e^{-1}, & k = 1, \\ e^{-\frac{1}{m_{k-1}}}, & k = 2, ..., K \end{cases}.$$

Then the following estimate

$$\|x_0 - x_{m,\alpha}\| \le \rho(\underbrace{\ln ... \ln}_{K\text{-}times}\frac{1}{\alpha})^{-p}$$

holds true, where $x_{m,\alpha}$ determined by (17).

Proof. First, we note that

$$\|x_0 - x_{m,\alpha}\| = \|[(\underbrace{\ln ... \ln}_{K\text{-times}}(A^*A)^{-1})^{-p}v -$$
$$\sum_{i=1}^{m} \alpha^{i-1}(A^*A + \alpha I)^{-i}A^*A(\underbrace{\ln ... \ln}_{K\text{-times}}(A^*A)^{-1})^{-p}v]\| \le$$
$$\le \rho\|[I - \sum_{i=1}^{m} \alpha^{i-1}(A^*A + \alpha I)^{-i}A^*A](\underbrace{\ln ... \ln}_{K\text{-times}}(A^*A)^{-1})^{-p}\| \le$$
$$\le \rho \sup_{0<\lambda\le m_K} |[I - \sum_{i=1}^{m} \alpha^{i-1}\frac{\lambda}{(\lambda+\alpha)^i}](\underbrace{\ln ... \ln}_{K\text{-times}}\frac{1}{\lambda})^{-p}| \le$$
$$\le \rho \sup_{0<\lambda\le m_K} |(\frac{\alpha}{\alpha+\lambda})^m(\underbrace{\ln ... \ln}_{K\text{-times}}\frac{1}{\lambda})^{-p}|.$$

To estimate the expression standing under sign of supremum we consider two events:

1) $\lambda \le \alpha$. As function $(\underbrace{\ln ... \ln}_{K\text{-times}}\frac{1}{\lambda})^{-p}$ monotonously decreases for λ, then

$$(\frac{\alpha}{\alpha+\lambda})^m(\underbrace{\ln ... \ln}_{K\text{-times}}\frac{1}{\lambda})^{-p} < (\underbrace{\ln ... \ln}_{K\text{-times}}\frac{1}{\alpha})^{-p}.$$

2) $\lambda \ge \alpha$. We consider the function

$$f(\lambda) = \frac{1}{\lambda^m}(\underbrace{\ln ... \ln}_{K\text{-times}}\frac{1}{\lambda})^{-p}, \quad \lambda \in (0; m_K].$$

It is easy to show that

$$f'(\lambda) = \lambda^{-m-1}(\underbrace{\ln ... \ln}_{K\text{-times}}\frac{1}{\lambda})^{-p-1}(\underbrace{\ln ... \ln}_{(K-1)\text{-times}}\frac{1}{\lambda})^{-1} ... (\ln\frac{1}{\lambda})^{-1}\times$$

$$\times[p - m\underbrace{\ln ... \ln}_{K\text{-times}}\frac{1}{\lambda}\underbrace{\ln ... \ln}_{(K-1)\text{-times}}\frac{1}{\lambda}\cdot ... \cdot \ln\frac{1}{\lambda}].$$

As $f'(\lambda) < 0$ for $p < m$ consequently $f(\lambda)$ monotonously decreases for $p < m, \ \ m > 0$. Thus,

$$f(\lambda) \le f(\alpha) \text{ for } \lambda \ge \alpha \text{ and}$$

$$\left(\frac{\alpha}{\alpha+\lambda}\right)^m(\underbrace{\ln ... \ln}_{K\text{-times}}\frac{1}{\lambda})^{-p} = \left(\frac{\alpha}{\alpha+\lambda}\right)^m\cdot\lambda^m\cdot\frac{1}{\lambda^m}(\underbrace{\ln ... \ln}_{K\text{-times}}\frac{1}{\lambda})^{-p} \le$$

$$\le \frac{\lambda^m}{(\alpha+\lambda)^m}(\underbrace{\ln ... \ln}_{K\text{-times}}\frac{1}{\alpha})^{-p} \le (\underbrace{\ln ... \ln}_{K\text{-times}}\frac{1}{\alpha})^{-p}.$$

Herewith, in general case we have

$$\|x_0 - x_{m,\alpha}\| \le \rho(\underbrace{\ln ... \ln}_{K\text{-times}}\frac{1}{\alpha})^{-p},$$

hence, the proof is completed. □

Lemma 3. *Let*

$$\|A\| \le M_K, \ M_K = m_K^{1/2}, \ m_k = \begin{cases} e^{-1}, & k = 1, \\ e^{-\frac{1}{m_{k-1}}}, & k = 2, ..., K \end{cases}.$$

Then the estimate

$$\|x_{m,\alpha} - x^h_{m,\alpha,n}\| \le \frac{\rho(m+\sqrt{m})(h+\varepsilon)}{\sqrt{\alpha}}$$

holds true, where $x_{m,\alpha}$ and $x^h_{m,\alpha,n}$ determened by (17), (18) correspondently.

Proof. Clearly, that

$$\|x_0\| = \|(\underbrace{\ln ... \ln}_{K\text{-times}}(A^*A)^{-1})^{-p}v\| \le$$

$$\le \rho \sup_{0<\lambda\le m_K}|(\underbrace{\ln ... \ln}_{K\text{-times}}\frac{1}{\lambda})^{-p}| \le \rho.$$

$$\|A - A_{h,n}\| \le \|A - A_h\| + \|A_h - A_{h,n}\| \le h + \varepsilon.$$

Further, we estimate the norm

$$\|x_{m,\alpha} - x^h_{m,\alpha,n}\| = \|g_{m,\alpha}(A^*A)A^*y - g_{m,\alpha}(A^*_{h,n}A_{h,n})A^*_{h,n}y\| =$$

$$= \|g_{m,\alpha}(A^*A)A^*Ax_0 - g_{m,\alpha}(A^*_{h,n}A_{h,n})A^*_{h,n}Ax_0\| \le$$

$$\le \rho\|g_{m,\alpha}(A^*A)A^*A - g_{m,\alpha}(A^*_{h,n}A_{h,n})A^*_{h,n}A\|.$$

We consider the expression standing under norm's sign:

$$g_{m,\alpha}(A^*A)A^*A - g_{m,\alpha}(A^*_{h,n}A_{h,n})A^*_{h,n}A =$$

$$= g_{m,\alpha}(A^*A)A^*A - g_{m,\alpha}(A^*_{h,n}A_{h,n})A^*_{h,n}A_{h,n}+$$

$$+g_{m,\alpha}(A_{h,n}^*A_{h,n})A_{h,n}^*A_{h,n} - g_{m,\alpha}(A_{h,n}^*A_{h,n})A_{h,n}^*A = I_1 + I_2,$$

where

$$I_1 := g_{m,\alpha}(A^*A)A^*A - g_{m,\alpha}(A_{h,n}^*A_{h,n})A_{h,n}^*A_{h,n},$$
$$I_2 := g_{m,\alpha}(A_{h,n}^*A_{h,n})A_{h,n}^*A_{h,n} - g_{m,\alpha}(A_{h,n}^*A_{h,n})A_{h,n}^*A.$$

Now we estimate each of summands I_1, I_2.

Thus,

$$I_1 = (I - \alpha^m(\alpha I + A^*A)^{-m} - (I - \alpha^m(\alpha I + A_{h,n}^*A_{h,n})^{-m}) =$$
$$= \alpha^m[(\alpha I + A_{h,n}^*A_{h,n})^{-m} - (\alpha I + A^*A)^{-m}].$$

We apply the formula (20) to expression standing in braces:

$$I_1 = \alpha^m \sum_{j=0}^{m-1}(\alpha I + A_{h,n}^*A_{h,n})^{-j} \cdot [(\alpha I + A_{h,n}^*A_{h,n})^{-1} - (\alpha I + A^*A)^{-1}]\times$$
$$\times(\alpha I + A^*A)^{-m+j+1} = \alpha^m \sum_{j=0}^{m-1}(\alpha I + A_{h,n}^*A_{h,n})^{-j-1}(A^*A - A_{h,n}^*A_{h,n})\times$$
$$\times(\alpha I + A^*A)^{-m+j} = \alpha^m \sum_{j=0}^{m-1}(\alpha I + A_{h,n}^*A_{h,n})^{-j-1}(A^* - A_{h,n}^*)A\times$$
$$\times(\alpha I + A^*A)^{-m+j} + \alpha^m \sum_{j=0}^{m-1}(\alpha I + A_{h,n}^*A_{h,n})^{-j-1}A_{h,n}^*(A - A_{h,n})\times$$
$$\times(\alpha I + A^*A)^{-m+j}.$$

Whence by Lemma 1 and estimates (21) we obtain

$$\|I_1\| \leq \sum_{j=0}^{m-1}[\|(\alpha I + A_{h,n}^*A_{h,n})^{-1}\|^{j-1}\|(\alpha I + AA^*)^{-m+j}A\| +$$
$$+\|(\alpha I + A_{h,n}^*A_{h,n})^{-j-1}A_{h,n}^*\| \cdot \|(\alpha I + A^*A)^{-1}\|^{m-j}]\times$$
$$\times\alpha^m\|A - A_{h,n}\| = \sum_{j=0}^{m-1}[\alpha^{-j-1}\|(\alpha I + A^*A)^{-m+j-1}\|\times$$
$$\times\|(\alpha I + A^*A)^{-1}A\| + \|(\alpha I + A_{h,n}^*A_{h,n})^{-1}\|^{j}\times$$
$$\times\|(\alpha I + A_{h,n}^*A_{h,n})^{-1}A_{h,n}^*\|\alpha^{-m+j}]\alpha^m\|A - A_{h,n}\| \leq$$
$$\leq \sum_{j=0}^{m-1}[\alpha^{-j-1} \cdot \alpha^{-m+j+1} \cdot \frac{1}{2\sqrt{\alpha}} + \alpha^{-j}\frac{1}{2\sqrt{\alpha}} \cdot \alpha^{-m+j}]\alpha^m\times$$
$$\times\|A - A_{h,n}\| = \sum_{j=0}^{m-1}\frac{1}{\sqrt{\alpha}}\|A - A_{h,n}\| \leq \frac{m}{\sqrt{\alpha}}(h + \varepsilon). \tag{22}$$

Then, due to (21) we find

$$\|I_2\| = \|g_{m,\alpha}(A_{h,n}^*A_{h,n})A_{h,n}^*(A_{h,n} - A)\| \leq$$

$$\leq \sup_{0<\lambda\leq m_K} |\sqrt{\lambda} g_{m,\alpha}(\lambda)| \cdot \|A_{h,n} - A\| \leq \frac{\sqrt{m}}{\sqrt{\alpha}}(h+\varepsilon). \tag{23}$$

Summarizing relations (22) and (23) we finally obtain

$$\|x_{m,\alpha,n} - x^h_{m,\alpha,n}\| \leq \rho\left(\frac{m}{\sqrt{\alpha}}(h+\varepsilon) + \frac{\sqrt{m}}{\sqrt{\alpha}}(h+\varepsilon)\right) =$$

$$\frac{\rho(m+\sqrt{m})(h+\varepsilon)}{\sqrt{\alpha}}.$$

Thus, Lemma is proved. □

Theorem 1. *Let*

$$\|A\| \leq M_K,\ M_K = m_K^{1/2},\ m_k = \begin{cases} e^{-1}, & k=1, \\ e^{-\frac{1}{m_{k-1}}}, & k=2,...,K \end{cases}$$

and $x_0 = A^{-1}y \in M^K_{p,\rho}(A)$.

Then the estimate

$$\|x_0 - x^{h,\delta}_{m,\alpha,n}\| \leq \rho((\underbrace{\ln ... \ln}_{K\text{-}times}\frac{1}{\alpha})^{-p} + \frac{\rho(m+\sqrt{m})(h+\varepsilon)}{\sqrt{\alpha}} + \frac{\delta\sqrt{m}}{\sqrt{\alpha}} \tag{24}$$

holds true, where $x^{h,\delta}_{m,\alpha,n}$ *is approximate solution determined by (19).*

Proof. Using triangle's rule we obtain

$$\|x_0 - x^{h,\delta}_{m,\alpha,n}\| \leq \|x_0 - x_{m,\alpha}\| + \|x_{m,\alpha} - x^h_{m,\alpha,n}\| + \|x^h_{m,\alpha,n} - x^{h,\delta}_{m,\alpha,n}\|.$$

We consider last summand:

$$\|x^h_{m,\alpha,n} - x^{h,\delta}_{m,\alpha,n}\| = \|g_{m,\alpha}(A^*_{h,n}A_{h,n})A^*_{h,n}y-$$

$$-g_{m,\alpha}(A^*_{h,n}A_{h,n})A^*_{h,n}y_\delta\| \leq \|g_{m,\alpha}(A^*_{h,n}A_{h,n})A^*_{h,n}\|\times$$

$$\times\|y-y_\delta\| \leq \sup_{0<\lambda\leq m_K}(\sqrt{\lambda}g_{m,\alpha}(\lambda)).$$

Thus, by inequality (19) we find

$$\|x^h_{m,\alpha,n} - x^{h,\delta}_{m,\alpha,n}\| \leq \frac{\delta\sqrt{m}}{\sqrt{\alpha}}. \tag{25}$$

And finally summarizing Lemma 2, Lemma 3 and relation (25) we obtain the assertion of Theorem. □

3. The balancing principle

The balancing principle consists in choosing a value of regularization parameter α such that to balance two functions which give accuracy estimation. In our case, these functions are represented by (see (24))

$$\Phi(\alpha) := \rho(\underbrace{\ln ... \ln}_{K\text{-times}}\frac{1}{\alpha})^{-p},$$

$$\Psi(\alpha) := \frac{\rho(m+\sqrt{m})(h+\varepsilon)+\delta\sqrt{m}}{\sqrt{\alpha}}.$$

Taking into account, that (see (16))

$$\varepsilon = \begin{cases} \delta\rho^{-1} & ,0 < h \le \delta, \\ h, & h > \delta \end{cases}$$

we can represent function $\Psi(\alpha)$ as

$$\Psi(\alpha) = \frac{\rho c_1 h + c_2\delta}{\sqrt{\alpha}},$$

where

$$c_1 = \begin{cases} m + \sqrt{m} & ,0 < h \le \delta, \\ 2(m + \sqrt{m}) & ,h > \delta \end{cases}, \quad c_2 = \begin{cases} m + 2\sqrt{m} & ,0 < h \le \delta, \\ \sqrt{m} & ,h > \delta \end{cases}.$$

Thus, we can rewrite (24) in the form

$$\|x_0 - x^{h,\delta}_{m,\alpha,n}\| \le \Phi(\alpha) + \Psi(\alpha). \tag{26}$$

Since $\phi(t) = (\underbrace{\ln ... \ln}_{K\text{-times}} \frac{1}{t})^{-p}$ is monotonously increasing function then for increasing α the function $\Phi(\alpha)$ increases. By other side, the function $\Psi(\alpha)$ is monotonously decreasing. According to behavior of functions Φ and Ψ (namely, their monotonicity and concavity) to choose a value of regularization parameter $\alpha = \widehat{\alpha}$ minimizing right-hand side of (26) we will balancing values $\Phi(\alpha)$ and $\Psi(\alpha)$, i.e.

$$\Phi(\widehat{\alpha}) = \Psi(\widehat{\alpha})$$

And, hence

$$\|x_0 - x^{h,\delta}_{m,\alpha,n}\| \le 2\Phi(\widehat{\alpha}).$$

But, since function ϕ is unknown (namely, parameter p is unknown), then such a priori choice of the best value $\widehat{\alpha}$ is impossible. Therefore in considering situation we need to make use of some a posteriori choice of α. For further studing we choice the balancing principle as such rule.

Let describe this principle according to our problem. Consider two sets

$$\Delta_N = \{\alpha_i = (q^2)^i\alpha_0, i = 1, 2, ..., N\}, \quad q > 1,$$
$$\alpha_0 = n(h + \delta)^2, \quad N : \alpha_N \asymp 1,$$

and

$$M^+(\Delta_N) = \{\alpha_i \in \Delta_N : \|x^{h,\delta}_{m,\alpha_i,n} - x^{h,\delta}_{m,\alpha_j,n}\| \le \le 4\Psi(\alpha_j), \quad j = 1, 2, ..., i\}. \tag{27}$$

Within the framework of balancing principle we take

$$\alpha = \alpha_+ := \max\{\alpha \in M^+(\Delta_N)\}. \tag{28}$$

as value of regularization parameter Moreover, consider auxiliary set

$$M(\Delta_N) := \{\alpha_i \in \Delta_N : \Phi(\alpha_i) \le \Psi(\alpha_i)\}$$

and auxiliary value

$$\alpha_* := \max\{\alpha \in M(\Delta_N)\}.$$

Without loss of generally we assume that

$$M(\Delta_N) \ne \emptyset \quad \text{and} \quad \Delta_N \setminus M(\Delta_N) \ne \emptyset.$$

And finally we can estimate closeness of exact and approximate solutions for value of regularization parameter $\alpha = \alpha_+$.

4. The main results

Theorem 2. *Assume that the regularization parameter is choosing according to (28). Then for any* $x_0 \in M^K_{p,\rho}(A)$, $0 < p \le p_1$, $K = 1, 2, ...,$ *the following estimate*

$$\|x_0 - x^{h,\delta}_{m,\alpha_+,n}\| \le 6q\rho(\underbrace{\ln ... \ln}_{K\text{-}times}\frac{1}{\widehat{\alpha}})^{-p}$$

is valid.

Proof. First, we show that $\alpha_* \le \alpha_+$. Due to (26), behavior of functions $\Phi(\alpha), \Psi(\alpha)$ and definition of the set $M(\Delta_N)$, for any $\alpha_j < \alpha_*$ we have

$$\|x^{h,\delta}_{m,\alpha_*,n} - x^{h,\delta}_{m,\alpha_j,n}\| \le \|x_0 - x^{h,\delta}_{m,\alpha_*,n}\| + \|x_0 - x^{h,\delta}_{m,\alpha_j,n}\| \le$$

$$\le \Phi(\alpha_*) + \Psi(\alpha_*) + \Phi(\alpha_j) + \Psi(\alpha_j) \le$$

$$\le 2\Phi(\alpha_*) + \Psi(\alpha_*) + \Psi(\alpha_j) \le$$

$$\le 3\Psi(\alpha_*) + \Psi(\alpha_j) \le 4\Psi(\alpha_j).$$

Thus, $\alpha_* \in M^+(\Delta_N)$. And, hence the inequality $\alpha_* \le \alpha_+$ holds true. Further, according to (26) for $\alpha = \alpha_*$ and also definition of sets $M^+(\Delta_N)$ and $M(\Delta_N)$ we have

$$\|x_0 - x^{h,\delta}_{m,\alpha_+,n}\| \le \|x_0 - x^{h,\delta}_{m,\alpha_*,n}\| + \|x^{h,\delta}_{m,\alpha_*,n} - x^{h,\delta}_{m,\alpha_+,n}\| \le 6\Psi(\alpha_*). \tag{29}$$

It is easy to see that from definition of function Ψ it follows

$$\Psi(q^2\alpha_*) = \frac{\rho c_1 h + c_2\delta}{\sqrt{q^2\alpha_*}} = \frac{1}{q}\frac{\rho c_1 h + c_2\delta}{\sqrt{\alpha_*}} = \frac{1}{q}\Psi(\alpha_*). \tag{30}$$

By other side, obviously $\alpha_* \le \widehat{\alpha} \le q^2\alpha_*$. According to (29) and (30) we obtain

$$\|x_0 - x^{h,\delta}_{m,\alpha_+,n}\| \le 6q\Psi(q^2\alpha_*) \le 6q\Psi(\widehat{\alpha}) =$$

$$= 6q\Phi(\widehat{\alpha}) = 6q\rho(\underbrace{\ln ... \ln}_{K\text{-times}}\frac{1}{\widehat{\alpha}})^{-p}.$$

Proof of Theorem 2 is completed. □

Theorem 3. *Let* $x_0 \in M^1_{p,\rho}(A)$, $0 < p \le p_1$, *and the condition of Theorem 2 is satisfies. Then for any* $\delta, h > 0$ *the estimate*

$$\|x_0 - x^{h,\delta}_{m,\alpha_+,n}\| \le c_p\left(\ln\frac{\rho}{\rho c_1 h + c_2\delta}\right)^{-p}$$

holds true, where $c_p = 6q\rho\left(\frac{2p+1}{2}\right)^p$.

Proof. According to $\Phi(\widehat{\alpha}) = \Psi(\widehat{\alpha})$ we find

$$\rho \ln^{-p} \frac{1}{\widehat{\alpha}} = \frac{\rho c_1 h + c_2 \delta}{\sqrt{\widehat{\alpha}}}.$$

Then

$$\widehat{\alpha} = \left(\frac{\rho c_1 h + c_2 \delta}{\rho} \right)^2 \ln^{2p} \frac{1}{\widehat{\alpha}}.$$

As for any $x > 0$ the relation $\ln x < x$ is valid, then

$$\widehat{\alpha} \leq \left(\frac{\rho c_1 h + c_2 \delta}{\rho} \right)^2 \left(\frac{1}{\widehat{\alpha}} \right)^{2p},$$

$$\widehat{\alpha} \leq \left(\frac{\rho c_1 h + c_2 \delta}{\rho} \right)^{\frac{2}{2p+1}}.$$

Hence, due to Theorem 2 we have

$$\|x_0 - x^{h,\delta}_{m,\alpha_+,n}\| \leq 6q\rho \left(\ln \left(\frac{\rho}{\rho c_1 h + c_2 \delta} \right)^{\frac{2}{2p+1}} \right)^{-p} =$$

$$= 6q\rho \left(\frac{2p+1}{2} \right)^p \left(\ln \frac{\rho}{\rho c_1 h + c_2 \delta} \right)^{-p}.$$

Denoting $c_p = 6q\rho \left(\frac{2p+1}{2} \right)^p$, we obtain the assertion of Theorem. □

Remark 2. *In the case $p_1 = 1$ and $h = 0$ the result of Theorem 3 was obtained earlier in* [13]. *Thus, Theorem 3 generalizes result of* [13] *for any $p_1 > 0$ and $h > 0$.*

Theorem 4. *Let $x_0 \in M^K_{p,\rho}(A)$, $\;0 < p \leq p_1$, $\;K = 2, 3, ...$ and the condition of Theorem 2 is fulfiled. Then, for sufficiently small $h, \delta > 0$ the estimate*

$$\|x_0 - x^{h,\delta}_{m,\alpha_+,n}\| \leq c_p \left[\underbrace{\ln ... \ln}_{K\text{-times}} \frac{\rho}{\rho c_1 h + c_2 \delta} \right]^{-p}$$

holds true, where $c_p = 2^p 6 q\rho$.

Proof. $\Phi(\widehat{\alpha}) = \Psi(\widehat{\alpha})$, then

$$\rho \left(\underbrace{\ln ... \ln}_{K\text{-times}} \frac{1}{\widehat{\alpha}} \right)^{-p} = \frac{\rho c_1 h + c_2 \delta}{\sqrt{\widehat{\alpha}}},$$

$$\widehat{\alpha} = \left(\frac{\rho c_1 h + c_2 \delta}{\rho} \right)^2 \left(\underbrace{\ln ... \ln}_{K\text{-times}} \frac{1}{\widehat{\alpha}} \right)^{2p}.$$

As for any $x > \underbrace{\exp(\exp(...(\exp}_{K\text{-times}}(1))))$ the inequality $\underbrace{\ln ... \ln}_{K\text{-times}} x < x$ is valid, then

$$\widehat{\alpha} \leq \left(\frac{\rho c_1 h + c_2 \delta}{\rho} \right)^{\frac{2}{2p+1}},$$

by that we have found the upper estimate for value of regularization parameter which theoretically minimizing accuracy.

Thus, by Theorem 2 we obtain

$$\|x_0 - x^{h,\delta}_{m,\alpha_+,n}\| \le 6q\rho \left(\underbrace{\ln \ldots \ln}_{K\text{-times}} \frac{1}{\widehat{\alpha}}\right)^{-p} \le$$

$$\le 6q\rho \left[\underbrace{\ln \ldots \ln}_{K\text{-times}} \left(\frac{\rho}{\rho c_1 h + c_2\delta}\right)^{\frac{2}{2p+1}}\right]^{-p}.$$

Further, we will find upper-bound estimate for

$$\left[\underbrace{\ln \ldots \ln}_{K\text{-times}} \left(\frac{\rho}{\rho c_1 h + c_2\delta}\right)^{\frac{2}{2p+1}}\right]^{-p}.$$

First, let $K = 2$, i.e. we will find upper-bound estimate for

$$\left[\ln\ln \left(\frac{\rho}{\rho c_1 h + c_2\delta}\right)^{\frac{2}{2p+1}}\right]^{-p}.$$

Obviously, that for any fixed p, $0 < p < \infty$, there exist such $h_0, \delta_0 > 0$ that for all $0 < h \le h_0$ and $0 < \delta \le \delta_0$ the inequality

$$\left(\frac{2p+1}{2}\right)^2 \le \ln \frac{\rho}{c_1\rho h + c_2\delta}$$

is fulfiled. Whence, from monotonicity of ln it follows

$$\ln \left(\frac{2p+1}{2}\right)^2 \le \ln\ln \frac{\rho}{c_1\rho h + c_2\delta},$$

$$\ln \left(\frac{2p+1}{2}\right) \le \frac{1}{2} \ln\ln \frac{\rho}{c_1\rho h + c_2\delta}.$$

$$\ln\ln \left(\frac{\rho}{\rho c_1 h + c_2\delta}\right)^{\frac{2}{2p+1}} = \ln\ln \frac{\rho}{\rho c_1 h + c_2\delta} - \ln \frac{2p+1}{2} \ge$$

$$\ge \frac{1}{2} \ln\ln \frac{\rho}{\rho c_1 h + c_2\delta}.$$

Hence,

$$\left[\ln\ln \left(\frac{\rho}{\rho c_1 h + c_2\delta}\right)^{\frac{2}{2p+1}}\right]^{-p} \le 2^p \left[\ln\ln \frac{\rho}{\rho c_1 h + c_2\delta}\right]^{-p}.$$

Further, in case of arbitrary $K > 2$ we will show, that for sufficiently small $h, \delta > 0$ the inequality

$$\underbrace{\ln \ldots \ln}_{K\text{-times}} \left(\frac{\rho}{\rho c_1 h + c_2\delta}\right)^{\frac{2}{2p+1}} \ge \frac{1}{2} \underbrace{\ln \ldots \ln}_{K\text{-times}} \left(\frac{\rho}{\rho c_1 h + c_2\delta}\right) \tag{31}$$

is fulfiled. For that reason we will carry out the proof by induction. Thus, for $K = 2$ the inequality (31) was proof earlier. Let assume now, that inequality (31) is fulfiled for $K-1, \quad K \geq 3$, i.e.

$$\underbrace{\ln \ldots \ln}_{(K-1)\text{-times}} \left(\frac{\rho}{\rho c_1 h + c_2\delta}\right)^{\frac{2}{2p+1}} \geq \frac{1}{2} \underbrace{\ln \ldots \ln}_{(K-1)\text{-times}} \left(\frac{\rho}{\rho c_1 h + c_2\delta}\right).$$

Then the relation

$$\underbrace{\ln \ldots \ln}_{K\text{-times}} \left(\frac{\rho}{\rho c_1 h + c_2\delta}\right)^{\frac{2}{2p+1}} \geq \ln \left[\frac{1}{2} \underbrace{\ln \ldots \ln}_{(K-1)\text{-times}} \frac{\rho}{\rho c_1 h + c_2\delta}\right]$$

holds true.

Further,

$$\ln \left[\frac{1}{2} \underbrace{\ln \ldots \ln}_{(K-1)\text{-times}} \frac{\rho}{\rho c_1 h + c_2\delta}\right] - \frac{1}{2} \underbrace{\ln \ldots \ln}_{K\text{-times}} \frac{\rho}{\rho c_1 h + c_2\delta} =$$

$$= \ln \left[\frac{\frac{1}{2} \underbrace{\ln \ldots \ln}_{(K-1)\text{-times}} \frac{\rho}{\rho c_1 h + c_2\delta}}{\left(\underbrace{\ln \ldots \ln}_{(K-1)\text{-times}} \frac{\rho}{\rho c_1 h + c_2\delta}\right)^{1/2}}\right] =$$

$$= \ln \left[\frac{1}{2} \left(\underbrace{\ln \ldots \ln}_{(K-1)\text{-times}} \frac{\rho}{\rho c_1 h + c_2\delta}\right)^{1/2}\right] > 0.$$

Hence,

$$\ln \left[\frac{1}{2} \underbrace{\ln \ldots \ln}_{(K-1)\text{-times}} \frac{\rho}{\rho c_1 h + c_2\delta}\right] \geq \frac{1}{2} \underbrace{\ln \ldots \ln}_{K\text{-times}} \frac{\rho}{\rho c_1 h + c_2\delta}.$$

Thus, inequality (31) holds true, then

$$\left[\underbrace{\ln \ldots \ln}_{K\text{-times}} \left(\frac{\rho}{\rho c_1 h + c_2\delta}\right)^{\frac{2}{2p+1}}\right]^{-p} \leq \left[\frac{1}{2} \underbrace{\ln \ldots \ln}_{K\text{-times}} \frac{\rho}{\rho c_1 h + c_2\delta}\right]^{-p} =$$

$$= 2^p \left[\underbrace{\ln \ldots \ln}_{K\text{-times}} \frac{\rho}{\rho c_1 h + c_2\delta}\right]^{-p}.$$

And it means, that due to Theorem 2

$$\|x_0 - x^{h,\delta}_{m,\alpha_+,n}\| \leq 6q\rho 2^p \left[\underbrace{\ln \ldots \ln}_{K\text{-times}} \frac{\rho}{\rho c_1 h + c_2\delta}\right]^{-p}.$$

Denoting $c_p = 2^p 6q\rho$ we complete the proof of Theorem. □

Remark 3. *In* [14] *for solving severely ill-posed problems (1)-(2) with perturbed right-hand sides* y_δ *and exactly given operators* A *a combination of standard Tikhonov regularization with Morozov's discrepancy principle was considered. This approach allows to achieve the accuracy* $O((\underbrace{\ln\ldots\ln}_{K\text{-}times}\frac{1}{\delta})^{-p})$ *among the set* $M^K_{p,\rho}(A), K\in\mathbb{N}$, *of solutions. Moreover, in* [14] *the lower bound* p_0 *of possible values for parameter* p $(p>p_0>0)$ *was used. By other side, in Theorem 4 was shown that the strategy (14), (27), (28) of solving severely ill-posed problems guarantees the same order of accuracy on the same set* $M^K_{p,\rho}(A)$ *of solutions. But in this case the upper bound of possible values for* p $(0<p\le p_1)$ *is used.*

Remark 4. *In* [14] *for solving problems (1) with perturbed right-hand sides only and with desired solutions from the set (2) for arbitrary* $K\in\mathbb{N}$ *was shown*

$$e(M^K_{p,\rho}(A),\delta)=O((\underbrace{\ln\ldots\ln}_{K\text{-}times}\frac{1}{\delta})^{-p}),$$

where

$$e(M^K_{p,\rho}(A),\delta):=\inf_{S:Y\to X}\quad\sup_{x_0\in M^K_{p,\rho}(A)}\quad\sup_{y_\delta\in Y:\|y-y_\delta\|\le\delta}\quad\|x_0-Sy_\delta\|.$$

Hence, $e(M^K_{p,\rho}(A),\delta)$ *determines the least possible accuracy of solving (1) on the set (2) among all approximate methods* $S:Y\to X$ *constructed on perturbed data* y_δ. *It means (see Theorem (4.1)* [14]*) that the value* $O((\underbrace{\ln\ldots\ln}_{K\text{-}times}\frac{1}{\delta})^{-p})$ *gives the order-optimal accuracy.*

On the other hand, it follows from Theorems 3, 4 when $h=0$ *the received accuracy of approximate solving (1) has the representation* $O((\underbrace{\ln\ldots\ln}_{K\text{-}times}\frac{1}{\delta})^{-p})$. *This, in its turn, means that in the case of exactly given operator* A *the suggested approach also provides the order-optimal accuracy of solving severely ill-posed problems.*

Bibliography

1. Bakushinski A. B. A general method of constructing regularizing algorithms for a linear ill-posed equationin Hilbert space / A. B. Bakushinski // Zh. Mat. Mat. Fiz.– 1967.– Vol. 7, № 3.– P. 672-677.
2. Bakushinski A. B. Iterative methods for solution of ill-posed operator equations with smooth operators / A. B. Bakushinski, M. Yu. Kokurin.– Moscow, Editorial URSS, 2002 (in Russian).
3. Bao G. Modeling and optimal design of diffractive optical structures / G. Bao, D. Dobson // Surv. Math. Ind.– 1998.– Vol. 8, № 1.– P. 37-62.
4. Freeden W. Regularization wavelets and multiresolution / W. Freeden, F. Schneider // Inverse Problems.– 1998.– Vol. 14, № 2.– P. 225-243.
5. Hettlich F. Schifter's theorem in inverse scattering for periodic structures / F. Hettlich, A. Kirsch // Inverse problems.– 1997.– Vol. 13, № 2.– P. 351-361.
6. Hohage T. Regularization of exponentially ill-posed problems / T. Hohage // Numer. Funct. Anal. Optim.– 2000.– Vol. 21, № 3-4.– P. 439-464.

7. Kokurin M.Yu. On necessary conditions for the qualified convergence of methods for solving linear ill-posed problems / M.Yu. Kokurin, N.A. Yusupova // Izv. Vyssh. Uchebn. Zaved. Mat.– 2001.– Vol. 45, №. 2, P. 39-47 (in Russian).
8. Mair B. A. Tikhonov regularization for finitely and infinitely smoothing operators / B. A. Mair // SIAM J. Math. Anal.– 1994.– Vol. 25, № 1.– P. 135-147.
9. Mathe P. Regularization of some linearill-posed problems with discretized random noisy data / P. Mathe, S. Pereverzev // Mathematics of Computation.– 2006.– Vol. 75, № 256.– P. 1913-1929.
10. Plato R. On the regularization of projectionmethods for solving ill-posed problems / R. Plato, G. Vainikko // Numer. Math.– 1990.– Vol. 57, № 1.– P. 63-79.
11. Rummel R. Gravity field determination from satellite gradiometry / R. Rummel, O. L. Colombo // Bull. Geod.– 1985.– Vol. 59, № 3.– P. 233-246.
12. Schock E. Morozov's discrepancy principle for Tikhonov regularization of severely ill-posed problems in finite-dimensional subspaces / E. Schock, S. V. Pereverzev // Numer. Funct. Anal. Optim.– 2000.– Vol. 21, №. 7-8.– P. 901-916.
13. Solodky S. G. On the a posteriori choice of regularization parametr in the solution of severelly ill-posed problems / S. G. Solodky, A. V. Grushevaya // Journal of Mathematical Science.– 2012.– Vol. 181, № 1.– P. 98-105.
14. Solodky S. G. Unsaturable methods for solving severely ill-posed problems / S. G. Solodky, A. V. Mosentsova // Int. J. Comput. Sci. Math.– 2009.– Vol. 2, № 3.– P. 229-242.
15. Vainikko G. M. Projection method and self-regularization in ill-posed problems / G. M. Vainikko, V. A. Khyamarik // Izv. Vyssh. Uchebn. Zaved. Mat.– 1985.– № 10.– P. 3-17 (in Russian).

Ganna Myleiko, Sergei Solodky,
Institute of Mathematics, National Academy of Sciences,
3, Tereschenkivska Str., Kiev, 01601, Ukraine

Received 23.05.2012

Журнал обчислювальної та прикладної математики | **2012, № 3 (109)** с. 89–104 | Journal of Numerical & Applied Mathematics

UDC 517.958+519.6

STREAMLINE DIFFUSION SCHEMES FOR SOLVING A NONLINEAR HYPERBOLIC BOUNDARY VALUE PROBLEM

Davood Rostamy, Fatemeh Zabihi

Резюме. В роботі вивчається метод скінченних елементів для розв'язування нелінійної гіперболічної крайової задачі. З'ясовано питання існування і єдиності розв'язку, а також оцінено апріорну та апостеріорну похибки. Отримано оцінку стійкості і оптимальні порядки збіжності, показано апріорну оцінку $O(h^{k+1/2})$, де h – крок сітки і k – степінь кусково-поліноміальних функцій на скінченних елементах, в областях, де точний розв'язок є гладкий або негладкий. Для пропонованого методу наведено результати чисельних експериментів.

Abstract. In this paper we study the streamline diffusion finite element method for treating a nonlinear hyperbolic boundary value problem. The existence and uniqueness are discussed. Also, a priori and a posteriori errors are estimated for this problem. We derive the stability estimate and optimal convergence rates, showing an a priori error estimate of order $\mathcal{O}(h^{k+1/2})$ in domains where the exact solution is smooth or non-smooth; here h is the mesh width and k is the degree of the piecewise polynomial functions spanning the finite element subspaces. Also, some numerical illustrations are given for the presented method.

AMS Subject Classification: 65M12, 65M15, 65M60, 82D10, 35L80

1. Introduction

In this paper we consider the following wave equation:

$$u_{tt} - u_{xx} = \lambda F(x,t,u), \quad (x,t) \in \Omega, \tag{1}$$

$$\alpha u(t,t) - \beta \frac{\partial u}{\partial n_1}(t,t) = \alpha u(1+t, 1-t) + \beta \frac{\partial u}{\partial n_2}(1+t,1-t), \quad 0 \le t \le 1, \tag{2}$$

$$u(x,0) = 0, \qquad 0 \le x \le 2. \tag{3}$$

Where Ω is as follows:

$$\Omega = \{(x,t) : 0 \le t \le 1,\ t \le x \le 2-t\}$$

and the parameters $\lambda, \alpha, \beta \in \mathbb{R}$ such that $\alpha^2 + \beta^2 \neq 0$. The two vectors n_1 and n_2 are the exterior unit normals and $\frac{\partial u}{\partial n_1}$, $\frac{\partial u}{\partial n_2}$ are the normal derivatives. Also, $F(x,t,u) \ge 0$ and $\frac{\partial F(x,t,u)}{\partial u}$ are arbitrary continuous in Ω. The above boundary value problem for mass-spring system has an analog the continuum case which was first formulated [21, 34] as above (see also ([23, 26, 27, 28])). Our problem is a generalization of the problems studied by Kalmenov [21], and [26, 27, 28, 38]. The purpose of this paper is to present extension of the *streamline diffusion*

[†]*Key words.* Streamline diffusion method, hyperbolic problems, wave equations, error estimate, finite element.

(Sd) method to a nonlinear mass-spring system. The mathematical study of the mass spring system with this triangle domain has been considered by several authors in various settings (see [11, 22, 24, 37, 39]). One of the applications of mass spring systems to arch structure railways and long bridge-like structures reduces the dynamic and static loads due to train. Also, we can see this system to simulate facial soft tissue of great interest to many medical forms and make visible to applications([18, 20]).

Streamline diffusion ideas carry out slightly better than the different finite element methods for smooth solutions and non-smooth solutions of the first order hyperbolic problems ([32, 34, 35]) which both is higher order accurate and has good stability properties (see [2, 3, 5, 13, 14, 15, 16, 17, 19, 25, 31]). Due to the fact that the added diffusion removes oscillations near boundary layers([4, 6, 7, 8, 9, 12]). Hughes and Brooks [25] introduced this idea in the case of stationary problems. The mathematical analysis of this method for linear problems, together with extensions to time-dependent problems using space-time elements, was started in Johnson and Navert [31] and was continued in [29, 30, 33, 36]. In this paper we shall go into the details for the nonlinear hyperbolic problem and a new version of Sd method for solving the problem is given. The remaining structure of this article is organized as follows:
The uniqueness of the problem is discussed in section 2. In Section 3, we present and analyze the Sd method. In Sect. 4, by using the Sd method, we investigate stability and obtain an a priori error estimations for this system. A posteriori error estimations are given in sections 5, 6 and 7. Finally, in Sect. 8 the paper would be completed by the inclusion of numerical results to provide experimental support for the theoretical results and show how the method performs in practice.

2. Existence and Uniqueness of the solution

In the following propositions, it is shown that there is a unique solution for (1)-(3) in linear and nonlinear form for $F(x,t,v)$ in Sobolev space ([1]) of $W_2^1(\Omega)\cap W_2^1(\partial\Omega)\cap C(\overline{\Omega})$. In [28] we observe that the linear problem is considered and in [42] existence theorems for some nonlinear hyperbolic equations are given, but in this section the uniqueness of nonlinear form is studied.

Proposition 1. *For $k = 0, 1, 2, ...$ given $\lambda, \alpha, \beta \in \mathbb{R}$ and $F \in H^k(\Omega)$, problem of (1)-(3) has a unique solution in the Hilbert space $u \in H^{k+2}(\Omega)$.*

Proof. We extend the proof of theorem's Iraniparst (see [28]) and we use some propositions and lemmas in [42] (see 2.3 and 4.3). We influence the change of variables $X = x - t$ and $Y = x + t$ into (1)-(3). Hence, we have

$$V_{XY} = \gamma\widehat{F}(X,Y,V(X,Y)), \quad (X,Y)\in\Omega', \tag{4}$$

$$\Omega' = \{(X,Y) : 0 \le Y \le 2, \quad 0 \le X \le Y\}$$

$$\alpha V(0,Y) + \beta V_X(0,Y) = \alpha V(Y,2) + \beta V_Y(Y,2), \quad 0 \le Y \le 2$$

$$V(X,X) = 0, \quad 0 \le X \le 2,$$

where $\widehat{F}(X, Y, V(X, Y)) = F(\frac{x+y}{2}, \frac{-x+y}{2}, u(\frac{x+y}{2}, \frac{-x+y}{2}))$ and $\gamma = \frac{-\lambda}{4}$. Integrating Eq. (4) and using the above boundary conditions we have

$$\begin{aligned} V(\xi, \eta) =& \gamma(\int\int_{\Omega'} G(\xi, \eta; X, Y)\widehat{F}(X, Y, V(X, Y))\, dX\, dY \\ & - \int_0^2 r(\xi, \eta, X)(\widehat{F}(0, X, V(0, X)) - \widehat{F}(2, X, V(X, 2))dX), \end{aligned}$$

such that η, $\xi \in \Omega'$ and

$$r(\xi, \eta, X) = \begin{cases} 0 & \text{if } 0 \le X \le \xi \\ \beta/(2\alpha) & \text{if } \xi \le X \le \eta \\ 0 & \text{if } \eta \le X \le 2. \end{cases}$$

In [21, 34, 35, 28] the Green's function, $G(\xi, \eta; X, Y)$ described. Also, in [42, 35, 28] we observe that the critical eigenvalues are extended based on spectral theory (see section 2 [42]). □

Proposition 2. *If* $F(x, t, v) = \int_\Omega k(x - t)v(x, t)d\Omega = k * v$ *and we have the above assumptions in proposition of 1 then problem* (1)-(3) *has a unique solution.*

Proof. By using the above proposition, [26, 27, 28, 42] and the Hilbert translations the proof is completed. □

3. The Streamline Diffusion Method

For simplify in (1)-(3) we assume $\lambda = 1$. We introduce variables $v = \partial u/\partial t$ and $\dot{v} = \partial v/\partial t$. Hence, we rewrite (1)-(3) to

$$\begin{cases} L\mathbf{w} \equiv \dot{\mathbf{w}}(x, t) + A\mathbf{w}(x, t) = f(u) & \text{in } \Omega \\ \mathbf{w}(x, 0) = \mathbf{0}, & 0 \le x \le 2 \\ B\mathbf{w'}(t, t) = C\mathbf{w''}(1 + t, 1 - t), & 0 \le t \le 1. \end{cases} \tag{5}$$

Here, we assume that $\mathbf{w}(x, t) = (u(x, t), v(x, t))^T$, $\dot{\mathbf{w}}(x, t) = (\dot{u}(x, t), \dot{v}(x, t))^T$, $\mathbf{w'} = (u, \frac{\partial u}{\partial n_1})^T$, $\mathbf{w''} = (u, \frac{\partial u}{\partial n_2})^T$, $A = \begin{pmatrix} 0 & -1 \\ -\frac{\partial^2}{\partial x^2} & 0 \end{pmatrix}$, $B = \begin{pmatrix} \alpha & -\beta \end{pmatrix}$, $C = \begin{pmatrix} \alpha & \beta \end{pmatrix}$ and $f(u) = (0, F(x, t, u))^T$.

In this section, we consider the Sd-method for solving (5). In this method, instead of using the standard Galerkin method, is usual in Finite Element Method, for the one variable (spatial or time) we used the Galerkin method simultaneously in space and time. That is, we use finite element and interpolation functions depend on time and space.

Space-time Sd-method can be used to improve stabilization, however used without care, this would lead to a very large linear system to be solved. One of the reasons for it is that in this technique the use of continuous (in time) test and trial functions in all levels of time. One way to avoid this difficultly, and decrease the size of the corresponding linear system, is to work in slabs of space-time, with the help of interpolation functions that will be continuous in the spatial variables but will be discontinuous in the time variables at the common frontier of every two slabs.

Sd-method for (5) is based on using finite element over the space-time domain Ω. To define this method, let $0 = t_0 < t_1 < \cdots < t_N = 1$ be a subdivision of the time interval $[0, 1]$ into intervals $I_n = (t_n, t_{n+1})$, with time steps $k_n = t_{n+1} - t_n$, $n = 0, 1, \cdots, N-1$ and introduce the corresponding space-time slabs (see Fig. 1.), i. e.,

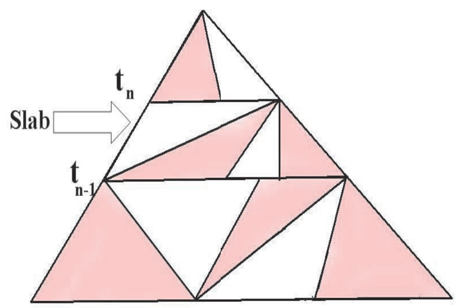


FIG. 1. The slabs on Ω

$$S_n = \Big\{(x,t) : \quad t_{n+1} \le x \le 2 - t_{n+1}, \qquad t \le x \le t_{n+1},$$
$$2 - t_{n+1} \le x \le 2 - t, \qquad t_n < t < t_{n+1}\Big\},$$

for $n = 0, 1, ..., N-2$ and

$$S_{N-1} = \{(x,t) : t \le x \le 2 - t, \quad t_{N-1} < t < t_N\}.$$

Further, for each n let $\mathbf{W}^{n,\alpha,\beta}$ be a finite element subspace of $H^1(S_n) \times H^1(S_n)$, based on triangulation of the slab S_n with elements of size h and let

$$\dot{\mathbf{W}}^{n,\alpha,\beta} = \Big\{\mathbf{w} \in \mathbf{W}^{n,\alpha,\beta} | B\mathbf{w}'(t,t) = C\mathbf{w}''(t+1, 1-t), \quad 0 \le t \le 1\Big\}.$$

Simplifying, we get boundary condition in $\dot{\mathbf{W}}^{n,\alpha,\beta}$ equal zero. We can formulate Sd-method on the slab S_n for (5), as follows:
For $n = 0, \cdots, N-1$, find $\mathbf{w}^n \in \dot{\mathbf{W}}^n$ such that

$$(\dot{\mathbf{w}}^{n,\alpha,\beta} + A\mathbf{w}^{n,\alpha,\beta}, g + \delta(\dot{g} + Ag))_n + \langle \mathbf{w}_+^n, g_+ \rangle_n + \langle \mathbf{w}_+^{n,\alpha,\beta}, g_+ \rangle_{\Gamma_n} = \tag{6}$$

$$= (f(u^n), g + \delta(\dot{g} + Ag))_n + \langle \mathbf{w}_-^{n,\alpha,\beta}, g_+ \rangle_n.$$

We have $g + \delta(\dot{g} + Ag)$, as a test function such that $\delta = \bar{C}h$ with $\bar{C}$ is a suitable chosen (sufficiently small, see [33]) positive constant. Further, we define the following notations for (6) and everywhere in the paper:

$$(\mathbf{u},\mathbf{v})_n = \int_{S_n} \mathbf{u}^T.\mathbf{v}\,dxdt,$$

$$(\mathbf{u},\mathbf{u})_n = \|\mathbf{u}\|_n^2,$$

$$\langle \mathbf{u},\mathbf{v}\rangle_n = \int_{t_n}^{2-t_n} \mathbf{u}^T(x,t_n).\mathbf{v}(x,t_n)dx,$$
$$\langle \mathbf{u},\mathbf{u}\rangle_n = |\mathbf{u}|_n^2,$$
$$\mathbf{v}_+(x,t) = \lim_{s\to 0^+} \mathbf{v}(x,t+s),$$
$$\mathbf{v}_-(x,t) = \lim_{s\to 0^-} \mathbf{v}(x,t+s),$$
$$\langle \mathbf{u}_+,\mathbf{v}_+\rangle_\Gamma = \int_\Gamma {\mathbf{u}_+}^T.\mathbf{v}_+ d\sigma,$$
$$\langle \mathbf{u}_+,\mathbf{v}_+\rangle_{\Gamma_n} = \int_{\Gamma_n} {\mathbf{u}_+}^T.\mathbf{v}_+ ds,$$

also $\|.\|=\|.\|_{L_2(\Omega)}$, $\|.\|_{\infty,\Omega} = \|.\|_{L_\infty(\Omega)}$, $\|.\|_s = \|.\|_{s,\Omega} = \|.\|_{H^s(\Omega)}$, $\Gamma = \partial\Omega$ and $\Gamma = \bigcup_{n=0}^{N-1}\Gamma_n$. The terms including $\langle \ , \ \rangle_{\Gamma,\Gamma_n}$ in the above formula is a jump conditions which imposes a weakly enforced continuity condition across the slab interfaces, at t_n and is the mechanism by which information is propagated from one slab to another. For more concisely, after summing over n and $f(\varpi) \simeq f(g) + (\varpi - g).\frac{\partial f(g)}{\partial u}$ (such that $\varpi = (\ u0\ )$), we get the function space $\prod_{n=0}^{N-1} \dot{\mathbf{W}}^{n,\alpha,\beta}$, therefore we may rewrite (6) as follow:

find $\mathbf{w} \in \prod_{n=0}^{N-1} \dot{\mathbf{W}}^{n,\alpha,\beta}$, such that

$$B(\mathbf{w},g) = L(g), \tag{7}$$

for $g \in \prod_{n=0}^{N-1} \dot{\mathbf{W}}^{n,\alpha,\beta}$. The bilinear form $B(.,.)$ and the linear form $L(.)$ are defined by

$$B(\mathbf{w},g) = \sum_{n=0}^{N-1}\{(\dot{\mathbf{w}}^{n,\alpha,\beta} + A\mathbf{w}^{n,\alpha,\beta} - \varpi^n.\frac{\partial f}{\partial u}(g), g + \delta(\dot{g} + Ag))_n + \langle \mathbf{w}_+^{n,\alpha,\beta}, g_+\rangle_{\Gamma_n}\}$$
$$+ \sum_{n=1}^{N-1}\langle [\mathbf{w}^{n,\alpha,\beta}], g_+\rangle_n + \langle \mathbf{w}_+^{n,\alpha,\beta}, g_+\rangle_0,$$
$$L(g) = \sum_{n=0}^{N-1}(f(g) - g.\frac{\partial f}{\partial u}(g), g + \delta(\dot{g} + Ag))_n$$

for $\mathbf{w} = (\mathbf{w}_1,\mathbf{w}_2)^T$ and $\varpi = (\mathbf{w}_1, 0)^T$. Also, we assume that $[\mathbf{w}_i] = \mathbf{w}_{i,+} - \mathbf{w}_{i,-}$, for $i = 1,2$, $[\mathbf{w}] = ([\mathbf{w}_1],[\mathbf{w}_2])^T$. Let T_h^n be a triangulation of the slab S_n into triangles K, for $h > 0$, and introduce

$$\mathbf{W}_h^{n,\alpha,\beta} = \Big\{\mathbf{w} \in \dot{\mathbf{W}}^{n,\alpha,\beta} : \mathbf{w}|_K \in [P_k(K)] \times [P_k(K)] \subseteq H^1(S_n) \times H^1(S_n), \\ K \in T_h^{n,\alpha,\beta}\Big\}$$

where $P_k(K)$ denotes the set of polynomials in K of degree less than or equal k and

$$\mathbf{W}_h = \prod_{n=0}^{N-1}\mathbf{W}_h^{n,\alpha,\beta}.$$

Thus (7) can be formulated as follows:

Find $\mathbf{w}_h = \begin{pmatrix} u_h \\ v_h \end{pmatrix} \in \mathbf{W}_h$ such that

$$B(\mathbf{w}_h, g) = L(g), \tag{8}$$

for $g \in \mathbf{W}_h$. Moreover, we know that the exact solution of (7) satisfies

$$B(\mathbf{w}, g) = L(g),$$

for $g \in \dot{\mathbf{W}}^{n,\alpha,\beta}$, and by subtraction we have the following error equation

$$B(e, g) = 0, \tag{9}$$

where $e = \mathbf{w} - \mathbf{w}_h$ and $\mathbf{w} \in \mathbf{W}_h$.

4. Stability for the S*d*-Method

Below, we derive the stability estimate for Sd-method (7). These estimate will be of crucial importance in proving the finite element analysis. We apply properties of the bilinear $B(.,.)$ and obtain stability estimate. For our problem, we have the following stability Proposition:

Proposition 3. *For any* $\boldsymbol{w} = \begin{pmatrix} u \\ v \end{pmatrix} \in \prod_{n=0}^{N-1} \boldsymbol{W}^{n,\alpha,\beta}$ *with assumptions* $uv \leq 0$ *and* $\frac{\partial u}{\partial x}\frac{\partial v}{\partial x} \geq 0$ *we have:*

$$B(\boldsymbol{w}, \boldsymbol{w}) \geq |||\boldsymbol{w}|||^2 = \frac{1}{2}\{\mid \boldsymbol{w}_- \mid_N^2 - \mid \boldsymbol{w}_+ \mid_0^2 + \delta \parallel \dot{\boldsymbol{w}} + A\boldsymbol{w} \parallel_\Omega^2\} + \mid \boldsymbol{w}_+ \mid_\Gamma^2 . \tag{10}$$

Proof. Using the definition of the bilinear form B and setting $g = \mathbf{w}$ it follows:

$$B(\mathbf{w}, \mathbf{w}) = (\dot{\mathbf{w}}, \mathbf{w})_\Omega + (A\mathbf{w}, \mathbf{w})_\Omega + \delta \parallel \dot{\mathbf{w}} + A\mathbf{w} \parallel_\Omega^2 + \mid \mathbf{w}_+ \mid_\Gamma^2 + \\ + \sum_{n=1}^{N-1} \langle [\mathbf{w}], \mathbf{w}_+ \rangle_n + \langle \mathbf{w}_+, \mathbf{w}_+ \rangle_0.$$

Integrating by parts yields

$$(\dot{\mathbf{w}}, \mathbf{w})_\Omega + \sum_{n=1}^{N-1} \langle [\mathbf{w}], \mathbf{w}_+ \rangle_n + \langle \mathbf{w}_+, \mathbf{w}_+ \rangle_0 = \\ = \frac{1}{2}\{\mid \mathbf{w}_- \mid_N^2 + \mid \mathbf{w}_+ \mid_0^2 + \sum_{n=1}^{N-1} \mid [\mathbf{w}] \mid_n^2\}.$$

Therefore, by using the assumptions of the proposition the proof is complete. □

We use the standard argument for finite element and introduce the linear nodal interpolate $I_h\mathbf{w} \in W_h$ of the exact solution $\mathbf{w}$ and we set $\zeta = \mathbf{w} - I_h\mathbf{w}$, $\xi = \mathbf{w}_h - I_h\mathbf{w}$. Thus, we have:

$$e := \mathbf{w} - \mathbf{w}_h = (\mathbf{w} - I_h\mathbf{w}) - (-I_h\mathbf{w} + \mathbf{w}_h) = \zeta - \xi.$$

Recalling the Galerkin orthogonality relation (9):

$$B(e, \mathbf{w}) = 0. \tag{11}$$

Now, we can prove the basic global error estimate by using proposition 3.

Proposition 4. *If* $\boldsymbol{w}_h \in \boldsymbol{W}_h$ *satisfies in* (8) *and* $\boldsymbol{w}$ *is exact solution converted mass-spring* (5), *and also*

$$\| A \|_{\infty,\Omega} \leq C,$$

then, there is a constant C *such that*

$$|||\boldsymbol{w} - \boldsymbol{w}_h||| \leq Ch^{k+1/2} \|\boldsymbol{w}\|_{k+1}.$$

Proof. Using the basic stability estimate (10) with $\mathbf{w} = e$ and (11), with $\mathbf{w} = \xi$, we get that

$$|||e|||^2 \leq B(e,e) = B(e,\zeta) - B(e,\xi) = B(e,\zeta) =$$

$$= (\dot{e} + Ae, \zeta + \delta(\dot{\zeta} + A\zeta))_\Omega + \sum_{n=0}^{N-1} \langle [e], \zeta_+ \rangle_n + \langle e_+, \zeta_+ \rangle_\Gamma.$$

Moreover, we use the inequality $2ab \leq \epsilon a^2 + \epsilon^{-1} b^2$ for a, b real numbers and $\epsilon > 0$. Therefore, we have:

$$B(e,\zeta) \leq \frac{\delta}{8} \| \dot{e} + Ae \|_\Omega^2 + \frac{2}{\delta} \| \zeta \|_\Omega^2 + \frac{\delta}{8} \| \dot{e} + Ae \|_\Omega^2 + 2\delta \| \dot{\zeta} + A\zeta \|_\Omega^2$$

$$+ \frac{1}{4} \sum_{n=1}^{N-1} | [e] |_n^2 + \sum_{n=1}^{N-1} | \zeta_+ |_n^2 + \frac{1}{4} | e_+ |_0^2 + | \zeta_+ |_0^2 + \frac{1}{4} \| e_+ \|_\Gamma^2 + \| \zeta_+ \|_\Gamma^2 .$$

According to the above proposition and (10), we can write

$$B(e,\zeta) \leq \frac{1}{4} |||e|||^2 +$$

$$+ \left\{ \frac{2}{\delta} \| \zeta \|_\Omega^2 + 2\delta \| \dot{\zeta} + A\zeta \|_\Omega^2 + \sum_{n=1}^{N-1} | \zeta_+ |_n^2 + | \zeta_+ |_0^2 + \| \zeta_+ \|_\Gamma^2 \right\}.$$

On the other hand, we have the inequality

$$\| \dot{\zeta} + A\zeta \|_\Omega \leq \| \dot{\zeta} \|_\Omega + \| A \|_{\infty,\Omega} \| \zeta \|_\Omega . \tag{12}$$

With using inverse estimate inequality, we have

$$\| \dot{\zeta} \|_\Omega \leq Ch^{-1} \| \zeta \|_\Omega . \tag{13}$$

Therefore, with (12), (13) and assumption $\delta = \bar{C}h$, we obtain:

$$|||e|||^2 \leq C \left\{ \| \zeta_+ \|_\Gamma^2 + h^{-1} \| \zeta \|_\Omega^2 + \sum_{n=0}^{N-1} | \zeta_+ |_n^2 + h \| \zeta \|_{1,\Omega}^2 \right\}.$$

Finally, by standard interpolation theory it follows that (see e.g. Ciarlet [12])

$$\left[h \| \zeta_+ \|_\Gamma^2 + \| \zeta \|_\Omega^2 + h \sum_{n=0}^{N-1} | \zeta_+ |_n^2 + h^2 \| \zeta \|_{1,\Omega}^2 \right]^{1/2} \leq Ch^{k+1} \| \mathbf{w} \|_{k+1,\Omega},$$

which proves the desired estimates. □

We observe in the remarked references that the corresponding optimal convergence rate for the popular numerical methods in the literatures such as conservative finite difference method, semi-implicit finite difference method, semi-discrete finite element method, the time-splitting spectral method or Galerkin method are of order $\mathcal{O}(h^k)$.

5. An a posteriori error estimate

In this section, we shall consider the following simplified version of Sd-method for (6) and (8) with $\delta = 0$:
Find $\mathbf{w}_h \in \mathbf{W}_h$, such that for $n = 0, 1, ..., N-1$:

$$(\dot{\mathbf{w}}_h + A\mathbf{w}_h, g)_n + \langle [\mathbf{w}_h], g_+\rangle_n = (f, g)_n, \qquad \forall g \in \mathbf{W}_h, \tag{14}$$

where $[\mathbf{w}_h] = \mathbf{w}^n_{h,+} - \mathbf{w}^n_{h,-}$ and $\mathbf{w}^0_{h,-} = 0$.
In order to obtain a representation of the error, we consider the following auxiliary problem, referred to as the linearized dual problem:
Find Φ such that

$$\begin{cases} L^*\Phi \equiv -\Phi_t + A^T\Phi = \psi^{-1}e, & \text{in} \quad \Omega, \\ \Phi(t,t) = 0, & t \in [0,1], \\ \Phi(1+t, 1-t) = 0, & t \in [0,1], \\ \Phi(x,1) = 0, & x \in [0,2] \end{cases} \tag{15}$$

and L^* denotes the adjoint of the operator L defined in (15) and ψ is a positive weight function. Note that this problem is computed "backward", but there is a corresponding change in sign. Further, we shall introduce the following notation:

$$\| e \|_{L_2^{\psi}(\Omega)} = (e, \psi e)_{\Omega}^{1/2}. \tag{16}$$

Multiplying (15) by e and integrating by parts, and summing over n, we obtain the following error representation formula:

$$\| e \|^2_{L_2^{\psi^{-1}}(\Omega)} = (e, \psi^{-1}e)_{\Omega} = (e, L^*\Phi) \tag{17}$$

$$= \sum_{n=0}^{N-1} (e, -\Phi_t + A^T\Phi)_n = \sum_{n=0}^{N-1} (e, -\Phi_t)_n + \sum_{n=0}^{N-1} (e, A^T\Phi)_n.$$

On the other hand, we have for $n = 0, 1, ..., N-2$:

$$(e, -\Phi_t)_n = \int_{S_n} (-e^T.\Phi_t)dxdt = \int_{t_n}^{t_{n+1}} \int_{t_n}^{x} (-e^T.\Phi_t)dtdx \tag{18}$$

$$+ \int_{t_{n+1}}^{2-t_{n+1}} \int_{t_n}^{t_{n+1}} (-e^T.\Phi_t)dtdx$$

$$+ \int_{2-t_{n+1}}^{2-t_n} \int_{t_n}^{2-x} (-e^T.\Phi_t)dtdx$$

$$= (e_t, \Phi)_n + \int_{t_n}^{2-t_n} e^T(x, t_n).\Phi(x, t_n)dx - \int_{t_{n+1}}^{2-t_{n+1}} e^T(x, t_{n+1}).\Phi(x, t_{n+1})dx,$$

and for $n = N-1$:

$$(e, -\Phi_t)_{N-1} = (e_t, \Phi)_{N-1} + \int_{t_{N-1}}^{2-t_{N-1}} e^T(x, t_{N-1}).\Phi(x, t_{N-1})dx. \qquad (19)$$

Hence, if we assume $e = \begin{pmatrix} e_1 \\ e_2 \end{pmatrix}$ and $\Phi = \begin{pmatrix} \Phi_1 \\ \Phi_2 \end{pmatrix}$ then, we obtain for $n = 0, 1, ..., N-1$:

$$(e, A^T\Phi)_n = \int_{S_n} e^T.A^T\Phi dxdt$$

$$= \int_{S_n} e^T.\begin{pmatrix} 0 & -\frac{\partial^2}{\partial x^2} \\ -1 & 0 \end{pmatrix}\begin{pmatrix} \Phi_1 \\ \Phi_2 \end{pmatrix} dxdt = \int_{S_n} (e_1, e_2).\begin{pmatrix} \frac{-\partial^2\Phi_2}{\partial x^2} \\ -\Phi_1 \end{pmatrix} dxdt.$$

Therefore by parts integrating and same as (18), we have:

$$\int_{S_n} (-e_1\frac{\partial^2\Phi_2}{\partial x^2} - e_2\Phi_1)\, dx\, dt = \int_{S_n} (-\Phi_2\frac{\partial^2 e_1}{\partial x^2} - e_2\Phi_1)dxdt = (Ae, \Phi)_n. \quad (20)$$

By using (18) and (19) in the following definition, we have:

$$J = \sum_{n=0}^{N-2}\left(\int_{t_n}^{2-t_n} e^T(x, t_n).\Phi(x, t_n)dx - \int_{t_{n+1}}^{2-t_{n+1}} e^T(x, t_{n+1}).\Phi(x, t_{n+1})dx\right) +$$

$$+ \int_{t_{N-1}}^{2-t_{N-1}} e^T(x, t_{N-1}).\Phi(x, t_{N-1})dx =$$

$$= (\langle e_-, \Phi_-\rangle_1 - \langle e_+, \Phi_+\rangle_0) + (\langle e_-, \Phi_-\rangle_2 - \langle e_+, \Phi_+\rangle_1) + ...$$

$$+(\langle e_-, \Phi_-\rangle_{N-1} - \langle e_+, \Phi_+\rangle_{N-2}) + (\langle e_-, \Phi_-\rangle_N - \langle e_+, \Phi_+\rangle_{N-1}).$$

We rearrange the above summation by putting $\Phi_- = \Phi_- - \Phi_+ + \Phi_+$, then we can write:

$$J = \langle e_-, \Phi_-\rangle_N + \langle e_+, \Phi_+\rangle_0 + \sum_{n=0}^{N-1}\langle [e], \Phi_+\rangle_n + \sum_{n=0}^{N-1}\langle e_-, [\Phi]\rangle_n.$$

According to (15), $\Phi(., t_N = 1) = 0$ and since $e_-^0 = [\mathbf{w}^0] = 0$, we get

$$J = \sum_{n=0}^{N-1}\langle [\mathbf{w}_h], \Phi_+\rangle_n. \qquad (21)$$

Therefore by replacing (18)-(21) in (17), we have:

$$\| e \|^2_{L_2^{\psi^{-1}}(\Omega)} = \sum_{n=0}^{N-1}(e_t, \Phi) + \sum_{n=0}^{N-1}(Ae, \Phi)_n - \sum_{n=0}^{N-1}\langle [\mathbf{w}_h], \Phi_+\rangle_n$$

$$= \sum_{n=0}^{N-1}((\mathbf{w} - \mathbf{w}_h)_t + A(\mathbf{w} - \mathbf{w}_h), \Phi)_n - \sum_{n=0}^{N-1}\langle [\mathbf{w}_h], \Phi_+\rangle_n$$

$$= \sum_{n=0}^{N-1}(f - \mathbf{w}_{h,t} - A\mathbf{w}_h, \Phi)_n - \sum_{n=0}^{N-1}\langle [\mathbf{w}_h], \Phi_+\rangle_n.$$

Hence, with recalling (5) and using the Galerkin orthogonality, we obtain

$$\begin{aligned} \| e \|^2_{L_2^{\psi^{-1}}(\Omega)} = & \sum_{n=0}^{N-1} (f - \mathbf{w}_{h,t} - A\mathbf{w}_h, \hat{\Phi} - \Phi)_n - \\ & - \sum_{n=0}^{N-1} \langle [\mathbf{w}_h], (\hat{\Phi} - \Phi)_+ \rangle_n \equiv I + II. \end{aligned} \tag{22}$$

Where $\hat{\Phi} \in W_h$ is an interpolation of Φ. The idea is now to estimate $\hat{\Phi} - \Phi$ in terms of $\psi^{-1}e$ using a strong stability estimates for solution Φ of the dual problem.

6. Interpolation estimates

In the following we consider two L_2-projections for $\hat{\Phi} \in \mathbf{W}_h$ in (9):

$$P_n : L_2([0,2]) \longmapsto \mathbf{W}_h^n,$$

$$\pi_n : L_2(S_n) \longmapsto \Pi_{0,n} = \{\mathbf{w} \in L_2(S_n) : \mathbf{w}(x,.) \text{ is constant on } I_n, x \in [0,2]\},$$

such that

$$\int_0^2 (P_n\Phi)^T.\mathbf{w}dx = \int_0^2 \Phi^T.\mathbf{w}dx, \qquad \forall \mathbf{w} \in \mathbf{W}_h^n,$$

$$\pi_n \mathbf{w} \mid_{S_n} = \frac{1}{k_n} \int_{I_n} \mathbf{w}(.,t)dt, \quad \forall \mathbf{w} \in \Pi_{0,n}.$$

Then, we can define $\hat{\Phi} \mid_{S_n} \in \mathbf{W}_h^n$ by letting

$$\hat{\Phi} \mid_{S_n} = P_n \pi_n \Phi = \pi_n P_n \Phi \in \mathbf{W}_h^n,$$

where $\Phi = \Phi \mid_{S_n}$ and we can observe that P_n and π_n are commuted. Moreover, if we introduce P and π defined by

$$(P\Phi) \mid_{S_n} = P_n(\Phi \mid_{S_n}),$$

and

$$(\pi\Phi) \mid_{S_n} = \pi_n(\Phi \mid_{S_n}),$$

then we can put :

$$\hat{\Phi} = P\pi\Phi = \pi P\Phi \in \mathbf{W}_h.$$

Now, we define the following residuals:

$$R_0 = f - \mathbf{w}_{h,t} - A\mathbf{w}_h,$$

$$R_1 = \frac{\mathbf{w}_{h,+}^n - \mathbf{w}_{h,-}^n}{k_n}, \qquad \text{on} \quad S_n,$$

$$R_2 = \frac{(P_n - I)\mathbf{w}_{h,-}^n}{k_n}, \qquad \text{on} \quad S_n,$$

where I is the identity operator.

In the end of this section, we shall give a lemma for some interpolation estimates by the projection operators P, leaving the overall of I and II to next section.

Lemma 1. *There is a constant C such that for residual $R \in L_2(\Omega)$,*

$$|(R, \Phi - P\Phi)_\Omega| \leq C \parallel h^2(I-P)R \parallel_{L_2^{\psi^{-1}}(\Omega)} \parallel \Phi_{xx} \parallel_{L_2^{\psi}(\Omega)} . \qquad (23)$$

Proof. See [41] and [40]. □

7. The completion of the proof of a posteriori error estimates

In this section we state and prove a posteriori error estimate by estimating of the terms I and II in the error representation formula (22). To this approach we introduce the stability factors(see [10]) associated with discretization in time and space, defined by

$$Y_e^t = \frac{\parallel \Phi_t \parallel_{L_2^{\psi}(\Omega)}}{\parallel e \parallel_{L_2^{\psi^{-1}}(\Omega)}}, \qquad (24)$$

and

$$Y_e^x = \frac{\parallel \Phi_{xx} \parallel_{L_2^{\psi}(\Omega)}}{\parallel e \parallel_{L_2^{\psi^{-1}}(\Omega)}} \qquad (25)$$

respectively. We now apply the result of the previous sections; using Cauchy-Schwartz inequality in (22) coupled with the interpolation estimate (23) and the strong stability factors (24) and (25), to derive the $L_2(L_2)$ a posteriori error estimates for the scheme (14).

Proposition 5. *The error $e = \boldsymbol{w} - \boldsymbol{w}_h$, where $\boldsymbol{w}$ is the solution of the continuous problems (5) and $\boldsymbol{w}_h$ that of (14), satisfies the following stability estimate:*

$$\parallel e \parallel_{L_2^{\psi^{-1}}(\Omega)} \leq CY_e^x \parallel h^2(I-P)R_0 \parallel_{L_2^{\psi^{-1}}(\Omega)} + CY_e^t \parallel k_n R_1 \parallel_{L_2^{\psi^{-1}}(\Omega)} +$$
$$+ Y_e^x \parallel h^2 R_2 \parallel_{L_2^{\psi^{-1}}(\Omega)} + Y_e^t \parallel k_n R_2 \parallel_{L_2^{\psi^{-1}}(\Omega)} .$$

Proof. Using the notation introduce above, we may write (22) as

$$\parallel e \parallel^2_{L_2^{\psi^{-1}}(\Omega)} = \sum_{n=0}^{N-1} (R_0, \hat{\Phi} - \Phi)_n + \sum_{n=0}^{N-1} \langle k_n \frac{[\mathbf{w}_h]}{k_n}, (\hat{\Phi} - \Phi)_+ \rangle_n = I + II.$$

Below we shall estimate the terms I and II separately. Splitting the interpolation error by writing $\hat{\Phi} - \Phi = \hat{\Phi} - P\Phi + P\Phi - \Phi$ and $\hat{\Phi}_n = \pi_n P\Phi$, we have:

$$I = \sum_{n=0}^{N-1} (R_0, \hat{\Phi}_n - P\Phi + P\Phi - \Phi)_n = \sum_{n=0}^{N-1} (R_0, \hat{\Phi}_n - P\Phi)_n +$$
$$+ \sum_{n=0}^{N-1} (R_0, P\Phi - \Phi)_n \leq C \parallel h^2(I-P)R_0 \parallel_{L_2^{\psi^{-1}}(\Omega)} \parallel \Phi_{xx} \parallel_{L_2^{\psi}(\Omega)} .$$

It remains to estimate the term II, to this end, we consider the following notation:

$$\Phi_+^n(x) = \Phi(x,t) - \int_{t_n}^{t} \frac{\partial}{\partial \tau} \Phi(x,\tau) d\tau,$$

hence, with integrating over I_n, we have:

$$k_n\Phi_+^n(x) = \int_{I_n}\Phi(x,t)dt - \int_{I_n}\int_{t_n}^{t}\Phi_\tau(x,\tau)d\tau dt \qquad (26)$$

where $\Phi_\tau = \frac{\partial\Phi}{\partial\tau}$ and $\Phi^n = \Phi(.,t_n)$.

$$II = \sum_{n=0}^{N-1}\langle k_n\frac{[\mathbf{w}_h]}{k_n}, (\hat{\Phi}-\Phi)_+\rangle_n = \sum_{n=0}^{N-1}\langle k_n\frac{[\mathbf{w}_h]}{k_n}, (\hat{\Phi}_n - P\Phi + P\Phi - \Phi)_+\rangle_n$$

$$= \sum_{n=0}^{N-1}\langle k_n\frac{[\mathbf{w}_h]}{k_n}, (\hat{\Phi}_n - P\Phi)_+\rangle_n + \sum_{n=0}^{N-1}\langle k_n\frac{[\mathbf{w}_h]}{k_n}, (P\Phi-\Phi)_+\rangle_n := II_1 + II_2.$$

To estimate II_1, we use (26) to get

$$II_1 = \sum_{n=0}^{N-1}\langle k_nR_1, (\hat{\Phi}_n)_+ - P\Phi_+\rangle_n = \sum_{n=0}^{N-1}\langle R_1, k_n\hat{\Phi}_n - Pk_n\Phi_+\rangle_n$$

$$= \sum_{n=0}^{N-1}\langle R_1, k_n\hat{\Phi}_n - \int_{I_n}P\Phi(.,t)dt + \int_{I_n}\int_{t_n}^{t}P\Phi_\tau(.,\tau)d\tau dt\rangle_n$$

$$= \sum_{n=0}^{N-1}\int_{I_n}\int_{t_n}^{t}\langle R_1, P\Phi_\tau(.,\tau)\rangle_n d\tau dt$$

by using (16), (17) and Hölder inequality, we have:

$$II_1 \leq \| k_nR_1 \|_{L_2^{\psi^{-1}}(\Omega)} \| P\Phi_t \|_{L_2^{\psi}(\Omega)} \leq \| k_nR_1 \|_{L_2^{\psi^{-1}}(\Omega)} \| \Phi_t \|_{L_2^{\psi}(\Omega)} .$$

As for the II_2-terms we can write

$$II_2 = \sum_{n=0}^{N-1}\langle k_n\frac{[\mathbf{w}_h]}{k_n}, (P\Phi-\Phi)_+\rangle_n = \sum_{n=0}^{N-1}\langle\frac{\mathbf{w}_{h,+}^n - \mathbf{w}_{h,-}^n}{k_n}, (P_n - I)k_n\Phi_+\rangle_n$$

$$= \sum_{n=0}^{N-1}\langle\frac{P_n\mathbf{w}_{h,-}^n - \mathbf{w}_{h,-}^n}{k_n}, (P_n-I)(\int_{I_n}\Phi(.,t)dt - \int_{I_n}\int_{t_n}^{t}\Phi_\tau(.,\tau)d\tau dt)\rangle_n$$

$$\leq \sum_{n=0}^{N-1}\int_{I_n}\langle\frac{(P_n-I)\mathbf{w}_{h,-}^n}{k_n}, (P_n-I)\Phi(.,t)\rangle_n dt$$

$$+\sum_{n=0}^{N-1}\int_{I_n}\int_{t_n}^{t}\langle\frac{(P_n-I)\mathbf{w}_{h,-}^n}{k_n}, (P_n-I)\Phi_\tau(.,t)d\tau dt\rangle_n$$

by using (16), (17) and Hölder inequality, we have:

$$II_2 \leq \| k_nR_2 \|_{L_2^{\psi^{-1}}(\Omega)} \| \Phi_{xx} \|_{L_2^{\psi}(\Omega)} + \| k_nR_2 \|_{L_2^{\psi^{-1}}(\Omega)} \| \Phi_t \|_{L_2^{\psi}(\Omega)} .$$

The a posteriori error estimate now follows immediately after collecting the terms and using the definition of the stability factors (24) and (25). □

8. Numerical results

At present, three numerical examples for testing Sd method are given. We carry out (7), by an AMD Opteron computer with 15 Gigabytes RAM memory with 2.2 GHz CPU. For each slab S_n, let x_i^n be a mesh, portioned into intervals $J_i^n = (x_{i-1}^n, x_i^n)$, with $h_i^n = x_i^n - x_{i-1}^n$. We define the time mesh function $k = k(t)$ by $k(t) = k_n$ for $t \in (t_n, t_{n+1})$. For $h > 0$ let T_h^n be a triangulation of the slab S_n into triangle K (cf. Figure 1.), satisfying as usual the minimum angle condition (see, e.g. [33]), and indexed by the parameter h representing the maximum diameter of the triangle $K \in T_h^n$. The triangulation of S_n may be chosen independently of that of S_{n-1}, but for the sake of simplicity it must satisfy quasi-uniformity conditions for finite element meshes [12]. To give numerical results obtained using the Sd method, we shall use finite element approximation on a space time slab with the trial function which are piecewise polynomials in space and linear in time; that is, for $(x,t) \in S_n$, we let $w_h^n(x,t) = (u_h^n(x,t), v_h^n(x,t))^T \in \mathbf{W}_h^n$ where

$$u_h^n = \sum_{i=1}^{M} \phi_i(x)(\theta_1(t)\widetilde{u_i^n} + \theta_2(t)u_i^{n+1}) \text{ and}$$

$$v_h^n = \frac{\partial u_h^n}{\partial t} = \sum_{i=1}^{M} \phi_i(x)(\theta_1'(t)\widetilde{v_i^n} + \theta_2'(t)v_i^{n+1})$$

such that $\{\phi_i(x_j) = \delta_{ij}\}$, $i,j = 0,\cdots,M$ are the spatial shape functions at node i and $\{\theta_1 = \frac{t_{n+1}-t}{k}, \theta_2 = \frac{t-t_n}{k}\}$ are the time linear interpolation functions. Moreover, we assume the nodal values of u for node i ant $(t_n)_+$ and $(t_{n+1})_+$ are denoted by $\widetilde{u_i^n}$ $(= \widetilde{v_i^n})$ and u_i^{n+1} $(= v_i^{n+1})$, respectively. Therefore, we consider the above algorithm for the following test problems.

Tabl. 1. Error $= \|w - w_h^n\|_\infty$ by Sd method at different δ.

(x,t)	$\delta = 0.15$	$\delta = 0.10$	$\delta = 0.05$	$\delta = 0.01$	$\delta = 0.005$
$(-1, 0.1)$	0.231e-6	0.212e-9	0.431e-8	0.751e-10	0.321e-9
$(0.0, 0.5)$	0.231e-5	0.761e-7	0.454e-7	0.983e-9	0.522e-10
$(1, 0.9)$	0.514e-7	0.634e-10	0.713e-10	0.761e-9	0.510e-9

Test problem 1. Streamline diffusion method is computed by given δ, $\beta = 0$, $M = 20$, $h = 0.1$, $k = 0.005$, $u(x,t) = \sin\xi(x+t)$ and $v(x,t) = \sin\zeta(x+t)$ such that we define $\xi(x) = \xi(x+2)$, $\zeta(x) = \zeta(x+2) + \pi$ and $\xi = \begin{cases} \pi/2 & x = 0 \\ x & x \neq 0 \end{cases}$ Therefore, we have the exact solution of (1) and in Table 1., we verify pointwise of the error $= \|w - w_h^n\|_\infty = \max\{|u(x,t) - u_h^n(x,t)|, |v(x,t) - v_h^n(x,t)|\}$. In this example we test how well the stability theory developed in Proposition 5 matches with computation by the stability factors that is (24) and (25). Therefore, this proposition guarantees computational stability for small time step.

Test problem 2. Streamline diffusion method is shown by given $\beta = 0$ and $F(x,t,u) = (u-1)^2$ in Figure 2 (in the first row). The results are given after 10 time step that is $n = 1, 2, ..., 10$ and $k_n = 0.1$. In this example, we haven't the exact solution but Proposition 5 guarantees computational stability.

Test problem 3. Streamline diffusion method is shown by given $\alpha = 0$ and $F(x,t,u) = (u-1)^2$ in Figure 2 (in the second row). The results are given after 10 time step that is $n = 1, 2, ..., 10$ and $k_n = 0.1$. In this example, we haven't the exact solution but Proposition 5 guarantees computational stability.

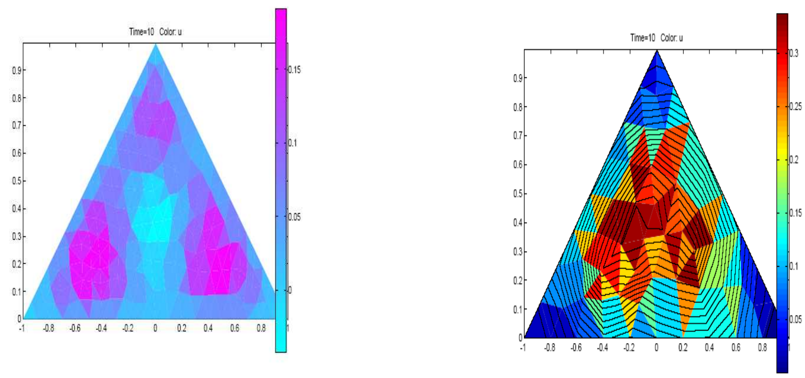


FIG. 2. The approximation solution of u for example 2 (in the first row) and example 3 (in the second row) when $\delta = 0.1$ and the stability factors $Y_e^t, Y_e^x \leq 10^{-3}$

9. Conclusion

To this end, a special nonlinear second order hyperbolic initial-boundary value problem is investigated. We use streamline diffusion method for this case of this wave equation and obtain a priori and a posteriori error estimates. A posteriori error estimate is a very powerful mathematical tool in this problem by Sd method. We try to obtain optimal bounds and the eigenvalues and eigenfunctions remains a challenge that deserves special attention and will be consideration elsewhere.

Acknowledgments

The authors express their gratitude to anonymous refere for many helpful remarks and suggestions.

Bibliography

1. Adams R. A. Sobolev Spaces / R. A. Adams.– Academic Press: New York, 1975.
2. Asadzadeh M. Streamline Diffusion Methods for the Vlasov-Poisson Equations / M. Asadzadeh // RAIRO Mathematical Modelling and Numerical Analysis.– 1990.– Vol. 24.– P. 177-196.
3. Asadzadeh M. Streamline Diffusion Methods for the Fermi and Fokker-Planck Equations / M. Asadzadeh // Transport Theory and Statistical Physics.– 1997.– P. 319-340.

4. Asadzadeh M. Convergence of a hp streamline diffusion scheme for the Vlasov-Fokker-Planck System / M. Asadzadeh, A. Sopasakis.– Preprint, 2003.
5. Asadzadeh M. Convergence of Streamline Diffusion Methods for the Vlasov-Poisson-Fokker-Planck System / M. Asadzadeh, P. Kowalczyk // Numer. Methods. in Partial Differential Eq.– 2005.– Vol. 21.– P. 472-495.
6. Asadzadeh M. A posteriori Error Estimates for the Fokker-Planck and Fermi Pencil Beam Equations / M. Asadzadeh // Math. Models and Methods in Appl. Sci.– 2000.– Vol. 10.– P. 737-769.
7. Asadzadeh M. A Posteriori Error Estimates for a Coupled Wave System with a Local Damping / M. Asadzadeh, D. Rostamy, F. Zabihi // Journal of Mathematical Sciences.– 2011.– Vol. 175.– P. 228-248.
8. Babuska I. Adaptive Approaches and Reliability Estimates in Finite element Analysis / I. Babuska, W.C. Rheinboldt // Comp. Meth. Appl. Mech. Eng.– 1979.– Vol. 17/18.– P. 519-540.
9. Brenner S. C. The Mathematical Theory of Finite Element Method, Springer-Verlag / S. C. Brenner, L. R. Scott.– New York, 1994.
10. Burman E. Adaptive Finite Element Methods for Compressible Two-Phase Flow / E. Burman // Math. Mod. Meth. App. Sci.– 2000.– Vol. 10.– P. 963-989.
11. Di Cairano S. Model predictive control of magnetically actuated mass spring dampers for automotive applications / S. Di Cairano, A. Bemporad, I. V. Kolmanovsky, D. Hrovat // Internat. J. Control.– 2007.– Vol. 80, № 11.– P. 1701-1716.
12. Ciarlet P. G. The Finite Element Method for Elliptic Problems / P. G. Ciarlet.– Amesterdam: North Holland, 1987.
13. Codina R. Finite element approximation of the hyperbolic wave equation in mixed form / R. Codina // Comput. Methods Appl. Mech. Engrg.– 2008.– Vol. 197.– P. 1305-1322.
14. Dubois F. Boundary conditions for nonlinear hyperbolic systems of conservation laws / F. Dubois, P. G. Le Floch // J. Differential Equations.– 1988.– Vol. 71.– P. 93-122.
15. Eriksson K. Adaptive Finite element methods for parabolic problems I: A linear model problem / K. Eriksson, C. Johnson // SIAM J. Numer. Anal.– 1991.– Vol. 28.– P. 43-77.
16. Eriksson K. Adaptive Finite element methods for parabolic problems II: Optimal error estimates in $L_\infty L_2$ and $L_\infty L_\infty$ / K. Eriksson, C. Johnson // SIAM J. Numer.Anal.– 1995.– Vol. 32.– P. 706-740.
17. Eriksson K. Adaptive Streamline Diffusion Finite Element Methods for Stationary Convection-Diffusion Problems / K. Eriksson, C. Johnson // Math. Comp.– 1993.– Vol. 60.– P. 167-188.
18. Evans L. C. Partial Differential Equations / L. C. Evans.– AMS, 1998.
19. Fuhrer C. An Adaptive Streamline Diffusion Finite Element Method for Hyperbolic Conservation Laws / C. Fuhrer, R. Rannacher // East-West J. Numer. Math.– 1997.– Vol. 5.– P. 145-162.
20. Folland G. B. Introduction to Partial Differential Equations / G. B. Folland.– Princeton University Press, 1976.
21. Kalmenov T. On the Spectrum of a Selfadjoint Problem for the Wave Equation / T. Kalmenov // Akad. Nauk. Kazakh. SSR, Vestnik.– 1983.– Vol. 1.– P. 63-66.
22. Kaveh A. Decomposition of symmetric mass-spring vibrating systems using groups graphs and linear algebra / A. Kaveh, M. Nikbakht // Comm. Numer. Methods Engrg.– 2007.– Vol. 23.– № 7.– P. 639-664.
23. Haws L. Symmetric Green's functions for certain hyperbolic problems / L. Haws // Comput. Math. Appl.– 1991.– Vol. 21, № 5.– P. 65–78.
24. Xiantong Huang Physical parameters reconstruction of a fixed–fixed mass-spring system from its characteristic data / Huang Xiantong, Hu Xiyan, Zhang Lei // J. Comput. Appl. Math.– 2007.– Vol. 206, № 2.– P. 645-655.
25. Hughes T. J. R. A Multidimensional upwind scheme with no Crosswind diffusion / T. J. R. Hughes, A. Brooks // Finite Element methods for Convection Dominated flows.– 1979.– AMD-vol-34.– P. 19-35.

26. Iraniparast N. A boundary value problem for the wave equation / N. Iraniparast // Int. J. Math. Math. Sci.– 1999.– Vol. 22, № 4.– P. 835–845.
27. Iraniparast N. A CIV boundary value problem for the wave equation / N. Iraniparast // Appl. Anal.– 2000.– Vol. 76, №3-4.– P. 261–271.
28. Iraniparast N. A selfadjoint hyperbolic boundery-value problem / N. Iraniparast // E. Jornal of Diffeential Equations.– 2003.– P. 153-161.
29. Izadi M. Streamline diffusion Finite Element Method for coupling equations of non-linear hyperbolic scalar conservation laws / M. Izadi // M. Sc Thesis, Supervizor Prof. Asadzadeh.– 2005.
30. Johnson C. Numerical solutions of partial differential equations by the Finite Element Method / C. Johnson.– Cambridge University, 1987.
31. Johnson C. Finite Element Method for Linear Hyperbolic Problems / C. Johnson, U. Navert, J. Pitkaranta // Comput. Methods Appl. Mech. Engrg.– 1984.– Vol. 45.– P. 285-312.
32. Johnson C. An Analysis of the Discontinuous Galerkin Method for a Scalar Hyperbolic. / C. Johnson, J. Pitkrata // Equation, Mathematics of Computation.– 1986.– Vol. 46.– P. 56-66.
33. Johnson C. Finite Element Methods for Convection-Diffusion Problems / R. Glowinski, J. L. Lions, C. Johnson // Computing Method in Engineering and Applied Sciences. Amsterdam.– 1981.
34. Kreith K. Symmetric Green Functions for a Class of CIV Boundary Value Problems / K. Kreith // Canadian Math. Bull.– 1988.– Vol. 31.– P. 272-279.
35. Kreith K. Mixed selfadjoint boundary conditions for the wave equation / K. Kreith // Differential equations and its applications. Budapest.– 1991.– P. 219–226, Colloq. Math. Soc. Janos Bolyai, 62, North-Holland, Amsterdam, 1991.
36. Navert U. A Finite Element Method for Convection-Diffusion Problems. Ph. D. Thesis Department of Computer Science / U. Navert.– Chalmers University of Technology, Gëoteborg, 1982.
37. Yitshak M. Physical parameters reconstruction of a free-free mass-spring system from its spectra / M. Ram Yitshak, James Caldwell // SIAM J. Appl. Math.– 1992.– Vol. 52, № 1.– P. 140-152.
38. Rostamy D. A streamline diffusion method for the mass-spring system / D. Rostamy, F. Zabihi // Journal of Numerical Mathematics and Stochastics.– 2010.– Vol. 2, № 1.– P. 64-76.
39. Santos F. C. The motion of two masses coupled to a finite mass spring / F. C. Santos, Y. A. Coutinho, L. Ribeiro-Pinto, A. C. Tort // European J. Phys.– 2006.– Vol. 27, № 5.– P. 1037-1051.
40. Sandboge R. Adaptive Finite Element Methods for Systems of Reaction- Diffusion Equations / R. Sandboge // Comput. Methods Appl. Mech. Engrg.– 1998.– Vol. 166.– P. 309-328.
41. Szepessy A. Adaptive Finite Element Methods for Conservation Laws Based on a posteriori Error Estimates / A. Szepessy, C. Johnson // Comm. Pure. Appl. Math.– 1995.– XIVIII, 48.– P. 199-234.
42. Stewart A. C. Existence theorems for some nonlinear hyperbolic equations on a waveguide. A dissertation submitted to the Johns Hopkins University in conformity with the requirements for the degree of Doctor of Philosophy / Ann C. Stewart.– Baltimore: Maryland May, 2006.

Davood Rostamy, Fatemeh Zabihi,
Department of Mathematics, Imam Khomeini International University, Qazvin, Iran

Received 11.04.2012

Журнал обчислювальної та прикладної математики
2012, № 3 (109)
с. 105–115
Journal of Numerical & Applied Mathematics

UDC 519.6

TWO-STEP METHOD FOR SOLVING NONLINEAR EQUATIONS WITH NONDIFFERENTIABLE OPERATOR

Stepan Shakhno, Halina Yarmola

Резюме. Запропоновано двокроковий метод для розв'язування нелінійних рівнянь з недиференційовним оператором, побудований на базі двох методів з порядком збіжності $1+\sqrt{2}$. Вивчено локальну та напівлокальну збіжність запропонованого методу та встановлено порядок збіжності. Проведено числове дослідження на тестових задачах та зроблено порівняння отриманих результатів.

Abstract. In this paper we propose a two-step method for solving nonlinear equations with a nondifferentiable operator. Its method is based on two methods of order of convergence $1+\sqrt{2}$. We study a local and a semilocal convergence of the proposed method and set an order of convergence. We apply our results to the numerical solution of a nonlinear equation and systems of nonlinear equations.

1. Introduction

We consider the equation

$$H(x) \equiv F(x) + G(x) = 0, \tag{1}$$

where F and G are nonlinear operators, defined on a convex subset D of a Banach space X with values in a Banach space Y. F is a Fréchet-differentiable operator, G is a continuous operator.

There are kinds of methods to find a solution of (1). In [1] Argyros studied the two-point iterative process

$$x_{n+1} = x_n - A_n^{-1}(F(x_n) + G(x_n)), \quad n = 0, 1, \dots, \tag{2}$$

where $A_n = A(x_{n-1}, x_n)$ is a bounded linear operator. There was provided a local and a semilocal convergence analysis for the method (2) and some cases where $A_n = F'(x_n)$, $A_n = F'(x_n) + G(x_{n-1}; x_n)$ were considered. Here $G(x; y)$ is a first order divided difference of the operator G at the points x and y. The convergence analysis for the case where $A_n = F'(x_n)$ was given by Zabrejko and Nguen [11]. In the paper [3] the convergence analysis results for modification of the method (2) for some cases of A_n were presented. There are studies in which there are considered difference methods, i.e., the secant method, the parametric secant method [5, 6] and the method based on the method of linear interpolation and the secant method [7]. In [4] Chen studied a Broyden-like method for solving (1). In [9] we researched a semilocal convergence of the

---

†*Key words.* Nondifferentiable operator, convergence order, local and semilocal convergence.

method (2) for $A_n = F'(x_n) + G(2x_n - x_{n-1}; x_{n-1})$. The Newton's method cannot be applied, as differentiability of operator H is required.

In this work we propose a two-step method which is based on the methods with the order of convergence $1+\sqrt{2}$ [8, 10],

$$\begin{aligned} x_{n+1} &= x_n - \Big[F'\Big(\frac{x_n+y_n}{2}\Big) + G(x_n;y_n)\Big]^{-1}(F(x_n)+G(x_n)),\\ y_{n+1} &= x_{n+1} - \Big[F'\Big(\frac{x_n+y_n}{2}\Big) + G(x_n;y_n)\Big]^{-1}(F(x_{n+1})+G(x_{n+1})),\\ & n=0,1,\ldots. \end{aligned} \tag{3}$$

Although the numbers of evaluations of the function values increases by one at each step for the proposed method (3), the convergence order is higher than for the one-step methods.

2. Convergence analysis

Definition 1. Let F be a nonlinear operator defined on a subset D of a linear space X with values in a linear space Y and let x, y be two points of D. A linear operator from X into Y, denoted as $G(x;y)$, which satisfies the condition

$$G(x;y)(x-y) = G(x) - G(y).$$

is called a divided difference of G at the points x and y.

Theorem 1. *Let F and G be nonlinear operators, defined on an open convex subset D of a Banach space X with values in a Banach space Y. F is a twice Fréchet-differentiable operator, G is a continuous operator. Let us suppose that equation (1) has a solution $x^* \in D$, G has a first order divided difference in D and there exist $[A(x,y)]^{-1} = \Big[F'\Big(\frac{x+y}{2}\Big) + G(x;y)\Big]^{-1}$ for all $x \neq y$ and $\|[A(x,y)]^{-1}\| \le B$. Let in D the following conditions fulfill*

$$\|F'(x) - F'(y)\| \le 2p_1\|x-y\|, \tag{4}$$

$$\|F''(x) - F''(y)\| \le p_2\|x-y\|^{\alpha},\ \alpha \in (0,1], \tag{5}$$

$$\|G(x;y) - G(u;v)\| \le q_1(\|x-u\| + \|y-v\|). \tag{6}$$

Suppose that $U = \{x:\ \|x-x^*\| < r_*\} \subset D$, where r_* is the smallest positive zero of equations

$$\begin{aligned} q(r) &= 1,\\ 3B(p_1+q_1)r\,q(r) &= 1, \end{aligned} \tag{7}$$

$$q(r) = B\Big[(p_1+q_1)r + \frac{p_2}{4(\alpha+1)(\alpha+2)}r^{1+\alpha}\Big].$$

Then the sequences $\{x_n\}_{n\ge0}$, $\{y_n\}_{n\ge0}$ generated by the iterative process (3) are well defined for all $x_0,\ y_0 \in U$, remain in U and converge to the solution x^*. Moreover, the following inequalities hold for all $n \ge 0$

$$\begin{aligned} \|x_{n+1}-x^*\| \le & B\Big[(p_1+q_1)\|y_n - x^*\| + \\ & + \frac{p_2}{4(\alpha+1)(\alpha+2)}\|x_n - x^*\|^{1+\alpha}\Big]\|x_n - x^*\|, \end{aligned} \tag{8}$$

$$\|y_{n+1} - x^*\| \le B(p_1 + q_1)\Big[\|y_n - x^*\| + \\ + \|x_n - x^*\| + \|x_{n+1} - x^*\|\Big]\|x_{n+1} - x^*\|. \tag{9}$$

Proof. Since the following equality holds for all $x, h \in D$ [10]

$$F(x+h) = F(x) + F'(x)h + \int_0^1 (1-t)F''(x+th)hhdt,$$

then

$$F(x_n) - F(x^*) - F'\Big(\frac{x_n + x^*}{2}\Big)(x_n - x^*) = \\ = F(x_n) - F\Big(\frac{x_n + x^*}{2}\Big) - F'\Big(\frac{x_n + x^*}{2}\Big)\frac{x_n - x^*}{2} - \\ - \Big[F(x^*) - F\Big(\frac{x_n + x^*}{2}\Big) - F'\Big(\frac{x_n + x^*}{2}\Big)\frac{x^* - x_n}{2}\Big] = \\ = \int_0^1 (1-t)F''\Big(\frac{x_n + x^*}{2} + t\frac{x_n - x^*}{2}\Big)\frac{x_n - x^*}{2}\frac{x_n - x^*}{2}dt - \\ - \int_0^1 (1-t)F''\Big(\frac{x_n + x^*}{2} + t\frac{x^* - x_n}{2}\Big)\frac{x_n - x^*}{2}\frac{x_n - x^*}{2}dt. \tag{10}$$

Using the condition (5) and the equality (10), we obtain

$$\Big\|F(x_n) - F(x^*) - F'\Big(\frac{x_n + x^*}{2}\Big)(x_n - x^*)\Big\| \le \\ \le \frac{p_2\|x_n - x^*\|^{2+\alpha}}{4}\int_0^1 (1-t)t^\alpha dt = \frac{p_2\|x_n - x^*\|^{2+\alpha}}{4(\alpha+1)(\alpha+2)}. \tag{11}$$

Let us choose $x_0 \in U$ and show that the sequences given in (3) are well defined. We denote $A_n = F'\Big(\frac{x_n + y_n}{2}\Big) + G(x_n; y_n)$. If $x_n,\ y_n \in U$, then from the definition of the first order divided difference and (4), (6), (11), we obtain

$$\|x_{n+1} - x^*\| = \|x_n - x^* - A_n^{-1}(F(x_n) + G(x_n) - F(x^*) - G(x^*))\| \le \\ \le \|A_n^{-1}\|\Big\|F(x_n) - F(x^*) - F'\Big(\frac{x_n + x^*}{2}\Big)(x_n - x^*)\Big\| + \\ + \|A_n^{-1}\|\Big\|F'\Big(\frac{x_n + x^*}{2}\Big) - F'\Big(\frac{x_n + y_n}{2}\Big)\Big\|\|x_n - x^*\| + \\ + \|A_n^{-1}\|\|G(x_n; x^*) - G(x_n; y_n)\|\|x_n - x^*\| \le \\ \le B\Big[(p_1 + q_1)\|y_n - x^*\| + \frac{p_2}{4(\alpha+1)(\alpha+2)}\|x_n - x^*\|^{1+\alpha}\Big]\|x_n - x^*\|$$

and

$$\|y_{n+1} - x^*\| = \|x_{n+1} - x^* - A_n^{-1}(F(x_{n+1}) + G(x_{n+1}) - F(x^*) - G(x^*))\| \le \\ \le \|A_n^{-1}\|\Big\|\int_0^1 \Big\{F'(x^* + t(x_{n+1} - x^*)) - F'\Big(\frac{x_n + y_n}{2}\Big)\Big\}dt\Big\|\|x_{n+1} - x^*\| +$$

$$+\|A_n^{-1}\|\|G(x_{n+1};x^*) - G(x_n;y_n)\|\|x_{n+1} - x^*\| \leq$$

$$\leq B(p_1+q_1)[\|y_n - x^*\| + \|x_n - x^*\| + \|x_{n+1} - x^*\|]\|x_{n+1} - x^*\|.$$

We prove that inequalities (8) and (9) are fulfilled. Taking $n = 0$ above, we obtain

$$\|x_1 - x^*\| < B\Big[(p_1+q_1)r_* + \frac{p_2}{4(\alpha+1)(\alpha+2)} r_*^{1+\alpha}\Big]\|x_0 - x^*\| \leq \|x_0 - x^*\| < r_*$$

and

$$\begin{aligned}\|y_1 - x^*\| <& 3B^2(p_1+q_1)\Big[(p_1+q_1)r_* + \frac{p_2}{4(\alpha+1)(\alpha+2)} r_*^{1+\alpha}\Big] r_*\|x_0 - x^*\| \leq \\ \leq& \|x_0 - x^*\| < r_*.\end{aligned}$$

Therefore, $x_1,\ y_1 \in U$. If $\|x_n - x^*\| < r_*$ and $\|y_n - x^*\| < r_*$ then from (7) – (9), it follows

$$\begin{aligned}\|x_{n+1} - x^*\| <& B\Big[(p_1+q_1)r_* + \frac{p_2}{4(\alpha+1)(\alpha+2)} r_*^{1+\alpha}\Big]\|x_n - x^*\| \leq \\ \leq& \|x_n - x^*\| < \ldots < r_*,\end{aligned}$$

$$\begin{aligned}\|y_{n+1} - x^*\| <& 3B^2(p_1+q_1)\Big[(p_1+q_1)r_* + \\ &+ \frac{p_2}{4(\alpha+1)(\alpha+2)} r_*^{1+\alpha}\Big] r_*\|x_n - x^*\| \leq \\ \leq& \|x_n - x^*\| < \ldots < r_*.\end{aligned}$$

So, iterative process (3) is well defined, the sequences $\{x_n\}_{n\geq 0}$, $\{y_n\}_{n\geq 0}$ belong to U. From the last inequalities and estimates (8) and (9) we can see that $\{x_n\}_{n\geq 0}$ and $\{y_n\}_{n\geq 0}$ converge to x^*. □

Corollary 2. *Let us suppose that the hypotheses of Theorem 1 hold. Then the iterative process (3) converges to a solution x^* of the equation (1) with the order of convergence $1+\sqrt{1+\alpha}$.*

Proof. We denote

$$a_n = \|x_n - x^*\|,\ b_n = \|y_n - x^*\|,\ C_1 = B(p_1+q_1),\ C_2 = \frac{Bp_2}{4(\alpha+1)(\alpha+2)}.$$

By (8) and (9), we get

$$a_{n+1} \leq C_1 a_n b_n + C_2 a_n^{2+\alpha},$$

$$\begin{aligned}b_{n+1} \leq C_1(a_{n+1} + a_n + b_n)a_{n+1} \leq C_1(2a_n + b_n)a_{n+1} \leq \\ \leq C_1(2a_n + C_1(2a_0+b_0)a_n)a_{n+1} = C_1(2 + C_1(2a_0+b_0))a_n a_{n+1},\end{aligned}$$

Then for large n and $a_{n-1} < 1$, from previous inequalities, we obtain

$$\begin{aligned}a_{n+1} &\leq C_1 a_n b_n + C_2 a_n^2 a_{n-1}^{\alpha} \leq \\ &\leq C_1^2(2 + C_1(2a_0+b_0))a_n^2 a_{n-1} + C_2 a_n^2 a_{n-1}^{\alpha} \leq \\ &\leq [C_1^2(2 + C_1(2a_0+b_0)) + C_2]a_n^2 a_{n-1}^{\alpha}.\end{aligned} \tag{12}$$

From (12) we can write down an equation of the convergence order of the iterative process (3): $t^2 - 2t - \alpha = 0$. The order of convergence is the unique positive solution $t^* = 1 + \sqrt{1+\alpha}$. If $\alpha = 1$, we get that the iterative process (3) converges to the solution of the equation (1) with the order $1 + \sqrt{2}$. $\square$

Theorem 2. *Let F and G be nonlinear operators, defined on an open convex subset D of a Banach space X with values in a Banach space Y. F is a Fréchet-differentiable operator, G is a continuous operator. We assume that $U_0 = \{x : \|x - x_0\| \le r_0\}$ is contained in D, the linear operator $A_0 = F'\Big(\dfrac{x_0 + y_0}{2}\Big) + G(x_0; y_0)$, where $x_0, \; y_0 \in D$, is invertible and the Lipschitz conditions are fulfilled*

$$\|A_0^{-1}(F'(x) - F'(y))\| \le 2p_0\|x - y\|, \tag{13}$$

$$\|A_0^{-1}(G(x; y) - G(u; v))\| \le q_0(\|x - u\| + \|y - v\|). \tag{14}$$

Let's a, c $(c > a)$, r_0 be non-negative numbers such that

$$\|x_0 - x_{-1}\| \le a, \quad \|A_0^{-1}(F(x_0) + G(x_0))\| \le c, \tag{15}$$

$$r_0 \ge c/(1-\gamma), \quad (p_0 + q_0)(2r_0 - a) < 1,$$

$$\gamma = \frac{(p_0 + q_0)(r_0 - a) + 0.5p_0r_0}{1 - (p_0 + q_0)(2r_0 - a)}, \quad 0 \le \gamma < 1.$$

Then the following inequalities hold for all $n \ge 0$

$$\|x_n - x_{n+1}\| \le t_n - t_{n+1}, \quad \|y_n - x_{n+1}\| \le s_n - t_{n+1}, \tag{16}$$

$$\|x_n - x^*\| \le t_n - t^*, \quad \|y_n - x^*\| \le s_n - t^*, \tag{17}$$

where

$$t_0 = r_0, \quad s_0 = r_0 - a, \quad t_1 = r_0 - c,$$

$$t_{n+1} - t_{n+2} = \frac{(p_0 + q_0)(s_n - t_{n+1}) + 0.5p_0(t_n - t_{n+1})}{1 - (p_0 + q_0)[(t_0 - t_{n+1}) + (s_0 - s_{n+1})]}(t_n - t_{n+1}), \tag{18}$$

$$t_{n+1} - s_{n+1} = \frac{(p_0 + q_0)(s_n - t_{n+1}) + 0.5p_0(t_n - t_{n+1})}{1 - (p_0 + q_0)[(t_0 - t_n) + (s_0 - s_n)]}(t_n - t_{n+1}), \tag{19}$$

$\{t_n\}_{n\ge0}$, $\{s_n\}_{n\ge0}$ are non-negative, decreasing sequences that converge to certain t^* such that $r_0 - c/(1-\gamma) \le t^* < t_0$; sequences $\{x_n\}_{n\ge0}$, $\{y_n\}_{n\ge0}$ generated by the iterative process (3) are well defined, remain in U_0 that converge to a solution x^* of equation (1).

Proof. Firstly, we prove, by mathematical induction, that the following inequalities hold for all $k \ge 0$

$$t_{k+1} \ge s_{k+1} \ge t_{k+2} \ge r_0 - \frac{c}{1-\gamma} \ge 0, \tag{20}$$

$$t_{k+1} - t_{k+2} \le \gamma(t_k - t_{k+1}), \quad t_{k+1} - s_{k+1} \le \gamma(t_k - t_{k+1}). \tag{21}$$

From (18), (19) for $k = 0$ we obtain

$$t_1 - t_2 = \frac{(p_0 + q_0)(s_0 - t_1) + 0.5p_0(t_0 - t_1)}{1 - (p_0 + q_0)[(t_0 - t_1) + (s_0 - s_1)]}(t_0 - t_1) \le \gamma(t_0 - t_1),$$

$$t_1 - s_1 = [(p_0 + q_0)(s_0 - t_1) + 0.5p_0(t_0 - t_1)](t_0 - t_1) \le \gamma(t_0 - t_1),$$

$$t_2 \geq r_0 - c - \frac{(p_0+q_0)s_0 + 0.5p_0t_0}{1-(p_0+q_0)[t_0+s_0]}c =$$
$$= r_0 - (1+\gamma)c = r_0 - \frac{(1-\gamma^2)c}{1-\gamma} \geq r_0 - \frac{c}{1-\gamma} \geq 0,$$
$$t_1 \geq t_2,\ s_1 \geq t_2,\ t_1 \geq s_1 \geq t_2 \geq r_0 - \frac{c}{1-\gamma} \geq 0.$$

Let us suppose that inequalities (20) and (21) hold for $k = 0, 1, \ldots, n-1$. Then for $k = n$ we obtain

$$t_{n+1} - t_{n+2} = \frac{(p_0+q_0)(s_n - t_{n+1}) + 0.5p_0(t_n - t_{n+1})}{1-(p_0+q_0)[(t_0-t_{n+1})+(s_0-s_{n+1})]}(t_n - t_{n+1}) \leq$$
$$\leq \frac{(p_0+q_0)s_n + 0.5p_0t_n}{1-(p_0+q_0)[t_0+s_0]}(t_n - t_{n+1}) \leq \gamma(t_n - t_{n+1}),$$

$$t_{n+1} - s_{n+1} = \frac{(p_0+q_0)(s_n - t_{n+1}) + 0.5p_0(t_n - t_{n+1})}{1-(p_0+q_0)[(t_0-t_n)+(s_0-s_n)]}(t_n - t_{n+1}) \leq$$
$$\leq \frac{(p_0+q_0)s_n + 0.5p_0t_n}{1-(p_0+q_0)[t_0+s_0]}(t_n - t_{n+1}) \leq \gamma(t_n - t_{n+1})$$

and

$$t_{n+1} \geq s_{n+1} \geq t_{n+2} \geq t_{n+1} - \gamma(t_n - t_{n+1}) \geq$$
$$\geq r_0 - \frac{1-\gamma^{n+2}}{1-\gamma}c \geq r_0 - \frac{c}{1-\gamma} \geq 0.$$

So, we prove, that sequences $\{t_n\}_{n\geq 0}$ and $\{s_n\}_{n\geq 0}$ are non-negative, decreasing sequences and converge to t^* such that $t^* \geq 0$.

Let us prove, by mathematical induction, that the iterative process (3) is well defined and inequalities (16) hold for all $n \geq 0$.

Using (15) and $t_0 - t_1 = c$, we prove that (16) hold for $n = 0$.

Let denote $A_n = F'\Big(\frac{x_n + y_n}{2}\Big) + G(x_n; y_n)$. Using Lipschitz conditions (13) and (14), we have

$$\|I - A_0^{-1}A_{n+1}\| = \|A_0^{-1}[A_0 - A_{n+1}]\| \leq$$
$$\leq \Big\|A_0^{-1}\Big[F'\Big(\frac{x_0+y_0}{2}\Big) - F'\Big(\frac{x_{n+1}+y_{n+1}}{2}\Big)\Big]\Big\| +$$
$$+ \|A_0^{-1}[G(x_0; y_0) - G(x_{n+1}; y_{n+1})]\| \leq$$
$$\leq 2p_0\Big(\frac{\|x_0 - x_{n+1}\|}{2} + \frac{\|y_0 - y_{n+1}\|}{2}\Big) + q_0(\|x_0 - x_{n+1}\| + \|y_0 - y_{n+1}\|) \leq$$
$$\leq (p_0+q_0)(\|x_0 - x_{n+1}\| + \|y_0 - y_{n+1}\|) \leq$$
$$\leq (p_0+q_0)(t_0 - t_{n+1} + s_0 - s_{n+1}) \leq$$
$$\leq (p_0+q_0)(t_0+s_0) = (p_0+q_0)(2r_0 - a) < 1.$$

By Banach lema on invertible operator, it follows that A_{n+1} is invertible and

$$\|A_{n+1}^{-1}A_0\| \le \Big[1-(p_0+q_0)(\|x_0-x_{n+1}\|+\|y_0-y_{n+1}\|)\Big]^{-1}.$$

Let us prove that iterative process (3) is well defined for $k=n+1$. From the definition of the first order divided difference and (13), (14), we obtain

$$\|A_0^{-1}(F(x_{n+1})+G(x_{n+1}))\| =$$

$$= \|A_0^{-1}\left[F(x_{n+1})+G(x_{n+1})-F(x_n)-G(x_n)-A_n(x_{n+1}-x_n)\right]\| \le$$

$$\le \Big\|A_0^{-1}\Big[\int_0^1 \Big\{F'(x_{n+1}+t(x_n-x_{n+1}))-F'\Big(\frac{x_n+y_n}{2}\Big)\Big\}dt\Big]\Big\|\|x_n-x_{n+1}\|+$$
$$+\left\|A_0^{-1}\left[G(x_n;y_n)-G(x_n;x_{n+1})\right]\right\|\|x_n-x_{n+1}\| \le$$

$$\le 2p_0\Big[\|x_n-x_{n+1}\|\int_0^1\Big|t-\frac{1}{2}\Big|dt+\frac{\|y_n-x_{n+1}\|}{2}\Big]\|x_n-x_{n+1}\|+$$
$$+q_0\|y_n-x_{n+1}\|\|x_n-x_{n+1}\| =$$

$$= (p_0+q_0)\|y_n-x_{n+1}\|\|x_n-x_{n+1}\|+0.5p_0\|x_n-x_{n+1}\|^2.$$

Hence, using (16), we have

$$\|x_{n+1}-x_{n+2}\| = \|A_{n+1}^{-1}(F(x_{n+1})+G(x_{n+1}))\| \le$$

$$\le \|A_{n+1}^{-1}A_0\|\|A_0^{-1}(F(x_{n+1})+G(x_{n+1}))\|\|x_n-x_{n+1}\| \le$$

$$\le \frac{(p_0+q_0)\|y_n-x_{n+1}\|+0.5p_0\|x_n-x_{n+1}\|}{1-(p_0+q_0)(\|x_0-x_{n+1}\|+\|y_0-y_{n+1}\|)}\|x_n-x_{n+1}\| \le$$
$$\le \frac{(p_0+q_0)(s_n-t_{n+1})+0.5p_0(t_n-t_{n+1})}{1-(p_0+q_0)[(t_0-t_{n+1})+(s_0-s_{n+1})]}(t_n-t_{n+1}) = t_{n+1}-t_{n+2},$$

$$\|x_{n+2}-y_{n+2}\| = \|A_{n+1}^{-1}(F(x_{n+2})+G(x_{n+2}))\| \le$$

$$\le \|A_{n+1}^{-1}A_0\|\|A_0^{-1}(F(x_{n+2})+G(x_{n+2}))\|\|x_n-x_{n+1}\| \le$$

$$\le \frac{(p_0+q_0)\|y_{n+1}-x_{n+2}\|+0.5p_0\|x_{n+1}-x_{n+2}\|}{1-(p_0+q_0)(\|x_0-x_{n+1}\|+\|y_0-y_{n+1}\|)}\|x_{n+1}-x_{n+2}\| \le$$
$$\le \frac{(p_0+q_0)(s_{n+1}-t_{n+2})+0.5p_0(t_{n+1}-t_{n+2})}{1-(p_0+q_0)[(t_0-t_{n+1})+(s_0-s_{n+1})]}(t_{n+1}-t_{n+2}) = s_{n+2}-t_{n+2}.$$

So, iterative process (3) is well defined and (15) holds for all $n \ge 0$. From this it follows

$$\|x_n-x_k\| \le t_n-t_k,\ \|y_n-x_k\| \le s_n-t_k, \|y_n-y_k\| \le s_n-s_k, \quad 0 \le n \le k, \tag{22}$$

i.e., $\{x_n\}_{n\ge0}$ and $\{y_n\}_{n\ge0}$ are fundamental sequences in a Banach space X. From (22) for $k \to \infty$ it follows inequalities (17). Let's show that x^* is solution of equation (1). Indeed,

$$\|A_0^{-1}(F(x_{n+1})+G(x_{n+1}))\| \le$$

$$\le (p_0+q_0)\|y_n-x_{n+1}\|\|x_n-x_{n+1}\|+0.5p_0\|x_n-x_{n+1}\|^2 \to 0, \quad n \to \infty.$$

So, $H(x^*) = 0$. $\square$

Remark 5. *If we choose* $F(x) = 0$, $p_1 = 0$, $p_2 = 0$ *then the estimates (8) and (9) reduce to similar ones in* [8] *for the case* $\alpha = 1$.

Remark 6. *If the divided difference of the operator* G *satisfies the condition (6), i.e. the operator* $G(x; y)$ *is Lipschitz continuous, then* G *is Fréchet-differentiable.*

3. Numerical experiments

For the numerical investigation we choose the equation and the systems of equations considered in [1, 4, 5, 6, 7].

Example 1.

$$e^{x-0.5} - 1.05 + 0.2x|x-1| = 0,$$
$$\boldsymbol{x}^* = 0.5.$$

Example 2.

$$3x^2y - y^2 - 1 + |x-1| = 0,$$
$$x^4 + xy^3 - 1 + |y| = 0,$$
$$(\boldsymbol{x}^*; \boldsymbol{y}^*) \approx (0.894655; 0.327827).$$

Example 3.

$$x^2 - y + 1 + \frac{1}{9}|x-1| = 0,$$
$$y^2 + x - 7 + \frac{1}{9}|y| = 0,$$
$$(\boldsymbol{x}^*; \boldsymbol{y}^*) \approx (1.15936; 2.36182).$$

Example 4.

$$z^2(1-y) - xy + |y - z^2| = 0,$$
$$z^2(x^3 - x) - y^2 + |3y^2 - z^2 + 1| = 0,$$
$$6xy^3 + y^2z^2 - xy^2z + |x + z - y| = 0,$$
$$(\boldsymbol{x}^*; \boldsymbol{y}^*; \boldsymbol{z}^*) = (-1; 2; 3).$$

Let $X = Y = \mathbb{R}^m$, $m = 1, 2, 3$. In this case the first order divided difference $G(x; y)$ is a matrix of dimension $m \times m$. Its elements are calculated as [8]

$$G(x;y)_{i,j} = \frac{G_i(x^1, \ldots, x^j, y^{j+1}, \ldots, y^m) - G_i(x^1, \ldots, x^{j-1}, y^j, \ldots, y^m)}{x^j - y^j},$$
$$i, j = \overline{1, m}.$$

In calculations we use the norm $\|x\|_\infty = \max\limits_{1 \le i \le m} |x^i|$. In the following Tables there are results obtained by methods (3) and (2) in particular, for such cases

$$x_{n+1} = x_n - [F'(x_n)]^{-1}(F(x_n) + G(x_n)), \quad n = 0, 1, \ldots, \tag{23}$$

$$x_{n+1} = x_n - [F'(x_n) + G(x_{n-1}; x_n)]^{-1}(F(x_n) + G(x_n)), \quad n = 0, 1, \ldots, \tag{24}$$

$$x_{n+1} = x_n - [H(x_{n-1}; x_n)]^{-1}(F(x_n) + G(x_n)), \quad n = 0, 1, \ldots. \tag{25}$$

TABL. 1. Numbers of iterations for solving equations with initial points $x_0 = 1 \cdot d$, $x_{-1} = y_0 = 2 \cdot d$ – for Example 1, $x_0 = (1, 0)d$, $x_{-1} = y_0 = (5, 5)d$ – for Example 2

d	ε	**Example 1**				**Example 2**			
		(23)	(24)	(25)	(3)	(23)	(24)	(25)	(3)
1	10^{-5}	5	5	6	5	11	4	5	5
	10^{-15}	6	7	8	6	33	6	9	6
10	10^{-5}	14	15	20	13	19	13	18	12
	10^{-15}	15	17	22	14	41	15	21	13
100	10^{-5}	104	105	–	88	27	21	30	19
	10^{-15}	105	107	–	89	49	23	32	20

The calculations were conducted in MATLAB 7.1. Iterations were stopped after conditions $\|x_{n+1} - x_n\|_\infty \le \varepsilon$ and $\|H(x_{n+1})\|_\infty \le \varepsilon$ were satisfied. Sign "'-"' means, that in this case the solution was not possible to be found. We examined the convergence of the considered method for such variants of choice of the additional initial approximation y_0: for Example 1 – $x_{-1} = y_0 = 2 \cdot d$, for Examples 2, 3 y_0 was chosen as x_{-1} in the works [1, 5, 6, 7] and $x_{-1}^i = y_0^i = x_0^i + 10^{-4}$, $i = 1, 2, 3$ – for Example 4.

The obtained results show that the methods (24) and (3) differ a little for the initial points that are close to the solution. But the method (3) converge faster than (2) for the initial points with $d = 100$. In this case $\|x_0 - x^*\|$ takes the largest value. The method (23) has the lowest speed of convergence.

TABL. 2. Numbers of iterations for solving equations with initial points $x_0 = (1, 1)d$, $x_{-1} = y_0 = (0.9, 1.1)d$ – for Example 3, $x_0 = (-2, 3, 5)d$, $x_{-1}^i = y_0^i = x_0^i + 10^{-4}$ – for Example 4

d	ε	**Example 3**				**Example 4**			
		(23)	(24)	(25)	(3)	(23)	(24)	(25)	(3)
1	10^{-5}	6	5	6	5	85	7	10	7
	10^{-15}	13	7	9	6	266	10	12	8
10	10^{-5}	8	7	9	6	102	10	25	14
	10^{-15}	15	9	11	7	284	20	27	16
100	10^{-5}	11	11	14	9	110	28	39	23
	10^{-15}	18	12	16	10	292	30	41	24

In Table 3 the numerical results are presented for the example 1 with $\varepsilon = 10^{-10}$, where n is the iteration number, x_n is the approximate value for x^*, TABL. 3. Numerical results for the Example 1: $x_0 = 1$, $y_0 = 2$

n	x_n	$\lvert x_n - x_{n-1}\rvert$	$\lvert H(x_n)\rvert$
0	1		0.5987212707001
1	0.8079964212227	0.1920035787772	0.3417237602029
2	0.5200746907444	0.28792173047835	0.02019694382837
3	0.5000182789519	0.02005641179247	$1.827905217970 \cdot 10^{-5}$
4	0.5000000000006	$1.827895124595 \cdot 10^{-5}$	$6.967343368913 \cdot 10^{-13}$
5	0.5	$6.967759702547 \cdot 10^{-13}$	$4.163336342344 \cdot 10^{-17}$

$|x_n - x_{n-1}|$ is the norm of correction and $|H(x_n)|$ is the norm of deviation on every step of the iterative process (3).

Now we verify whether the hypothesis of Theorem 2 are satisfied. The research are carried out for the example 1. Since $m = 1$ than $\|\cdot\|_\infty = |\cdot|$. In [9] we showed that the following estimates hold for all $x,\ y \in [0; 1]$

$$
\begin{aligned}
|A_0^{-1}(F'(x) - F'(y))| \le& |A_0^{-1}||(F'(x) - F'(y))| \le \frac{e^{0.5}}{|A_0|}|x - y|,\\
|A_0^{-1}(G(x,y) - G(u,v))| \le& |A_0^{-1}||(G(x,y) - G(u,v))| \le\\
\le& \frac{1}{5|A_0|}(|x-u| + |y-v|).
\end{aligned}
$$

Hence $p_0 = \dfrac{e^{0.5}}{2|A_0|}$ and $q_0 = \dfrac{1}{5|A_0|}$. Let us choose $x_0 = 0.43$, $y_0 = 0.47$. Then we get

$$
\frac{1}{|A_0|} = 1.049985813745361, \quad p_0 = 0.8655669725276801,
$$
$$
q_0 = 0.2099971627490723, \quad c = 0.07201451611773883, \quad a = 0.04.
$$

Let us choose $r_0 = 0.1$. Then, according to formulas (18) and (19), we get

$$
\begin{aligned}
& t_0 = 0.1000000000000000, \quad s_0 = 0.06000000000000000,\\
& t_1 = 0.0798548388226117, \quad s_1 = 0.02326130579394141,\\
& t_2 = 0.0226355142098747, \quad s_2 = 0.02261740817032270,\\
& t_3 = 0.02261727501017343, \dots, \ t^* \approx 0.02261727484294557,\\
& 0.01720355125317807 < t^* < 0.1, \quad \gamma = 0.1302221628134378 < 1.
\end{aligned}
$$

The solution x^* is obtained in 3 iterations with $\varepsilon = 10^{-5}$.

TABL. 4. Numerical results for the Example 1

n	$\lvert x_{n-1} - x_n\rvert$	$t_{n-1} - t_n$	$\lvert y_{n-1} - x_n\rvert$	$s_{n-1} - t_n$
1	$7.0617898 \cdot 10^{-2}$	$7.2014516 \cdot 10^{-2}$	$3.0617898 \cdot 10^{-2}$	$3.2014516 \cdot 10^{-2}$
2	$6.1790108 \cdot 10^{-4}$	$5.3499697 \cdot 10^{-3}$	$1.8418431 \cdot 10^{-5}$	$6.2579158 \cdot 10^{-4}$
3	$3.4257955 \cdot 10^{-9}$	$1.8239200 \cdot 10^{-5}$	$6.1617378 \cdot 10^{-13}$	$1.3316015 \cdot 10^{-7}$

Thus for the given values hypothesis of the Theorem 2 are satisfied (See Tabl. 4). According to this theorem, the iterative process (2) is well-defined, remains in U_0 and converges to the solution $x^* \in U_0$.

Bibliography

1. Argyros I. K. A unifying local-semilocal convergence analysis and applications for two-point Newton-like methods in Banach space / I. K. Argyros // J. Math. Anal. Appl.– 2004.– Vol. 298.– P. 374-397.
2. Argyros I.K. Improving the rate of convergence of Newton methods on Banach spaces with a convergence structure and applications / I. K. Argyros // Appl. Math. Lett.– 1997.– Vol. 6.– P. 21-28.
3. Argyros I. K. On the convergence of modified Newton methods for solving equations containing a non-differentiable term / I. K. Argyros, H. Ren // J. Comp. App. Math.– 2009.– Vol. 231.– P. 897-906.
4. Chen X. On the convergence of Broyden-like methods for nonlinear equations with non-differentiable terms / X. Chen // Ann. Inst. Statist. Math.– 1990.– Vol. 42, № 2.– P. 387-401.
5. Hernandez M. A. A uniparametric family of iterative processes for solving nondiffrentiable operators / M. A. Hernandez, M. J. Rubio // J. Math. Anal. Appl.– 2002.– Vol. 275.– P. 821-834.
6. Hernandez M. A. The Secant method for nondifferentiable operators / M. A. Hernandez, M. J. Rubio // Appl. Math. Lett.– 2002.– Vol. 15.– P. 395-399.
7. Ren H. A new semilocal convergence theorem for a fast iterative method with nondifferentiable operators / H. Ren, I. K. Argyros // J. Appl. Math. Comp.– 2010.– Vol. 34, № 1-2.– P. 39-46.
8. Shakhno S. M. On an iterative algorithm with superquadratic convergence for solving nonlinear operator equations / S. M. Shakhno // J. Comp. App. Math.– 2009.– Vol. 231.– P. 222-235.
9. Shakhno S. M. Twopoint method for solving nonlinear equation with nondifferentiable operator / S. M. Shakhno, H. P. Yarmola // Matematychni Studii.– 2010.– Vol. 36, № 2.– P. 213–220 (in Ukrainian).
10. Werner W. Über ein Verfarhren der Ordnung $1+\sqrt{2}$ zur Nullstellenbestimmung / W. Werner // Numer. Math.– 1979.– Vol. 32.– P. 333–342.
11. Zabrejko P. P. The majorant method in the theory of Newton-Kantorovich approximations and the Pta'k error estimates / P. P. Zabrejko, D. F. Nguen // Numerical Functional Analysis and Optimization.– 1987.– Vol. 9, № 5-6.– P. 671-684.

Stepan Shakhno, Halina Yarmola,
Ivan Franko National University of Lviv,
1, Universytets'ka Str., Lviv, 79000, Ukraine

Received 12.04.2012

Журнал обчислювальної та прикладної математики | 2012, № 3 (109) с. 116–124 | Journal of Numerical & Applied Mathematics

UDC 519.6

ON THE RECOVERY OF CONTINUOUS FUNCTIONS OF TWO VARIABLES FROM NOISY FOURIER COEFFICIENTS

Kosnazar Sharipov

Резюме. Розглянуто некоректну задачу відновлення гладких функцій двох змінних по наближено заданим коефіцієнтам Фурьє. Ця задача розглянута для двох модельних класів функцій скінченної гладкості: функцій соболевського типу гладкості та функцій з домінуючою змішаною частинною похідною.

Abstract. We consider the ill-posed problem of the recovery of smooth functions of two variables from noisy Fourier coefficients.This problem is considered for two model classes of function of finite smoothness: functions of Sobolev type of smoothness and functions with dominating mixed partial derivative.

1. Introduction

Let $L_2 = L_2(Q_n)$ be the space of square integrable real-valued functions of n variables on a cube $Q_n = [0,1]^n$. Denote by $C = C(Q_n)$ the space of continuous functions on Q_n.

This paper is dedicated to the problem of summation of Fourier series of continuous functions with inaccurately given coefficients. Note that almost all previously known results on this problem were obtained mainly for classes of functions of one variable ($n = 1$).

Let us briefly consider the history of the problem under investigation. Assume that the system of functions $\{\varphi_k(t)\}_{k=1}^{\infty}$ is orthonormal in $L_2(Q_1)$ with respect to the standard scalar product $\langle \cdot, \cdot \rangle$, and $\sum_{k=1}^{\infty} y_k \cdot \varphi_k(t)$ is a Fourier series of the function $y(t) \in C$. Suppose that instead of Fourier coefficients their approximate values $y_{\delta,k}$ are given : the condition

$$\sum_{k=1}^{\infty} (y_k - y_{\delta,k})^2 \le \delta^2$$

is fulfilled.

It is well known (see, for example [6],[7]) that the problem of summation of Fourier series of a continuous function $y(t)$ with approximately given coefficients $\{y_{\delta,k}\}_{k=1}^{\infty}$ on some orthonormal system $\{\varphi_k(t)\}_{k=1}^{\infty}$ is ill-posed, since deviation of a function $y(t) \in C$ of the amount of its series $\sum_{k=1}^{\infty} y_{\delta,k} \cdot \varphi_k(t)$ in the metric of the space C can be arbitrary large.

Papers of the many authors, see, example [1]-[8] are dedicated to the solution of this problem.

†*Key words.* Orthonormal system, stable summation, Fourier series, regularization method.

For the first time to solve this problem A.N. Tikhonov proposed regularization method [7], having the form

$$T^{\alpha,s}(y_\delta)(t) = \sum_{k=1}^{\infty} \frac{y_{\delta,k}}{1+\alpha\cdot k^{2s}}\cdot\varphi_k(t) \tag{1}$$

where α is a regularization parameter, and s characterizes the smoothness of the function to be recovered. Convergence and stability of the method (1) to small perturbations of Fourier coefficients on any orthonormal system $\{\varphi_k(t)\}_{k=1}^{\infty}$ on the class of continuous functions, satisfying the condition

$$\sum_{k=1}^{\infty} |\langle y,\varphi_k\rangle|^2\cdot\psi_k<\infty$$

were proved in [8]. In the condition $\{\psi_k\}_{k=1}^{\infty}$ is a sequence of positive numbers, the order of which is not less than $k^{2+\varepsilon}$, $\varepsilon>0$.

Later V.A. Il'in and E.G. Poznjak [3] in the case of the trigonometric system and B.Aliev [1] in the case of any orthonormal uniform boundary systems for the special classes of functions have obtained the estimate

$$\left\|y(t)-T^{\alpha,s}(y_\delta)(t)\right\|_C \le C\cdot\left(\sqrt{\alpha}+\frac{\delta}{\alpha}\right). \tag{2}$$

We note that one of the major topics within the theory of ill-posed problems is the optimal choice of the regularization parameter α, or the discretization level n, depending on the level of error δ. From (2) one can see that the optimal choice for α is $\alpha_0=\delta^{2/3}$ for which

$$\left\|y(t)-T^{\alpha,s}(y_\delta)(t)\right\|_C \le C\cdot\delta^{\frac{1}{3}}.$$

Later,in [4] P.Mathe and S.V.Pereverzev have considered a general method of summation which is defined as follows

$$T_n^{\lambda}(y_\delta)(t)=\sum_{k=1}^{n}\lambda_k\cdot y_{\delta,k}\cdot\varphi_k(t) \tag{3}$$

where for a triangular array $\lambda=\{\lambda_k=\lambda_k^n,\quad k=1,2,...n,\quad n\in N\}$ it is assumed that there exists a constant C and some number $\theta>0$, such that the condition

$$|1-\lambda_k|\le C\cdot\left(\frac{k}{n}\right)^{\theta}$$

is satisfied. In this case we say that the method of summation (3) is of degree θ.

Error estimates of the method (3) in [4] were obtained for the class

$$W_2^{\mu}=\left\{y\in L_2(Q_1):\quad \|y\|_{\mu}^2=\sum_{k=1}^{n}k^{2\mu}\cdot|\langle y,\varphi_k\rangle|^2<\infty\right\}$$

in the cases of arbitrary orthonormal systems that satisfy various conditions. In particular, in the case of systems of functions $\{\varphi_k(t)\}_{k=1}^{\infty}$ satisfying condition

$$\|\varphi_k\|_C \asymp k^{\beta},\quad \beta\ge 0 \tag{4}$$

the estimate was obtained

$$\|y - T_n^\lambda(y_\delta)\|_C \le C \cdot \delta^{\frac{\mu-\beta-\frac{1}{2}}{\mu}}. \tag{5}$$

In [5] the last result was generalized to the case of a class of continuous functions W_2^ψ related to a given orthonormal system $\{\varphi_k(t)\}_{k=1}^\infty$, satisfying the condition (4) as follows

$$W_2^\psi = \left\{ y \in L_2(Q_1) : \quad \|y\|_\psi^2 = \sum_{k=1}^\infty \psi^2(k) \cdot |\langle y, \varphi_k \rangle|^2 < \infty \right\}.$$

where $\psi(k)$ is some monotone increasing function. At the same,for the method of summation (3) from [4] on a class of functions W_2^ψ the estimate

$$\left\|y - T_n^\lambda(y_\delta)\right\|_C \le C \cdot \delta \cdot \left[\psi^{-1}\left(\frac{1}{\delta}\right)\right]^{\beta+1/2}.$$

was obtained.

The aim in this paper is to obtain results on this problem for some classes of continuous functions of two variables ($n = 2$). Bellow we will consider two model classes of functions of finite smoothness: functions of Sobolev type class and a class of functions with dominating mixed partial derivative.

2. Generalized class of functions with dominating mixed partial derivative

Let $\{\varphi_k(t)\}_{k=1}^\infty$ be an orthonormal system of functions in $L_2(Q_1)$ for which the condition (4) is fulfilled, and

$$\sum_{i=1}^\infty \sum_{j=1}^\infty y_{ij} \cdot \varphi_i(t) \cdot \varphi_j(\tau), \quad y_{ij} = \langle y, \varphi_i \varphi_j \rangle,$$

is Fourier series of a function $y(t, \tau) \in C(Q_2)$.

Suppose that instead of Fourier coefficients $\{y_{ij}\}_{i,j=1}^\infty$ their inaccurate values are given, i.e. a sequence of numbers $y_\delta := \{y_{\delta,i,j}\}_{i,j=1}^\infty$ is given, such that

$$y_{\delta,i,j} = y_{i,j} + \delta \cdot \xi_{i,j}, \quad i, j = 1, 2, ... \tag{6}$$

where $\xi = \{\xi_{i,j}\}_{i,j=1}^\infty$ is a noise. It is assumed that $\delta \in (0, 1)$ and

$$\|\xi\|_{l_2} = \left(\sum_{i=1}^\infty \sum_{j=1}^\infty |\xi_{i,j}|^2\right)^{\frac{1}{2}}.$$

Consider the two-dimensional analogue of the summation method (3) from [4] and [5], which has the form

$$T_n^\lambda(y_\delta)(t, \tau) = \sum_{i=1}^n \sum_{j=1}^n \lambda_{i,j} \cdot y_{\delta,i,j} \cdot \varphi_i(t)\varphi_j(\tau). \tag{7}$$

The quality of the method $T_n^\lambda(y_\delta)$ depends on truncation level n and on the properties of the set $\lambda = \{\lambda_{i,j} = \lambda_{i,j}^n : \quad i, j = 1, 2, ... \quad n \in N\}$. We will assume that there exists a constant C and some $\theta > 0$, such that

$$|1-\lambda_{i,j}| \le C \cdot \left(\frac{ij}{n^2}\right)^{\theta}. \tag{8}$$

In this section we study regularization properties of the summation methods (7) on the class

$$L_2^{\mu} = \left\{ y(t,\tau) \in L_2(Q_2): \quad \|y\|_{\mu,2}^2 = \sum_{i=1}^{\infty}\sum_{j=1}^{\infty} (ij)^{2\mu} \cdot |\langle y, \varphi_i\varphi_j\rangle|^2 < \infty \right\}.$$

It is easy to see that the functions from L_2^{μ} are generalization of a class of functions with dominating mixed partial derivative of degree 2μ.

Lemma 1. *At $\mu > \beta + \frac{1}{2}$ we have the following estimates*

$$\left\| \sum_{i=n+1}^{\infty}\sum_{j=1}^{n} y_{i,j} \cdot \varphi_i(t) \cdot \varphi_j(\tau) \right\|_C \le C \cdot n^{-\mu+\beta+\frac{1}{2}} \cdot \|y\|_{\mu,2},$$

$$\left\| \sum_{i=1}^{n}\sum_{j=n+1}^{\infty} y_{i,j} \cdot \varphi_i(t) \cdot \varphi_j(\tau) \right\|_C \le C \cdot n^{-\mu+\beta+\frac{1}{2}} \cdot \|y\|_{\mu,2}, \tag{9}$$

$$\left\| \sum_{i=n+1}^{\infty}\sum_{j=n+1}^{\infty} y_{i,j} \cdot \varphi_i(t) \cdot \varphi_j(\tau) \right\|_C \le C \cdot n^{-\mu+\beta+\frac{1}{2}} \cdot \|y\|_{\mu,2}. \tag{10}$$

Proof. An application of the Cauchy-Schwarz inequality provides

$$\left\| \sum_{i=n+1}^{\infty}\sum_{j=1}^{n} y_{i,j} \cdot \varphi_i(t) \cdot \varphi_j(\tau) \right\|_C = \left\| \sum_{i=n+1}^{\infty}\sum_{j=1}^{n} (i\cdot j)^{\mu} \cdot y_{i,j} \cdot \frac{\varphi_i(t)\cdot\varphi_j(\tau)}{(i\cdot j)^{\mu}} \right\|_C \le$$

$$\le \left\| \left\{ \sum_{i=n+1}^{\infty}\sum_{j=1}^{n} (i\cdot j)^{2\mu} \cdot |y_{i,j}|^2 \right\}^{\frac{1}{2}} \cdot \left\{ \sum_{i=n+1}^{\infty}\sum_{j=1}^{n} \frac{|\varphi_i(t)\cdot\varphi_j(\tau)|^2}{(i\cdot j)^{2\mu}} \right\}^{\frac{1}{2}} \right\|_C \le$$

$$\le C \cdot \left\| \left\{ \sum_{i=n+1}^{\infty}\sum_{j=1}^{n} \frac{(i\cdot j)^{2\beta}}{(i\cdot j)^{2\mu}} \right\}^{\frac{1}{2}} \right\|_C \cdot \|y\|_{\mu,2} \le$$

$$\le C \cdot \left(\sum_{i=n+1}^{\infty} \frac{1}{i^{2\mu-2\beta}} \right)^{\frac{1}{2}} \cdot \left(\sum_{j=1}^{n} \frac{1}{j^{2\mu-2\beta}} \right)^{\frac{1}{2}} \|y\|_{\mu,2} \le C \cdot n^{-\mu+\beta+\frac{1}{2}} \cdot \|y\|_{\mu,2}\,.$$

The relations (9) and (10) can be proved in the same way. Respectively we have

$$\left\|\sum_{i=1}^{n}\sum_{j=n+1}^{\infty} y_{i,j}\cdot\varphi_i(t)\cdot\varphi_j(\tau)\right\|_C \le$$
$$\le C\cdot\left(\sum_{i=1}^{n}\frac{1}{i^{2\mu-2\beta}}\right)^{\frac{1}{2}}\cdot\left(\sum_{j=n+1}^{\infty}\frac{1}{j^{2\mu-2\beta}}\right)^{\frac{1}{2}}\|y\|_{\mu,2} \le C\cdot n^{-\mu+\beta+\frac{1}{2}}\cdot\|y\|_{\mu,2}$$

and

$$\left\|\sum_{i=n+1}^{\infty}\sum_{j=n+1}^{\infty} y_{i,j}\cdot\varphi_i(t)\cdot\varphi_j(\tau)\right\|_C \le$$
$$\le C\cdot\left(\sum_{i=1}^{n}\frac{1}{i^{2\mu-2\beta}}\right)^{\frac{1}{2}}\cdot\left(\sum_{j=n+1}^{\infty}\frac{1}{j^{2\mu-2\beta}}\right)^{\frac{1}{2}}\|y\|_{\mu,2} \le$$
$$\le C\cdot n^{-2\mu+2\beta+1}\cdot\|y\|_{\mu,2}.$$

The main result of this section is given in the following theorem. □

Theorem 1. *Let for an orthonormal system* $\{\varphi_k(t)\}_{k=1^\infty}$ *the condition (5) is fulfilled. Assume that we have a sequence of noisy values (6) and a priori it is known that* $y\in L_2^\mu(Q_2)$ *at* $\mu>\beta+1/2$ *. Then for the summation method* $T_n^\lambda(y_\delta)$ *of degree* $\theta>\mu$ *at* $n\asymp\delta^{-\frac{2}{2\mu+2\beta+1}}$ *we have the following estimate*

$$\sup_{\|y\|_{\mu,2}\le 1}\ \sup_{\|\xi\|_{l_2}\le 1}\left\|y(t,\tau)-T_n^\lambda(y_\delta)(t,\tau)\right\|_C \le C\cdot\delta^{\frac{\mu-\beta-\frac{1}{2}}{\mu+\beta+\frac{1}{2}}}.$$

Proof. Taking into consideration (6) for $T_n^\lambda(y_\delta)$ we have

$$\left\|y(t,\tau)-T_n^\lambda(y_\delta)(t,\tau)\right\|_C =$$
$$=\left\|\sum_{i=1}^{\infty}\sum_{j=1}^{\infty} y_{i,j}\cdot\varphi_i(t)\cdot\varphi_j(\tau)-\sum_{i=1}^{n}\sum_{j=1}^{n}\lambda_{i,j}\cdot y_{\delta,i,j}\cdot\varphi_i(t)\cdot\varphi_j(\tau)\right\|_C \le$$
$$\le\left\|\sum_{i=n+1}^{\infty}\sum_{j=1}^{n} y_{i,j}\cdot\varphi_i(t)\cdot\varphi_j(\tau)\right\|_C+\left\|\sum_{i=1}^{n}\sum_{j=n+1}^{\infty} y_{i,j}\cdot\varphi_i(t)\cdot\varphi_j(\tau)\right\|_C+$$
$$+\left\|\sum_{i=n+1}^{\infty}\sum_{j=n+1}^{\infty} y_{i,j}\cdot\varphi_i(t)\cdot\varphi_j(\tau)\right\|_C+ \tag{11}$$
$$+\left\|\sum_{i=1}^{n}\sum_{j=1}^{n}(1-\lambda_{i,j})\cdot y_{i,j}\cdot\varphi_i(t)\cdot\varphi_j(\tau)\right\|_C+$$
$$+\delta\cdot\left\|\sum_{i=1}^{n}\sum_{j=1}^{n}\lambda_{i,j}\cdot\xi_{i,j}\cdot\varphi_i(t)\cdot\varphi_j(\tau)\right\|_C.$$

First we estimate the fourth summand of the relation (11)

$$\Big\| \sum_{i=1}^{n}\sum_{j=1}^{n}(1-\lambda_{i,j})\cdot y_{i,j}\cdot\varphi_i(t)\cdot\varphi_j(\tau)\Big\|_C \le$$

$$\le \Big\|\Big\{\sum_{i=1}^{n}\sum_{j=1}^{n}(1-\lambda_{i,j})^2\cdot|y_{i,j}|^2\Big\}^{\frac12}\Big\{\sum_{i=1}^{n}\sum_{j=1}^{n}|\varphi_i(t)\cdot\varphi(\tau)|^2\Big\}^{\frac12}\Big\|_C \le$$

$$\Big\|\Big\{\sum_{i=1}^{n}\sum_{j=1}^{n}(ij)^{2\mu}\cdot|y_{i,j}|^2\cdot(1-\lambda_{i,j})^2\cdot\frac{1}{(ij)^{2\mu}}\Big\}^{\frac12}\Big\|_C\cdot\Big\{\sum_{i=1}^{n}\sum_{j=1}^{n}(ij)^{2\beta}\Big\}^{\frac12} \le$$

$$\le \Big\|\Big\{\sum_{i=1}^{n}\sum_{j=1}^{n}(ij)^{2\mu}\cdot|y_{i,j}|^2\cdot\max_{1\le i,j\le n}\Big[\Big(\frac{ij}{n^2}\Big)^{2\theta}\cdot\frac{1}{(ij)^{2\mu}}\Big]\Big\}^{\frac12}\Big\|_C\times \tag{12}$$

$$\times\Big(\sum_{i=1}^{n}i^{2\beta}\Big)\cdot\Big(\sum_{j=1}^{n}j^{2\beta}\Big)\le$$

$$\le C\cdot n^{-2\mu}\cdot\|y\|_{\mu,2}\cdot n^{2\beta+1}=C\cdot n^{-2\mu+2\beta+1}\cdot\|y\|_{\mu,2}\,.$$

For the last summand of the relation (11) we have

$$\delta\cdot\Big\|\sum_{i=1}^{n}\sum_{j=1}^{n}\lambda_{i,j}\cdot\xi_{i,j}\cdot\varphi_i(t)\cdot\varphi_j(\tau)\Big\|_C\le$$

$$\le\delta\cdot\Big\|\Big\{\sum_{i=1}^{n}\sum_{j=1}^{n}|\lambda_{i,j}|^2\cdot|\xi_{i,j}|^2\Big\}^{\frac12}\cdot\Big\{\sum_{i=1}^{n}\sum_{j=1}^{n}|\varphi_i(t)\cdot\varphi(\tau)|^2\Big\}^{\frac12}\Big\|_C\le \tag{13}$$

$$\le C\cdot\delta\cdot\Big\|\Big\{\sum_{i=1}^{n}\sum_{j=1}^{n}|\xi_{i,j}|^2\Big\}^{\frac12}\Big\|_C\cdot n^{2\beta+1}\le C\cdot\delta\cdot n^{2\beta+1}\,.$$

Summarizing the estimates (12), (13) and the results of the Lemma 1 for $\|y\|_{\mu,2}\le 1$ we have

$$\big\|y(t,\tau)-T_n^{\lambda}(y_\delta)(t,\tau)\big\|_C\le C\cdot n^{2\beta+1}(n^{-\mu-\beta-\frac12}+\delta)\,.$$

If we choose n such that $n\asymp\delta^{-\frac{2}{2\mu+2\beta+1}}$, it follows

$$\big\|y(t,\tau)-T_n^{\lambda}(y_\delta)(t,\tau)\big\|_C\le C\cdot\delta^{\frac{\mu-\beta-\frac12}{\mu+\beta+\frac12}}\,.$$

The theorem is proved. □

3. Class of functions of Sobolev type of smoothness

In this section we study the approximating properties of the summation method (7) on the class of functions of Sobolev type, which has the following form

$$W_2^{\mu}(Q_2):=\Big\{y\in L_2(Q_2):\quad \|y\|^2_{W_2^{\mu}}=\sum_{i=1}^{\infty}\sum_{j=1}^{\infty}\big(i^{2\mu}+j^{2\mu}\big)\cdot|y_{i,j}|^2<\infty\Big\}.$$

To prove the main result of this section we need the following lemma.

Lemma 2. *For* $y \in W_2^\mu$ *at* $\mu > 2\beta + 1$ *we have the following estimates*

$$\left\| \sum_{i=n+1}^{\infty} \sum_{j=1}^{n} y_{i,j} \cdot \varphi_i(t) \cdot \varphi_j(\tau) \right\|_C \leq C \cdot n^{-\frac{\mu}{2}+\beta+\frac{1}{2}} \cdot \|y\|_{W_2^\mu},$$

$$\left\| \sum_{i=1}^{n} \sum_{j=n+1}^{\infty} y_{i,j} \cdot \varphi_i(t) \cdot \varphi_j(\tau) \right\|_C \leq C \cdot n^{-\frac{\mu}{2}+\beta+\frac{1}{2}} \cdot \|y\|_{W_2^\mu}, \tag{14}$$

$$\left\| \sum_{i=n+1}^{\infty} \sum_{j=n+1}^{\infty} y_{i,j} \cdot \varphi_i(t) \cdot \varphi_j(\tau) \right\|_C \leq C \cdot n^{-\mu+2\beta+\frac{1}{2}} \cdot \|y\|_{W_2^\mu}. \tag{15}$$

Proof. Applying Cauchy-Schwarz inequality for $y \in W_2^\mu(Q_2)$ at $\mu > 2\beta + 1$ we have the following

$$\left\| \sum_{i=n+1}^{\infty} \sum_{j=1}^{n} y_{i,j} \cdot \varphi_i(t) \cdot \varphi_j(\tau) \right\|_C =$$

$$= \left\| \sum_{i=n+1}^{\infty} \sum_{j=1}^{n} \left(i^{2\mu} + j^{2\mu}\right)^{\frac{1}{2}} \cdot y_{i,j} \cdot \frac{\varphi_i(t) \cdot \varphi_j(\tau)}{\left(i^{2\mu} + j^{2\mu}\right)^{\frac{1}{2}}} \right\|_C \leq$$

$$\leq \left\| \left\{ \sum_{i=n+1}^{\infty} \sum_{j=1}^{n} \left(i^{2\mu} + j^{2\mu}\right) \cdot |y_{i,j}|^2 \right\}^{\frac{1}{2}} \left\{ \sum_{i=n+1}^{\infty} \sum_{j=1}^{n} \frac{|\varphi_i(t) \cdot \varphi_j(\tau)|}{\left(i^{2\mu} + j^{2\mu}\right)} \right\}^{\frac{1}{2}} \right\|_C \leq$$

$$\leq C \cdot \|y\|_{W_2^\mu} \cdot \left\| \left\{ \sum_{i=n+1}^{\infty} \sum_{j=1}^{n} \frac{i^{2\beta} j^{2\beta}}{2 i^\mu j^\mu} \right\}^{\frac{1}{2}} \right\|_C \leq$$

$$\leq C \cdot \|y\|_{W_2^\mu} \cdot \left\| \left(\sum_{i=n+1}^{\infty} \frac{1}{i^{\mu-2\beta}} \right)^{\frac{1}{2}} \cdot \left(\sum_{j=1}^{n} \frac{1}{j^{\mu-2\beta}} \right)^{\frac{1}{2}} \right\|_C \leq C \cdot n^{-\frac{\mu}{2}+\beta+\frac{1}{2}} \cdot \|y\|_{W_2^\mu}.$$

Similarly, for proofs of (14) and (15) respectively we have

$$\left\| \sum_{i=1}^{n} \sum_{j=n+1}^{\infty} y_{i,j} \cdot \varphi_i(t) \cdot \varphi_j(\tau) \right\|_C \leq$$

$$\leq C \cdot \|y\|_{W_2^\mu} \cdot \left\| \left(\sum_{i=1}^{n} \frac{1}{i^{\mu-2\beta}} \right)^{\frac{1}{2}} \cdot \left(\sum_{j=n+1}^{\infty} \frac{1}{j^{\mu-2\beta}} \right)^{\frac{1}{2}} \right\|_C \leq$$

$$\leq C \cdot n^{-\frac{\mu}{2}+\beta+\frac{1}{2}} \cdot \|y\|_{W_2^\mu}$$

and

$$\left\| \sum_{i=n+1}^{\infty} \sum_{j=n+1}^{\infty} y_{i,j} \cdot \varphi_i(t) \cdot \varphi_j(\tau) \right\|_C \leq$$

$$\leq C \cdot \|y\|_{W_2^\mu} \cdot \left\| \left(\sum_{i=n+1}^{\infty} \frac{1}{i^{\mu-2\beta}} \right)^{\frac{1}{2}} \cdot \left(\sum_{j=n+1}^{\infty} \frac{1}{j^{\mu-2\beta}} \right)^{\frac{1}{2}} \right\|_C \leq$$

$$\leq C \cdot n^{-\mu+2\beta+1} \cdot \|y\|_{W_2^\mu}.$$

The lemma is proved. □

Now we state the main result of the section.

Theorem 2. *Let for an orthonormal system* $\{\varphi_k(t)\}_{k=1^\infty}$ *the condition (5) is fulfilled. Assume that we have a sequence of noisy values (6) and a priori it is known that* $y \in W_2^\mu(Q_2)$ *for* $\mu > 2\beta + 1$ *. Then for the summation method* $T_n^\lambda(y_\delta)$ *of degree* $\theta > \frac{\mu}{2}$ *at* $n \asymp \delta^{-\frac{2}{\mu+2\beta+1}}$ *we have the estimate*

$$\sup_{\|y\|_{\mu,2}\leq 1} \sup_{\|\xi\|_{l_2}\leq 1} \left\| y(t,\tau) - T_n^\lambda(y_\delta)(t,\tau) \right\|_C \leq C \cdot \delta^{\frac{\mu-2\beta-1}{\mu+2\beta+1}}.$$

Proof. To prove the theorem we need to estimate the summands of the relations (11) for $y \in W_2^\mu(Q_2)$, $\mu > 2\beta+1$. Using the fact that for $T_n^\lambda(y_\delta)$ the condition (8) is fulfilled and $\theta > \mu/2$ we estimate the fourth summand of the relation (11)

$$\left\| \sum_{i=1}^{n} \sum_{j=1}^{n} (1-\lambda_{i,j}) \cdot y_{i,j} \cdot \varphi_i(t) \cdot \varphi_j(\tau) \right\|_C \leq$$

$$\leq \left\| \left\{ \sum_{i=1}^{n} \sum_{j=1}^{n} (1-\lambda_{i,j})^2 \cdot |y_{i,j}|^2 \right\}^{\frac{1}{2}} \left\{ \sum_{i=1}^{n} \sum_{j=1}^{n} |\varphi_i(t) \cdot \varphi(\tau)|^2 \right\}^{\frac{1}{2}} \right\|_C \leq$$

$$\leq \left\| \left\{ \sum_{i=1}^{n} \sum_{j=1}^{n} (i^{2\mu}+j^{2\mu}) |y_{i,j}|^2 \frac{|1-\lambda_{i,j}|^2}{(i^{2\mu}+j^{2\mu})} \right\}^{\frac{1}{2}} \right\|_C \times$$

$$\times \left\| \left\{ \sum_{i=1}^{n} \sum_{j=1}^{n} |\varphi_i(t) \cdot \varphi(\tau)|^2 \right\}^{\frac{1}{2}} \right\|_C \leq \tag{16}$$

$$\leq C \cdot \left\| \left\{ \sum_{i=1}^{n} \sum_{j=1}^{n} (i^{2\mu}+j^{2\mu}) |y_{i,j}|^2 \cdot \max_{1\leq i,j\leq n} \left[\frac{1}{(ij)^\mu} \cdot \left(\frac{ij}{n^2} \right)^{2\theta} \right] \right\}^{\frac{1}{2}} \right\|_C \times$$

$$\times \left\| \left(\sum_{i=1}^{n} i^{2\beta} \right)^{\frac{1}{2}} \left(\sum_{j=1}^{n} j^{2\beta} \right)^{\frac{1}{2}} \right\|_C \leq$$

$$\leq C \cdot n^{-\mu} \cdot \left\| \left\{ \sum_{i=1}^{n} \sum_{j=1}^{n} (i^{2\mu}+j^{2\mu}) |y_{i,j}|^2 \right\}^{\frac{1}{2}} \right\|_C \cdot n^{2\beta+1} \leq C \cdot n^{-\mu+2\beta+1} \cdot \|y\|_{W_2^\mu}.$$

Taking into consideration that from (8) it follows that $|\lambda_{i,j}| < \infty$, then for the last summand of the relation (11) we have

$$\delta \cdot \left\| \sum_{i=1}^{n} \sum_{j=1}^{n} \lambda_{i,j} \cdot \xi_{i,j} \cdot y_{i,j} \cdot \varphi_i(t) \cdot \varphi_j(\tau) \right\|_C \le C \cdot \cdot \delta n^{2\beta+1}. \qquad (17)$$

Summarizing the estimates (16), (17) and the estimates from the Lemma 3.1 for $y \in W_2^{\mu}(Q_2)$ we obtain

$$\left\| y(t,\tau) - T_n^{\lambda}(y_\delta)(t,\tau) \right\|_C \le C \cdot n^{2\beta+1} \left(n^{-\frac{\mu}{2}-\beta-\frac{1}{2}} + \delta \right).$$

When choosing $n \asymp \delta^{-\frac{2}{\mu+2\beta+1}}$ it follows that

$$\left\| y(t,\tau) - T_n^{\lambda}(y_\delta)(t,\tau) \right\|_C \le C \cdot \delta^{\frac{\mu-2\beta-1}{\mu+2\beta+1}}.$$

Thus, the theorem is proved. □

Bibliography

1. Aliev B. Estimation of the regularization method for the problem of the Fourier Series // Dokl. Akad. Nauk Tadzhik. SSR.– 1978.– Vol. 21.– P. 3-6.
2. Bakushinskii A. N. On the Problem of Stable summation of Fourier Series / A. V. Bakushinskii // Dokl. Akad. Nauk.– 2009.– Vol. 425, № 2.– P. 151-154.
3. Il'in V. A. Foundation of Mathematical Analysis / V. A. Il'in, E. G. Poznjak.– Nauka: Moscow, 1973
4. Mathe P. Stable Summation of othogonal Series with Noisy Coefficints / P. Mathe, S. V. Pereverzev // Journal of Approximation Theory.– 2002.– Vol. 118.– P. 2060-2076.
5. Sharipov K. On the Recovery of Continuous Functions from Noisy Fourier Coefficints / P. Sharipov // Computational Methods in Applied Mathematics.– 2011.– Vol. 11.– P. 75-82.
6. Tikhonov A. N. On stable methods of summation of the Fourier Series / A. N. Tikhonov // Dokl. Akad. Nauk SSSR.– 1964.–Vol. 156.–No. 2. P. 268-271.
7. Tikhonov A. N. On regularization of ill-posed Problems / A. N. Tikhonov // Dokl. Akad. Nauk SSSR.– 1963.– Vol. 153, № 1.– P. 49-52.
8. Tikhonov A. N. Solutions of ill-posed Problems / A. N. Tikhonov, V. Y. Arsenin.– Nauka: Moscow, 1974.

Kosnazar Sharipov,
Institute of Mathematics, National Academy of Sciences,
3, Tereschenkivska Str., Kiev, 01601, Ukraine

Received 04.04.2012

Журнал обчислювальної та прикладної математики | **2012, № 3 (109)** с. 125–132 | Journal of Numerical & Applied Mathematics

UDC 519.642

ON ACCURACY OF SOLVING SEMIDISCRETE ILL-POSED PROBLEMS IN SOBOLEV SPACES WITH ν-METHODS

Sergei Solodky, Evgeny Volynets

Резюме. Для розв'язування некоректної задачі у соболівських шкалах, отриманої в результаті застосування методу колокації до інтегрального рівняння Фредгольма першого роду, використано поєднання ν-методів з принципом балансу як апостеріорним правилом вибору параметра регуляризації.

Abstract. To solve ill-posed problem in Sobolev scales appearing as a result of application by a collocation method to Fredholm integral equation of the first kind a combination of ν-methods with balancing principle as an a-posteriori regularization parameter choice rule is used.

1. Introduction

Let us consider an equation

$$Af = g \tag{1}$$

with integral operator A defined as

$$Af(x) := \int_\Omega k(x,t)f(t)dt, \qquad x \in \Omega.$$

Here $\Omega \subset \mathbb{R}^d$ is a bounded domain with a Lipschitz continuous boundary and kernel $k(x,t) : \Omega \times \Omega \to \mathbb{R}$ is such that A is compact operator with infinite dimensional range acting from $L_2 = L_2(\Omega)$ into L_2. Without loss of generality we may assume that $\|A\| \leq 1$.

To guarantee a stable solution some regularization method should be used. In the paper we use ν-methods, but regularization process will be done not for original problem (1) but for semi-discrete equation obtained from it by collocation scheme. Let $X = \{x_1, \dots, x_n\} \subset \Omega$ be some set of pairwise distinct points. Consider an equation

$$A_X f = \bar{g}, \tag{2}$$

where $\bar{g} = \{g_1, \dots, g_n\}^T$, $g_j = g(x_j)$, and A_X is defined as

$$(A_X f)_j = Af(x_j), \quad 1 \leq j \leq n,$$

i.e. A_X is the restriction of A to set X ($A_X f = Af|_X$). To obtain good approximation to exact solution in the framework of ν-methods it is important to choose regularization parameter in properly way. In this case regularization parameter is the number of iteration step. As a rule we use balancing principle (see [5], [7]).

[†]*Key words.* Inverse problems, ν-methods, Sobolev scales, collocation method, a-posteriori parameter choice, error bound.

In practice exact right-hand side of (1) is usually unavailable and only noisy data vector $\bar{g}^\delta = \{g_1^\delta, \dots, g_n^\delta\}^T$ such that

$$|g_j - g_j^\delta| \le \delta, \;\; j = \overline{1,n}$$

is known. Let n-dimensional Euclidean space $\mathbb{R}^n$ provided with standard norm $\|\cdot\|_{\mathbb{R}^n}$ and corresponding inner product $\langle \cdot, \cdot \rangle_{\mathbb{R}^n}$. Then the whole data error can be estimated as

$$\|\bar{g} - \bar{g}^\delta\|_{\mathbb{R}^n} \le \delta\sqrt{n}.$$

Our aim is stable recovery of unknown solution of (2) from noisy values $\bar{g}^\delta$.

2. ν-methods in Sobolev scales

Following [2] we assume that A acts along scale of Sobolev spaces $\mathcal{H}^\tau$, $\tau \ge d/2$, with step $\alpha > 0$ i.e. there are constants $c'' \ge c' \ge 0$ such that for fixed $\alpha \in \mathbb{R}$

$$c'\|f\|_\tau \le \|Af\|_{\tau+\alpha} \le c''\|f\|_\tau. \tag{3}$$

Recall that Sobolev space $H^\tau = H^\tau(\Omega)$ is completion in norm of space of square-summable function in Ω together with derivatives of order τ, and $\mathcal{H}^0 = L_2(\Omega)$.

For the first time ill-posed problems in Hilbert scales was considered in [6]. But in the paper we consider the case of discretization by projection methods. The first result in Hilbert scales for the case of discretization by collocation method was obtained recently in [2] where a-priori rule is used for regularization parameter choice. We consider aposteriory rule for choosing the parameter, i.e. without information about smoothing of exact solution.

Let f^* be an exact solution of original problem (1). Then f^* also solves semi-discrete problem (2) and can be represented in the form

$$f^* = f_\delta + v_0,$$

where $f_\delta = A_X^\dagger \bar{g}$, $A_X^\dagger$ is the Moore-Penrose generalized inverse of A_X, and v_0 belongs to the null space of A_X.

We will obtain approximation to solution f_δ. Since A_X acts from $\mathcal{H}^\tau$ into $\mathbb{R}^n$ than $f_\delta \in \mathcal{H}^\tau$ for some $\tau > 0$.

In [3] was shown that always exists some continuous increasing index function $\phi(\lambda)$, $\lambda \in [0,1]$, such that $\phi(0) = 0$ and

$$f_\delta = \phi(A_X^* A_X)v, \tag{4}$$

where $v \in \mathcal{H}^\tau$, $\|v\|_\tau \le \rho$, $\rho > 0$, and $A_X^* : \mathbb{R}^n \to \mathcal{H}^\tau$ is the adjoint of A_X. Later we assume that (4) is fulfills.

Recall that ν-methods is the process of successive computation of elements f_k^δ, $k = 1, 2, \dots$ by the rule

$$f_k^\delta = p_k(A_X^* A_X)A_X^* \bar{g}^\delta,$$

where $\{p_k\}$ is some series of the polynomials of order $k-1$. Consider one more polynomial:

$$r_k(\lambda) := 1 - \lambda p_k(\lambda).$$

It is easy to obtain that for $f_k = p_k(A_X^* A_X)A_X^* \bar{g}$ we have

$$f_k - f_k^\delta = p_k(A_X^* A_X)A_X^*(\bar{g} - \bar{g}^\delta),$$

$$f_\delta - f_k = r_k(A_X^* A_X) f^\dagger,$$

$$\sup_{0\le\lambda\le1} \sqrt{\lambda} p_k(\lambda) \le 2k, \tag{5}$$

$$\sup_{0\le\lambda\le1} \lambda p_k(\lambda) \le 2, \tag{6}$$

$$|\lambda^\mu r_k(\lambda)| \le c_\mu k^{-2\mu}, \tag{7}$$

$$|r_k(\lambda)| \le 1, \tag{8}$$

where $\lambda \in [0,1]$, $c_\mu > 0$ is some constant, $0 < \mu \le \nu$.

3. Auxiliary assertions

Lemma 1. *If $\frac{\phi(t)}{t^{\nu-1/2}}$ is the decreasing function then estimations*

$$\|f_\delta - f_k\|_\tau \le \varkappa\rho\phi(k^{-2}), \tag{9}$$

$$\|A_X f_\delta - A_X f_k\|_{\mathbb{R}^n} \le c_\nu \rho k^{-1}\phi(k^{-2}), \tag{10}$$

are hold, where c and c_ν are some constants.

Proof. In [1, Theorem 6.15] the estimate

$$\|f_\delta - f_k\| \le \varkappa\|f_\delta - f_{\gamma_k,\nu}\|,$$

is obtained, where $f_{\gamma_k,\nu} = \sum_{i=1}^{\nu} \gamma_k^{i-1}(A_X^* A_X + \gamma_k I)^{-i} A_X^* \bar{g}$ is the approximate solution obtained by iterated Tikhonov method of order ν (ν is integer), $\gamma_k \in [(k+1)^{-2}, k^{-2}]$, and $\varkappa$ is a constant. On the other hand, it is easy to show that

$$\|f_\delta - f_{\gamma_k,\nu}\|_\tau \le \rho\phi(\gamma_k).$$

So, first statement of the Lemma is proved.

Further

$$\|A_X f_\delta - A_X f_k\|_{\mathbb{R}^n} = \|A_X(f_\delta - f_k)\|_{\mathbb{R}^n} = \|A_X r_k(A_X^* A_X) f_\delta\|_{\mathbb{R}^n} =$$
$$= \|A_X r_k(A_X^* A_X)\phi(A_X^* A_X) v\|_{\mathbb{R}^n} \le \|v\|_\tau \sup_{0\le\lambda\le1} \sqrt{\lambda} r_k(\lambda)\phi(\lambda).$$

To estimate expression in the right-hand side we consider two cases.

1. $\lambda \le k^{-2}$. Due to (8) and increase of the function ϕ we have

$$\sqrt{\lambda} r_k(\lambda)\phi(\lambda) \le k^{-1}\phi(k^{-2}).$$

2. $k^{-2} \le \lambda$. Due to decrease of the function $\frac{\phi(t)}{t^{\nu-1/2}}$ and (7) we obtain

$$\sqrt{\lambda} r_k(\lambda)\phi(\lambda) = \sqrt{\lambda} r_k(\lambda)\lambda^{\nu-1/2}\frac{\phi(\lambda)}{\lambda^{\nu-1/2}} \le \sqrt{\lambda} r_k(\lambda)\lambda^{\nu-1/2}\frac{\phi(k^{-2})}{k^{-2(\nu-1/2)}} \le$$
$$\le \frac{\lambda^\nu r_k(\lambda)}{k^{-2\nu}} k^{-1}\phi(k^{-2}) \le c_\nu k^{-1}\phi(k^{-2}).$$

Hence,

$$\|A_X f_\delta - A_X f_k\|_{\mathbb{R}^n} \le c_\nu \rho k^{-1}\phi(k^{-2})$$

and Lemma is proved. □

Remark 7. *It is follows from decreasing of* $\frac{\phi(t)}{t^{\nu-1/2}}$ *that in the case of* $\phi(\gamma)=\gamma^{\beta}$ *the restriction is arisen* β*:* $0\le\beta\le\nu-1/2$*.*

Lemma 2. *Following estimations*

$$\|f_k-f_k^\delta\|_\tau\le 2k\delta\sqrt{n}, \tag{11}$$

$$\|A_Xf_k-A_Xf_k^\delta\|_{\mathbb{R}^n}\le 2\delta\sqrt{n}. \tag{12}$$

are hold.

Proof. Due to (5) we obtain the following estimation

$$\|f_k-f_k^\delta\|_\tau=\|p_k(A_X^*A_X)A_X^*(\bar g-\bar g^\delta)\|_\tau\le$$
$$\le\|\bar g-\bar g^\delta\|_{\mathbb{R}^n}\sup_{0\le\lambda\le1}\sqrt{\lambda}p_k(\lambda)\le 2k\delta\sqrt{n}$$

and the first statement of Lemma is proved.

Later due to (6) we have

$$\|A_Xf_k-A_Xf_k^\delta\|_{\mathbb{R}^n}=\|A_Xp_k(A_X^*A_X)A_X^*(\bar g-\bar g^\delta)\|_{\mathbb{R}^n}\le$$
$$\le\|\bar g-\bar g^\delta\|_{\mathbb{R}^n}\sup_{0\le\lambda\le1}\lambda p_k(\lambda)\le 2\delta\sqrt{n}.$$

Thus, Lemma is proved. □

Define data density of the set X in domain Ω as

$$h:=\sup_{x\in\Omega}\ \min_{x_i\in X}\ \|x-x_i\|_{\mathbb{R}^d}.$$

Below we need the sampling inequality obtained in [2, Theorem 4.8]. Namely, for arbitrary function $u\in\mathcal{H}^\theta=\mathcal{H}^\theta(\Omega)$, $\theta>d/2$ and sufficiently small h it is true

$$\|u\|_\sigma\le\kappa\left(h^{\theta-\sigma}\|u\|_\theta+h^{\frac{d}{2}-\sigma}\|u|_X\|_{\mathbb{R}^n}\right), \tag{13}$$

where $\sigma\in[0,\lfloor\theta\rfloor)$, and κ is some constant, doesn't depending on u and h.

4. Error estimate

Theorem 1. *Let (3) is true. Then for any discrete set* X *with sufficiently small data density* h *there is constant* $c_1>0$ *such that*

$$\|f_\delta-f_k^\delta\|_{L_2}\le c_1(h^\tau\left(\varkappa\rho\phi(k^{-2})+2\delta k\sqrt{n}\right)+ \tag{14}$$
$$+h^{\frac{d}{2}-\alpha}\left(c_\nu\rho k^{-1}\phi(k^{-2})+2\delta\sqrt{n}\right)).$$

Proof. First of all we estimate $\|f_\delta-f_k^\delta\|_\tau$. Due to (9) and (11) we have

$$\|f_\delta-f_k^\delta\|_\tau\le\|f_\delta-f_k\|_\tau+\|f_k-f_k^\delta\|_\tau\le\varkappa\rho\phi(k^{-2})+2k\delta\sqrt{n}. \tag{15}$$

Using the sampling inequality (13) with $u=A(f_\delta-f_k^\delta)$, $\sigma=\alpha$ and $\theta=\tau+\alpha$ we obtain

$$\|A(f_\delta-f_k^\delta)\|_\alpha\le\kappa\left(h^\tau\|A(f_\delta-f_k^\delta)\|_{\tau+\alpha}+h^{\frac{d}{2}-\alpha}\|A(f_\delta-f_k^\delta)|_X\|_{\mathbb{R}^n}\right).$$

Now we apply condition (3) to last inequality

$$c'\|f_\delta-f_k^\delta\|_{L_2}\le\kappa\left(c''h^\tau\|f_\delta-f_k^\delta\|_\tau+h^{\frac{d}{2}-\alpha}\|A(f_\delta-f_k^\delta)|_X\|_{\mathbb{R}^n}\right).$$

Taking into account that $Af|_X = A_X f$, we obtain

$$\|f_\delta - f_k^\delta\|_{L_2} \le c_1 \left(h^\tau \|f_\delta - f_k^\delta\|_\tau + h^{\frac{d}{2}-\alpha} \|A_X(f_\delta - f_k^\delta)\|_{\mathbb{R}^n} \right),$$

where $c_1 = \frac{\kappa}{c'} \max\{1, c''\}$.

Considering estimates (10), (12), (15) we have Theorem's statement. □

Let partition of the set X is uniform, i.e. $h = \chi n^{-\frac{1}{d}}$ for some constant χ. Then inequality (14) can be rewritten as

$$\|f_\delta - f_k^\delta\|_{L_2} \le \Phi(k) + \Psi(k),$$

where

$$\Phi(k) := c_2 \rho \left(\chi^{\frac{d}{2}-\alpha} n^{\frac{\alpha}{d}-\frac{1}{2}} k^{-1} \phi(k^{-2}) + \chi^\tau n^{-\frac{\tau}{d}} \phi(k^{-2}) \right),$$

$$\Psi(k) := c_2 \left(\chi^{\frac{d}{2}-\alpha} n^{\frac{\alpha}{d}} \delta + \chi^\tau n^{-\frac{\tau}{d}} k \delta \sqrt{n} \right),$$

and $c_2 = c_1 \max\{2, \varkappa, c_\nu\}$.

It is obvious that due to the monotonicity of ϕ the function Φ is increasing and Ψ is decreasing. Herewith optimal value of the regularization parameter $\gamma = \gamma_{opt}$ balances functions Φ and Ψ, i.e. $\Phi(\gamma_{opt}) = \Psi(\gamma_{opt})$ and

$$\|f_\delta - f_{k_{opt}}^\delta\|_{L_2} \le 2\Phi(k_{opt}).$$

In the case of unknown function ϕ such apriory rule for choosing regularization parameter is inapplicable so it is necessary to use one of the aposteriory rules. As a rule we use balancing principle.

Take into consideration following sets

$$\Delta_N = \left\{1, \dots, N, \qquad N \asymp (\delta\sqrt{n})^{-1}\right\}, \tag{16}$$

and

$$M^+(\Delta_N) = \left\{ k \in \Delta_N : \|f_k^\delta - f_l^\delta\|_{L_2} \le 4\Psi(l), \quad l = k, \dots, N \right\}.$$

To obtain approximate solution we use as regularization parameter such element

$$k = k_+ := \min\left\{k \in M^+(\Delta_N)\right\}.$$

Let us consider one more set

$$M(\Delta_N) := \{k \in \Delta_N : \quad \Phi(k) \le \Psi(k)\}$$

and define

$$k_* := \min\{k \in M(\Delta_N)\}.$$

Without loss of generality we assume that $M(\Delta_N) \ne \emptyset$ and $\Delta_N \backslash M(\Delta_N) \ne \emptyset$.

Theorem 2. *Let the set Δ_N is defined as (16). Then for regularization parameter $k = k_+$ following estimate*

$$\|f_\delta - f_{k_+}^\delta\|_{L_2} \le 6q\Phi(k_{opt}), \tag{17}$$

holds, where $2 \ge q \ge \frac{k_+}{k_+ - 1}$.

Proof. From the beginning we show that $k_* \le k_+$. For any element $l > k_*$ we have

$$\begin{aligned} \|f^\delta_{k_*} - f^\delta_l\|_{L_2} &\le \|f_\delta - f^\delta_{k_*}\|_{L_2} + \|f_\delta - f^\delta_l\|_{L_2} \\ &\le \Phi(k_*) + \Psi(k_*) + \Phi(l) + \Psi(l) \\ &\le 2\Phi(k_*) + \Psi(k_*) + \Psi(l) \\ &\le 3\Psi(k_*) + \Psi(l) \le 4\Psi(l). \end{aligned}$$

So, $k_* \in M^+(\Delta_N)$ and by the definition $k_* \ge k_+$.

Define the unknown norm using $\Psi(k_*)$

$$\begin{aligned} \|f_\delta - f^\delta_{k_+}\|_{L_2} &\le \|f_\delta - f^\delta_{k_*}\|_{L_2} + \|f^\delta_{k_*} - f^\delta_{k_+}\|_{L_2} \\ &\le 6\Psi(k_*). \end{aligned}$$

Due to monotonicity of the function Ψ for $2 \ge q \ge \frac{k_+}{k_+-1} > 1$ we have

$$\begin{aligned} \Psi(k_*) &= c_2\left(\chi^{\frac{d}{2}-\alpha} n^{\frac{\alpha}{d}}\delta + \chi^\tau n^{-\frac{\tau}{d}} k_*\delta\sqrt{n}\right) \\ &\le qc_2\left(\chi^{\frac{d}{2}-\alpha} n^{\frac{\alpha}{d}}\delta + \chi^\tau n^{-\frac{\tau}{d}} \frac{k_*}{q}\delta\sqrt{n}\right) \\ &= q\Psi(\tfrac{k_*}{q}). \end{aligned}$$

It follows from the definitions of the elements k_*, k_{opt} that $k_* \ge k_{opt} \ge k_* - 1$. Then

$$\|f_\delta - f^\delta_{\gamma_+}\|_{L_2} \le 6\Psi(k_*) \le 6q\Psi(k_*/q) \le 6q\Psi(k_{opt}) = 6q\Phi(k_{opt})$$

and Theorem is proved. □

Corollary 3. *For $\theta(k) = \phi(k^{-2})k^{-1}$ the estimate*

$$\|f_\delta - f^\delta_{m,\gamma_+}\|_{L_2} \le 6q\Phi\left(\theta^{-1}\left(\frac{\delta\sqrt{n}}{\rho}\right)\right),$$

is true. In particular for $\phi(\gamma) = \gamma^\beta$ with $0 < \beta \le \nu - 1/2$

$$\|f_\delta - f^\delta_k\|_{L_2} \le 6qc_2\left(\chi^{\frac{d}{2}-\alpha}\delta n^{\frac{\alpha}{d}} + \chi^\tau \rho^{\frac{1}{2\beta+1}} n^{-\frac{\tau}{d}} (\delta\sqrt{n})^{\frac{2\beta}{2\beta+1}}\right). \tag{18}$$

Proof. By the definition of k_{opt} it holds that $\Phi(k_{opt}) = \Psi(k_{opt})$, i.e.

$$\rho k^{-1}\phi(k^{-2})\left(\chi^{\frac{d}{2}-\alpha} n^{\frac{\alpha}{d}-\frac{1}{2}} + \chi^\tau n^{-\frac{\tau}{d}} k\right) = \delta\sqrt{n}\left(\chi^{\frac{d}{2}-\alpha} n^{\frac{\alpha}{d}-1/2} + \chi^\tau n^{-\frac{\tau}{d}} k\right).$$

Then $k_{opt}^{-2} = \theta^{-1}\left(\frac{\delta\sqrt{n}}{\rho}\right)$.

Taking into account that for $\phi(\gamma) = \gamma^\beta$ we have $\theta^{-1}(\gamma) = \gamma^{\frac{2}{2\beta+1}}$, then from (17) we obtain (18). □

Remark 8. *In view of the data error estimation*

$$\|\bar{g} - \bar{g}^{\delta}\|_{\mathbb{R}^n} \leq \delta\sqrt{n}$$

it is natural to assume that $\delta\sqrt{n} \ll 1$*, or, what is the same,* $n \ll \delta^{-2}$*. If* n *can be chosen at will, then, as it has been shown in* [2, Corollary 4.13]*, under the condition* $\alpha + \tau > d/2$*, an optimal choice is* $n \simeq \delta^{-\frac{d}{\alpha+\tau}}$*. However, it is very often, that the amount of available noisy data is limited such that one should deal with*

$$n \ll \delta^{-\frac{d}{\alpha+\tau}}.$$

For such n *using a-priori parameter choice* $\tilde{\gamma} = \delta n^{-\frac{\alpha+\tau-d}{d}}$ *suggested in* [2, Corollary 4.11] *one has the following error bound*

$$\begin{aligned} \|f_\delta - f_{\tilde{\gamma}}^{\delta}\|_{L_2} &\leq \tilde{C}\left(n^{-\frac{\tau}{d}} + \delta n^{\frac{\alpha}{d}} + \sqrt{\delta} n^{\frac{\alpha-\tau}{2d}}\right) \\ &= O(n^{-\frac{\tau}{d}}). \end{aligned}$$

At the same time, from Corollary 1 it follows that a-posteriori parameter choice $k = k_+$ *allows a higher order error bound. Indeed, keeping in mind that*

$$n^{-\frac{\tau}{d}} \gg \delta n^{-\frac{\alpha}{d}}, \quad n^{-\frac{\tau}{d}} \gg \sqrt{\delta} n^{-\frac{\alpha-\tau}{2d}}$$

from (18) we have

$$\|f_\delta - f_{k_+}^{\delta}\|_{L_2} \ll n^{-\frac{\tau}{d}}.$$

Remark 9. *Recall that we are looking for the solution* f^+ *of a normally solvable problem (2). It is well known (see, for example,* [1, Section 3.3]*) that in such situation the error bound for direct reconstruction of* f^+ *from noisy data is determined by* $\frac{\varepsilon}{\lambda_n}$*, where* ε *is a given data error level of the right-hand side and* λ_n *is the smallest singular value of* A_X*. In view of the condition (3) it is natural to assume that in our case it holds* $\lambda_n \sim n^{-\frac{\alpha}{d}}$*. Then, keeping in mind* $\varepsilon = \delta\sqrt{n}$ *we obtain*

$$\frac{\varepsilon}{\lambda_n} \sim \delta n^{\frac{\alpha}{d}+\frac{1}{2}}. \tag{19}$$

At the same time, from (18) it follows that for $\delta^{-1} \leq n^{\frac{1}{2}+\frac{(2\beta+1)(\alpha+\tau)}{d}}$

$$\|f_\delta - f_{k_+}^{\delta}\|_{L_2} \leq O(\delta n^{\frac{\alpha}{d}+\frac{1}{2}}). \tag{20}$$

Comparing (19) and (20) one can conclude that, if the amount n *of available discrete data is sufficiently large such that* $n \ll \delta^{-2}$ *but*

$$\delta^{-\frac{2d}{2(2\beta+1)(\alpha+\tau)+d}} \ll n,$$

or (see Remark 2)

$$\delta^{-\frac{2d}{2(2\beta+1)(\alpha+\tau)+d}} \ll n \ll \delta^{-\frac{d}{\alpha+\tau}}$$

then the regularized solution $f_{k_+}^{\delta}$ *allows a better error bound (in the sense of order) than the direct reconstruction.*

Bibliography

1. Engl H. W. Regularization of Inverse Problems / H. W. Engl, M. Hanke, A. Neubauer.– Kluwer academic publishers, 1996.
2. Krebs J. Sobolev error estimates and a priori parameter selection for semi-discrete Tikhonov regularization / J. Krebs, A. Louis, H. Wendland // J. Inv. Ill-Posed Problems.– 2009.– Vol. 17.– P. 845-869.
3. Mathe P. How general are general source conditions ? / P. Mathe, B. Hofmann // Inverse Problems.– 2008.– Vol. 24, № 1.– P. 1-5.
4. Mathe P. Optimal discretization of inverse problems in Hilbert scales. Regularization and self-regularization of projection methods, / P. Mathe, S. Pereverzev // SIAM J. Numer. Analysis.– 2001.– Vol. 38, № 6.– P. 1999-2021.
5. Mathe P. Regularization of some linear ill-posed problems with discretized random noisy data / P. Mathe, S. Pereverzev // Mathematics of Computation.– 2006.– Vol. 75, № 256.– P. 1913-1929.
6. Neubauer A. An a posteriori choice for Tikhonov regularization in Hilbert scales leading to optimal convergence rates / A. Neubauer // SIAM J. Numer. Anal.– 1988.– Vol. 25, № 6.– P. 1313-1326.
7. Pereverzev S. On the adaptive selection of the parameter in regularization of ill-posed problems. / S. Pereverzev, E. Schock // SIAM J. Numer. Anal.– 2005.– Vol. 43, № 5.– P. 2060-2076.
8. Solodky S. G. The optimal approximation for stationary iterated Tikhonov method / S. G. Solodky // East Journal of Approximation.– 2004.– Vol. 10, № 1-2.– P. 189-200.

Sergei Solodky, Evgeny Volynets,
Institute of Mathematics, National Academy of Sciences,
3, Tereschenkivska Str., Kiev, 01601, Ukraine

Received 23.04.2012

Журнал обчислювальної та прикладної математики | 2012, № 3 (109) с. 133–138 | Journal of Numerical & Applied Mathematics

UDC 517.9

THE MIXED DIRICHLET-NEUMANN PROBLEM FOR THE ELLIPTIC EQUATION OF THE SECOND ORDER IN DOMAIN WITH THIN INCLUSION

Yuriy Sybil

Резюме. Розглянуто змішану задачу Діріхле-Неймана для еліптичного рівняння другого порядку в обмеженій тривимірній Ліпшицевій області з тонким включенням, яке моделюється розімкнутою поверхнею. Гранична умова Діріхле задана на одній стороні цієї поверхні, а умова Неймана – на іншій. Введено функціональні простори в області із включенням та оператори сліду на розімкнутій Ліпшицевій поверхні. Доведено еквівалентність задачі у диференціальному формулюванні та відповідної варіаційної задачі. Досліджено питання існування та єдиності розв'язку поставленої задачі з неоднорідними граничними умовами у відповідних функціональних просторах.

Abstract. We consider Dirichlet-Neumann mixed boundary value problem for elliptic equation of the second order in three dimensional domain with thin inclusion which is presented by an open Lipschitz surface. The Dirichlet condition is posed on one side of the surface and the Neumann condition on the other side. Functional spaces in the domain with inclusion and corresponding trace operators on an open Lipschitz surface are introduced. We prove the equivalence of initial mixed boundary value problem and connected variational problem. As a result we obtain existence and uniqueness of solution of the posed problem with nonhomogeneous boundary conditions in appropriate functional spaces.

Introduction

Mixed boundary value problems for the second order elliptic equations in the case when on one part of closed boundary are given conditions of Dirichlet type and on another one conditions of Neumann type were considered in [2, 5, 9]. Boundary value problems in domains with thin inclusion as well as crack in solid bodies have a grate interest in applications. It's pretty convenient to present this thin object as an open double sided surface. Then for a mixed boundary value problem in unregular domain we have the Dirichlet conditions on one side of the open surface and the Neumann condition on the other one. Such kind of problems were considered in [3, 7] where the posed problems were reduced to systems of integral equations over the open boundary.

So far as domain with open surface is essentially unregular we have additional problems connected with definitions of corresponding trace maps and appropriate functional spaces [1, 2].

†*Key words.* Mixed boundary value problem; elliptic operator; variational problem; open surface.

In present paper we use a variational formulation of the posed mixed boundary value problem which gives us opportunity to obtain the existence and uniqueness of solution.

1. Functional spaces and trace operators

Let $\Omega_+ \subset \mathbb{R}^3$ be a bounded connected Lipschitz domain. This means that its boundary Σ is locally the graph of a Lipschitz function [1, 2]. Let us note that Σ can be piecewise smooth and have edges and corners. $\overline{\Omega}_+ = \Omega_+ \cup \Sigma$. We suppose that S is an open Lipschitz surface bounded by closed curve Γ, $\overline{S} = S \cup \Gamma$ and $\overline{S} \subset \Omega_+$. We denote $\Omega = \Omega_+ \setminus \overline{S}$ and consider S as a part of a some closed bounded Lipschitz surface $\Sigma_0 = \overline{S} \cup S_0$, $\Sigma_0 \subset \Omega_+$.

Since Σ and S are the Lipschitz surfaces almost everywhere we can define outward pointing vector of the normal $\vec{n}_x$, $x \in \Sigma$, and depend on the direction of $\vec{n}_x$, $x \in S$, we consider S as a double sided surface with sides S_+ and S_-.

In Ω_+ we consider the elliptic operator of the second order

$$Lu = -\sum_{i,j=1}^{3} \frac{\partial}{\partial x_i}\left(a_{ij}\frac{\partial u}{\partial x_j}\right) + a_0 u,$$

and connected bilinear form

$$a(u,v) = \int_{\Omega} \{\sum_{i,j=1}^{3} a_{ij}\frac{\partial v}{\partial x_i}\frac{\partial u}{\partial x_j} + a_0 uv\} dx.$$

Here $a_{ij}, a_0 \in C^1(\overline{\Omega}_+)$ are real functions which satisfy the following conditions for $x \in \Omega_+$:

$$\sum_{i,j=1}^{3} a_{ij} t_i t_j \geq c_1 \sum_{i=1}^{3} t_i^2, \quad t_i \in \mathbb{R},\ i = \overline{1,3}\,, c_1 > 0, \quad a_0(x) \geq c_2 > 0.$$

We use the Hilbert spaces $H^1(\Omega_+)$ and $H^1(\Omega_+, L)$ of real functions with norms and inner products

$$\|u\|^2_{H^1(\Omega_+)} = \int_{\Omega_+} \left\{|\nabla u|^2 + u^2\right\} dx, \quad (u,v)_{H^1(\Omega_+)} = \int_{\Omega_+} \{(\nabla u, \nabla v) + uv\} dx,$$

$$\|u\|^2_{H^1(\Omega_+,L)} = \|u\|^2_{H^1(\Omega_+)} + \|Lu\|^2_{L_2(\Omega_+)},$$

$$(u,v)_{H^1(\Omega_+,L)} = (u,v)_{H^1(\Omega_+)} + (Lu, Lv)_{L_2(\Omega_+)}.$$

The following trace operators $\gamma^+_{0,\Sigma} : H^1(\Omega_+) \to H^{1/2}(\Sigma)$ and $\gamma^+_{1,\Sigma} : H^1(\Omega_+, L) \to H^{-1/2}(\Sigma)$ are continuous and surjective [1, 4]. Here $\gamma^+_{1,\Sigma} u \in H^{-1/2}(\Sigma)$ and coincides with $\frac{\partial u}{\partial n_x}$ for $u \in C^1(\overline{\Omega}_+)$ where $\frac{\partial}{\partial n_x} = \sum_{i,j=1}^{3} \cos(\vec{n}_x, \vec{x}_i) a_{ij} \frac{\partial}{\partial x_j}$ is a conormal derivative, $\cos(\vec{n}_x, x_i)$ are the coordinates of the almost everywhere defined outward pointing vector of the normal $\vec{n}_x$ to Σ.

Let us denote by $C_0^\infty(\Omega)$ the class of infinitely differentiable functions with compact support in Ω. We introduce the Hilbert spaces $H^1(\Omega)$ and $H^1(\Omega, L)$

of real functions with norms

$$\|u\|^2_{H^1(\Omega)} = \int_\Omega \left\{|\nabla u|^2 + u^2\right\} dx, \tag{1}$$

$$\|u\|^2_{H^1(\Omega,L)} = \|u\|^2_{H^1(\Omega)} + \|Lu\|^2_{L_2(\Omega)},$$

where derivatives $\frac{\partial u}{\partial x_i} \in L_2(\Omega)$ are defined as

$$\left(\frac{\partial u}{\partial x_i}, \varphi\right)_{L_2(\Omega)} = -\int_\Omega u \frac{\partial \varphi}{\partial x_i} dx = -\left(u, \frac{\partial \varphi}{\partial x_i}\right)_{L_2(\Omega)}$$

for all $\varphi \in C_0^\infty(\Omega)$.

We consider some trace maps in Ω. We denote $\gamma^\pm_{0,S}$ and $\gamma^\pm_{1,S}$ the restrictions of trace maps $\gamma^\pm_{0,\Sigma_0}$ and $\gamma^\pm_{1,\Sigma_0}$ on S respectively [6]. Then we have $\gamma^\pm_{0,S} : H^1(\Omega) \to H^{1/2}(S)$ and $\gamma^\pm_{1,S} : H^1(\Omega, L) \to H^{-1/2}(S)$.

$$H^1_0(\Omega) = \left\{u \in H^1(\Omega) : \gamma^\pm_{0,S} u = 0,\ \gamma^+_{0,\Sigma} u = 0\right\}, \quad H^{-1}(\Omega) = (H^1_0(\Omega))'.$$

We also have that $H^1_0(\Omega)$ is a closure of $C_0^\infty(\Omega)$ in the norm (1).

In what follows we use the next trace maps: $[\gamma_{0,S}] = \gamma^+_{0,S} - \gamma^-_{0,S}$, $[\gamma_{1,S}] = \gamma^+_{1,S} - \gamma^-_{1,S}$. As it was shown in [6, 7] $[\gamma_{0,S}] : H^1(\Omega) \to H^{1/2}_{00}(S)$ and $[\gamma_{1,S}] : H^1(\Omega, L) \to H^{-1/2}_{00}(S)$, where $H^{1/2}_{00}(S) = \{g \in H^{1/2}(S) : p_0 g \in H^{1/2}(\Sigma_0)\}$. Here $p_0 g$ is extension by zero of the function g on S_0. The norm in $H^{1/2}_{00}(S)$ is given as $\|g\|_{H^{1/2}_{00}(S)} = \|p_0 g\|_{H^{1/2}(\Sigma_0)}$. $H^{-1/2}_{00}(S) = (H^{1/2}(S))'$, $H^{-1/2}(S) = (H^{1/2}_{00}(S))'$.

Let us denote $H^1_S(\Omega) = \{u \in H^1(\Omega) : \gamma^-_{0,S} u = 0\}$. If $u \in H^1_S(\Omega)$ then $\gamma^+_{0,S} u \in H^{1/2}_{00}(S)$ [8].

In [6] we obtained the first Green's formula for domain with an open surface which in presented case for $u \in H^1(\Omega, L)$ and $v \in H^1(\Omega)$ has the following form:

$$a(u,v) = (Lu, v)_{L_2(\Omega)} + \langle \gamma^+_{1,S} u, [\gamma_{0,S}] v\rangle + \langle [\gamma_{1,S}] u, \gamma^-_{0,S} v\rangle + \langle \gamma^+_{1,\Sigma} u, \gamma^+_{0,\Sigma} v\rangle. \tag{2}$$

Here $\langle\cdot,\cdot\rangle$ are relations of duality between $H^{1/2}_{00}(S)$ and $H^{-1/2}(S)$, $H^{1/2}(S)$ and $H^{-1/2}_{00}(S)$, $H^{1/2}(\Sigma)$ and $H^{-1/2}(\Sigma)$ respectively.

We assume that Ω_1 is a Lipschitz domain bounded by the closed surface Σ_0. $\overline{\Omega}_1 = \Omega_1 \cup \Sigma_0$, $\Omega_2 = \Omega_+ \setminus \overline{\Omega}_1$. We denote by u_i the restriction of $u \in H^1(\Omega)$ to Ω_i, $i = 1, 2$. It's obviously that $u_i \in H^1(\Omega_i)$, $i = 1, 2$.

In ([8], Lemma 5) we obtained the next proposition.

Lemma 1. *Let $u \in H^1(\Omega)$. Then the norm (1) can be presented in the following form:*

$$\|u\|^2_{H^1(\Omega)} = \|u_1\|^2_{H^1(\Omega_1)} + \|u_2\|^2_{H^1(\Omega_2)}$$

and this norm doesn't depend on the choice of S_0.

Lemma 2. *The operator $\gamma_{0,S} = (\gamma^+_{0,\Sigma}, \gamma^+_{0,S}) : H^1_S(\Omega) \to H^{1/2}(\Sigma) \times H^{1/2}_{00}(S)$ is continuous and surjective.*

Proof. Let $g \in H^{1/2}(\Sigma)$, $g_0 \in H^{1/2}_{00}(S)$ are arbitrary functions. We denote by $\tilde{g}_0 \in H^{1/2}(\Sigma_0)$ the extension g_0 by zero on S_0. Operator $\gamma^+_{0,\Sigma_0} : H^1(\Omega_1) \to H^{1/2}(\Sigma_0)$ is continuous and surjective [1]. Thus we have function $u_1 \in H^1(\Omega_1)$ with trace meaning $\gamma^+_{0,\Sigma_0} u_1 = \tilde{g}_0$ and

$$\|\tilde{g}_0\|_{H^{1/2}(\Sigma_0)} = \|g_0\|_{H^{1/2}_{00}(S)} \leq c_1 \|u_1\|_{H^1(\Omega_1)}. \tag{3}$$

Analogously there exists $u_2 \in H^1(\Omega_2)$ that $\gamma^-_{0,\Sigma_0} u_2 = 0$, $\gamma^+_{0,\Sigma} u_2 = g$ and

$$\|g\|_{H^{1/2}(\Sigma)} \leq c_2 \|u_2\|_{H^1(\Omega_2)}. \tag{4}$$

As a result we obtained the function $u \in L_2(\Omega)$ with restrictions $u_i \in H^1(\Omega_i)$ to Ω_i, $i = 1, 2$. Then $[\gamma_{0,S_0}]u = \gamma^+_{0,S_0} u_1 - \gamma^-_{0,S_0} u_2 = 0$ and by ([8], Lemma 4) we have $u \in H^1(\Omega)$. Since $\gamma^-_{0,S} u = 0$ it follows that $u \in H^1_S(\Omega)$.

In order to prove continuity of the trace map $\gamma_{0,S}$ we consider function $u \in H^1_S(\Omega)$ with $\gamma^+_{0,\Sigma} u = g \in H^{1/2}(\Sigma)$ and $\gamma^+_{0,S} u = g_0 \in H^{1/2}_{00}(S)$.

Then from (3), (4) and Lemma 1 we obtain

$$\|g\|_{H^{1/2}(\Sigma)} + \|g_0\|_{H^{1/2}_{00}(S)} \leq c_1 \|u_1\|_{H^1(\Omega_1)} + c_2 \|u_2\|_{H^1(\Omega_2)} \leq c \|u\|_{H^1(\Omega)}.$$

Here c, c_1, c_2 are some positive constants. □

2. Mixed boundary value problem and it's variational formulation

Let us state a Dirichlet-Neumann mixed boundary value problem in domain Ω.

Problem M. Find a function $u \in H^1(\Omega, L)$ that satisfies

$$Lu = h, \quad \gamma^-_{0,S} u = g, \quad \gamma^+_{1,S} u = f, \quad \gamma^+_{1,\Sigma} u = z.$$

Here $h \in L_2(\Omega)$, $g \in H^{1/2}(S)$, $f \in H^{-1/2}(S)$, $z \in H^{-1/2}(\Sigma)$ are given functions.

A partial case of the problem M when $g = 0$ we denote as problem M_0.

With problem M_0 it's closely connected the following variational problem.

Problem VM_0. Find a function $u \in H^1_S(\Omega)$ that satisfies

$$a(u, v) = l(v) \tag{5}$$

for every $v \in H^1_S(\Omega)$.

Here

$$l(v) = (h, v)_{L_2(\Omega)} + \langle f, \gamma^+_{0,S} v \rangle + \langle z, \gamma^+_{0,\Sigma} v \rangle, \tag{6}$$

$h \in L_2(\Omega)$, $f \in H^{-1/2}(S)$, $z \in H^{-1/2}(\Sigma)$ are given functions.

Lemma 3. *Bilinear form $a(u, v) : H^1_S(\Omega) \times H^1_S(\Omega) \to \mathbb{R}$ is continuous and $H^1_S(\Omega)$-elliptic.*

Proof. Since $H^1_S(\Omega)$ is a subspace of $H^1(\Omega)$ this lemma is a corollary of ([8], Lemma 7). □

Theorem 1. *Problems M_0 and VM_0 are equivalent.*

Proof. Let $u \in H^1_S(\Omega)$ be a solution of problem M_0. Then from the first Green's formula (1) for any $v \in H^1_S(\Omega)$ we have

$$a(u,v) = (h,v)_{L_2(\Omega)} + \langle f, \gamma^+_{0,S} v\rangle + \langle z, \gamma^+_{0,\Sigma}\rangle.$$

Thus u is a solution of problem VM_0.

Let now $u \in H^1_S(\Omega)$ be a solution of problem VM_0. Since $H^1_0(\Omega)$ is a subspace of $H^1_S(\Omega)$ for any $v \in H^1_0(\Omega)$ from (5) we obtain

$$a(u,v) = (h,v)_{L_2(\Omega)}.$$

But as it was shown in [6, 8] for any $u \in H^1(\Omega)$ and $v \in H^1_0(\Omega)$ we have $a(u,v) = \langle Lu, v\rangle$, where $\langle\cdot,\cdot\rangle$ is relations of duality between $H^1_0(\Omega)$ and $H^{-1}(\Omega)$. Thus $\langle Lu, v\rangle = \langle h, v\rangle$ or $\langle Lu - h, v\rangle = 0$ for any $v \in H^1_0(\Omega)$. It means that $Lu = h$. So far as $f \in L_2(\Omega)$ we get $u \in H^1(\Omega, L)$.

In Lemma 2 we showed that the trace operator $\gamma_{0,S} = (\gamma^+_{0,\Sigma}, \gamma^+_{0,S}) : H^1_S(\Omega) \to H^{1/2}(\Sigma) \times H^{1/2}_{00}(S)$ is surjective. Using (2), (5) and $Lu = h$ we have

$$\langle \gamma^+_{1,S} u - f, \gamma^+_{0,S} v\rangle + \langle \gamma^+_{1,\Sigma} u - z, \gamma^+_{0,\Sigma} v\rangle = 0$$

which is valid for an arbitrary $v \in H^1_S(\Omega)$. Thus $\gamma^+_{1,S} u = f$ and $\gamma^+_{1,\Sigma} u = z$. It gives us that u is a solution of problem M_0. □

Theorem 2. *Problem VM_0 has a unique solution for arbitrary $h \in L_2(\Omega)$, $f \in H^{-1/2}(S)$, $z \in H^{-1/2}(\Sigma)$.*

Proof. Lemma 3 gives us that the bilinear form $a(u,v) : H^1_S(\Omega) \times H^1_S(\Omega) \to \mathbb{R}$ is continuous and $H^1_S(\Omega)$-elliptic. Let's show that the functional $l : H^1_S(\Omega) \to \mathbb{R}$ given by (6) is continuous. If $v \in H^1_S(\Omega)$ then using Lemma 2 we have:

$$|l(v)| \le \|h\|_{L_2(\Omega)} \|v\|_{L_2(\Omega)} + \|f\|_{H^{-1/2}(S)} \|\gamma^+_{0,S} v\|_{H^{1/2}_{00}(S)} +$$

$$\|z\|_{H^{-1/2}(\Sigma)} \|\gamma^+_{0,\Sigma} v\|_{H^{1/2}(\Sigma)} \le \|h\|_{L_2(\Omega)} \|v\|_{H^1(\Omega)} + c_1 \|f\|_{H^{-1/2}(S)} \|v\|_{H^1(\Omega)} +$$

$$c_2 \|z\|_{H^{-1/2}(\Sigma)} \|v\|_{H^1(\Omega)} \le c\|v\|_{H^1(\Omega)},$$

where c, c_1, c_2 some positive constants which do not depend on v. Then by the Lax-Milgram Lemma we obtain what was to be proved. □

Corollary 4. *Problem M_0 has a unique solution for arbitrary $h \in L_2(\Omega)$, $f \in H^{-1/2}(S)$, $z \in H^{-1/2}(\Sigma)$.*

Lemma 4. *For every $g \in H^{1/2}(S)$ there exists function $w \in H^1(\Omega, L)$ that $\gamma^-_{0,S} w = g$.*

Proof. From [8] it follows that for every $g \in H^{1/2}(S)$ there exists $u_1 \in H^1(\Omega)$ and $Lu_1 = 0$, $\gamma^-_{0,S} u_1 = g$. Analogously for any $h_2 \in L_2(\Omega)$ we have a function $u_2 \in H^1(\Omega)$ that $Lu_2 = h_2$ and $\gamma^-_{0,S} u_2 = 0$. Thus we obtain a class of function $w = u_1 + u_2$ that $w \in H^1(\Omega, L)$ and $\gamma^-_{0,S} w = g$. □

Now we consider problem M which differs from problem M_0 only by non-homogeneous boundary condition on S_-. Lemma 4 gives us the function $w \in H^1(\Omega, L)$ which satisfies boundary condition $\gamma_{0,S}^- w = g$. Let the function u_1 be a solution of problem M_0:

$$Lu_1 = h_1, \quad \gamma_{0,S}^- u_1 = 0, \quad \gamma_{1,S}^+ u_1 = f_1, \quad \gamma_{1,\Sigma}^+ u_1 = z_1,$$

where $h_1 = h - Lw$, $f_1 = f - \gamma_{1,S}^+ w$, $z_1 = z - \gamma_{1,\Sigma}^+ w$. Then the function $u = u_1 + w$ is a solution of problem M. The preceding considerations imply the following assertion.

Theorem 3. *Problem M has a unique solution for arbitrary $h \in L_2(\Omega)$, $g \in H^{1/2}(S)$, $f \in H^{-1/2}(S)$, $z \in H^{-1/2}(\Sigma)$.*

Bibliography

1. Costabel M. Boundary integral operators on Lipschitz domains: elementary results / M. Costabel // SIAM J. Math. Anal.– 1988.– Vol. 19.– P. 613-626.
2. Grisvard P. Boundary value problems in non-smooth domains / P. Grisvard.– London: Pitman, 1985.– 410 p.
3. Krutitskii P. A. The method of integral equations in the mixed Dirichlet-Neumann problem for the Laplace equation in the exterior of cuts in the plane / P. A. Krutitskii, A. I. Sgibnev // Differential Equations.– 2001.– Vol. 37, № 10.– P. 1363-1375.
4. Lions J. L. Nonhomogeneous boundary-value problems and applications. Vol. 1. / J. L. Lions, E. Magenes.– Berlin: Springer-Verlag, 1972.
5. Stephan E. P. Boundary integral equations for the mixed boundary value problems in R^3 / E. P. Stephan // Math. Nachr.– 1987.– Vol. 134.– P. 21-53.
6. Sybil Yu. Three-dimensional elliptic boundary value problems for an open Lipschitz surface / Yu. Sybil // Matematychni Studii.– 1997.– Vol. 8, № 1.– P. 79-96.
7. Sybil Yu. Mixed boundary value problems for elliptic equations of the second order in tree-dimensional nonregular domains / Yu. Sybil // Matematychni Studii.– 2007.– Vol. 28, № 1.– P. 191-205.
8. Sybil Yu. On variational formulation of the Dirichlet problem for elliptic equation of the second order in region with thin inclusion / Yu. Sybil // Visnyk Lviv. Universytetu. Ser. Appl. Math. Comp. Sci.– 2011.– № 17.– P. 94-102 (in Ukrainian).
9. Wendland W. L. On the integral equation method for the plane mixed boundary value problem of the Laplacian / W. L. Wendland, E. P. Stephan, G. C. Hsiao // Math. Meth. Appl. Sci.– 1979.– Vol. 1.– P. 265-321.

Yuriy Sybil,
Ivan Franko National University of Lviv,
1, Universytets'ka Str., Lviv, 79000, Ukraine

Received 17.07.2012

Журнал обчислювальної та прикладної математики | 2012, № 3 (109) с. 139–149 | Journal of Numerical & Applied Mathematics

UDC 519.6

NONLOCAL PROBLEM FOR AN EVOLUTION FIRST ORDER EQUATION IN BANACH SPACE

Vitaliy Vasylyk

Резюме. Розглянуто двоточкову нелокальну задачу для диференціального еволюційного рівняння першого порядку з операторним коефіцієнтом у банаховому просторі. Запропоновано і обґрунтовано експонентціально збіжний алгоритм у припущенні, що операторний коефіцієнт є строго позитивний і виконуються деякі умови існування і єдиності. Алгоритм приводить до системи лінійних рівнянь, які можна розв'язати методом простої ітерації. Алгоритм забезпечує експонетціальну збіжність за часом, що в поєднанні з швидкими алгоритмами за просторовими змінними може бути ефективним для розв'язування таких задач. Ефективність пропонованих алгоритмів продемонстрована на чисельних експериментах.

Abstract. Two-points nonlocal problem for the first order differential evolution equation with an operator coefficient in a Banach space X is considered. An exponentially convergent algorithm is proposed and justified under the assumption that the operator coefficient is strongly positive and some existence and uniqueness conditions hold. This algorithm leads to a system of linear equations that can be solved by fixed-point iteration. The algorithm provides exponentially convergence in time that in combination with fast algorithms on spatial variables can be efficient for solving such problems. The efficiency of the proposed algorithms is demonstrated through numerical examples.

AMS Subject Classification: 65J10, 65M12, 65M15, 46N20, 46N40, 47N20, 47N40

1. Introduction

The m-point initial (nonlocal) problem for a differential equation with the nonlocal condition

$$u(t_0) + g(t_1; \ldots ; t_p; u) = u_0$$

and a given function g on a given point set $P = \{0 = t_0 < t_1 < \cdots < t_p\}$ is one of the important topics in the study of differential equations. Interest in such problems originates mainly from some physical problems with a control of the solution at P. For example, when the function $g(t_1; \ldots ; t_p; u)$ is linear we will have a periodic problem $u(t_0) = u(t_1)$. Problems with nonlocal conditions arise in the theory of physics of plasma [15], nuclear physics [10], mathematical chemistry [11], waveguides [8] etc. Two-point problem is also useful for considering the finale value problem [18].

Differential equations with operator coefficients in a Hilbert or Banach space can be considered as meta-models for systems of partial or ordinary differential equations and are suitable for investigating using the tools of the functional

---

†*Key words.* First order differential evolution equations in Banach space, nonlocal problem, unbounded operator coefficient, operator exponential, exponentially convergent algorithms.

analysis (see e.g. [4, 9]). Nonlocal problems can also be considered within this framework [2, 3].

Discretization methods for differential equations in Banach and Hilbert spaces were intensively studied in the last decade (see e.g. [5, 7, 12, 13, 16, 17, 22, 23] and the references therein). Methods from [7, 12, 13, 17, 22, 23] possess an exponential convergence rate, i.e. the error estimate in an appropriate norm is of the type $O(\mathrm{e}^{-N^{\alpha}})$, $\alpha > 0$ with respect to a discretization parameter $N \to \infty$. For a given tolerance ε such discretization provides optimal or nearly optimal computational complexity [7].

In the present paper we consider the problem

$$\begin{aligned} &\frac{du(t)}{dt} + A_1(t)u(t) = f_1(t), \\ &u(0) + \alpha u(1) = \varphi, \end{aligned} \tag{1}$$

where $A_1(t)$ is a densely defined closed (unbounded) operator with the domain $D(A_1)$ independent of t in a Banach space X, φ is a given vector and $f_1(t)$ is a given vector-valued function, $\alpha \in \mathbb{R}$. We suppose that the operator $A_1(t)$ is strongly positive; i.e. there exists a positive constant M_R independent of t such that on the rays and outside a sector $\Sigma_\theta = \{z \in \mathbb{C} : 0 \le arg(z) \le \theta, \theta \in (0, \pi/2)\}$ the following estimate for a resolvent holds:

$$\|(zI - A_1(t))^{-1}\| \le \frac{M_R}{1 + |z|}. \tag{2}$$

This assumption implies that there exists a positive constant c_κ such that (see [6], p.103)

$$\|A_1^{\kappa}(t) e^{-sA_1(t)}\| \le c_\kappa s^{-\kappa}, \quad s > 0, \quad \kappa \ge 0. \tag{3}$$

Our further assumption is that there exists a real positive ω such that

$$\|e^{-sA_1(t)}\| \le e^{-\omega s} \quad \forall s,\ t \in [0, 1] \tag{4}$$

(see [14], Corollary 3.8, p.12, for corresponding assumptions on $A_1(t)$). Let us also assume that the following conditions are valid

$$\|[A_1(t) - A_1(s)]A_1^{-\gamma}(t)\| \le L_{1,\gamma}|t - s| \quad \forall t,\ s,\ 0 \le \gamma < 1, \tag{5}$$

$$\|A_1^{\gamma}(t)A_1^{-\gamma}(s) - I\| \le L_\gamma |t - s| \quad \forall t,\ s \in [0, 1]. \tag{6}$$

We suppose also that

$$f_1(t) \in C(0, 1; X). \tag{7}$$

The aim of this paper is to construct an exponentially convergent approximation for a solution to problem (1). The paper is organized as follows. In Section 2 we discuss the existence and uniqueness of the solution as well as its representation through input data. A numerical algorithm is presented in section 3. The main result of this section is theorem 1 about the convergence rate of the proposed discretization. In the next section 4 we present a numerical example which confirm theoretical results from the previous sections.

2. Existence and uniqueness of the solution

It is well known, that for $\alpha = 0$ the problem (1) has a unique solution under the assumptions (2)-(7) (se e.g. [14, 9]). This solution can be written down as follows:

$$u(t) = U(t,0)u(0) + \int_0^t U(t,s)f_1(s)ds = U(t,0)\varphi + \int_0^t U(t,s)f_1(s)ds, \quad (8)$$

where $U(t,s)$ is an evolution operator that corresponds to (1) for $\alpha = 0$.

Let us study conditions when there is a unique solution to the two-points problem (1). We have from (8)

$$u(1) = U(1,0)u(0) + \int_0^1 U(1,s)f_1(s)ds.$$

Substituting this expression into the nonlocal condition we obtain

$$u(0) = [I + \alpha U(1,0)]^{-1}\left[\varphi - \alpha \int_0^1 U(1,s)f_1(s)ds\right],$$

and for $u(t)$ we have

$$u(t) = U(t,0)\,[I + \alpha U(1,0)]^{-1}\left[\varphi - \alpha \int_0^1 U(1,s)f_1(s)ds\right] + \\ + \int_0^t U(t,s)f_1(s)ds.$$

It is necessary to establish conditions on α for the existence of $u(t)$. In fact, we have to explore when exists $[I + \alpha U(1,0)]^{-1}$. So, we obtain using estimate for $U(t,s)$ (see e.g. [14, 9]).

$$\left\|[I + \alpha U(1,0)]^{-1}\right\| \le [1 - |\alpha|\,\|U(1,0)\|]^{-1} \le [1 - |\alpha| M]^{-1} \le C,$$

for small enough α ($\alpha < M^{-1}$).

3. Numerical algorithm

We use the approach developed in [7] and [21] to construct numerical method for solving problem (1). First of all we change variable in (1) by $t \to \frac{1+t}{2}$ and for $v(t) = u\left(\frac{1+t}{2}\right)$ we have

$$\begin{aligned} &\frac{dv(t)}{dt} + A(t)v(t) = f(t), \\ &v(-1) + \alpha v(1) = \varphi, \end{aligned} \quad (9)$$

where $A(t) = \frac{1}{2}A_1\left(\frac{1+t}{2}\right)$, $f(t) = \frac{1}{2}f_1\left(\frac{1+t}{2}\right)$,

We choose a mesh $\omega_n = \{t_k,\ k = 0, ..., n\}$ of $n+1$ various points on $[-1,1]$ that are Chebyshev-Gauss-Lobatto nodes $t_k = \cos\left(\frac{n-k}{n}\pi\right)$ and set $\tau_k = t_k - t_{k-1}$. Let

$$\begin{aligned} \overline{A}(t) =& A_k = A(t_k), t \in (t_{k-1}, t_k], \quad k = \overline{1,n}, \\ & A_0 = A(-1). \end{aligned}$$

Let us rewrite the problem (9) in the equivalent form

$$\begin{aligned}&\frac{dv}{dt}+\overline{A}(t)v=[\overline{A}(t)-A(t)]v(t)+f(t),\quad t\in(-1,1)\\&v(-1)=\varphi-\alpha v(1).\end{aligned}\tag{10}$$

Note that now all operators on the left hand side of these equations are constant on each subinterval and piece-wise constant on the whole interval $[-1,1]$.

On each subinterval we can write down the equivalent to (10) integral equation

$$\begin{aligned}v(t)=&\mathrm{e}^{-A_k(t-t_{k-1})}v(t_{k-1})+\int_{t_{k-1}}^{t}\mathrm{e}^{-A_k(t-s)}\left[A_k-A(t)\right]v(s)ds+\\&+\int_{t_{k-1}}^{t}\mathrm{e}^{-A_k(t-s)}f(s)ds,\quad t\in[t_{k-1},t_k],\quad k=\overline{2,n},\\v(t)=&\mathrm{e}^{-A_1(t+1)}\left[\varphi-\alpha v(1)\right]+\int_{-1}^{t}\mathrm{e}^{-A_1(t-s)}\left[A_1-A(t)\right]v(s)ds+\\&+\int_{-1}^{t}\mathrm{e}^{-A_1(t-s)}f(s)ds,\quad t\in[-1,t_1].\end{aligned}\tag{11}$$

Let

$$P_n(t;v)=P_nv=\sum_{j=0}^{n}v(t_j)L_{j,n}(t),$$

be the interpolation polynomial for $v(t)$ on the mesh ω_n, $x=(x_0,...,x_n)$, $x_i\in X$ given vector and

$$P_n(t;y)=P_nx=\sum_{j=0}^{n}x_jL_{j,n}(t)$$

the polynomial that interpolates x where

$$L_{j,n}(s)=\frac{T_n'(s)(1-s^2)}{\frac{d}{ds}[(1-s^2)T_n'(s)]_{s=s_j}(s-s_j)},\quad j=0,...,n$$

are the Lagrange fundamental polynomials. Substituting $P_n(s;x)$ for $v(s)$, x_k for $v(t_k)$ and then setting $t=t_k$ in (11) we obtain the following system of linear equations with respect to the unknown x_k :

$$\begin{aligned}&x_0+\alpha x_n=\varphi,\\&x_k=\mathrm{e}^{-A_k\tau_k}x_{k-1}+\sum_{j=0}^{n}\alpha_{kj}x_j+\phi_k,\quad k=\overline{1,n},\end{aligned}\tag{12}$$

which represents our algorithm. Here we use the notations

$$\begin{aligned}&\alpha_{kj}=\int_{t_{k-1}}^{t_k}\mathrm{e}^{-A_k(t_k-s)}[A_k-A(s)]L_{j,n}(s)ds,\\&\phi_k=\int_{t_{k-1}}^{t_k}\mathrm{e}^{-A_k(t_k-s)}f(s)ds,\quad k=\overline{1,n},\quad j=\overline{0,n},\end{aligned}$$

and suppose that we have an algorithm to compute these coefficients.

For the error $z = (z_1, ..., z_n)$, with $z_k = v(t_k) - x_k$ we have the relations

$$
\begin{aligned}
&z_0 + \alpha z_n = 0, \\
&z_k = \mathrm{e}^{-A_k \tau_k} z_{k-1} + \sum_{j=0}^{n} \alpha_{kj} z_j + \psi_k, \quad k = \overline{1,n},
\end{aligned} \tag{13}
$$

where

$$
\psi_k = \int_{t_{k-1}}^{t_k} \mathrm{e}^{-A_k(t_k - s)} [A_k - A(s)][v(s) - P_n(s; v)] ds, \quad k = \overline{1,n},
$$

In order to represent algorithm (12) in a block-matrix form we introduce the matrix

$$
S = \begin{pmatrix}
I & 0 & 0 & \cdot & \cdot & \cdot & 0 & \alpha\sigma_0 \\
-\sigma_1 & I & 0 & \cdot & \cdot & \cdot & 0 & 0 \\
0 & -\sigma_2 & I & \cdot & \cdot & \cdot & 0 & 0 \\
\cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\
0 & 0 & 0 & \cdot & \cdot & \cdot & -\sigma_n & I
\end{pmatrix}, \tag{14}
$$

where $\sigma_0 = A_0^{\gamma} A_n^{-\gamma}$, $\sigma_k = \mathrm{e}^{-A_k \tau_k} A_k^{\gamma} A_{k-1}^{-\gamma}$, $k = \overline{1,n}$, the matrix $B = \{\tilde{\alpha}_{k,j}\}_{k,j=0}^{n}$ with $\tilde{\alpha}_{k,j} = A_k^{\gamma} \alpha_{k,j} A_j^{-\gamma}$, $k = \overline{1,n}$, $j = \overline{0,n}$, and $\tilde{\alpha}_{0,j} = 0$, $j = \overline{0,n}$, the vectors

$$
\tilde{x} = \begin{pmatrix} A_0^{\gamma} x_0 \\ A_1^{\gamma} x_1 \\ \cdot \\ \cdot \\ A_n^{\gamma} x_n \end{pmatrix}, \
\phi = \begin{pmatrix} A_0^{\gamma} \varphi \\ A_1^{\gamma} \phi_1 \\ \cdot \\ \cdot \\ A_n^{\gamma} \phi_n \end{pmatrix}, \
\tilde{z} = \begin{pmatrix} A_0^{\gamma} z_0 \\ A_1^{\gamma} z_1 \\ \cdot \\ \cdot \\ A_n^{\gamma} z_n \end{pmatrix}, \
\psi = \begin{pmatrix} 0 \\ A_1^{\gamma} \psi_1 \\ \cdot \\ \cdot \\ A_n^{\gamma} \psi_n \end{pmatrix}. \tag{15}
$$

It is easy to check that for the (left) inverse

$$
S^{-1} = \delta \left(R_1 - R_2 \right),
$$

where

$$
\delta = \left(I + \alpha\sigma_0 \sigma_1 \ldots \sigma_n \right)^{-1},
$$

$$
R_1 = \begin{pmatrix}
I & 0 & \cdots & 0 & 0 \\
\sigma_1 & I & \cdots & 0 & 0 \\
\sigma_2\sigma_1 & \sigma_2 & \cdots & 0 & 0 \\
\cdot & \cdot & \cdots & \cdot & \cdot \\
\sigma_n \cdots \sigma_1 & \sigma_n \cdots \sigma_2 & \cdots & \sigma_n & I
\end{pmatrix},
$$

$$
R_2 = \alpha s_0 \begin{pmatrix}
0 & \sigma_n \ldots \sigma_2 & \sigma_n \ldots \sigma_3 & \cdots & \sigma_n & I \\
0 & 0 & \sigma_1 \sigma_n \ldots \sigma_3 & \cdots & \sigma_1 \sigma_n & \sigma_1 \\
\cdot & \cdot & \cdot & \cdots & \cdot & \cdot \\
0 & 0 & 0 & \cdots & 0 & \sigma_{n-1} \ldots \sigma_1 \\
0 & 0 & 0 & \cdots & 0 & 0
\end{pmatrix}.
$$

Remark 10. *Using results of* [7] *one can get a parallel and sparse approximations with an exponential convergence rate of the operator exponentials contained in* S^{-1} *and as a consequence a parallel and sparse approximation of* S^{-1}.

We multiply the equations in (12) and the equation in (13) by A_k^γ, $k=\overline{0,n}$ and obtain

$$\begin{aligned} &A_0^\gamma x_0 + \alpha A_0^\gamma x_n = A_0^\gamma \varphi, \\ &A_k^\gamma x_k = \mathrm{e}^{-A_k\tau_k} A_k^\gamma x_{k-1} + \sum_{j=0}^{n} \tilde{\alpha}_{kj} A_j^\gamma x_j + A_k^\gamma \phi_k, \quad k=\overline{1,n}, \end{aligned} \tag{16}$$

$$\begin{aligned} &A_0^\gamma z_0 + \alpha A_0^\gamma z_n = 0, \\ &A_k^\gamma z_k = \mathrm{e}^{-A_k\tau_k} A_k^\gamma z_{k-1} + \sum_{j=0}^{n} \tilde{\alpha}_{kj} A_j^\gamma z_j + A_k^\gamma \psi_k, \quad k=\overline{1,n}, \end{aligned} \tag{17}$$

Then systems (16), (17) can be written down in the matrix form using notations (14), (15) as

$$\begin{aligned} S\tilde{x} &= B\tilde{x} + \phi, \\ S\tilde{z} &= B\tilde{z} + \psi. \end{aligned} \tag{18}$$

Next, for a vector $v=(v_1,v_2,...,v_n)^T$ and a block operator matrix $A=\{a_{ij}\}_{i,j=1}^n$ we introduce a vector norm

$$|||v||| \equiv |||v|||_1 = \max_{1\le k\le n} \|v_k\|,$$

and the consistent matrix norm

$$|||A||| \equiv |||A|||_1 = \max_{1\le i\le n} \sum_{j=1}^{n} \|a_{i,j}\|.$$

Due to (6) we have

$$|||A_k^\gamma A_{k-1}^{-\gamma}||| = |||A_k^\gamma A_{k-1}^{-\gamma} - I + I||| \le 1 + L_\gamma \tau_k,$$

$$\|\sigma_0\| = \|A_0^\gamma A_n^{-\gamma}\| \le 1 + L_\gamma T.$$

In our case $T=2$. So, we have the following, using these estimates

$$\|\sigma_k\| = \|\mathrm{e}^{-A_k\tau_k} A_k^\gamma A_{k-1}^{-\gamma}\| \le \mathrm{e}^{-\omega\tau_k} \|A_k^\gamma A_{k-1}^{-\gamma}\| \le \mathrm{e}^{-\omega\tau_k}\left(1+L_\gamma\tau_k\right),$$

$$\|\delta\| = \left\| \left(I + \alpha\sigma_0\sigma_1\dots\sigma_n\right)^{-1} \right\| \le \Big(1 - |\alpha|\, \|\sigma_0\|\, \|\sigma_1\|\, \|\sigma_2\| \dots \|\sigma_n\|\Big)^{-1} \le$$

$$\le \Big(1 - |\alpha|\, \left(1+2L_\gamma\right) \mathrm{e}^{-\omega\tau_1} \left(1+L_\gamma\tau_1\right) \mathrm{e}^{-\omega\tau_2} \left(1+L_\gamma\tau_2\right) \dots \mathrm{e}^{-\omega\tau_n} \left(1+L_\gamma\tau_n\right)\Big)^{-1}$$

$$\le \left(1 - |\alpha|\, \left(1+2L_\gamma\right) \mathrm{e}^{-2\omega} \left(1 + \frac{2L_\gamma}{n}\right)^n\right)^{-1} \le$$

$$\le \Big(1 - |\alpha|\, \left(1+2L_\gamma\right) \mathrm{e}^{-2\omega}\mathrm{e}^{2L_\gamma}\Big)^{-1} \le c,$$

for α small enough.

In order to estimate the norm of matrix S we must estimate the norms of matrices R_1, R_2. In [7] it was proved that for a matrix similar to R_1 the estimate $|||R_1||| \le cn$ holds true. Let us estimate the norm of matrix R_2.

$$|||R_2||| \leq (1+2c)\left(1+e^{-\omega\tau}(1+c\tau)+\cdots+[e^{-\omega\tau}(1+c\tau)]^{n-1}\right) \leq$$
$$\leq (1+2c)\left(1+(1+c\tau)+\cdots+(1+c\tau)^{n-1} \leq \frac{(1+c\tau)^n-1}{c\tau}\right) \leq$$
$$\leq (1+2c)\,\frac{e^{2c}}{c\tau} \leq cn.$$

Using these estimates we obtain that

$$|||S^{-1}||| \leq cn. \tag{19}$$

It was proved an estimate for the matrix B in [7]:

$$|||B||| \leq cn^{\gamma-2}\ln(n). \tag{20}$$

So we can formulate the following assertion

Lemma 1. *Let assumptions* (2)-(6) *are fulfilled. Then estimates* (19), (20) *hold true.*

Using (18) we have

$$\begin{aligned} \tilde{x} &= \left[E - S^{-1}B\right]^{-1} S^{-1}\phi, \\ \tilde{z} &= \left[E - S^{-1}B\right]^{-1} S^{-1}\psi, \end{aligned} \tag{21}$$

where E is a diagonal matrix with unit operators I on diagonal. Using lemma 1 we obtain that

$$|||S^{-1}B||| \leq cn^{\gamma-1}\ln(n) \to 0,\ n \to \infty. \tag{22}$$

It means that for n large enough there exists the matrix $\left[E - S^{-1}B\right]^{-1}$ and

$$\left|\left|\left|\left[E - S^{-1}B\right]^{-1}\right|\right|\right| \leq c.$$

Consequently we obtain the following stability estimates from (21) using lemma 1:

$$\begin{aligned} |||\tilde{x}||| &\leq cn|||\phi|||, \\ |||\tilde{z}||| &\leq cn|||\psi|||. \end{aligned} \tag{23}$$

Let Π_n be the set of all polynomials in t with vector coefficients of degree less or equal than n. In complete analogy with [1, 19, 20] the following Lebesgue inequality for vector-valued functions can be proved

$$\|u(t) - P_n(t;u)\|_{C[-1,1]} \equiv \max_{t\in[-1,1]} \|u(t) - P_n(t;u)\| \leq (1+\Lambda_n)E_n(u),$$

with the error of the best approximation of u by polynomials of degree not greater than n

$$E_n(u) = \inf_{p\in\Pi_n} \max_{t\in[-1,1]} \|u(t) - p(t)\|.$$

Now, we can go over to the main result of this section.

Theorem 1. *Let the assumptions of Lemma 1 with $\gamma < 1$ hold, then there exists a positive constant c such that*

1. *For n large enough it holds*
$$|||\tilde{z}||| \leq cn^{\gamma-1} \cdot \ln n \cdot E_n(A_0^\gamma v),$$
where v is the solution of (9)*;*
2. *The first equation in* (18) *can be written in the form*
$$\tilde{x} = S^{-1}B\tilde{x} + S^{-1}\phi,$$
and can be solved by the fixed point iteration
$$\tilde{x}^{(k+1)} = S^{-1}B\tilde{x}^{(k)} + S^{-1}\phi,\ k = 0, 1, ...;\ \tilde{x}^{(0)} - arbitrary,$$
with the convergence rate of an geometrical progression with the denominator $q \leq cn^{\gamma-1}\ln(n) < 1$ for n large enough.

Proof. For $\tilde{z}$ we have the second estimate in (23). The norm of the first summand on the right hand side of this inequality can be estimated in the following way

$$|||\psi||| = \max_{1\leq k\leq n}\left\| \int_{t_{k-1}}^{t_k} \left\{ A_k^\gamma \mathrm{e}^{-A_k(t_k-s)}[A_k - A(s)]\times \right.\right.$$
$$\left.\left. \times A_k^{-\gamma}(A_k^\gamma A_0^{-\gamma})(A_0^\gamma v(s) - P_n(s; A_0^\gamma v)) \right\} ds \right\| \leq$$
$$\leq c \max_{1\leq k\leq n} \int_{t_{k-1}}^{t_k} |t_k - s|^{-\gamma} |t_k - s|\, \|A_0^\gamma v(s) - P_n(s; A_0^\gamma v)\| ds \leq$$
$$\leq c\tau_{max}^{2-\gamma} \|A_0^\gamma u(s) - P_n(\cdot; A_0^\gamma v)\|_{C[-1,1]} \leq c\tau_{max}^{2-\gamma}(1+\Lambda_n)E_n(A_0^\gamma v).$$

So, we obtain
$$|||\psi||| \leq cn^{\gamma-2} \cdot \ln n \cdot E_n(A_0^\gamma u), \tag{24}$$
Now, the first assertion of the theorem follows from (23), (24). The second one follows from (18) and (22). □

Tabl. 1. The error in the case $n = 4$, $x = 0.5$

Point t	ε
-1	0.00005276
-0.70710678	0.00097645
0	0.00063440
0.70710678	0.00029592
1	0.00010552

4. Examples

Let us consider the following problem
$$\begin{aligned} &\frac{\partial u(x,t)}{\partial t} - \frac{\partial^2 u(x,t)}{\partial x^2} + q(x,t)u(x,t) = f(x,t), \\ &u(0,t) = u(1,t) = 0, \\ &u(x,-1) + \alpha u(x,1) = \varphi(x), \end{aligned} \tag{25}$$

TABL. 2. The error in the case $n = 6$, $x = 0.5$

Point t	ε
-1	8.12568908E-7
-0.86602540	0.00010146
-0.5	0.00030932
0	0.00022136
0.5	0.00013419
0.86602540	0.00007182
1	0.00000162

TABL. 3. The error in the case $n = 8$, $x = 0.5$

Point t	ε
-1	0.00000117
-0.92387953	0.00000613
-0.70710678	0.00004544
-0.38268343	0.00005753
0	0.00004745
0.38268343	0.00003362
0.70710678	0.00002096
0.92387953	0.00000846
1	0.00000235

TABL. 4. The error in the case $n = 12$, $x = 0.5$

Point t	ε
-1	0.49451310e-8
-0.96592582	0.14687232e-7
-0.86602540	0.23393074e-6
-0.70710678	0.54494052e-6
-0.5	0.76722515e-6
-0.25881904	0.82803283e-6
0	0.76362937e-6
0.25881904	0.63174173e-6
0.5	0.47173110e-6
0.70710678	0.30381367e-6
0.86602540	0.14341583e-6
0.96592582	0.21271757e-7
1	0.98902621e-8

with $f(x,t) = \mathrm{e}^{-\pi^2(1+t)} \sin(\pi x)(1+t)$, $\alpha = 0.5$, $\varphi(x) = \left(1 + 0.5\mathrm{e}^{-2\pi^2}\right) \sin(\pi x)$, $q(x,t) = 1+t$. Then, the solution of this problem is $u(x,t) = \mathrm{e}^{-\pi^2(1+t)} \sin(\pi x)$.

TABL. 5. The error in the case $n = 16$, $X = 0.5$

Point t	ε
-1	0.20628738e-11
-0.98078528	0.28602854e-10
-0.92387953	0.48425552e-9
-0.83146961	0.14258845e-8
-0.70710678	0.25968220e-8
-0.55557023	0.36339719e-8
-0.38268343	0.42916820e-8
-0.19509032	0.44975339e-8
0	0.43045006e-8
0.19509032	0.38169887e-8
0.38268343	0.31414290e-8
0.55557023	0.23686579e-8
0.70710678	0.15787207e-8
0.83146961	0.85640040e-9
0.92387953	0.30309439e-9
0.98078528	0.16809109e-10
1	0.41257476e-11

The problem (25) can be written down in the form (9) where the operator $A(t)$ is defined by

$$D(A(t)) = D(A) = \{v \in H^2(0,1) : v(0) = 0,\, v(1) = 0\},$$
$$A(t)v = -\frac{\partial^2 v}{\partial x^2} + (1+t)v.$$

Coefficients of the system (16) were calculated by using the Fourier series expansion. The results of calculation presented in tables confirm our theory above.

BIBLIOGRAPHY

1. Babenko K. I. Foundations of the numerical analysis / K. I. Babenko. Nauka: Moscow, 1986 (in Russian).
2. Byszewski L. Theorems about the existence and uniqueness of solution of a semilinear evolution nonlocal cauchy problem. / L. Byszewski. // J. Math. Anal. Appl.– 1991.– Vol. 162, № 5.– P. 494-505.
3. Byszewski L. Application of properties of the right-hand sides of evolution equations to an investigation of nonlocal evolution problems. / L. Byszewski // Nonlinear Anal.– 1998.– Vol. 33, № 5.– P. 413-426.
4. Clement Ph. One-parameter semigroups / Ph. Clement, H. J. A. M. Heijmans, S. Angenent, C. J. van Duijn, B. de Pagter Amsterdam: CWI Monographs, 5. North-Holland Publishing Co., 1987.
5. Fernandez M. L. Fast Runge-Kutta approximation of inhomogeneous parabolic equations. / M. L. Fernandez, Ch. Lubich, C. Palencia, A. Schädle // Numerische Mathematik.– 2005.– P. 1-17.
6. Fujita H. Operator Theory and Numerical Methods / H. Fujita, N. Saito, T. Suzuki Heidelberg: Elsevier, 2001.

7. Gavrilyuk I. Exponentially convergent algorithms for abstract differential equations Frontiers in Mathematics / I. Gavrilyuk, V. Makarov, V. Vasylyk.– Basel: Birkhäuser / Springer Basel AG, 2011.
8. Gordeziani D. Investigation of the nonlocal initial boundary value problems for some hyperbolic equations / D. Gordeziani, G. Avalishvili // Hiroshima Math. J.– 2001.– Vol. 31, № 3.– P. 345-366.
9. Krein S. G. Linear Differential Operators in Banach Spaces / S. G. Krein.– New York: Amer. Math. Soc., 1971.
10. Leung A. W. Optimal control of multigroup neutron fission systems / A. W. Leung, G.-S. Chen // J. Appl. Math. and Opt.– 1999.– Vol. 40, № 1.– P. 39-60.
11. Leung A. W. Existence and monotone scheme for time-periodic nonquasimonotone reaction-diffusion systems: application to autocatalytic chemistry / A. W. Leung, L. A. Ortega // J. Math. Anal. Appl.– 1998.– Vol. 221, № 2.– P. 712-733.
12. López-Fernández M. On the numerical inversion of the laplace transform of certain holomorphic mappings / M. López-Fernández, C. Palencia // Applied Numerical Mathematics.– 2004.– Vol. 51.– P. 289-303.
13. López-Fernández M. A spectral order method for inverting sectorial laplace transforms / M. López-Fernández, C. Palencia, A. Schädle // SIAM J. Numer. Anal.– 2006.– Vol. 44.– P. 1332-1350.
14. Pazy A. Semigroups of linear operator and applications to partial differential equations / A. Pazy.– New York, Berlin, Heidelberg: Springer Verlag, 1983.
15. Samarskii A. A. Some problems of the theory of differential equations / A. A. Samarskii // Differentsial'nye Uravneniya (Differential equations).– 1980.– Vol. 16, № 11.– P. 1925-1935.
16. Sheen D. A parallel method for time-discretization of parabolic problems based on contour integral representation and quadrature / D. Sheen, I. H. Sloan, V. Thomée // Math.Comp.– 2000.– Vol. 69.– P. 177-195.
17. Sheen D. A parallel method for time-discretization of parabolic equations based on laplace transformation and quadrature / D. Sheen, I. H. Sloan, V. Thomée // IMA Journal of Numerical Analysis.– 2003.– Vol. 23.– P. 269-299.
18. Showalter R. E. Initial and final value problems for degenerate parabolic evolution systems / R. E. Showalter // Indiana Univ. Math. J.– 1979.– Vol. 28, № 6.– P. 883-893.
19. Szegö G. Orthogonal Polynomials / G. Szegö.– New York: American Mathematical Society, 1959.
20. Szegö G. Orthogonal Polynomials (with an Introduction and a Complement by J. L. Geronimus) / G. Szegö.– Moscow: State Publishing House of Physical and Mathematical Literature, 1962 (in Russian).
21. Bohonova T. Ju. Exponentially convergent duhamel-like algorithms for differential equations with an operator coefficient possessing a variable domain in a banach space / T. Ju. Bohonova, I. P. Gavrilyuk, V. L. Makarov, V. B. Vasylyk // SIAM J. Numer. Anal.– 2007/08.– Vol. 46, № 5.– P. 365-396.
22. Vasylyk V. Uniform exponentially convergent method for the first order evolution equation with unbounded operator coefficient / V. Vasylyk // Journal of Numerical and Applied Mathematics (ISSN 0868-6912).– 2003.– Vol. 1.– P. 99-104.
23. Vasylyk V. Approximate solution of the cauchy problem for differential equation with sectorial operator in banach space / V. Vasylyk // Bulletin of Lviv University. Physical and Mathematical sciences.– 2004.– Vol. 9.– P. 34-46 (in Ukrainian).

Vitaliy Vasylyk,
Institute of Mathematics, National Academy of Sciences,
3, Tereschenkivska Str., Kiev, 01601, Ukraine

Received 03.07.2012

Taras Shevchenko National University of Kyiv

Journal of Numerical and Applied

2012 M A T H E M A T I C S 3 (109)

Founded in 1965

C O N T E N T S