ISSN 1997-1397 (Print)
ISSN 2313-6022 (Online)

Журнал Сибирского федерального университета

Математика и физика

Journal of Siberian Federal University

Mathematics & Physics

2020 13 (5)

ISSN 1997-1997-1397 (Print)

ISSN 2313-6022 (Online)

ЖУРНАЛ СИБИРСКОГО ФЕДЕРАЛЬНОГО УНИВЕРСИТЕТА
Математика и Физика

2020 13 (5)

JOURNAL OF SIBERIAN FEDERAL UNIVERSITY
Mathematics & Physics

Издание индексируется Scopus (Elsevier), Emerging Sources Citation Index (WoS, Clarivate Analytics), Российским индексом научного цитирования (НЭБ), представлено в международных и российских информационных базах: Ulrich's periodicals directiory, ProQuest, EBSCO (США), Google Scholar, MathNet.ru, КиберЛенинке.

Включено в список Высшей аттестационной комиссии «Рецензируемые научные издания, входящие в международные реферативные базы данных и системы цитирования».

Все статьи представлены в открытом доступе http://journal.sfu-kras.ru/en/series/mathematics_physics.

Журнал Сибирского федерального университета. Математика и физика.
Journal of Siberian Federal University. Mathematics & Physics.

Учредитель: Федеральное государственное автономное образовательное учреждение высшего образования "Сибирский федеральный университет"(СФУ)
Главный редактор: А.М. Кытманов. Редакторы: В.Е. Зализняк, А.В. Щуплев.
Компьютерная верстка: Г.В. Хрусталева
№ 5. 26.10.2020. Индекс: 42327. Тираж: 1000 экз. Свободная цена
Адрес редакции и издательства: 660041 г. Красноярск, пр. Свободный, 79, оф. 32-03.
Отпечатано в типографии Издательства БИК СФУ
660041 г. Красноярск, пр. Свободный, 82а.
Свидетельство о регистрации СМИ ПИ № ФС 77-28724 от 27.06.2007 г., выданное Федеральной службой по надзору в сфере массовых коммуникаций, связи и охраны культурного наследия
http://journal.sfu-kras.ru
Подписано в печать 15.10.20. Формат 84×108/16. Усл.печ. л. 11,9.
Уч.-изд. л. 11,6. Бумага тип. Печать офсетная.
Тираж 1000 экз. Заказ 11246
Возрастная маркировка в соответствии с Федеральным законом № 436-ФЗ:16+

CONTENTS

СОДЕРЖАНИЕ

Journal of Siberian Federal University. Mathematics & Physics 2020, 13(5), 519–532

DOI: 10.17516/1997-1397-2020-13-5-519-532
УДК 517.55

A Perturbation of the de Rham Complex

Ihsane Malass*
Nikolai Tarkhanov†
University of Potsdam
Potsdam, Germany

Received 16.05.2020, received in revised form 07.06.2020, accepted 16.07.2020

Abstract. We consider a perturbation of the de Rham complex on a compact manifold with boundary. This perturbation goes beyond the framework of complexes, and so cohomology does not apply to it. On the other hand, its curvature is "small", hence there is a natural way to introduce an Euler characteristic and develop a Lefschetz theory for the perturbation. This work is intended as an attempt to develop a cohomology theory for arbitrary sequences of linear mappings.

Keywords: De Rham complex, cohomology, Hodge theory, Neumann problem.

Citation: I. Malass, N. Tarkhanov, A Perturbation of the de Rham Complex, J. Sib. Fed. Univ. Math. Phys., 2020, 13(5), 519–532. DOI: 10.17516/1997-1397-2020-13-5-519-532.

Introduction

De Rham cohomology is a tool belonging both to algebraic topology and to differential topology, capable of expressing basic topological information about smooth manifolds in a form particularly adapted to computation and the concrete representation of cohomology classes. It is a cohomology theory based on the existence of differential forms with prescribed properties. We tracked out its analytic ingredients in [5].

In this paper we consider a perturbation of the de Rham complex on a compact manifold with boundary $\mathcal{X}$ of dimension n. That is

$$0 \longrightarrow \Omega^0(\mathcal{X}) \xrightarrow{d+a} \Omega^1(\mathcal{X}) \xrightarrow{d+a} \dots \xrightarrow{d+a} \Omega^n(\mathcal{X}) \longrightarrow 0, \tag{0.1}$$

where $\Omega^i(\mathcal{X})$ stands for the space of all differential forms of degree i with C^∞ coefficients on $\mathcal{X}$, by d is meant the exterior differentiation of forms, and a is a given C^∞ one-form on $\mathcal{X}$. The differential in (0.1) is defined by $(d+a)u = du + a \wedge u$ for all $u \in \Omega^i(\mathcal{X})$ whence

$$\begin{aligned}(d+a)^2 u &= d^2 u + da \wedge u - a \wedge du + a \wedge du + (a \wedge a) \wedge u = \\ &= da \wedge u.\end{aligned}$$

The differential square is sometimes referred to as the curvature of sequence (0.1) and it is “small” in some relevant sense. This enables one to introduce the Euler characteristic and prove a Lefschetz fixed point formula for (0.1), see [11, 13].

However, $(d+a)^2$ need not be zero and for this reason the standard cohomology construction does not apply to (0.1). To introduce cohomology in (0.1) we use a construction of mathematical

*ihsane malass@hotmail.com
†tarkhanov@math.uni-potsdam.de

folklore. To wit, we factorise the null-space $\mathcal{Z}^i$ of $d+a$ in $\Omega^i(\mathcal{X})$ by the (part of the) range of $d+a$ over $\Omega^{i-1}(\mathcal{X})$ belonging to the space $\mathcal{Z}^i$. Thus, we define

$$H^i(\Omega^\cdot(\mathcal{X}), d+a) := \frac{\mathcal{Z}^i}{(d+a)(\Omega^{i-1}(\mathcal{X})) \cap \mathcal{Z}^i}$$

for $i = 0, 1, \ldots, n$.

If $f = (d+a)u$ belongs to $\mathcal{Z}^i$, then $(d+a)f = (d+a)^2 u = 0$, i.e., $\omega \wedge u = 0$ in $\mathcal{X}$, where $\omega = da$ is a given C^∞ two-form on $\mathcal{X}$. Denote by $\Omega^i_\omega(\mathcal{X})$ the subspace of $\Omega^i(\mathcal{X})$ consisting of all forms $u \in \Omega^i(\mathcal{X})$ which satisfy $\omega \wedge u = 0$ in $\mathcal{X}$. By the very definition, $\Omega^i_\omega(\mathcal{X})$ coincides with $\Omega^i(\mathcal{X})$ for $i = n-1, n$. Given any $u \in \Omega^i_\omega(\mathcal{X})$, we get

$$\begin{aligned} \omega \wedge (d+a)u &= \omega \wedge du + a \wedge (\omega \wedge u) = \\ &= (d+a)(\omega \wedge u) = 0, \end{aligned}$$

for ω is a closed two-form on $\mathcal{X}$. Hence it follows that $d+a$ maps $\Omega^i_\omega(\mathcal{X})$ into $\Omega^{i+1}_\omega(\mathcal{X})$ for all $i = 0, 1, \ldots, n-1$. We have thus associated the complex of linear mappings

$$0 \longrightarrow \Omega^0_\omega(\mathcal{X}) \xrightarrow{d+a} \Omega^1_\omega(\mathcal{X}) \xrightarrow{d+a} \ldots \xrightarrow{d+a} \Omega^n_\omega(\mathcal{X}) \longrightarrow 0 \qquad (0.2)$$

to sequence (0.1).

Complex (0.2) gains in interest if we realise that its cohomology just amounts to $H^i(\Omega^\cdot(\mathcal{X}), d+a)$ at each step i.

If the differential form a is closed in all of $\mathcal{X}$, then (0.1) is actually a complex and both (0.1) and (0.2) coincide. Otherwise $\Omega^i_\omega(\mathcal{X})$ need not be a space of sections of some vector bundle over $\mathcal{X}$. Indeed, the forms u of $\Omega^i_\omega(\mathcal{X})$ are described by the equation $\omega \wedge u = 0$ on $\mathcal{X}$. At any fixed point $x \in \mathcal{X}$ the equation reduces to a system of C^{i+2}_n linear equations for C^i_n unknowns which are the coefficients of $u(x)$. The dimension of the space of solutions to this system just amounts to C^i_n minus the rank of the bundle homomorphism $\Lambda^i T^*\mathcal{X} \to \Lambda^{i+2}T^*\mathcal{X}$ determined by the exterior multiplication with $\omega(x) = da(x)$. Hence, the dimension varies with x and the family of subspaces of $\Lambda^i T^*\mathcal{X}$ defined pointwise by $\omega \wedge u = 0$ fails to constitute a subbundle of $\Lambda^i T^*\mathcal{X}$ in general. However, if the rank of the bundle homomorphism $da(x)\wedge$ is constant, i.e., independent of $x \in \mathcal{X}$, then $\Omega^i_\omega(\mathcal{X})$ is specified within the framework of smooth sections of a subbundle V^i of $\Lambda^i T^*\mathcal{X}$. In this case (0.2) is a complex of first order differential operators between sections of vector bundles over $\mathcal{X}$. However, it is not elliptic, and so the Neumann problem after Spencer [9] does not apply to study the cohomology of complex (0.2).

Example 0.1. If a is an one-form on an open set $U \in \mathbb{R}^3$ with non-vanishing coefficient $\omega_{1,2}$ of $\omega = da$, then (0.2) reduces to

$$0 \longrightarrow 0 \longrightarrow C^\infty(U, \mathbb{R}^2) \xrightarrow{A^1} C^\infty(U, \mathbb{R}^3) \xrightarrow{A^2} C^\infty(U) \longrightarrow 0,$$

the principal symbols of A^1 and A^2 being

$$\begin{aligned} \sigma^1(A^1)(x,\xi) &= \begin{pmatrix} q_1(x)\xi_2 & q_2(x)\xi_2 - \xi_3 \\ -q_1(x)\xi_1 + \xi_3 & -q_2(x)\xi_1 \\ -\xi_2 & \xi_1, \end{pmatrix}, \\ \sigma^1(A^2)(x,\xi) &= (\xi_1, \xi_2, \xi_3) \end{aligned}$$

up to the factor $\imath = \sqrt{-1}$. Here,

$$q_1 = -\frac{\omega_{2,3}}{\omega_{1,2}}, \quad q_2 = \frac{\omega_{1,3}}{\omega_{1,2}}.$$

The rank of $\sigma^1(A^1)(x,\xi)$ is equal to 1 for all ξ on the line $\xi_1 = \xi_2$, $\xi_3 = (q_1 + q_2)\xi_2$ in $\mathbb{R}^3$.

Locally any closed differential form a is exact. If $a = dp$ for some smooth function p on $\mathcal{X}$, then an easy calculation shows that $(d+a)u = e^{-p}\, d\,(e^{p}u)$ for all forms $u \in \Omega^i(\mathcal{X})$. In this particular case complex (0.2) is obtained by a similarity transformation from the de Rham complex. This conjugation was used in the work [15] who proved Morse inequality using some spectral information of the Laplacian of the complex. The approach of [4] to solving the $\bar{\partial}$ problem is basically the same, and it is from 1965, predating [15] by a couple of decades. By varying the "weight" function p one can get L^2 estimates for the solution of $\bar{\partial}$ problem. The paper [4] does not make explicit that it is perturbing the Dolbeault complex, but that is exactly what it is doing. For a further development of this approach to Morse theory we refer the reader to [6] and the references given there.

In this work we focus on the cohomology of complex (0.2) in the case where the form a is not closed. Our basic assumption is that the rank of the bundle homomorphism $da(x)\wedge$ is constant.

1. The cohomology of the associated complex

A complex is said to be Fredholm if its cohomology is finite-dimensional at each step. This concept can be extended within the framework of nonstandard cohomology to arbitrary sequences (0.1). However, it should be noted that it differs from the concept of Fredholm quasicomplexes studied in [11]. Since (0.1) is a "small" perturbation of the de Rham complex on $\mathcal{X}$, it is a Fredholm quasicomplex in the sense of [11]. This allows one to introduce an Euler characteristic of (0.1). Still, the paper [11] does not contain any definition of cohomology for Fredholm quasicomplexes.

Lemma 1.1. *As defined above, the cohomology of complex (0.2) at step i coincides with $H^i(\Omega^{\cdot}(\mathcal{X}), d+a)$.*

Proof. For any $f \in \Omega^i_\omega(\mathcal{X}) \cap \mathcal{Z}^i$, we write $[f]$ for the equivalence class of f in $H^i(\Omega^i_\omega(\mathcal{X}))$. Let ι be the embedding of $\Omega^i_\omega(\mathcal{X}) \cap \mathcal{Z}^i$ into $\Omega^i(\mathcal{X}) \cap \mathcal{Z}^i$. Consider the mapping

$$\iota_* : H^i(\Omega^i_\omega(\mathcal{X})) \to H^i(\Omega^i_\omega(\mathcal{X}), d+a)$$

which assigns to any class $[f] \in H^i(\Omega^i_\omega(\mathcal{X}))$ the class of $\iota(f)$ in $H^i(\Omega^{\cdot}(\mathcal{X}), d+a)$. This definition is correct, for if $f = (d+a)u$ with some form $u \in \Omega^{i-1}_\omega(\mathcal{X})$, then $(d+a)f = (d+a)^2 u = \omega \wedge u = 0$, i.e., the class of $(d+a)f$ in $H^i(\Omega^{\cdot}(\mathcal{X}), d+a)$ is zero.

The mappng ι_* ist injective. For let $f \in \Omega^i_\omega(\mathcal{X}) \cap \mathcal{Z}^i$ satisfy $\iota_*[f] = 0$. Then there is a form $u \in \Omega^{i-1}(\mathcal{X})$ such that $f = (d+a)u$. Hence it follows immediately that

$$\begin{aligned} \omega \wedge u &= (d+a)^2 u = \\ &= (d+a)f = 0, \end{aligned}$$

i.e., $[f] = 0$, as desired.

It remains to show that ι_* ist surjective. To this end, pick an arbitrary form $f \in \mathcal{Z}^i$. Then $\omega \wedge f = (d+a)^2 f = 0$, and so $f \in \Omega^i_\omega(\mathcal{X})$ and $\iota_*[f]$ is the class of f in $H^i(\Omega^{\cdot}(\mathcal{X}), d+a)$. □

The sequence of symbol mappings in (0.1) is exact away from the zero section of $T^*\mathcal{X}$, and so (0.1) is an elliptic sequence in the interior of $\mathcal{X}$. Moreover, the sequence of boundary symbols is exact away from the zero section of $T^*\partial\mathcal{X}$, both conditions are usually referred to as ellipticity on a manifold with boundary. However, they characterise the Fredholm property in Sobolev spaces while we go beyond these spaces in (0.2).

2. Perturbations on 2-dimensional manifolds

To learn complex (0.2) we consider the particular case $n = 2$, i.e., $\mathcal{X}$ is a manifold of dimension 2.

If $da(x_0) \neq 0$ at some point $x_0 \in \mathcal{X}$, then by continuity $da(x) \neq 0$ holds for all x in a neighbourhood U of x_0 on $\mathcal{X}$. We restrict our attention to some local chart U in $\mathcal{X}$ with this property.

Write $a = a_1 dx^1 + a_2 dx^2$ in the coordinates of U, then $da = (\partial_1 a_2 - \partial_2 a_1) dx^1 \wedge dx^2$ is different from zero in U. By definition, we get

$$\Omega^0_\omega(U) \;\; = \;\; \{u \in C^\infty(U) : (\partial_1 a_2 - \partial_2 a_1) u \, dx^1 \wedge dx^2 = 0\} = \{0\}$$

and $\Omega^i_\omega(U) = \Omega^i(U)$ for $i = 1, 2$ since $da \wedge u$ has degree at least 3, if $u \in \Omega^i(U)$. Complex (0.2) thus becomes

$$0 \longrightarrow 0 \longrightarrow \Omega^1(U) \stackrel{d+a}{\longrightarrow} \Omega^2(U) \longrightarrow 0. \tag{2.1}$$

The symbol sequence of (2.1) over a point $(x, \xi) \in T^*U$ reduces immediately to the complex

$$0 \longrightarrow 0 \longrightarrow \Lambda^1 T^*_x U \stackrel{\xi\wedge}{\longrightarrow} \Lambda^2 T^*_x U \longrightarrow 0$$

which is not exact at the term $\Lambda^1 T^*_x U$, for $\xi \wedge \xi = 0$ for all $\xi \in T^*_x U$. The exactness at the term $\Lambda^2 T^*_x U$ is well known for all $\xi \in \mathbb{R}^2 \setminus \{0\}$. It follows that complex (2.1) fails to be elliptic.

We now turn to the cohomology of (2.1). Namely, given any form $f \in \Omega^2(U)$, we look for a solution $u \in \Omega^1(U)$ to the inhomogeneous equation $(d + a)u = f$ in U. On writing

$$\begin{aligned} u &= u_1 dx^1 + u_2 dx^2, \\ f &= f_{12} dx^1 \wedge dx^2 \end{aligned}$$

we reduce the equation to

$$(\partial_1 + a_1) u_2 - (\partial_2 + a_2) u_1 = f_{12}$$

in U.

An important class of solutions to this equation is constituted by the so-called potential solutions, i.e., those of the form $u = dp$ where p is a smooth function in U satisfying $a \wedge dp = f$, i.e.,

$$a_1 \partial_2 p - a_2 \partial_1 p = f_{12}.$$

This first order partial differential equation is known to have a unique solution p with prescribed data on any curve $\mathcal{S}$ in U which is not characteristic, i.e., the vector field $(-a_2, a_1)$ is tangent to $\mathcal{S}$ at no point. Hence it follows that the cohomology of complex (2.1) is infinite dimensional at the term $\Omega^1(U)$ and zero at the term $\Omega^2(U)$.

Thus, sequence (0.1) fails to be Fredholm in the sense of Section 1. unless the form $a(x)$ is closed. A substantial theory is hardly expected for the case of nonclosed differential forms a, cf. [6]. If a is a closed one-form, then complex (0.1) falls into a useful general construction in homological algebra called the Koszul complex, see Section 1.2.8 of [10].

3. Quasicomplexes

In this section we recall some basic facts about complexes and quasicomplexes in Hilbert spaces. For the theory of quasicomplexes of Banach spaces we refer to [2] where quasicomplexes are called essential complexes.

Let us consider the sequence

$$(H^{\cdot}, d): \ 0 \longrightarrow H^0 \xrightarrow{d^0} H^1 \xrightarrow{d^1} \dots \xrightarrow{d^{N-1}} H^N \longrightarrow 0$$

where H^i are Hilbert spaces and d^i are continuous linear maps. The sequence $(H^{\cdot}, d)$ is called a complex if $d^i d^{i-1} = 0$ for all $i = 1, \dots, N$. The elements of the spaces $Z^i(H^{\cdot}, d) = \ker d^i$ and $B^i(H^{\cdot}, d) = \operatorname{im} d^{i-1}$ are called cocycles and coboundaries, respectively. The quotient space $H^i(H^{\cdot}, d) = \ker d^i / \operatorname{im} d^{i-1}$ is the cohomology of the complex $(H^{\cdot}, d)$ at step i. The complex $(H^{\cdot}, d)$ is said to be Fredholm if its cohomology is finite dimensional at each step $i = 0, \dots, N$.

It is well known that "small" perturbations of Fredholm operators do not affect the Fredholm property. For example, perturbing a single Fredholm operator by a compact operator gives us a Fredholm operator. It would be natural to have the same property for Fredholm complexes. However, a "small" perturbation of a Fredholm complex need not be even a complex anymore, i.e., the operators no longer satisfy $d^i d^{i-1} = 0$.

Note that perturbing an elliptic complex by lower order terms does not change the complex of principal symbols which remains to be exact. Hence, instead of complexes it is natural to consider sequences $(H^{\cdot}, d)$ with the property that the compositions $d^i d^{i-1}$ are "small" in some sense. By "small" operators one usually means compact operators. Let us denote by $\mathcal{K}(H, \tilde{H})$ the subspace of $\mathcal{L}(H, \tilde{H})$ consisting of compact operators.

Definition 3.1. *A sequence $(H^{\cdot}, d)$ of Hilbert spaces H^i and continuous linear maps d^i is a quasicomplex if $d^i d^{i-1} \in \mathcal{K}(H^{i-1}, H^{i+1})$ for all $i = 1, \dots, N$.*

For $d^1, d^2 \in \mathcal{L}(H, \tilde{H})$, we write $d^1 \sim d^2$ if $d^1 - d^2 \in \mathcal{K}(H, \tilde{H})$. It is known that an operator $d \in \mathcal{L}(H, \tilde{H})$ is Fredholm if and only if its image in the Calkin algebra $\mathcal{L}(H, \tilde{H}) / \mathcal{K}(H, \tilde{H})$ is invertible. Hence, the idea of Fredholm quasicomplexes is to pass in a given quasicomplex to quotients modulo spaces of compact operators and require exactness. To make the definition precise we introduce a functor ϕ_Σ studied in [7].

Taking an arbitrary Hilbert space Σ, we set $\phi_\Sigma(H^i) = \mathcal{L}(\Sigma, H^i) / \mathcal{K}(\Sigma, H^i)$ for each Hilbert space H^i. Then, for any map $d^i \in \mathcal{L}(H^i, H^{i+1})$, we introduce a map $\phi_\Sigma(d^i) \in \mathcal{L}(\phi_\Sigma(H^i), \phi_\Sigma(H^{i+1}))$ by

$$\phi_\Sigma(d^i)(A + \mathcal{K}(\Sigma, H^i)) = d^i A + \mathcal{K}(\Sigma, H^{i+1})$$

for all $A \in \mathcal{L}(\Sigma, H^i)$. Obviously, this operator is well defined. It is easily seen that $\phi_\Sigma(d^i d^{i-1}) = \phi_\Sigma(d^i) \phi_\Sigma(d^{i-1})$ and that ϕ_Σ vanishes on compact operators for every Hilbert space Σ. Hence, if $(H^{\cdot}, d)$ is a quasicomplex then $(\phi_\Sigma(H^{\cdot}), \phi_\Sigma(d))$ is a complex for each Hilbert space Σ.

Definition 3.2. *A quasicomplex $(H^{\cdot}, d)$ is Fredholm if the complex $(\phi_\Sigma(H^{\cdot}), \phi_\Sigma(d))$ is exact for each Hilbert space Σ.*

Let $(H^{\cdot}, d)$ and $(H^{\cdot}, \tilde{d})$ be two quasicomplexes of Hilbert spaces, such that $d^i \sim \tilde{d}^i$ for any $i = 0, 1, \dots, N$. Then the complexes $(\phi_\Sigma(H^{\cdot}), \phi_\Sigma(d))$ and $(\phi_\Sigma(H^{\cdot}), \phi_\Sigma(\tilde{d}))$ coincide. Hence, the quasicomplexes $(H^{\cdot}, d)$ and $(H^{\cdot}, \tilde{d})$ are Fredholm simultaneously. Thus, any compact perturbation of a Fredholm quasicomplex is a Fredholm quasicomplex.

A sequence

$$(H^{\cdot}, \pi): \ 0 \longleftarrow H^0 \xleftarrow{\pi^1} H^1 \xleftarrow{\pi^2} \dots \xleftarrow{\pi^N} H^N \longleftarrow 0$$

with $\pi^i \in \mathcal{L}(H^i, H^{i-1})$ is said to be a parametrix of the quasicomplex $(H^{\cdot}, d)$, provided

$$\pi^{i+1} d^i + d^{i-1} \pi^i = I_{H^i} - \varkappa^i$$

for all $i = 0, 1, \dots, N$, where $\varkappa^i \in \mathcal{K}(H^i)$.

It is well known that a complex of Hilbert spaces is Fredholm if and only if it has a parametrix. The same property is also true for quasicomplexes, see [11].

Theorem 3.3. *A quasicomplex* $(H^\cdot, d)$ *is Fredholm if and only if it possesses a parametrix.*

Obviously, if a parametrix $(H^\cdot, \pi)$ of a quasicomplex $(H^\cdot, d)$ is a quasicomplex itself, then $(H^\cdot, d)$ is in turn a parametrix of $(H^\cdot, \pi)$.

It should be noted that Theorem 3.3 does not extend to arbitrary complexes of Banach spaces. The advantage of using Hilbert spaces lies in the fact that any quasicomplex of Hilbert spaces admits the so-called adjoint quasicomplex. By this is meant

$$(H^\cdot, d^*): \ 0 \longleftarrow H^0 \overset{d^{0*}}{\longleftarrow} H^1 \overset{d^{1*}}{\longleftarrow} \ldots \overset{d^{N-1*}}{\longleftarrow} H^N \longleftarrow 0,$$

where $d^{i*} \in \mathcal{L}(H^{i+1}, H^i)$ stands for the adjoint of d^i in the sense of Hilbert spaces. Since $d^{i*}d^{i+1*} = (d^{i+1}d^i)^*$ are compact operators, $(H^\cdot, d^*)$ is a quasicomplex indeed. The selfadjoint operators $\Delta^i = d^{i-1}d^{i-1*} + d^{i*}d^i$ are called the Laplacians of the quasicomplex $(H^\cdot, d)$. The null-space of Δ^i consists of all $h \in H^i$ satisfying $d^i h = 0$ and $d^{i-1*}h = 0$, as is easy to see.

Theorem 3.4. *A quasicomplex* $(H^\cdot, d)$ *is Fredholm if and only if all its Laplacians* Δ^i *are Fredholm operators.*

Proof. See Lemma 4.2 of [12]. □

As is proved in [11], every Fredholm quasicomplex can actually be transformed into a complex. Another way of stating this theorem is to say that each Fredholm quasicomplex is a perturbation of a Fredholm complex by compact operators.

Theorem 3.5. *For any Fredholm quasicomplex* $(H^\cdot, d)$ *there are operators* $D^i \in \mathcal{L}(H^i, H^{i+1})$ *satisfying* $D^i \sim d^i$ *and* $D^i D^{i-1} = 0$ *for all* i.

4. A parametrix of the perturbation

We now return to the perturbation of the de Rham complex on $\mathcal{X}$ defined in (0.1). In order to rewrite it in the context of Hilbert spaces, we choose any integer number $s \geqslant n$ and set $s_i = s - i$ for $i = 0, 1, \ldots, n$. Consider the sequence of linear mappings

$$0 \longrightarrow H^{s_0}(\mathcal{X}, \Lambda^0) \overset{d+a}{\longrightarrow} H^{s_1}(\mathcal{X}, \Lambda^1) \overset{d+a}{\longrightarrow} \ldots \overset{d+a}{\longrightarrow} H^{s_n}(\mathcal{X}, \Lambda^n) \longrightarrow 0, \tag{4.1}$$

where $\Lambda^i = \Lambda^i T^* \mathcal{X}$ is the bundle of exterior forms of degree i over $\mathcal{X}$ and by $H^{s_i}(\mathcal{X}, \Lambda^i)$ is meant the space of all differential forms of degree i with coefficients of the Sobolev class $H^{s_i} = W^{s_i,2}$ on $\mathcal{X}$. We fix a unitary structure in each of these spaces, thus obtaining a sequence of Hilbert spaces and their continuous linear mappings. By the above, the curvature of (4.1) just amounts to the bundle homomorphism of Λ^{i-1} to Λ^{i+1} defined via the exterior multiplication by the two-form da. On applying the Rellich theorem we conclude readily that it is a compact operator from $H^{s_{i-1}}(\mathcal{X}, \Lambda^{i-1})$ to $H^{s_{i+1}}(\mathcal{X}, \Lambda^{i+1})$ for all $i = 1, \ldots, n$. Hence, (4.1) is a quasicomplex.

We next show that this quasicomplex is Fredholm. By Theorem 3.3, for this purpose it suffices to construct a parametrix of (4.1). To this end, we use the parametrix of the de Rham complex on $\mathcal{X}$ obtained from the Neumann problem after Spencer, see Section 4 of [5]. More precisely, there are operators G^i of order -2 in Boutet de Monvel's algebra of pseudodifferential operators acting in $\Omega^i(X)$, such that

$$f = Hf + (d^*G)df + d(d^*G)f \tag{4.2}$$

for all $f \in \Omega^i(\mathcal{X})$. Here, H stands for the orthogonal projection in $L^2(\mathcal{X}, \Lambda^i)$ onto the finite-dimensional subspace of harmonic forms, i.e., those $h \in \Omega^i(\mathcal{X})$ which satisfy $dh = 0$, $d^*h = 0$ in $\mathcal{X}$ and $n(f) = 0$ on the boundary of $\mathcal{X}$, where $n(f)$ is the normal part of f on $\partial\mathcal{X}$. By d^* is meant the formal adjoint for the exterior derivative with respect to the L^2 scalar product in

$\Omega^i(\mathcal{X})$. The operator G^i satisfies $n(Gf) = 0$ and $n(dGf) = 0$ on the boundary of $\mathcal{X}$ for any $f \in \Omega^i(\mathcal{X})$. It is usually referred to as the Green operator of the Hodge theory on manifolds with boundary. Thus, on sufficiently smooth forms $P = d^*G$ is a very special parametrix of the de Rham complex.

Lemma 4.1. *Let s be an arbitrary nonnegative integer. As defined above, the Green operator G extends to a continuous mapping of $H^s(\mathcal{X}, \Lambda^i)$ into $H^{s+2}(\mathcal{X}, \Lambda^i)$.*

Proof. Since $u = Gf$ gives a solution to the Neumann problem after Spencer for the de Rham complex on $\mathcal{X}$ and this latter problem is elliptic, the desired assertion follows from the regularity theorem for elliptic boundary value problems in Sobolev spaces, see [1] and elsewhere. □

The Rellich theorem on compact embeddings of Sobolev spaces implies that sequence (4.1) is a compact perturbation of the de Rham complex evaluated in Sobolev spaces. The shortest way to derive the Fredholm property of (4.1) from here is given by the next theorem.

Theorem 4.2. *The sequence of pseudodifferential operators $P^i = d^*G^i$ of order -1 defines a parametrix of sequence (4.1).*

Proof. Using equality (4.2), we get

$$\begin{aligned} P^{i+1}(d+a)f + (d+a)P^i f &= (P^{i+1}df + dP^i f) + (P^{i+1}(a \wedge f) + a \wedge P^i f) = \\ &= f - K^i f \end{aligned} \tag{4.3}$$

for all $f \in H^{s_i}(\mathcal{X}, \Lambda^i)$, where

$$K^i f = H^i f - \left(P^{i+1}(a \wedge f) + a \wedge P^i f\right).$$

The projector H^i is an operator with smooth Schwartz kernel on the product $\mathcal{X} \times \mathcal{X}$, and so it is a compact operator on $H^{s_i}(\mathcal{X}, \Lambda^i)$. On the other hand, the operator $P^{i+1}(a \wedge \cdot) + a \wedge P^i \cdot$ acts on $H^{s_i}(\mathcal{X}, \Lambda^i)$ through the compact embedding $H^{s_i+1}(\mathcal{X}, \Lambda^i) \hookrightarrow H^{s_i}(\mathcal{X}, \Lambda^i)$, which is due to the Rellich theorem. Hence, this operator is compact as well. On summing up we conclude that $K^i \in \mathcal{K}(H^{s_i}(\mathcal{X}, \Lambda^i))$, as desired. □

On applying the paper [11] we are in a position to introduce the Euler characteristic of Fredholm quasicomplex (4.1). To wit, $\chi(H^{s\cdot}(\mathcal{X}, \Lambda^\cdot), d + a)$ is defined to be equal to the Euler characteristic $\chi(\mathcal{X})$ of $\mathcal{X}$, i.e., to that of the de Rham complex on $\mathcal{X}$.

Our next goal is to improve the parametrix $P^i = d^*G^i$ of quasicomplex (4.1). It is surprising that the standard procedure using the formal Neumann series for $(I - K^i)^{-1}$ no longer works to do this modulo smoothing operators. By abuse of notation we omit the indices of P^i, K^i, etc., thus using the graded operators P and K defined by $Pf = P^i f$ and $Kf = K^i f$ for $f \in \Omega^i(\mathcal{X})$, respectively. According to (4.3) we get

$$K = -Pa - aP$$

modulo the harmonic projection.

Lemma 4.3. *For any $k = 0, 1, \ldots$, the commutator $[K^k, d+a] := K^k(d+a) - (d+a)K^k$ just amounts to*

$$[K^k, d + a] = -\sum_{j=0}^{k-1} K^j \,[P, (da)]\, K^{k-1-j}. \tag{4.4}$$

Recall that by a is meant the operator on differential forms given by $f \mapsto a \wedge f$. In contrast to the composition $da = d \circ a$ we write (da) for the operator defined by the differential of a.

Proof. For $k = 0$ the assertion is obvious, so we start with $k = 1$. We first observe that

$$\begin{aligned}[][K, d] &= [H - Pa - aP, d] = \\ &= -[Pa + aP, d],\end{aligned}$$

for $[H, d] = 0 - 0$ is zero by the very definition of harmonic projection H. Hence it follows that

$$\begin{aligned}[][K, d] &= dPa + (da) \wedge P - adP - Pad - aPd = \\ &= (dP + Pd)a - Pda - Pad + (da) \wedge P - a(dP + Pd) = \\ &= (I - H)a - [P, (da)] - a(I - H) = \\ &= -[H, a] - [P, (da)]\end{aligned}$$

the commutator $[H, a]$ being smoothing and of finite rank. On the other hand, we get

$$[K, a] = (H - Pa - aP)a - a(H - Pa - aP) = [H, a],$$

for $a \wedge a = 0$. Thus, $[K, d + a] = -[P, (da)]$, as desired.

For arbitrary integer $k > 1$ we proceed successively using the equality for $k = 1$. To wit,

$$\begin{aligned}[][K^k, d + a] &= K^{k-1}K(d + a) - (d + a)K^k = \\ &= K^{k-1}(d + a)K - K^{k-1}[P, (da)] - (d + a)K^k = \\ &= K^{k-2}(d + a)K^2 - K^{k-2}[P, (da)]K - K^{k-1}[P, (da)] - (d + a)K^k,\end{aligned}$$

etc., which proves (4.4). □

Note that each summand on the right-hand side of (4.4) is a pseudodifferential operator of order $-k$.

Theorem 4.4. *Given any* $N = 0, 1, \ldots$, *the operators* $P_N^i = \Big(\sum_{k=0}^{N} (K^{i-1})^k\Big) P^i$ *satisfy*

$$P_N\,(d + a) + (d + a)\,P_N = I - K^{N+1} + \sum_{k=0}^{N} \Big(\sum_{j=0}^{k-1} K^j\,[P, (da)]\,K^{k-1-j}\Big) P.$$

Proof. On multiplying both sides of (4.3) by $\sum_{k=0}^{N} K^k$ from the left we immediately obtain

$$P_N\,(d + a) + (d + a)\,P_N = I - K^{N+1} - \Big[\sum_{k=0}^{N} K^k, d + a\Big]\,P.$$

Substituting the expressions for $[K^k, d + a]$ of (4.4) into the latter equality yields the desired formula. □

In particular, if the differential form a is closed, then the operators P_N^i constitute a parametrix of complex (4.1) up to a remainder of order $-N - 1$, more precisely, K^{N+1}. For perturbations of nonzero curvature there is an additional residual term depending linearly on da.

5. Local calculation of the Laplacian

Denote by $\Delta_a = (d + a)^*(d + a) + (d + a)(d + a)^*$ the Laplacian of the perturbed complex. A trivial verification shows that

$$\Delta_a = \Delta_0 + (d^*a + a^*d + da^* + ad^*) + (a^*a + aa^*),$$

where Δ_0 is the Laplacian of the de Rham complex on $\mathcal{X}$. From this equality we deduce immediately that Δ_a is an elliptic operator of order two at each step $i = 0, 1, \ldots, n$.

The Laplacian Δ_a has especially simple form $\Delta_0 + (a^*a + aa^*)$ if the first order differential operator $d^*a + a^*d + da^* + ad^*$ vanishes. This latter is the case if and only if the one-form a satisfies an overdetermined system of first order partial differential equations on $\mathcal{X}$. Since

$$\begin{aligned} d^* &= (-1)^q *^{-1} d*, \\ a^* &= (-1)^{q-1} *^{-1} a* \end{aligned}$$

holds on q-forms, where by $*$ is meant the Hodge star operator related to a Riemannian metric on $\mathcal{X}$, the system for a reduces to $[d, *^{-1}a*] + [*^{-1}d*, a] = 0$ up to the inessential multiple $(-1)^{q-1}$ on q-forms. It is linear over $\mathbb{R}$ but fails to be so over $\mathbb{C}$.

To see if the system possesses solutions a among nonclosed one-forms a, we consider it in a local chart U on $\mathcal{X}$ with coordinates $x = (x^1, \ldots, x^n)$. In these coordinates the form a can be written as $a = a_1 dx^1 + \ldots + a_n dx^n$, where $a_1, \ldots, a_n$ are smooth functions of x. Moreover, any bundle Λ^q is trivial over U under the representation of a form

$$u(x) = \sum_{\substack{I=(i_1,\ldots,i_q) \\ 1\leqslant i_1<\ldots<i_q\leqslant n}} u_I(x)\, dx^{i_1} \wedge \ldots \wedge dx^{x_q}$$

by the k_q-column $(u_I(x))$ of its coefficients, k_q being the binomial coefficient C_n^q. The operator $d+a$ on q-forms is represented by a $(k_{q+1} \times k_q)$-matrix of first order partial differential operators. On assuming the canonical metric on $\mathbb{C}^{k_q}$ we get the formula

$$\Delta_a = \Delta_0 - \sum_{j=1}^{n} \Big((a_j(x) - \overline{a_j(x)}) \frac{\partial}{\partial x^j} + \frac{1}{2} \frac{\partial}{\partial x^j} (a_j(x) - \overline{a_j(x)}) \Big) + |a(x)|^2$$

provided the coefficients of a satisfy

$$\frac{\partial}{\partial x^k} a_j + \frac{\partial}{\partial x^j} \overline{a}_k = 0 \tag{5.1}$$

for all $1 \leqslant j \leqslant k \leqslant n$.

Example 5.1. For $i = 1, \ldots, n$, choose $a_i(x) = a_{i,1}x^1 + \ldots + a_{i,n}x^n + c_i$, where $A = (a_{i,j})$ is an $(n \times n)$-matrix of complex numbers and c_i complex numbers independent of x. Then system (5.1) is fulfilled if and only if $a_{j,k} + \overline{a}_{k,j} = 0$ for all $1 \leqslant j \leqslant k \leqslant n$, i.e., the matrix A is skew-Hermitean. An easy calculation shows that

$$da = - \sum_{1\leqslant i<j\leqslant n} \frac{\partial}{\partial x^j} (a_i(x) + \overline{a_i(x)})\, dx^i \wedge dx^j = - \sum_{1\leqslant i<j\leqslant n} (a_{i,j} + \overline{a}_{i,j})\, dx^i \wedge dx^j,$$

which need not vanish.

Under conditions (5.1), if moreover the differential form a is real valued, then the Laplacian Δ_a reduces to $\Delta_0 + |a(x)|^2$.

6. Analytic torsion

For elliptic complexes on compact manifolds the cohomology is represented by harmonic sections, i.e., those belonging to the null spaces of Laplacians. Hence, the harmonic spaces might substitute for the cohomology of Fredholm quasicomplexes, provided the Hodge theory holds. For manifolds with boundary the Hodge theory reduces to the Neumann problem after Spencer, and so the question arises if it is elliptic.

The calculations of Section 2 show that the Laplace operators under the homogeneous Neumann conditions actually remain the only efficient tool to reveal resolving properties of an arbitrary sequence of differential operators on a manifold with boundary. The null spaces of the Laplace operators may substitute for the cohomology of such a sequence while their dimensions can be thought of as generalised Betti numbers. This agrees completely with the definition of the Euler characteristics for quasicomplexes of Banach spaces given in [11]. Moreover, the Fredholm property of the Neumann problem after Spencer allows one to introduce the concept of analytic torsion for a sequence. If it is independent on the choice of Hermitean structure of the sequence and thus reveals a topological nature should be a subject of special treatments. In this section we outline the concept of analytic torsion following to [8].

The Reidemeister torsion is a global invariant of a cell decomposition of a manifold and of an acyclic representation of its fundamental group. It is an invariant of the piecewise linear structure of the manifold. The Reidemeister torsion for an arbitrary finite-dimensional unimodular representation of the fundamental group can be defined as a canonical norm on the determinant line of the cohomology of a manifold. It is a multiplicative analogue of the Euler characteristic in the case of manifolds of odd dimension. (The Euler characteristic of a closed manifold is trivial in the odd-dimensional case.) Formulas for the Reidemeister torsion of the Cartesian product of two manifolds are similar to the multiplicative property of the Euler characteristic. The analytic torsion was introduced in [8] for a closed Riemannian manifold $\{\mathcal{X}, g\}$ with an acyclic orthogonal representation of the fundamental group $\pi_1(\mathcal{X})$. It is equal to the product of the corresponding powers of the determinants of the Laplace operators on differential forms of $\Omega^{\cdot}(\mathcal{X})$. These determinants are regularised with the help of the zeta-functions of the Laplacians. (The Reidemeister torsion can also be written by analogous formula where the Riemannian Laplacians are replaced by the combinatorial ones.) The analytic torsion is defined with the help of a Riemannian metric g on $\mathcal{X}$. However, it is independent of g in the acyclic case, see [8]. So it is an invariant of the smooth structure on $\mathcal{X}$ and behaves in much the same way as the Reidemeister torsion. As was conjectured in [8], for any compact closed manifold $\mathcal{X}$ and acyclic representation ρ of the fundamental group $\pi_1(\mathcal{X})$ the Reidemeister torsion of $\{\mathcal{X}, \rho\}$ has proven to be equal to the analytic torsion of $\{\mathcal{X}, \rho\}$.

On returning to sequence (0.1) we observe that it coicides with the de Rham complex up lower order operators. Hence, in the algebra of boundary value problems on $\mathcal{X}$ the sequence bears the same principal symbol structure as the de Rham complex. In particular, the tangential and normal components of a differential form $u \in \Omega^i(\mathcal{X})$ on the boundary of $\mathcal{X}$ with respect to sequence (0.1) coincide with those with respect to the de Rham complex. They are denoted by $t(u)$ and $n(u)$, respectively, so that $u = t(u) + d\varrho \wedge n(u)$ on $\partial\mathcal{X}$ where ϱ is a defining function of the boundary with $|d\varrho| = 1$ on $\partial\mathcal{X}$, see Section 3.2.2 in [10]. The Neumann problem after Spencer for sequence (0.1) at step i consists in finding, given any $f \in \Omega^i(\mathcal{X})$, a form $u \in \Omega^i(\mathcal{X})$ satisfying

$$\begin{array}{rcll} \Delta_a u & = & f & \text{in} \quad \mathcal{X}, \\ n(u) & = & 0 & \text{on} \quad \partial\mathcal{X}, \\ n((d+a)u) & = & 0 & \text{on} \quad \partial\mathcal{X}. \end{array} \tag{6.1}$$

By the above, the Laplace operator Δ_a^i is elliptic in $\mathcal{X}$. The boundary conditions are coercive for the Laplacian, see e.g. [5]. Thus, (6.1) is a classical elliptic boundary value problem in $\mathcal{X}$. Moreover, it is formally selfadjoint with respect to the Green formula.

Elliptic theory applies well to problem (6.1). All solutions $u \in H^2(\mathcal{X}, \Lambda^i)$ of the homogeneous problem corresponding to (6.1) belong actually to the space $\Omega^i(\mathcal{X})$, and they form a finite dimensional space $\mathcal{H}_a^i(\mathcal{X})$. The elements of $\mathcal{H}_a^i(\mathcal{X})$ are called harmonic forms. Denote by H_a the orthogonal projection of $L^2(\mathcal{X}, \Lambda^i)$ onto $\mathcal{H}_a^i(\mathcal{X})$. Given any $f \in L^2(\mathcal{X}, \Lambda^i)$, the equation $\Delta_a u = f - H_a f$ has a unique solution u in $H^2(\mathcal{X}, \Lambda^i)$ orthogonal to $\mathcal{H}_a^i(\mathcal{X})$. The operator $f \mapsto G_a f := u$ is called the Green (or Neumann after Spencer) operator and it maps $L^2(\mathcal{X}, \Lambda^i)$

continuously into $H^2(\mathcal{X}, \Lambda^i)$. The Green operator is known to be a selfadjoint operator in $L^2(\mathcal{X}, \Lambda^i)$. Its smoothness properties are deduced from the fact that G_a is an operator of order -2 in the algebra of boundary value problems on $\mathcal{X}$ outlined in [3]. By the above,

$$f = H_a f + P_a(d+a)f + (d+a)P_a f + (d+a)^*[d+a, G_a]f$$

for all $f \in L^2(\mathcal{X}, \Lambda^i)$, where $P_a = (d+a)^* G_a$. The commutator $[G_a, d+a]$ need not vanish unless a is a closed one-form.

Lemma 6.1. *Suppose that the perturbation satisfies $n(da) = 0$ on the boundary of $\mathcal{X}$. Then*

$$[d+a, G_a] = G_a[(d+a)^*, da]G_a. \tag{6.2}$$

Proof. Denote by $\mathcal{N}_a^i(\mathcal{X})$ the subspace of $\Omega^i(\mathcal{X})$ consisting of those differential forms u which satisfy the boundary conditions in (6.1), i.e., $n(u) = 0$ and $n((d+a)u) = 0$ on $\partial\mathcal{X}$.

We first show that $A = d + a$ maps $\mathcal{N}_a^i(\mathcal{X})$ continuously into $\mathcal{N}_a^{i+1}(\mathcal{X})$. If $u \in \mathcal{N}_a^i(\mathcal{X})$, then $n(Au) = 0$ and it remains to check if $n(A^2 u) = 0$ holds. We get $A^2 u = da \wedge u$ and

$$\begin{aligned} da \wedge u &= (t(da) + d\varrho \wedge n(da)) \wedge (t(u) + d\varrho \wedge n(u)) = \\ &= t(da) \wedge t(u) + d\varrho \wedge (t(da) \wedge n(u) + n(da) \wedge t(u)) \end{aligned}$$

on $\partial\mathcal{X}$ whence $n(da \wedge u) = 0$, as desired.

On applying the operator A to the equality $I = H_a + \Delta_a G_a$ on $\Omega^i(\mathcal{X})$ from the left and from the right we see that $A\Delta_a G_a = \Delta_a G_a A$, for both AH_a and $H_a A$ vanish. Thus,

$$\Delta_a[A, G_a] = [\Delta_a, A]G_a$$

and so the equality $G_a \Delta_a [A, G_a]u = G_a[\Delta_a, A]G_a u$ holds for all $u \in \Omega^i(\mathcal{X})$.

From the construction of the Green operator we deduce that $(G_a \Delta_a)g = g - H_a g$ for all $g \in \mathcal{N}_a^{i+1}(\mathcal{X})$. Now, if $u \in \Omega^i(\mathcal{X})$ then $g = [A, G_a]u = A(G_a u) - G_a(Au)$ belongs to the space $\mathcal{N}_a^{i+1}(\mathcal{X})$, for A maps $\mathcal{N}_a^i(\mathcal{X})$ into $\mathcal{N}_a^{i+1}(\mathcal{X})$. Hence it follows that

$$G_a \Delta_a [A, G_a] = (I - H_a)[A, G_a]$$

is valid on all of $\Omega^i(\mathcal{X})$. As $H_a A = 0$ and $H_a G_a = 0$, we get $G_a \Delta_a [A, G_a] = [A, G_a]$, on the one hand.

On the other hand, an easy calculation gives

$$\begin{aligned} [\Delta_a, A] &= (A^*A + AA^*)A - A(A^*A + AA^*) = \\ &= A^*A^2 - A^2A^* = \\ &= [A^*, A^2] \end{aligned}$$

showing (6.2). □

From Lemma (6.1) it follows that $P_a(d+a)f + (d+a)P_a f = f - K_a f$ for all $f \in \Omega^i(\mathcal{X})$, where $K_a = H_a + A^* G_a [A^*, da] G_a$ is a pseudodifferential operator of order -2 on $\mathcal{X}$. Hence, P_a is a parametrix of sequence (4.1) whose remainder K_a is "smaller" than the remainder of the parametrix constructed in Theorem 4.2. It is to be expected that the standard procedure using the formal Neumann series for $(I - K_a)^{-1}$ works to construct a parametrix modulo smoothing operators but we will not develop this point here.

Given any nonnegative linear mapping L of a unitary space V, one defines the zeta function of L by $\zeta_L(s) = \operatorname{tr} L^{-s}$, where tr stands for the functional trace. Thus, we get

$$\zeta_L(s) = \sum_j \lambda_j^{-s},$$

where the sum is over all nonzero eigenvalues of L. For selfadjoint mappings L determined by elliptic operators or boundary value problems, it is known that the eigenvalues of L can be arranged in a monotone increasing sequence converging to infinity. Moreover, $\zeta_L(s)$ has a meromorphic extension to the entire complex plane which is regular at $s = 0$. On formally differentiating the series term-by-term one obtains

$$\zeta_L'(s) = -\sum_j \frac{\log \lambda_j}{\lambda_j^s},$$

and so if the functional determinant is well defined and different from zero, then it should be given by $\det L = \exp(-\zeta_L'(0))$. Since the analytic continuation of the zeta function is regular at zero, this can be rigorously adopted as a definition of the determinant.

For a sequence

$$(H^\cdot, d): \ 0 \longrightarrow H^0 \stackrel{d^0}{\longrightarrow} H^1 \stackrel{d^1}{\longrightarrow} \ldots \stackrel{d^{N-1}}{\longrightarrow} H^N \longrightarrow 0$$

of unitary spaces, the analytic torsion (or determinant) is now introduced by the formula

$$\begin{aligned} \log T(H^\cdot, d) &= \frac{1}{2}\sum_{i=0}^{N} (-1)^{i+1} i \log\det \Delta^i = \\ &= \sum_{i=0}^{N} \log\left(\det \Delta^i\right)^{(-1)^{i+1} i/2} = \\ &= \log \prod_{i=0}^{N} \left(\det \Delta^i\right)^{(-1)^{i+1} i/2} \end{aligned}$$

or

$$T(H^\cdot, d) = \prod_{i=0}^{N} \left(\det \Delta^i\right)^{(-1)^{i+1} i/2}, \qquad (6.3)$$

where $\Delta^i = d^{i-1} d^{i-1\,*} + d^{i\,*} d^i$ are the Laplacians of the sequence $(H^\cdot, d)$. It is easy to check that for short sequences (i.e, $N = 1$) the analytic torsion reduces to $|\det d^0|$.

Example 6.2. Consider the sequence of symbols corresponding to perturbation (0.1). It is the complex

$$(\Lambda^\cdot T_x^*\mathcal{X}, \sigma(\xi)): \ 0 \longrightarrow \Lambda^0 T_x^*\mathcal{X} \stackrel{\sigma(\xi)}{\longrightarrow} \Lambda^1 T_x^*\mathcal{X} \stackrel{\sigma(\xi)}{\longrightarrow} \ldots \stackrel{\sigma(\xi)}{\longrightarrow} \Lambda^n T_x^*\mathcal{X} \longrightarrow 0$$

parametrised by a point $x \in \mathcal{X}$ and a vector $\xi \in T_x^*\mathcal{X}$ different from zero, where $\sigma(\xi)v = \xi \wedge v$ for all $v \in \Lambda^i T_x^*\mathcal{X}$. Choosing an orthonormal basis in each tangent space $T_x\mathcal{X}$, we endow the spaces $\Lambda^i T_x^*\mathcal{X}$ with a unitary structure in the usual manner. The dimension of $\Lambda^i T_x^*\mathcal{X}$ is equal to the binomial coefficient $k_i := C_n^i$. The Laplacians of the complex $(\Lambda^\cdot T_x^*\mathcal{X}, \sigma(\xi))$ reduce to $\Delta^i = |\xi|^2 I^i$, where I^i is the identity mapping of $\Lambda^i T_x^*\mathcal{X}$. Hence it follows that $\det \Delta^i = |\xi|^{2k_i}$, and so the analytic torsion is

$$\begin{aligned} T(\Lambda^\cdot T_x^*\mathcal{X}, \sigma(\xi)) &= \prod_{i=0}^{n} \left(|\xi|^{2k_i}\right)^{(-1)^{i+1} i/2} = \\ &= |\xi|^{-0\,k_0 + 1\,k_1 - \ldots + (-1)^{n-1} n\,k_n}. \end{aligned}$$

If $n = 1$, then the analytic torsion just amounts to $|\xi|$. For all $n > 1$, it reduces to 1 since the exponent is $n(1-1)^{n-1}$ by the binomial formula.

Formula (6.3) can be used to introduce the concept of analytic torsion for quasicomplex (0.1). To this end, as Δ^i one takes the Laplacians Δ_a^i of (0.1) under the homogeneous Neumann (after Spencer) boundary conditions, cf. (6.1). Of course, the product of the eigenvalues of these Laplacians is infinite, so the notion of determinant must be regularised by means of $\det \Delta_a^i = \exp\left(-\zeta'_{\Delta_a}(0)\right)$. If defined in this way, the analytic torsion is a subtle spectral invariant. It is easy to check that the zeta function for the Laplacian on the simplest manifold with boundary, the interval, is a slight variation of the Riemann zeta function. Geometers can perhaps study the dependence of analytic torsion on the particular choice of Riemannian metric on $\mathcal{X}$, a technique usually unavailable to number theorists due to rigidity results.

The first author gratefully acknowledges the financial support of the Islamic Center Association for Guidance and Higher Education.

References

[1] S.Agmon, A.Douglis, L.Nirenberg, Estimates near the boundary for solutions of elliptic partial differential equations satisfying general boundary conditions. Pt. I, *Comm. Pure Appl. Math.*, **12**(1959), 623–727.

[2] C.-G.Ambrozie, F.-H.Vasilescu, Banach Space Complexes, Mathematics and its Applications, Vol. 334, Kluwer Academic Publishers, Dordrecht, NL, 1995.

[3] L.Boutet de Monvel, Boundary problems for pseudodifferential operators, *Acta Math.*, **126**(1971), 11–51.

[4] L.Hörmander, L^2 estimates and existence theorems for the $\overline{\partial}$ operator, *Acta Math.*, **113**(1965), no. 1–2, 89–152.

[5] I.Malass, N.Tarkhanov, *J. Siber. Fed. Univ., Math. and Phys.*, **12**(2019), no. 4, 455–465. DOI: 10.17516/1997-1397-2019-12-4-455-465

[6] S.P.Novikov, On Metric-Independent Exotic Homology, *Proc. Steklov Inst. Math.*, **251**(2005), 206–212.

[7] M.Putinar, Some invariants for semi- Fredholm systems of essentially commuting operators, *J. Operator Theory*, **8**(1982), 65–90.

[8] D.B.Ray, I.M.Singer, Reidemeister torsion and the Laplacian on Riemannian manifolds, *Adv. in Math.*, **7**(1971), 145–210.

[9] D.C.Spencer, Harmonic integrals and Neumann problems associated with linear partial differential equations, In: Outlines of Joint Soviet-American Symposium on Partial Differential Equations, Novosibirsk, 1963, 253–260.

[10] N.Tarkhanov, Complexes of Differential Operators, Kluwer Academic Publishers, Dordrecht, NL, 1995.

[11] N.Tarkhanov, Euler characteristic of Fredholm quasicomplexes, *Funct. Anal. and its Appl.*, **41**(2007),318–322.

[12] K.Krupchyk, N.Tarkhanov, J.Tuomela, Elliptic quasicomplexes in Boutet de Monvel algebra, *J. of Funct. Anal.*, **247**(2007), 202–230.

[13] D.Wallenta, A Lefschetz fixed point formula for elliptic quasicomplexes, *Integr. Equ. Oper. Theory*, **78**(2014), no. 4, 577–587.

[14] R.O.Wells, Differential Analysis on Complex Manifolds, Springer-Verlag, New York, 1980.

[15] E.Witten, Supersymmetry and Morse theory, *J. Diff. Geom.*, **17**(1982), 661–692.

Возмущение комплекса де Рама

Исан Малас
Николай Тарханов
Потсдамский университет
Потсдам, Германия

Аннотация. Рассмотрим возмущение комплекса де Рама на компактном многообразии с краем. Это возмущение выходит за рамки комплексов, и поэтому когомологии к нему не относятся. С другой стороны, его кривизна "мала", поэтому существует естественный способ ввести характеристику Эйлера и разработать теорию Лефшеца для возмущения. Данная работа предназначена для попытки разработать теорию когомологий для произвольных последовательностей линейных отображений.

Ключевые слова: комплекс де Рама, когомологии, теория Ходжа, проблема Неймана.

DOI: 10.17516/1997-1397-2020-13-5-533-546
УДК 517.956.223

Nonlocal Problem for a Three-dimensional Elliptic Equation with Singular Coefficients in a Rectangular Parallelepiped

Kamoliddin T. Karimov*
Ferghana State University
Ferghana, Uzbekistan

Received 20.05.2020, received in revised form 14.06.2020, accepted 08.07.2020

Abstract. The nonlocal problem for an elliptic equation with two singular coefficients in a rectangular parallelepiped is studied. The uniqueness of the solution of the problem is proved with the use of the method of energy integrals. The spectral Fourier method based on the separation of variables is used to prove the existence of solutions. The solution of the problem is constructed as double Fourier series in terms of a sum of trigonometric and Bessel functions. Under some conditions on parameters and given functions the uniform convergence of the constructed series and its derivatives up to the second order inclusive is proved.

Keywords: elliptic type equation, nonlocal problem, singular coefficient, spectral method, parallelepiped.

Citation: K.T. Karimov, Nonlocal Problem for a Three-dimensional Elliptic Equation with Singular Coefficients in a Rectangular Parallelepiped, J. Sib. Fed. Univ. Math. Phys., 2020, 13(5), 533–546. DOI: 10.17516/1997-1397-2020-13-5-533-546.

Introduction. Formulation of the problem

Let us consider a three-dimensional elliptic equation with two singular coefficients

$$Lu \equiv u_{xx} + u_{yy} + u_{zz} + \frac{2\beta}{y} u_y + \frac{2\gamma}{z} u_z = 0, \tag{1}$$

in a rectangular parallelepiped $\Omega = \{(x, y, z) : \ 0 < x < a, \ 0 < y < b, \ 0 < z < c\}$, where β and γ are real numbers with β, $\gamma < 1/2$; $u = u(x, y, z)$ is an unknown function.

Consider the following problem and study its unique solvability.

Dezin's problem. Find a function $u(x, y, z) \in C(\bar{\Omega}) \cap C^1(\bar{\Omega} \cap (\{x = 0\} \cup \{x = a\})) \cap C^2(\Omega)$ that satisfies equation (1) in the domain Ω and the following conditions

$$u(0, y, z) = u(a, y, z), \ \ u_x(0, y, z) = u_x(a, y, z), \quad 0 \leqslant y \leqslant b, \ 0 \leqslant z \leqslant c, \tag{2}$$

$$u(x, y, 0) = 0, \quad u(x, y, c) = 0, \quad 0 \leqslant x \leqslant a, \ 0 \leqslant y \leqslant b, \tag{3}$$

$$u(x, 0, z) = f_1(x, z), \quad 0 \leqslant x \leqslant a, \ 0 \leqslant z \leqslant c, \tag{4}$$

$$u(x, b, z) = f_2(x, z), \quad 0 \leqslant x \leqslant a, \ 0 \leqslant z \leqslant c, \tag{5}$$

where $f_1(x, z)$ and $f_2(x, z)$ are given continuous functions.

*karimovk80@mail.ru http://orcid.org/0000-0002-9098-4116

Desin studied equation

$$(d/dt)\,u - Au = f,\ \ 0 \leqslant t \leqslant a$$

with the boundary condition $b\,u|_{t=0} - u|_{t=a} = g$ [1]. Here, it is assumed that function $u\,(t)$ takes values in a complex Banach space B when $t \in [0, a]$, $A : B \to B$ is an unbounded linear operator commuting with d/dt and a density domain of definition, and b is a complex number. It is also explained that the given conditions are "nonlocal" in the sense that they determine the relationship between the values of the unknown function at different points of the boundary.

Nonlocal boundary value problems are very interesting problems. They generalize the classical problems and, at the same time, they are naturally obtained when constructing mathematical models of real processes and phenomena in physics, engineering, etc (for an extended discussion see [2–6]). Problems with nonlocal conditions for partial differential equations have been studied by many authors. Below an overview of problems close to the Dezin problem is given. They are formulated and studied in two-dimensional domains.

Frankl [7] considered the flow around a finite symmetric profile by a subsonic velocity stream and formulated the problem for the Chaplygin equation in a mixed domain with a nonlocal condition of the form $u(0, y) = u(0, -y)$. In addition, the local condition $u_x(0, y) = 0$ was fulfilled. Ionkin [8] proved the existence of a solution to a nonlocal problem with conditions $u_x(0, y) = u_x(1, y)$, $u(0, y) = 0$, $0 \leqslant y \leqslant T$ and $u(x, 0) = \tau(x)$, $0 \leqslant x \leqslant 1$ for the heat equation using the spectral analysis method. The uniqueness of the solution of this problem was proved [9]. Such conditions are encountered, for example, in problems of particle diffusion in turbulent plasma and in problems of heat propagation in a thin heated rod if the law of change of the total amount of heat of the rod is given. Ionkin and Moiseev [10] proved the unique solvability of the problem for the heat equation with conditions

$$a_1 u_x\,(0, t) + b_1 u_x\,(1, t) + a_0 u\,(0, t) + b_0 u\,(1, t) = 0,$$

$$c_1 u_x\,(0, t) + d_1 u_x\,(1, t) + c_0 u\,(0, t) + d_0 u\,(1, t) = 0,$$

where $a_j, b_j, c_j, d_j,\ j = \overline{0, 1}$ are given constants.

Lerner and Repin [11] studied the following problem in half-strip $D = \{(x, y) : 0 < x < 1,\, y > 0\}$. Find a function $u\,(x, y)$ with properties

$$u\,(x, y) \in C\left(\bar{D}\right) \cap C^1\,(D \cup \{x = 0\}) \cap C^2\,(D)\,;$$

$$y^m u_{xx} + u_{yy} = 0,\ \ (x, y) \in D,\ \ m > -1;$$

$$u\,(x, y) \to 0\ \text{ at }\ y \to +\infty\ \text{ uniformly in } x \in [0, 1]\,;$$

$$u\,(0, y) - u\,(1, y) = \varphi_1\,(y)\,,\quad u_x\,(0, y) = \varphi_2\,(y)\,,\ \ y \geqslant 0;\ \ u\,(x, 0) = \tau\,(x)\,,\ 0 \leqslant x \leqslant 1,$$

where $\tau(x)$, $\varphi_1\,(y)$ and $\varphi_2\,(y)$ are given sufficiently smooth functions, and $\tau\,(x)$ is orthogonal to the system of functions 1, $\cos\,(2n+1)\,\pi x$, $n = 0, 1, 2, \ldots$. The similar problem was studied in the half-strip D for equation [12]

$$u_{xx} + u_{yy} + \frac{2p}{y} u_y - b^2 u = 0,\ \ p, b \in R,$$

Assuming that $\varphi_1\,(y) \equiv 0$ and $\varphi_2\,(y) \equiv 0$. The uniqueness of the solution of this problem is proved on the basis of the extremum principle. Using the methods of separation of variables and integral transforms, the solvability of the problem in question was established. Moiseev [13] studied the following nonlocal boundary-value problem in the half-strip D for degenerate elliptic equation of the form

$$y^m u_{xx} + u_{yy} = 0, \quad m > -2;$$

$$u(x,0) = f(x), \quad 0 \leqslant x \leqslant 1; \quad u(0,y) = u(1,y), \quad u_x(0,y) = 0, \; y \geqslant 0;$$

$$f(x) \in C^{2+\alpha}[0,1], \; f(0) = f(1), \quad f'(0) = 0.$$

Using the spectral analysis method, the uniqueness and existence of the solution of this problem were proved in the class of functions $u(x,y) \in C(\bar{D}) \cap C^2(D)$. Functions tend to zero or are bounded at infinity. Moreover, the solution of the problem was constructed in the form of the sum of the biorthogonal series. These results are also applicable to equations $y^m u_{xx} + u_{yy} - b^2 y^m u = 0$, $m, b \in R$, with $b \geqslant 0$, $m > 0$ [14]. Equation $y^m u_{xx} - -u_{yy} - b^2 y^m u = 0$ was studied in rectangular domain $\{(x,y) : 0 < x < 1, 0 < y < T\}$ [15], where $m > 0$, $b \geqslant 0$, $T > 0$ are given real numbers. Initial conditions $u(x,0) = \tau(x)$, $u_y(x,0) = \nu(x)$, $0 \leqslant x \leqslant 1$ and nonlocal boundary conditions $u(0,y) = u(1,y)$, $u_x(0,y) = 0$ or $u_x(0,y) = u_x(1,y)$, $u(1,y) = 0$ at $0 \leqslant y \leqslant T$ were assumed. The uniqueness and existence theorems are proved with the use of the spectral analysis method. Equation

$$u_{xx} + \operatorname{sgny} u_{yy} + \frac{2p}{|y|} u_y + ku = 0, \quad p \geqslant 1/2, \quad k \in R$$

was considered in domain $D = \{(x,y) : 0 < x < 1, y < \alpha\}$, $\alpha > 0$ [16] and the following problem was studied

$$u \in C(\bar{D}) \cap C^2(D \backslash \{y = 0\}), \quad Lu = 0;$$

$$u(0,y) = u(1,y), \quad u_x(0,y) = 0, y < \alpha; \quad u(x,\alpha) = \varphi(x), \; 0 < x < 1,$$

where $\varphi(x)$ is the given continuous function that satisfies condition $\varphi(0) = \varphi(1)$.

Nonlocal problems for inhomogeneous Lavrentev-Bitsadze equation and for equation of mixed elliptic-hyperbolic type with power degeneration were studied in the rectangular domains [17–21].

However, nonlocal problems for equations with singular coefficients in three-dimensional domains remain poorly understood.

1. Construction of eigenfunctions

To find a solution to the Dezin problem we apply the Fourier method [22]. Let us find nontrivial solutions of problem (1)–(3). Using separation of variables $u(x,y,z) = = W(x,z)\, Q(y)$, we obtain from equation (1)

$$Q''(y) + \frac{2\beta}{y} Q'(y) - \lambda Q(y) = 0, \quad 0 < y < b, \tag{6}$$

$$W_{xx} + W_{zz} + W_x + \frac{2\gamma}{z} W_z + \lambda W = 0, \quad 0 < x < a, \quad 0 < z < c, \tag{7}$$

where $\lambda \in R$ is the separation constant.

Taking into account conditions (2) and (3), we obtain for equation (7) the following eigenvalue problem in the domain $\Pi = \{(x,z) : \; 0 < x < a, \; 0 < z < c\}$: find the values of parameter λ and the corresponding nontrivial solutions $W(x,z) \in C(\bar{\Pi}) \cap C^1(\bar{\Pi} \cap (\{x = 0\} \cup \{x = a\})) \cap C^2(\Omega)$ of equation (7) in $\bar{\Pi}$ that satisfy conditions $W(0,z) = W(a,z)$, $0 \leqslant z \leqslant c$; $W(x,0) = 0$, $W(x,c) = 0$, $0 \leqslant x \leqslant a$.

By separating variables $W(x,z)=X(x)Z(z)$, this problem reduces to the following eigenvalue problem for the ordinary differential equation:

$$L_{\lambda-\mu}^{\gamma}Z(z)=Z''(z)+(2\gamma/z)Z'(z)+(\lambda-\mu)Z(z)=0,\ Z(0)=0,\quad Z(c)=0; \tag{8}$$

$$L_{\mu}^{0}X(x)=0,\quad X(0)=X(a),\quad X'(0)=X'(a), \tag{9}$$

where $\mu\in R$ is the separation constant.

Let us find first a solution of problem (9). It is easy to verify that for $\mu<0$ problem (9) has only trivial solutions. At $\mu=0$ the solution of problem (9) is $X(x)=d_0$ ($d_0\neq 0$ is some constant). Consider now the case $\mu>0$. Since boundary conditions in problem (9) are periodic conditions, the problem is regular. In addition, it is easy to verify that boundary-value problem (9) is a self-adjoint problem [23].

Substituting the general solution of equation $L_{\mu}^{0}X(x)=0$

$$X(x)=d_1\sin\sqrt{\mu}\,x+d_2\cos\sqrt{\mu}\,x \tag{10}$$

into nonlocal conditions $X(0)=X(a)$ and $X'(0)=X'(a)$, we obtain the following system of equations

$$\begin{cases} d_1\sin\sqrt{\mu}\,a+d_2\left(\cos\sqrt{\mu}\,a-1\right)=0,\\ d_1\left(\cos\sqrt{\mu}\,a-1\right)-d_2\sin\sqrt{\mu}\,a=0. \end{cases} \tag{11}$$

Setting the main determinant of this system to zero, we find $\cos\sqrt{\mu}\,a-1=0$. Solving this equation, we find the eigenvalues of the problem: $\mu_n=(2\pi n/a)^2$, $n\in N$. Let us substitute $\mu=\mu_n$ into (10) and (11). It is easy to verify that the eigenvalue μ_0 is simple, and it corresponds to one normalized eigenfunction $1/\sqrt{a}$. Eigenvalues μ_n, $n\in N$ have multiplicity 2 and they correspond to two normalized eigenfunctions $\sqrt{2/a}\cos(2\pi nx/a)$, $\sqrt{2/a}\sin(2\pi nx/a)$. Therefore, the eigenvalues and the eigenfunctions of problem (9) that correspond to these eigenvalues can be represented in the form of $\mu_n=(2\pi n/a)^2$, $n\in N\cup\{0\}$ and

$$X_0(x)=\frac{1}{\sqrt{a}},\ \ X_{2n-1}(x)=\sqrt{\frac{2}{a}}\sin\left(\frac{2\pi nx}{a}\right),\ \ X_{2n}(x)=\sqrt{\frac{2}{a}}\cos\left(\frac{2\pi nx}{a}\right),\ \ n\in N. \tag{12}$$

The system of eigenfunctions (12) is orthogonal and complete in the space $L_2[0,a]$, and it forms orthonormal basis in it (see, for example, [20, 27]).

Now we turn to the study of problem (8). We consider equation $L_{\lambda-\mu}^{\gamma}Z(z)=0$ and find its general solution at $\mu=\mu_n$. Introducing $Z(z)=\left(t/\sqrt{\lambda-\mu_n}\right)^{1/2-\gamma}p(t)$, where $t=\sqrt{\lambda-\mu_n}\,z$, $\lambda>\mu_n$ (at $\lambda\leqslant\mu_n$, problem (8) has only trivial solutions), from equation $L_{\lambda-\mu_n}^{\gamma}Z(z)=0$ we obtain the Bessel equation [24]:

$$t^2p''(t)+tp'(t)+\left[t^2-(1/2-\gamma)^2\right]p(t)=0.$$

Taking into account the form of the general solution of the last equation [24] and the introduced notation, we obtain the general solution of equation $L_{\lambda-\mu_n}^{\gamma}Z(z)=0$ in the form

$$Z_n(z)=d_{1n}z^{1/2-\gamma}J_{1/2-\gamma}\left(\sqrt{\lambda-\mu_n}\,z\right)+d_{2n}z^{1/2-\gamma}Y_{1/2-\gamma}\left(\sqrt{\lambda-\mu_n}\,z\right), \tag{13}$$

where d_1, d_2 are arbitrary constants, $J_l(z)$ and $Y_l(z)$ are the Bessel functions of the first and second kind of the order l, respectively [24].

It follows from (13) that a solution of equation $L^{\gamma}_{\lambda-\mu_n} Z(z) = 0$ that satisfies condition $Z(0) = 0$ is $Z_n(z) = d_1 z^{1/2-\gamma} J_{1/2-\gamma}\left(\sqrt{\lambda-\mu_n}\, z\right)$. Substituting this into condition $Z(c) = 0$, we obtain the condition for the existence of a nontrivial solution of problem (8):

$$J_{1/2-\gamma}\left(\sqrt{\lambda-\mu_n}\, c\right) = 0, \ \ n \in N. \tag{14}$$

It is known that when $l > -1$ the Bessel function $J_l(z)$ has a countable number of zeros. They are real and has pairwise opposite signs [24]. Since $1/2 - \gamma > 0$ then equation (14) has a countable number of real roots. Denoting the mth positive root of equation (14) at $n = k$ by δ_{km}, we have values of parameter λ for which there are nontrivial solutions of problem (8): $\lambda_{nm} = \mu_n + (\delta_{nm}/c)^2$, $n = 0, 1, 2, \ldots$, $m \in N$.

Assuming in (13) that $\lambda = \lambda_{nm}$, $d_{1n} = 1$, $d_{2n} = 0$, $n = 0, 1, 2, \ldots$, $m \in N$, we obtain nontrivial solutions (eigenfunctions) of problem (8) up to a constant factor

$$Z_{nm}(z) = z^{1/2-\gamma} J_{1/2-\gamma}(\delta_{nm} z/c), \ \ n = 0, 1, 2, \ldots, \ \ m \in N. \tag{15}$$

According to [24], for each n the system of eigenfunctions (15) is complete in space $L_2[0, c]$ with the weight $z^{2\gamma}$.

Assuming in equation (6) $\lambda = \lambda_{nm}$, we find its general solution

$$Q_{nm}(y) = a_{nm} y^{1/2-\beta} I_{1/2-\beta}\left(\sqrt{\lambda_{nm}}\, y\right) + b_{nm} y^{1/2-\beta} K_{1/2-\beta}\left(\sqrt{\lambda_{nm}}\, y\right), \ \ 0 \leqslant y \leqslant b, \tag{16}$$

here a_{nm} and b_{nm} are arbitrary constants, $I_l(y)$ and $K_l(y)$ are the modified Bessel functions of the first and third kind of the order l, respectively [24].

2. The uniqueness of the solution

The proof of the uniqueness of the solution of the Dezin problem is based on the lemma given below.

Lemma 1. *If $\beta, \gamma < 1/2$, function $u(x, y, z)$ is the solution of equation (1), and it satisfies conditions $u(x, 0, z) = 0$ and $u(x, y, 0) = 0$ then inequalities $\left|\lim\limits_{y\to 0} y^{2\beta} u_y(x, y, z)\right| < +\infty$ and $\left|\lim\limits_{z\to 0} z^{2\gamma} u_z(x, y, z)\right| < +\infty$ are satisfied.*

Proof. Separating variables by the formula $u(x, y, z) = X(x)\, Q(y)\, Z(z)$, from equation (1) with the variables y and z we obtain the ordinary differential equations (6) and $L^{\gamma}_{\lambda-\mu} Z(z) = 0$. Using the general solutions of these equations, it is easy to verify that the solutions of equations (6) and $L^{\gamma}_{\lambda-\mu} Z(z) = 0$ that satisfy conditions $Q(0) = 0$ and $Z(0) = 0$ respectively at $\beta, \gamma < 1/2$ have the form (up to a constant factor) $Q(y) = y^{1/2-\beta} I_{1/2-\beta}\left(\sqrt{\lambda}\, y\right)$ and $Z(z) = z^{1/2-\gamma} J_{1/2-\gamma}\left(\sqrt{\lambda-\mu}\, z\right)$. Taking the first-order derivative of these functions using relations [24]

$$\frac{d}{dx}\left[x^{\pm\nu} J_\nu(x)\right] = \pm x^{\pm\nu} J_{\nu\mp 1}(x), \quad \frac{d}{dx}\left[x^{\pm\nu} I_\nu(x)\right] = x^{\pm\nu} J_{\nu\mp 1}(x), \tag{17}$$

we have $Q'(y) = \sqrt{\lambda}\, y^{1/2-\beta} I_{-1/2-\beta}\left(\sqrt{\lambda}\, y\right)$, $Z'(z) = \sqrt{\lambda-\mu}\, z^{1/2-\gamma} J_{-1/2-\gamma}\left(\sqrt{\lambda-\mu}\, z\right)$. It follows that $\left|\lim\limits_{y\to 0} y^{2\beta} Q'(y)\right| < +\infty$, $\left|\lim\limits_{z\to 0} z^{2\gamma} Z'(z)\right| < +\infty$. Taking into account that $u(x, y, z) =$

$= X(x)\,Q(y)\,Z(z)$, we have that $\left|\lim\limits_{y\to 0} y^{2\beta}u_y(x,y,z)\right| < +\infty$ and $\left|\lim\limits_{z\to 0} z^{2\gamma}u_z(x,y,z)\right| < +\infty$. Lemma 1 is proved. □

Now we turn to the proof of the uniqueness of the solution of the Dezin problem.

Theorem 1. *The Dezin problem does not have more than one solution.*

Proof. Let $V_1(x,y,z)$ and $V_2(x,y,z)$ be solutions of the Dezin problem. Then function $u(x,y,z) = V_1(x,y,z) - V_2(x,y,z)$ satisfies equation (1), conditions (2), (3) and homogeneous boundary conditions corresponding to (4) and (5). We prove that $u(x,y,z) \equiv 0$ in $\bar{\Omega}$.

In domain Ω we have the identity

$$y^{2\beta}z^{2\gamma}uLu = \left(y^{2\beta}z^{2\gamma}uu_x\right)_x + \left(y^{2\beta}z^{2\gamma}uu_y\right)_y + \left(y^{2\beta}z^{2\gamma}uu_z\right)_z - y^{2\beta}z^{2\gamma}\left(u_x^2+u_y^2+u_z^2\right) = 0.$$

Integrating this identity over domain

$$\Omega_{\varepsilon_1\varepsilon_2\varepsilon_3\varepsilon_4}^{\varepsilon_5\varepsilon_6} = \left\{(x,y,z):\ \varepsilon_1 < x < a-\ \varepsilon_2,\ \varepsilon_3 < y < b-\varepsilon_4,\ \varepsilon_5 < z < c-\varepsilon_6\right\},$$

where $\varepsilon_j,\ j=\overline{1,6}$ are sufficiently small positive numbers, we have

$$\iiint\limits_{\Omega_{\varepsilon_1\varepsilon_2\varepsilon_3\varepsilon_4}^{\varepsilon_5\varepsilon_6}} \left[\left(y^{2\beta}z^{2\gamma}uu_x\right)_x + \left(y^{2\beta}z^{2\gamma}uu_y\right)_y + \left(y^{2\beta}z^{2\gamma}uu_z\right)_z\right]dxdydz =$$
$$= \iiint\limits_{\Omega_{\varepsilon_1\varepsilon_2\varepsilon_3\varepsilon_4}^{\varepsilon_5\varepsilon_6}} \left[y^{2\beta}z^{2\gamma}\left(u_x^2+u_y^2+u_z^2\right)\right]dxdydz. \tag{18}$$

Obviously, if $\varepsilon_j,\ j=\overline{1,6}$ tends to zero then $\Omega_{\varepsilon_1\varepsilon_2\varepsilon_3\varepsilon_4}^{\varepsilon_5\varepsilon_6} \to \Omega$.

Applying the Gauss-Ostrogradsky formula [22] to the left hand side of equality (18), we obtain after some transformations that

$$\int\limits_{\varepsilon_5}^{c-\varepsilon_6}\int\limits_{\varepsilon_3}^{b-\varepsilon_4} y^{2\beta}z^{2\gamma}\left[u_x(a-\varepsilon_2,y,z) - u_x(\varepsilon_1,y,z)\right]u(a-\varepsilon_2,y,z)\,dydz +$$
$$+ \int\limits_{\varepsilon_5}^{c-\varepsilon_6}\int\limits_{\varepsilon_3}^{b-\varepsilon_4} y^{2\beta}z^{2\gamma}\left[u(a-\varepsilon_2,y,z) - u(\varepsilon_1,y,z)\right]u_x(\varepsilon_1,y,z)\,dydz +$$
$$+ \int\limits_{\varepsilon_5}^{c-\varepsilon_6}\int\limits_{\varepsilon_1}^{a-\varepsilon_2} z^{2\gamma}\left[(b-\varepsilon_4)^{2\beta}u(x,b-\varepsilon_4,z)\,u_y(x,b-\varepsilon_4,z) - \varepsilon_3^{2\beta}u(x,\varepsilon_3,z)\,u_y(x,\varepsilon_3,z)\right]dxdz +$$
$$+ \int\limits_{\varepsilon_3}^{b-\varepsilon_4}\int\limits_{\varepsilon_1}^{a-\varepsilon_2} y^{2\beta}\left[(c-\varepsilon_6)^{2\gamma}u(x,y,c-\varepsilon_6)\,u_z(x,y,c-\varepsilon_6) - \varepsilon_5^{2\gamma}u(x,y,\varepsilon_5)\,u_z(x,y,\varepsilon_5)\right]dxdy =$$
$$= \iiint\limits_{\Omega_{\varepsilon_1\varepsilon_2\varepsilon_3\varepsilon_4}^{\varepsilon_5\varepsilon_6}} \left[y^{2\beta}z^{2\gamma}\left(u_x^2+u_y^2+u_z^2\right)\right]dxdydz.$$

Hence, taking the limit at $\varepsilon_j \to 0,\ j=\overline{1,6}$ and taking into account the conditions of Lemma 1, $|u_y(x,b,z)| < +\infty$, $|u_z(x,y,c)| < +\infty$ and homogeneous boundary conditions, we obtain

$$\iiint\limits_{\Omega} \left[y^{2\beta}z^{2\gamma}\left(u_x^2+u_y^2+u_z^2\right)\right]dxdydz = 0.$$

Therefore $u_x(x,y,z) \equiv u_y(x,y,z) \equiv u_z(x,y,z) \equiv 0$, $(x,y,z) \in \Omega$. Then $u(x,y,z) \equiv$const, $(x,y,z) \in \Omega$. Since $u(x,y,z) \in C(\bar{\Omega})$ and $u(x,0,z) \equiv 0$ then $u(x,y,z) \equiv 0$, $(x,y,z) \in \bar{\Omega}$. The theorem is proved. □

3. Construction and justification of the solution of the Dezin problem

Let us assume that solution of the Dezin problem in domain Ω has the form

$$
\begin{aligned}
u(x,y,z) = &\sum_{n=0}^{\infty}\sum_{m=1}^{\infty} X_n(x) Q_{nm}(y) Z_{nm}(z) = \\
&= \frac{1}{\sqrt{a}} \sum_{m=1}^{\infty} z^{1/2-\gamma} J_{1/2-\gamma}\left(\frac{\delta_{0m} z}{c}\right) y^{1/2-\beta}\left[a_{0m} I_{1/2-\beta}(\delta_{0m} y) + b_{0m} K_{1/2-\beta}(\delta_{0m} y)\right] + \\
&+ \sqrt{\frac{2}{a}} \sum_{n=1}^{\infty}\sum_{m=1}^{\infty}\left(\sin\frac{2\pi n x}{a} + \cos\frac{2\pi n x}{a}\right) z^{1/2-\gamma} J_{1/2-\gamma}\left(\frac{\delta_{nm} z}{c}\right) \times \\
&\times y^{1/2-\beta}\left[a_{nm} I_{1/2-\beta}\left(\sqrt{\lambda_{nm}}\, y\right) + b_{nm} K_{1/2-\beta}\left(\sqrt{\lambda_{nm}}\, y\right)\right]. \quad (19)
\end{aligned}
$$

Each term of series (19) satisfies equation (1) and conditions (2) and (3). Assuming that this series converges absolutely and uniformly, we find constants a_{nm} and b_{nm} from the requirement that function (19) must satisfy boundary conditions (4) and (5). First, substituting it in conditions (4), we obtain

$$
\frac{F_0(z)}{\sqrt{a}} + \sqrt{\frac{2}{a}} \sum_{n=1}^{\infty} F_n(z)\left(\sin\frac{2\pi n x}{a} + \cos\frac{2\pi n x}{a}\right) = z^{\gamma-1/2} f_1(x,z), \quad (20)
$$

where

$$
F_n(z) = 2^{-1/2-\beta}\Gamma(1/2-\beta) \sum_{m=1}^{\infty} J_{1/2-\gamma}\left(\frac{\delta_{nm} z}{c}\right)\left(\sqrt{\lambda_{nm}}\right)^{\beta-1/2} b_{nm}. \quad (21)
$$

Series (20) and (21) are called the Fourier series of functions $z^{\gamma-1/2} f_1(x,z)$ and $F_n(z)$ expanded in the system of the trigonometric and Bessel functions, respectively. The Fourier coefficients are determined from (20) as follows

$$
F_0(z) = \frac{1}{\sqrt{a}} \int_0^a z^{\gamma-1/2} f_1(x,z)\, dx, \quad (22)
$$

$$
F_n(z) = \frac{1}{\sqrt{2a}} \int_0^a \left(\sin\frac{2\pi n x}{a} + \cos\frac{2\pi n x}{a}\right) z^{\gamma-1/2} f_1(x,z)\, dx. \quad (23)
$$

Since $F_0(z)$ and $F_n(z)$ are known we substitute them into (21), and unequivocally find coefficients b_{0m} and b_{nm}:

$$
b_{0m} = \frac{2^{3/2+\beta}\delta_{0m}^{1/2-\beta}}{\sqrt{a}\left[c J_{3/2-\gamma}(\delta_{0m})\right]^2 \Gamma(1/2-\beta)} f_{0m},
$$

$$
b_{nm} = \frac{2^{1+\beta}\left(\sqrt{\lambda_{nm}}\right)^{1/2-\beta}}{\sqrt{a}\left[c J_{3/2-\gamma}(\delta_{nm})\right]^2 \Gamma(1/2-\beta)} f_{nm},
$$

where

$$f_{nm} = \int\limits_0^c \int\limits_0^a \left(\sin\frac{2\pi nx}{a} + \cos\frac{2\pi nx}{a}\right) z^{1/2+\gamma} J_{1/2-\gamma}\left(\frac{\delta_{nm} z}{c}\right) f_1(x,z)dxdz,\ n+1,\quad m \in N.$$

Now, substituting function (19) into condition (5), we have

$$\frac{G_0(z)}{\sqrt{a}} + \sqrt{\frac{2}{a}} \sum_{n=1}^{\infty} G_n(z) \left(\sin\frac{2\pi nx}{a} + \cos\frac{2\pi nx}{a}\right) = z^{\gamma-1/2} b^{\beta-1/2} f_2(x,z), \tag{24}$$

where

$$G_n(z) = \sum_{m=1}^{\infty} J_{1/2-\gamma}\left(\frac{\delta_{nm} z}{c}\right) \left[a_{nm} I_{1/2-\beta}\left(\sqrt{\lambda_{nm}} b\right) + b_{nm} K_{1/2-\beta}\left(\sqrt{\lambda_{nm}} b\right)\right]. \tag{25}$$

From relation (24) we find

$$G_0(z) = \frac{1}{\sqrt{a}} \int_0^a z^{\gamma-1/2} b^{\beta-1/2} f_2(x,z)dx,$$

$$G_n(z) = \frac{1}{\sqrt{2a}} \int_0^a \left(\sin\frac{2\pi nx}{a} + \cos\frac{2\pi nx}{a}\right) z^{\gamma-1/2} b^{\beta-1/2} f_2(x,z)\, dx.$$

Substituting $G_0(z)$ and $G_n(z)$ into (25), we obtain

$$a_{0m} I_{1/2-\beta}(\delta_{0m} b) + b_{0m} K_{1/2-\beta}(\delta_{0m} b) = \frac{2b^{\beta-1/2}}{\sqrt{a}\left[cJ_{3/2-\gamma}(\delta_{0m})\right]^2} g_{0m},$$

$$a_{nm} I_{1/2-\beta}\left(\sqrt{\lambda_{nm}}\, b\right) + b_{nm} K_{1/2-\beta}\left(\sqrt{\lambda_{nm}}\, b\right) = \frac{\sqrt{2} b^{\beta-1/2}}{\sqrt{a}\left[cJ_{3/2-\gamma}(\delta_{nm})\right]^2} g_{nm},$$

where

$$g_{nm} = \int\limits_0^c \int\limits_0^a \left(\sin\frac{2\pi nx}{a} + \cos\frac{2\pi nx}{a}\right) z^{1/2+\gamma} J_{1/2-\gamma}\left(\frac{\delta_{nm} z}{c}\right) f_2(x,z)\, dxdz,\quad n+1,\ m \in N.$$

Since b_{0m} and b_{nm} are known then from the last system of equations we unequivocally find coefficients a_{0m} and a_{nm}:

$$a_{0m} = \frac{2b^{\beta-1/2} g_{0m}}{\sqrt{a}\left[cJ_{3/2-\gamma}(\delta_{0m})\right]^2 I_{1/2-\beta}(\delta_{0m} b)} - \frac{2^{3/2+\beta} \delta_{0m}^{1/2-\beta} K_{1/2-\beta}(\delta_{0m} b) f_{0m}}{\sqrt{a}\left[cJ_{3/2-\gamma}(\delta_{0m})\right]^2 \Gamma(1/2-\beta) I_{1/2-\beta}(\delta_{0m} b)},$$

$$a_{nm} = \frac{\sqrt{2} b^{\beta-1/2} g_{nm}}{\sqrt{a}\left[cJ_{3/2-\gamma}(\delta_{nm})\right]^2 I_{1/2-\beta}\left(\sqrt{\lambda_{nm}}\, b\right)} - \frac{2^{1+\beta}\left(\sqrt{\lambda_{nm}}\right)^{1/2-\beta} K_{1/2-\beta}\left(\sqrt{\lambda_{nm}}\, b\right) f_{nm}}{\sqrt{a}\left[cJ_{3/2-\gamma}(\delta_{nm})\right]^2 \Gamma\left(\frac{1}{2}-\beta\right) I_{1/2-\beta}\left(\sqrt{\lambda_{nm}}\, b\right)}.$$

Substituting the values of coefficients a_{0m}, a_{nm}, b_{0m} and b_{nm} into (19), we find the formal solution of the Dezin problem in the form

$$u(x,y,z) = \frac{2}{a} \sum_{n=0}^{\infty} \sum_{m=1}^{\infty} \left(\sin\frac{2\pi nx}{a} + \cos\frac{2\pi nx}{a}\right) \frac{z^{1/2-\gamma} J_{1/2-\gamma}(\delta_{nm} z/c)}{\left[cJ_{3/2-\gamma}(\delta_{nm})\right]^2} \omega_{nm}(y), \tag{26}$$

where

$$\omega_{nm}(y) = \frac{y^{1/2-\beta} I_{1/2-\beta}\left(\sqrt{\lambda_{nm}}\, y\right)}{b^{1/2-\beta} I_{1/2-\beta}\left(\sqrt{\lambda_{nm}}\, b\right)} \left[g_{nm} - \bar{K}_{1/2-\beta}\left(\sqrt{\lambda_{nm}}\, b\right) f_{nm}\right] + \\ + \bar{K}_{1/2-\beta}\left(\sqrt{\lambda_{nm}}\, y\right) f_{nm}, \tag{27}$$

where function $\bar{K}_\nu(x)$ [25] has the form: $\bar{K}_\nu(x) = 2^{1-\nu} x^\nu K_\nu(x) / \Gamma(\nu)$, $\bar{K}_\nu(0) = 1$ $(\nu > 0)$.

Each member of this series satisfies all conditions of the Dezin problem. Let us note that when $\beta, \gamma < 1/2$ the denominator of the coefficients of series (26) has no zeros. If we prove that series (26) and serieses u_{xx}, $(u_{yy} + (2\beta/y) u_y)$, $(u_{zz} + (2\gamma/z) u_z)$, obtained from it by differentiation converge absolutely and uniformly in corresponding domains then its sum is the solution of the Dezin problem. Moreover, we need the following lemmas.

Lemma 2. *For any positive integers n, m and $\forall y \in [0, b]$ the estimates*

$$|\omega_{nm}(y)| \leqslant 2\left(|g_{nm}| + |f_{nm}|\right), \tag{28}$$

$$\left|y^{-2\beta}\left[y^{2\beta}\omega'_{nm}(y)\right]'\right| \leqslant 2\lambda_{nm}\left(|g_{nm}| + |f_{nm}|\right) \tag{29}$$

are valid.

Proof. Obviously, for any values of λ_{nm} and $y \in [0, b]$ the inequality

$$\left|\bar{K}_{1/2-\beta}\left(\sqrt{\lambda_{nm}}\, y\right)\right| \leqslant 1 \tag{30}$$

is valid. Since $y \in [0, b]$, and $I_{1/2-\beta}\left(\sqrt{\lambda_{nm}}\, y\right)$ is increasing function, then

$$\frac{y^{1/2-\beta} I_{1/2-\beta}\left(\sqrt{\lambda_{nm}}\, y\right)}{b^{1/2-\beta} I_{1/2-\beta}\left(\sqrt{\lambda_{nm}}\, b\right)} \leqslant 1. \tag{31}$$

We obtain from (27) that

$$y^{-2\beta}\left[y^{2\beta}\omega'_{nm}(y)\right]' = \lambda_{nm}\omega_{nm}(y). \tag{32}$$

If inequalities (30) and (31) are taken into account then estimates (28) and (29) immediately follow from (27) and (32). Lemma 2 is proved. □

Lemma 3. *The following estimates hold for $n \in N$:*

$$\left|\sin\frac{2\pi n x}{a} + \cos\frac{2\pi n x}{a}\right| \leqslant \sqrt{2}, \tag{33}$$

$$\left|\frac{d}{dx}\left(\sin\frac{2\pi n x}{a} + \cos\frac{2\pi n x}{a}\right)\right| \leqslant \sqrt{2}\,\mu_n. \tag{34}$$

The validity of estimates (33)–(34) follows from the properties of trigonometric functions.

Lemma 4. *For all natural numbers n and $z \in [0, c]$ the following estimates hold for sufficiently large m,:*

$$|Z_{nm}(z)| \leqslant C_1, \tag{35}$$

$$\left|z^{-2\gamma}\left[z^{2\gamma}Z'_{nm}(z)\right]'\right| \leqslant C_1(\delta_{nm}/c)^2, \tag{36}$$

where C_1 is some positive constant.

Proof. It is obvious that $Z_{nm}(z) \in C[0,a]$. For sufficiently large ξ the asymptotic formula holds [26]

$$J_\nu(\xi) \approx \left(\frac{2}{\pi\xi}\right)^{1/2} \cos\left(\xi - \frac{\nu\pi}{2} - \frac{\pi}{4}\right). \tag{37}$$

Therefore, estimate (35) is true.

Using the first formula (17), we find from (15) that

$$z^{2\gamma} Z'_{nm}(z) = (\delta_{nm}/c)\, z^{1/2+\gamma} J_{-1/2-\gamma}(\delta_{nm} z/c). \tag{38}$$

Let us take the first-order derivative of function (38) using formula (17). Then, multiplying it by $z^{-2\gamma}$, we obtain

$$z^{-2\gamma}\left[z^{2\gamma} Z'_{nm}(z)\right]' = -(\delta_{nm}/c)^2 z^{1/2-\gamma} J_{1/2-\gamma}(\delta_{nm} z/c) = -(\delta_{nm}/c)^2 Z_{nm}(z).$$

Then validity of estimate (36) follows from (35). □

Lemma 5. *The following estimate holds for each fixed $n \in N$ and sufficiently large positive integer m:*

$$J^2_{3/2-\gamma}(\delta_{nm}) \geqslant \frac{C_2}{\delta_{nm}}, \tag{39}$$

where C_2 is some positive constant.

Proof. Since δ_{nm} is a zero of function $J_{1/2-\gamma}(x)$ then relation

$$\int_0^c z J^2_{1/2-\gamma}\left(\frac{\delta_{nm} z}{c}\right) dz = \frac{c^2}{2} J^2_{3/2-\gamma}(\delta_{nm}).$$

is true. It follows from this relation that

$$J^2_{3/2-\gamma}(\delta_{nm}) = \frac{2}{c^2}\int_0^c z J^2_{1/2-\gamma}\left(\frac{\delta_{nm} z}{c}\right) dz = \frac{2}{\delta^2_{nm}} \int_0^{\delta_{nm}} \xi J^2_{1/2-\gamma}(\xi)\, d\xi. \tag{40}$$

Taking into account asymptotic formula (37), there exists a sufficiently large number $c_0 > 0$ such that for $\xi > c_0$ equality

$$\xi J^2_{1/2-\gamma}(\xi) \approx \frac{2}{\pi} \sin^2\left(\xi + \frac{\gamma\pi}{2}\right).$$

is true. If we assume that δ_{nm} is a sufficiently large number and $\delta_{nm} > 2(c_0+1)$ then

$$\int_0^{\delta_{nm}} \xi J^2_{1/2-\gamma}(\xi)\, d\xi > \int_{c_0}^{\delta_{nm}} \xi J^2_{1/2-\gamma}(\xi)\, d\xi \geqslant \frac{2}{\pi}\int_{c_0}^{\delta_{nm}} \sin^2\left(\xi + \frac{\gamma\pi}{2}\right) d\xi =$$
$$= \frac{1}{\pi}\delta_{nm} - \frac{1}{\pi}\left[c_0 + \cos(\delta_{nm} + c_0 + \gamma\pi)\sin(\delta_{nm} - c_0)\right] \geqslant \frac{1}{2\pi}\delta_{nm}.$$

Taking this into account, we obtain that estimate (39) follows from (40). Lemma 5 is proved. □

Lemma 6. *Let the following conditions be satisfied*

$$\lim_{x\to 0} \frac{\partial^k}{\partial x^k} f_j(x,z) = \lim_{x\to a} \frac{\partial^k}{\partial x^k} f_j(x,z), \quad k = \overline{0,2}, \quad j = \overline{1,2}, \tag{41}$$

$$f_j(x,0) = 0, \quad f_j(x,c) = 0, \quad f_{jz}(x,z) \in C([0,a]\times[0,c]), \; j = \overline{1,2}, \tag{42}$$

$$\lim_{z\to 0}\left[f_{jzz}(x,z) + \frac{2\gamma}{z} f_{jz}(x,z)\right] = 0, \; \lim_{z\to c}\left[f_{jzz}(x,z) + \frac{2\gamma}{z} f_{jz}(x,z)\right] = 0, \; j = \overline{1,2}, \tag{43}$$

$$\frac{\partial}{\partial z}\left[f_{jzz}(x,z)+\frac{2\gamma}{z}f_{jz}(x,z)\right]\in C([0,a]\times[0,c]),\ \ j=\overline{1,2}, \tag{44}$$

$$\int\limits_0^c\int\limits_0^a\left|z^{-1/2-\gamma}\frac{\partial}{\partial z}\left\{z^{2\gamma}\frac{\partial}{\partial z}\left[z^{-2\gamma}\frac{\partial}{\partial z}\left(z^{2\gamma}f_{jxxxz}(x,z)\right)\right]\right\}\right|dxdz<+\infty,\ \ j=\overline{1,2}. \tag{45}$$

Then, for large n and m we have estimates

$$|f_{nm}|\leqslant\frac{C_3}{n^{3+\varepsilon_7}\delta_{nm}^{4+\varepsilon_8}},\quad |g_{nm}|\leqslant\frac{C_4}{n^{3+\varepsilon_7}\delta_{nm}^{4+\varepsilon_8}}, \tag{46}$$

where $\varepsilon_7,\varepsilon_8$, C_3, C_4 are positive constants.

Proof. Using formulas (17), coefficients f_{nm} are presented in the form

$$f_{jnm}=\frac{ac}{2\pi n\delta_{nm}}\int\limits_0^c\int\limits_0^a\frac{d}{dx}\left(\cos\frac{2\pi nx}{a}-\sin\frac{2\pi nx}{a}\right)\frac{d}{dz}\left[z^{1/2+\gamma}J_{-1/2-\gamma}\left(\frac{\delta_{nm}z}{c}\right)\right]f_1(x,z)\,dxdz.$$

Taking into account conditions (41)–(44) at $j=1$ and applying the rule of integration by parts three times for variable x and four times for variable z, we obtain

$$f_{nm}=\frac{a^3c^4}{(2\pi n)^3\delta_{nm}^4}\int\limits_0^c\int\limits_0^a\frac{\partial}{\partial z}\left\{z^{2\gamma}\frac{\partial}{\partial z}\left[z^{-2\gamma}\frac{\partial}{\partial z}\left(z^{2\gamma}f_{1xxxz}(x,z)\right)\right]\right\}\times$$
$$\times\left(\sin\frac{2\pi nx}{a}+\cos\frac{2\pi nx}{a}\right)z^{1/2-\gamma}J_{1/2-\gamma}\left(\frac{\delta_{nm}z}{c}\right)dxdz. \tag{47}$$

It is known [27] that if $f(x)$ is an absolutely integrable function on $[a,b]$ then equalities

$$\lim_{n\to\infty}\int_a^b f(x)\cos nx\,dx=\lim_{n\to\infty}\int_a^b f(x)\sin nx\,dx=0. \tag{48}$$

are true. There is an analog of properties (48) [27]:

$$\lim_{n\to\infty}\int_0^1 xf(x)J_p(\lambda_n x)\,dx=0, \tag{49}$$

here $f(x)$ is an absolutely integrable function on $[0,1]$, and λ_n, $n\in N$ are positive zeros of function $J_p(x)$ $(p>-1)$ numbered in ascending order.

Since conditions (45) are satisfied then by virtue of (48)–(49) the following equality takes place

$$\lim_{n,m\to\infty}\int\limits_0^c\int\limits_0^a\frac{\partial}{\partial z}\left\{z^{2\gamma}\frac{\partial}{\partial z}\left[z^{-2\gamma}\frac{\partial}{\partial z}\left(z^{2\gamma}f_{1xxxz}(x,z)\right)\right]\right\}\times$$
$$\times\left(\sin\frac{2\pi nx}{a}-\cos\frac{2\pi nx}{a}\right)z^{1/2-\gamma}J_{1/2-\gamma}\left(\frac{\delta_{nm}z}{c}\right)dxdz=0.$$

Using (47) and the latter relation and assuming that n and m are sufficiently large, estimate (46) is obtained. The second estimate in (46) is proved similarly. Lemma 6 is proved. □

Now we turn to the study of convergence of series. Differentiating (26), we obtain

$$u_{xx}=\frac{2}{a}\sum_{n=0}^{\infty}\sum_{m=1}^{\infty}\left(\sin\frac{2\pi nx}{a}+\cos\frac{2\pi nx}{a}\right)''\frac{Z_{nm}(z)\,\omega_{nm}(y)}{\left[cJ_{3/2-\gamma}(\delta_{nm})\right]^2}, \tag{50}$$

$$u_{yy}+\frac{2\beta}{y}u_y=\frac{2}{a}\sum_{n=0}^{\infty}\sum_{m=1}^{\infty}\left(\sin\frac{2\pi nx}{a}+\cos\frac{2\pi nx}{a}\right)\frac{y^{-2\beta}\left[y^{2\beta}\omega'_{nm}\left(y\right)\right]'}{\left[cJ_{3/2-\gamma}\left(\delta_{nm}\right)\right]^2}Z_{nm}\left(z\right),\tag{51}$$

$$u_{zz}+\frac{2\gamma}{z}u_z=\frac{2}{a}\sum_{n=0}^{\infty}\sum_{m=1}^{\infty}\left(\sin\frac{2\pi nx}{a}+\cos\frac{2\pi nx}{a}\right)\frac{z^{-2\gamma}\left[z^{2\gamma}Z'_{nm}\left(z\right)\right]'}{\left[cJ_{3/2-\gamma}\left(\delta_{nm}\right)\right]^2}\omega_{nm}\left(y\right).\tag{52}$$

Taking into account estimates (28), (33), (35) and (39), for sufficiently large natural numbers n and m, series (26) for any $(x,y,z)\in[0,a]\times[0,\infty)\times[0,c]$ is majorized by the number series

$$\sum_{n=0}^{\infty}\sum_{m=1}^{\infty}C_5\delta_{nm}\left(|g_{nm}|+|f_{nm}|\right).\tag{53}$$

Series (50)–(52) on each compact $K\subset\Omega$ are majorized respectively by the following number series:

$$\sum_{n=1}^{\infty}\sum_{m=1}^{\infty}C_6n^2\delta_{nm}\left(|g_{nm}|+|f_{nm}|\right),\tag{54}$$

$$\sum_{n=0}^{\infty}\sum_{m=1}^{\infty}C_7\delta_{nm}\left(n^2+\delta_{nm}^2\right)\left(|g_{nm}|+|f_{nm}|\right),\tag{55}$$

$$\sum_{n=0}^{\infty}\sum_{m=1}^{\infty}C_8\delta_{nm}^3\left(|g_{nm}|+|f_{nm}|\right),\tag{56}$$

where $C_j,\ j=\overline{5,8}$ are some positive constants.

According to [27, p. 276, formula (4.3)], for $\forall n\in N$ and sufficiently large positive integers m inequality $(1/\delta_{nm})\leqslant(2/m)$ is true.

Consequently, according to Lemma 6, for sufficiently large n and m series (53)–(56) are estimated respectively by numerical series

$$\sum_{n=1}^{\infty}\sum_{m=1}^{\infty}\frac{C_9}{n^{3+\varepsilon_7}m^{3+\varepsilon_8}},\ \sum_{n=1}^{\infty}\sum_{m=1}^{\infty}\frac{C_{10}}{n^{1+\varepsilon_7}m^{3+\varepsilon_8}},\ \sum_{n=1}^{\infty}\sum_{m=1}^{\infty}\frac{(n^2+m^2)C_{11}}{n^{3+\varepsilon_7}m^{3+\varepsilon_8}},\ \sum_{n=1}^{\infty}\sum_{m=1}^{\infty}\frac{C_{12}}{n^{3+\varepsilon_7}m^{1+\varepsilon_8}},\tag{57}$$

where $C_j,\ j=\overline{9,12}$ are some positive constants.

It is not difficult to establish that series (57) converge. Therefore, series (53)–(56) also converge. Then, according to the Weierstrass criterion, series (26) converges absolutely and uniformly in $\bar{\Omega}$, and series (50)–(52) converges on each compact $K\subset\Omega$. Therefore, function $u\left(x,y,z\right)$ defined by (26) satisfies all conditions of the Dezin problem.

Thus, the following theorem is proved.

Theorem 2. *Let function $f_j\left(x,z\right),\ j=\overline{1,2}$ satisfies conditions (41)–(45). Then a solution of the Dezin problem exists, it is unique, and it is determined by formula (26).*

References

[1] A.A.Dezin, Operators involving a first derivative with respect to time and nonlocal boundary conditions, *Math. USSR-Izv.*, **1**(1967), no. 1, 57-79.

[2] A.M.Nakhushev, Equations of Mathematical Biology, Moscow, Visshaya shkola, 1995 (in Russian).

[3] A.A.Guetter, A free boundary problem in plasma containment, *SIAM J. Appl. Math.*, **49**(1989), 99–115.

[4] C.V.Pao, Reaction diffusion equations with nonlocal boundary and nonlocal initial conditions, *Journal of Mathematical Analysis and Applications*, **195**(1995), no. 3, 702–718. DOI: 10.1006/jmaa.1995.1384

[5] E.Obolashvili, Nonlocal problems for some partial differential equations, *Applicable Analysis*, **45**(1992), 269–280. DOI: 10.1080/00036819208840101

[6] J.I.Diaz, J.M.Rakotoson, On a nonlocal stationary free boundary problem arising in the confinement of a plasma in a Stellarator geometry, *Archive for Rational Mechanics and Analysis*, **134**(1996), 53–95. https://doi.org/10.1007/BF00376255

[7] F.I.Frankl, Flowing profiles by a stream of subsonic velocity with a supersonic zone ending with a direct surge of seal, *PMM*, **20**(1956), no. 2, 196–202.

[8] N.I.Ionkin, Solving one regional problem of thermal conductivity theory with a non-classical edge condition, *Differ. equations*, **13**(1977), no. 2, 294–304. zbmath.org/0349.35040

[9] N.I.Ionkin, On the sustainability of one problem theory of thermal conductivity with a non-classical edge condition, *Differ. equations*, **15**(1979), no. 7, 1279–1283. zbmath.org/0418.35053

[10] N.I.Ionkin, E.I.Moiseyev, About the task for the equation of thermal conductivity with two-point edge conditions, *Differ. equations*, **15**(1979), no. 7, 1284–1295. zbmath.org/0415.35032

[11] M.E.Lerner, O.A.Repin, On Frankl'-type problems for some elliptic equations with degeneration of various types, *Differ. equations*, **35**(1999), no. 8, 1098–1104. mi.mathnet.ru/de9976

[12] M.E.Lerner, O.A.Repin, Nonlocal boundary value problems in a vertical half-strip for a generalized axisymmetric Helmholtz Equation, *Differ. equations*, **37**(2001), no. 11, 1562–1564. mi.mathnet.ru/de10495

[13] E.I.Moiseev, On the solution of a nonlocal boundary value problem by the spectral method, *Differ. equations*, **35**(1999), no. 8, 1105–1112. mi.mathnet.ru/de9977

[14] E.I.Moiseev, Solvability of a Nonlocal Boundary Value Problem, *Differ. equations*, **37**(2001), no. 11, 1643–1646. https://doi.org/10.1023/A:1017937403853

[15] Yu.K.Sabitova, Nonlocal initial-boundary-value problems for a degenerate hyperbolic equation, *Russian Math. (Iz. VUZ)*, **53**(2009), no. 12, 41–49. mi.mathnet.ru/ivm6024

[16] A.A.Abashkin, *Vestn. Samar. Gos. Tekhn. Univ., Ser. Fiz.-Mat. Nauki (J. Samara State Tech. Univ., Ser. Phys. Math. Sci.)*, **36**(2014), no. 3, 7–20 (in Russian). https://doi.org/10.14498/vsgtu1328

[17] A.M.Nakhushev, Offset problems for partial differential equations, Moscow, Nauka, 2006 (in Russian).

[18] Z.A.Nakhusheva, Nonlocal boundary value problems for basic and mixed types of differential equations, Nalchik, 2011 (in Russian).

[19] Z.A.Nakhusheva, *Differ. equations*, **45**(2009), no. 8, 1199–2003.
DOI: 10.1134/S0012266109080151

[20] K.B.Sabitov, V.A.Novikova, *Russian Math. (Iz. VUZ)*, **60**(2016), no. 6, 52–62.
DOI: 10.3103/S1066369X16060074

[21] K.B.Sabitov, *Differ. equations*, **55**(2019), no. 10, 1384–1389.
DOI: 10.1134/S0012266119100136

[22] A.N.Tikhonov, A.A.Samarskiy, Equations of mathematical physics, Moscow, Nauka, 1972 (in Russian).

[23] M.A.Naimark, Linear differential operators, Moscow, Nauka, 1969 (in Russian).

[24] G.N.Watson, Theory of Bessel functions, Moscow, Part. 1, Izd. IL., 1949 (in Russian).

[25] M.B.Kapilevich, On one equation of mixed elliptic-hyperbolic type, *Matematicheskii Sbornik.*, **30:72**(1952), 11-38 (in Russian).

[26] N.N.Lebedev, Special functions and their applications, Moscow, Fizmatlit, 1963 (in Russian).

[27] G.P.Tolstov, Fourier series, Moscow, Nauka, 1980 (in Russian).

Нелокальная задача для трехмерного эллиптического уравнения с сингулярными коэффициентами в прямоугольном параллелепипеде

Камолиддин Т. Каримов
Ферганский государственный университет
Фергана, Узбекистан

Аннотация. Исследована нелокальная задача для эллиптического уравнения с двумя сингулярными коэффициентами в прямоугольном параллелепипеде. Доказательство единственности решения и его построение проведены спектральным методом с использованием разложения в ряд Фурье и Фурье-Бесселя. При некоторых условиях относительно параметров и заданных функций доказана равномерная сходимость построенного ряда.

Ключевые слова: уравнения эллиптического типа, нелокальная задача, сингулярный коэффициент, спектральный метод, параллелепипед.

DOI: 10.17516/1997-1397-2020-13-5-547-558
УДК 517.9

On Initial Boundary Value Problem for Parabolic Differential Operator with Non-coercive Boundary Conditions

Alexander N. Polkovnikov*
Siberian Federal University
Krasnoyarsk, Russian Federation

Received 10.05.2020, received in revised form 02.06.2020, accepted 20.07.2020

Abstract. We consider initial boundary value problem for uniformly 2-parabolic differential operator of second order in cylinder domain in $\mathbb{R}^n$ with non-coercive boundary conditions. In this case there is a loss of smoothness of the solution in Sobolev type spaces compared with the coercive situation. Using by Faedo-Galerkin method we prove that problem has unique solution in special Bochner space.

Keywords: non-coercive problem, parabolic problem, Faedo-Galerkin method.

Citation: A. N. Polkovnikov, On Initial Boundary Value Problem for Parabolic Differential Operator with Non-coercive Boundary Conditions, J. Sib. Fed. Univ. Math. Phys., 2020, 13(5), 547–558. DOI: 10.17516/1997-1397-2020-13-5-547-558.

Initial boundary value problems for parabolic (by Petrovsky) differential operators with coercive boundary conditions are well studied (see, for instance, [1–4]). However the problem with non-coercive boubdary conditions are also appeared in both theory and applications, see, for instance, pioneer work in this direction [5] and papers [6, 7] and [8] for such problems in the Elasticity Theory. Recent results in Fredholm operator equations, induced by boundary value problems for elliptic differential operators with non-coercive boundary conditions (see, for instance, [9–12]) allows us to apply these one for studying the parabolic problem. Consideration of such problems essentially extends variety of boundary operators, but there is a loss of regularity of the solution (see [13] for elliptic case). Namely, let Ω_T be a cylinder,

$$\Omega_T = \Omega \times (0, T),$$

where Ω is a bounded domain in $\mathbb{R}^n$.

Consider a second order differential operator

$$A(x, t, \partial) = -\sum_{i,j=1}^{n} \partial_i (a_{i,j}(x) \partial_j \cdot) + \sum_{j=1}^{n} a_j(x) \partial_l + a_0(x) + \frac{\partial}{\partial t}$$

of divergence form in the domain Ω_T. The coefficients $a_{i,j}$, a_j are assumed to be complex-valued functions of class $L^\infty(\Omega)$. We suppose that the matrix $\mathfrak{A}(x) = (a_{i,j}(x))_{\substack{i=1,\dots,n \\ j=1,\dots,n}}$ is Hermitian and satisfies

$$\sum_{i,j=1}^{n} a_{i,j}(x) \overline{w}_i w_j \geqslant 0 \quad \text{for all} \quad (x, w) \in \overline{\Omega} \times \mathbb{C}^n, \tag{1}$$

*paskaattt@yandex.ru https://orcid.org/0000-0002-7066-3219.

$$\sum_{i,j=1}^{n} a_{i,j}(x)\xi_i\xi_j \geqslant m\,|\xi|^2 \text{ for all } (x,\xi) \in \overline{\Omega} \times (\mathbb{R}^n \setminus \{0\}), \tag{2}$$

where m is a positive constant independent of x and ξ. Estimate (2) is nothing but the statement that the operator $A(x,t,\partial)$ is uniformly 2-parabolic.

We note that, since the coefficients of the operator and the functions under consideration are complex-valued, inequalities (1) and (2) are weaker than

$$\sum_{i,j=1}^{n} a_{i,j}(x)\,\overline{w}_i w_j \geqslant m\,|w|^2 \tag{3}$$

for all $(x,w) \in \overline{\Omega} \times (\mathbb{C}^n \setminus \{0\})$. Inequality (3) means that correspondent Hermitian form (see form (4)) is coercive.

Consider boundary operator of Robin type:

$$B(x,\partial) = b_1(x) \sum_{i,j=1}^{n} a_{i,j}(x)\,\nu_i\partial_j + b_0(x),$$

where b_0, b_1 are bounded functions on $\partial\Omega$ and $\nu(x) = (\nu_1(x),\dots,\nu_n(x))$ is the unit outward normal vector of $\partial\Omega$ at $x \in \partial\Omega$. Let S be an open connected subset of $\partial\Omega$ with piecewise smooth boundary ∂S. We allow the function $b_1(x)$ to vanish on S. In this case we assume that $b_0(x)$ does not vanish for $x \in S$.

Consider now the following mixed initial-boundary problem in a bounded domain Ω_T with Lipschitz boundary $\partial\Omega_T$.

Problem 1. *Find a distribution $u(x,t)$, satisfying the problem*

$$\begin{cases} A(x,t,\partial)u &= f \quad in \quad \Omega_T, \\ B(x,\partial)u &= 0 \quad on \quad \partial\Omega \times (0,T), \\ u(x,0) &= u_0 \quad on \quad \Omega \end{cases}$$

with given data $f \in \Omega_T$.

For solving the problem we have to define appropriate functional spaces. Denote by $C^1(\overline{\Omega},S)$ the subspace of $C^1(\overline{\Omega})$ consisting of those functions whose restriction to the boundary vanishes on $\overline{S}$. Let $H^1(\Omega,S)$ be the closure of $C^1(\overline{\Omega},S)$ in $H^1(\Omega)$. Since on S the boundary operator reduces to $B = b_0(x)$ and $b_0(x) \neq 0$ for $x \in S$, then the functions $u \in H^1(\Omega)$ satisfying $Bu = 0$ on $\partial\Omega$ belong to $H^1(\Omega,S)$.

Split now both $a_0(x)$ and $b_0(x)$ into two parts

$$a_0 = a_{0,0} + \delta a_0,$$

$$b_0 = b_{0,0} + \delta b_0,$$

where $a_{0,0}$ is a non-negative bounded function in Ω and $b_{0,0}$ is a such function that $b_{0,0}/b_1$ is non-negative bounded function on S. Then, under reasonable assumptions, the Hermitian form

$$(u,v)_+ = \int_\Omega \sum_{i,j=1}^{n} a_{i,j}\partial_j u\overline{\partial_i v}\,dx + (a_{0,0}u,v)_{L^2(\Omega)} + (b_{0,0}/b_1\,u,v)_{L^2(\partial\Omega\setminus S)} \tag{4}$$

defines the scalar product on $H^1(\Omega,S)$. Denote by $H^+(\Omega)$ the completion of the space $H^1(\Omega,S)$ with respect to the corresponding norm $\|\cdot\|_+$. From now on we assume that the space $H^+(\Omega)$ is continuously embedded into the Lebesgue space $L^2(\Omega)$, i.e. there is a constant $c > 0$, independent of u, such that

$$\|u\|_{L^2(\Omega)} \leqslant c\|u\|_+ \text{ for all } u \in H^+(\Omega).$$

It is true, if there exist a positive constant c_1 such that

$$a_{0,0} \geqslant c_1 \text{ in } \Omega.$$

Actually we can get more subtle embedding for the space $H^+(\Omega)$.

Theorem 2. *Let the coefficients $a_{i,j}$ be C^∞ in a neighbourhood of the closure of Ω, inequalities (1), (2) hold and*

$$\frac{b_{0,0}}{b_1} \geqslant c_2 \text{ at } \partial\Omega \setminus S, \tag{5}$$

with some constant $c_2 > 0$. Then the space $H^+(\Omega)$ is continuously embedded into $H^{1/2-\varepsilon}(\Omega)$ for any $\varepsilon > 0$ if there is a positive constant c_1, such that

$$a_{0,0} \geqslant c_1 \text{ in } \Omega \tag{6}$$

or the operator A is strongly elliptic in a neighborhood X of $\overline{\Omega}$ and

$$\int_X \sum_{i,j=1}^n a_{i,j}\partial_j u \overline{\partial_i u}\, dx \geqslant m\, \|u\|^2_{L^2(X)} \tag{7}$$

for all $u \in C^\infty_{\rm comp}(X)$, with $m > 0$ a constant independent of u.

Proof. See [12, Theorem 2.5]. □

Of course, under coercive estimate (3), the space $H^+(\Omega)$ is continuously embedded into $H^1(\Omega)$. However, in general, the embedding, described in Theorem 2 is rather sharp (see [12, Remark 5.1]).

The absence of coerciveness does not allows to consider arbitrary derivatives $\partial_j u$ for an element $u \in H^+(\Omega)$. To cope with this difficulty we note that the matrix $\mathfrak{A}(x) = (a_{i,j}(x))_{\substack{i=1,\dots,n\\ j=1,\dots,n}}$ admits a factorisation, i.e. there is an $(m \times n)$-matrix $\mathfrak{D}(x) = (\mathfrak{D}_{i,j}(x))_{\substack{i=1,\dots,m\\ j=1,\dots,n}}$ of bounded functions in Ω, such that

$$(\mathfrak{D}(x))^*\mathfrak{D}(x) = \mathfrak{A}(x) \tag{8}$$

for almost all $x \in D$ (see, for instance, [14]). For example, one could take the standard non-negative self-adjoint square root $\mathfrak{D}(x) = \sqrt{\mathfrak{A}(x)}$ of the matrix $\mathfrak{A}(x)$. Then

$$\sum_{i,j=1}^n a_{i,j}\partial_j u\overline{\partial_i v} = (\mathfrak{D}\nabla v)^*\mathfrak{D}\nabla u = \sum_{l=1}^m \overline{\mathfrak{D}_l v}\,\mathfrak{D}_l u,$$

for all smooth functions u and v in Ω, where ∇u is thought of as n-column with entries $\partial_1 u, \dots, \partial_n u$, and $\mathfrak{D}_l u := \sum\limits_{s=1}^n \mathfrak{D}_{l,s}(x)\partial_s u$, $l = 1, \dots, m$. From now on we may confine ourselves with first order summand of the form

$$\sum_{l=1}^m \tilde{a}_l(x)\mathfrak{D}_l, \quad \tilde{a}_l(x) \in L^\infty(\Omega),$$

instead of
$$\sum_{j=1}^{n} a_j(x)\partial_j.$$
Since the coefficients δa_0, $\tilde{a}_l$ belong to $L^\infty(\Omega)$ for all $l = 0, \dots, m$, it follows from Cauchy inequality that
$$\left| \left(\Big(\sum_{l=1}^{m} \tilde{a}_l(x)\mathfrak{D}_l + \delta a_0 \Big) u, v \right)_{L^2(\Omega)} \right| \leqslant c\, \|u\|_+ \, \|v\|_+. \tag{9}$$

Let now $H^-(\Omega)$ stand for the dual space for the space $H^+(\Omega)$ with respect to the pairing $< \cdot, \cdot >$ induced by the scalar product $(\cdot, \cdot)_{L^2(\Omega)}$, see [2, 15] and elsewhere. It is a Banach space with the norm
$$\|u\|_- = \sup_{\substack{v \in H^+(\Omega) \\ v \neq 0}} \frac{|(v, u)_{L^2(\Omega)}|}{\|v\|_+}.$$
The space $L^2(\Omega)$ is continuously embedded into $H^-(\Omega)$, if the space $H^+(\Omega)$ is continuously embedded into $L^2(\Omega)$ (see [9]). We denote by $i' : L^2(\Omega) \to H^-(\Omega)$ and $i : H^+(\Omega) \to L^2(\Omega)$ the operators of correspondent continuously embeddings. Thus we have a triple of the functional spaces
$$H^+(\Omega) \overset{i}{\hookrightarrow} L^2(\Omega) \overset{i'}{\hookrightarrow} H^-(\Omega),$$
where each embeddings is compact under the hypothesis of Theorem 2.

Denote by $L^2(0, T; H^+(\Omega))$ the Bochner space of L^2-functions
$$u(t) : [0, T] \to H^+(\Omega).$$
It is a Banach space with the norm
$$\|u\|^2_{L^2(0,T;H^+(\Omega))} = \int_0^T \|u(t)\|_+^2 \, dt.$$

Then an integration by parts in Ω leads to a weak formulation of Problem (1):

Problem 3. *Given $f \in L^2(0, T; H^-(\Omega))$ and $u_0 \in L^2(\Omega)$, find $u \in L^2(0, T; H^+(\Omega))$, such that*
$$(u, v)_+ + \Big(\big(\sum_{l=1}^{m} \tilde{a}_l(x)\mathfrak{D}_l + \delta a_0 \big) u, v \Big)_{L^2(\Omega)} + \frac{\partial}{\partial t} (u, v)_{L^2(\Omega)} =< f, v > \tag{10}$$
for all $v \in H^+(\Omega)$, and
$$u(0) = u_0. \tag{11}$$

In general case the condition (11) have no sense for functions $u \in L^2(0, T; H^+(\Omega))$. But we will see below that function $u(t) \in L^2(0, T; H^+(\Omega))$, satisfying (10), is continuous and (11) have a sense.

We want to apply the Faedo-Galerkin method for solving the Problem 3 (see, for instance, [2, 4]). For this purpose we need some complete system of vectors in the space $H^+(\Omega)$. As this system we take the set of eigenvectors of an operator, induced by the weak statement of elliptic selfadjoint problem, corresponding to the parabolic Problem 3. Namely, for given $f \in H^-(\Omega)$, find $u \in H^+(\Omega)$, such that
$$(u, v)_+ + \left(\Big(\sum_{l=1}^{m} \tilde{a}_l(x)\mathfrak{D}_l + \delta a_0 \Big) u, v \right)_{L^2(\Omega)} =< f, v > . \tag{12}$$

Equality (12) induces a bounded linear operator $L : H^+(\Omega) \to H^-(\Omega)$,

$$(u,v)_+ + \Big(\Big(\sum_{l=1}^{m} \tilde{a}_l(x)\mathfrak{D}_l + \delta a_0\Big)u, v\Big)_{L^2(\Omega)} =< Lu, v > . \tag{13}$$

Denote by L_0 the operator L in the case, when $\delta a_0 = a_l = 0$ for all $l = 1, \dots, m$,

$$(u,v)_+ =< L_0 u, v > . \tag{14}$$

The operator $L_0 : H^+(\Omega) \to H^-(\Omega)$ is continuously invertible and $\|L_0\| = \|L_0^{-1}\| = 1$ (see [12, Lemma 2.6]). According to [12, Lemma 3.1], there is a system $\{h_j\}$ of eigenvectors of the compact positive selfadjoint operator $L_0^{-1} i' i : H^+(\Omega) \to H^+(\Omega)$, which is an orthonormal bases in $H^+(\Omega)$ and an orthogonal bases in $L^2(\Omega)$ and $H^-(\Omega)$.

Let now function $u \in L^2(0,T; H^+(\Omega))$ satisfies (10). We have from (13)

$$\Big(\frac{\partial u}{\partial t}, v\Big)_{L^2(\Omega)} =< \frac{\partial u}{\partial t}, v >=< f - Lu, v > .$$

Since $f \in L^2(0,T; H^-(\Omega))$ and operator $L : H^+(\Omega) \to H^-(\Omega)$ is bounded, then $\frac{\partial u}{\partial t} \in L^2(0,T; H^-(\Omega))$. It means, that

$$u \in C(0,T; L^2(\Omega)) \tag{15}$$

(see, for instance, [2] or [16]).

Using by the standard Faedo-Galerkin method (see, for instance, [1,2,4]) we get next Theorem.

Theorem 4. *Under the hypothesis of Theorem 2, the Problem 3 has at least one solution $u(t)$, and, moreover, $u(t) \in C(0,T; L^2(\Omega))$.*

Proof. For each k we are looking for approximate solution of Problem 3 on the next form

$$u_k(t) = \sum_{j=1}^{k} g_{jk}(t) h_j, \tag{16}$$

and function u_k satisfies

$$(u_k, h_i)_+ + \Big(\Big(\sum_{l=1}^{m} \tilde{a}_l(x)\mathfrak{D}_l + \delta a_0\Big)u_k, h_i\Big)_{L^2(\Omega)} + \Big(\frac{\partial u_k}{\partial t}, h_i\Big)_{L^2(\Omega)} =< f, h_i >, \tag{17}$$

$$u_k(0) = \sum_{j=1}^{k} \frac{(u_0, h_j)_{L^2(\Omega)}}{\|h_j\|^2_{L^2(\Omega)}} h_j, \tag{18}$$

for each $j = 1, \dots, k$, where $\{h_j\}$ is the orthonormal bases in $H^+(\Omega)$. It means that (17) takes the form

$$g_{ik}(t) + \sum_{j=1}^{k} \Big(\Big(\sum_{l=1}^{m} \tilde{a}_l(x)\mathfrak{D}_l + \delta a_0\Big)h_j, h_i\Big)_{L^2(\Omega)} g_{jk}(t) + g'_{ik}(t)\|h_i\|^2_{L^2(\Omega)} =< f, h_i >, \tag{19}$$

where $i = 1, \dots, k$. It is a system of linear differential equations of first order with initial conditions

$$g_{ik}(0) = \frac{(u_0, h_i)_{L^2(\Omega)}}{\|h_i\|^2_{L^2(\Omega)}}, \quad i = 1, \dots, k. \tag{20}$$

Since $< f(t), h_i >$ is measurable function for all $i = 1, \dots, k$, then there is unique function $g_{ik}(t)$ for each $i = 1, \dots, k$, satisfying (19) and (20) for all $t \in [0, T]$ (see, for instance, [17]). Note, as the function $u(t)$ is complex-valued, then the functions $\{g_{ik}(t)\}$ may be complex-valued too and the system (19) consists $2k$ real-valued equations in general case.

Now we have to get some priori estimates for function $u_k(t)$ independent of k. Multiplying the equality (17) by the $\overline{g_{ik}(t)}$ and summing by $i = 1, \dots, k$ we get

$$\|u_k\|_+^2 + \left(\frac{\partial u_k}{\partial t}, u_k\right)_{L^2(\Omega)} =< f, u_k > -\left(\left(\sum_{l=1}^m \tilde{a}_l(x)\mathfrak{D}_l + \delta a_0\right) u_k, u_k\right)_{L^2(\Omega)}. \tag{21}$$

Hence, by the Cauchy inequality,

$$\begin{aligned} &2\left|\|u_k\|_+^2 + \left(\frac{\partial u_k}{\partial t}, u_k\right)_{L^2(\Omega)}\right| = \\ =&2\left| < f, u_k > -\left(\sum_{l=1}^m \tilde{a}_l(x)\mathfrak{D}_l u_k, u_k\right)_{L^2(\Omega)} - (\delta a_0 u_k, u_k)_{L^2(\Omega)}\right| \leqslant \\ \leqslant&\|f\|_-^2 + \|u_k\|_+^2 + 2c_1\|u_k\|_+\|u_k\|_{L^2(\Omega)} + 2c_2\|u_k\|_{L^2(\Omega)}^2 \leqslant \\ \leqslant&\|f\|_-^2 + \frac{3}{2}\|u_k\|_+^2 + (2c_2 + 2c_1^2)\|u_k\|_{L^2(\Omega)}^2 \end{aligned} \tag{22}$$

for some positive constants c_1 and c_2. As the norm $\|u_k\|_+^2$ is a real-valued function, we have

$$\begin{aligned} &2\left|\|u_k\|_+^2 + \left(\frac{\partial u_k}{\partial t}, u_k\right)_{L^2(\Omega)}\right| = \\ &= 2\left|\|u_k\|_+^2 + \mathfrak{Re}\left(\left(\frac{\partial u_k}{\partial t}, u_k\right)_{L^2(\Omega)}\right) + i\mathfrak{Im}\left(\left(\frac{\partial u_k}{\partial t}, u_k\right)_{L^2(\Omega)}\right)\right| \geqslant \\ &\geqslant 2\|u_k\|_+^2 + 2\mathfrak{Re}\left(\left(\frac{\partial u_k}{\partial t}, u_k\right)_{L^2(\Omega)}\right), \end{aligned} \tag{23}$$

where $\mathfrak{Re}(g)$ and $\mathfrak{Im}(g)$ denote real and imaginary parts of function g respectively. On the other hand,

$$\begin{aligned} \frac{d}{dt}\|u_k\|_{L^2(\Omega)}^2 &= \left(\frac{\partial u_k}{\partial t}, u_k\right)_{L^2(\Omega)} + \left(u_k, \frac{\partial u_k}{\partial t}\right)_{L^2(\Omega)} = \\ &= 2\mathfrak{Re}\left(\left(\frac{\partial u_k}{\partial t}, u_k\right)_{L^2(\Omega)}\right). \end{aligned} \tag{24}$$

It follows from (22), (23) and (24) that

$$\frac{1}{2}\|u_k(t)\|_+^2 + \frac{d}{dt}\|u_k(t)\|_{L^2(\Omega)}^2 \leqslant \|f(t)\|_-^2 + (2c_2 + 2c_1^2)\|u_k\|_{L^2(\Omega)}^2. \tag{25}$$

Now, integrating (25) by t from 0 till some $s \in (0, T)$ we get

$$\begin{aligned} &\frac{1}{2}\int_0^s \|u_k(t)\|_+^2 dt + \|u_k(s)\|_{L^2(\Omega)}^2 - \|u_k(0)\|_{L^2(\Omega)}^2 \leqslant \\ &\leqslant \int_0^s \|f(t)\|_-^2 dt + (2c_2 + 2c_1^2)\int_0^s \|u_k\|_{L^2(\Omega)}^2 dt. \end{aligned}$$

Since the sequence $\{u_k(0)\}$ seeks to u_0 with $k \to \infty$ strongly in $L^2(\Omega)$, it follows from Gronwall type lemma (see [18] or [19]), that

$$\|u_k(s)\|^2_{L^2(\Omega)} \leqslant \left(\|u_0\|^2_{L^2(\Omega)} + \int_0^T \|f(t)\|^2_- dt \right) e^{(2c_2+2c_1^2)s}.$$

Hence

$$\sup_{s\in[0,T]} \|u_k(s)\|^2_{L^2(\Omega)} \leqslant \left(\|u_0\|^2_{L^2(\Omega)} + \int_0^T \|f(t)\|^2_- dt \right) e^{(2c_2+2c_1^2)T}. \tag{26}$$

The right side of (26) independent of k, therefore the sequence $\{u_k(t)\}$ is bounded in $L^\infty(0,T;L^2(\Omega))$. Then there is a subsequence $\{u_{k'}(t)\}$ of the sequence $\{u_k(t)\}$ and an element $u(t) \in L^\infty(0,T;L^2(\Omega))$ such that $u_{k'}(t) \to u(t)$ in the weak-* topology of $L^\infty(0,T;L^2(\Omega))$, namely

$$\lim_{k'\to\infty} \int_0^T (u_{k'}(t) - u(t), v(t))_{L^2(\Omega)} dt = 0 \tag{27}$$

for all $v \in L^1(0,T;L^2(\Omega))$.

Integrating again (25) by t from 0 till T and applying Gronwall type lemma we have

$$\begin{gathered} \frac{1}{2}\int_0^T \|u_k(t)\|^2_+ dt + \|u_k(T)\|^2_{L^2(\Omega)} \leqslant \\ \leqslant \left(\|u_0\|^2_{L^2(\Omega)} + \int_0^T \|f(t)\|^2_- dt \right) e^{(2c_2+2c_1^2)T}. \end{gathered} \tag{28}$$

It means that the sequence $\{u_k(t)\}$ is bounded in $L^2(0,T;H^+(\Omega))$. In particular, the sequence $\{u_{k'}(t)\}$ is bounded in $L^2(0,T;H^+(\Omega))$ too. Hence there is a subsequence $\{u_{k''}(t)\}$ of the sequence $\{u_{k'}(t)\}$ and an element $\widetilde{u}(t) \in L^2(0,T;H^+(\Omega))$ such that $u_{k''}(t) \to u(t)$ in the weak topology of $L^2(0,T;H^+(\Omega))$,

$$\lim_{k''\to\infty} \int_0^T (u_{k''}(t), v)_+ \, dt = \int_0^T (u(t), v)_+ \, dt \tag{29}$$

for all $v \in L^2(0,T;H^+(\Omega))$ and

$$\lim_{k''\to\infty} \int_0^T < u_{k''}(t) - \widetilde{u}(t), v(t) > dt = 0 \tag{30}$$

for all $v \in L^2(0,T;H^-(\Omega))$. In particular

$$\lim_{k''\to\infty} \int_0^T (u_{k''}(t), v(t))_{L^2(\Omega)} dt = \int_0^T (\widetilde{u}(t), v(t))_{L^2(\Omega)} dt \tag{31}$$

for all $v \in L^2(0,T;L^2(\Omega))$.

From (27) and (31) we have

$$\int_0^T (u(t) - \widetilde{u}(t), v(t))_{L^2(\Omega)} dt = 0 \tag{32}$$

for all $v \in L^2(0,T;L^2(\Omega))$. Hence

$$u(t) = \widetilde{u}(t) \in L^\infty(0,T;L^2(\Omega)) \cap L^2(0,T;H^+(\Omega)). \tag{33}$$

From now on we denote by $\{u_k(t)\}$ the subsequence $\{u_{k''}(t)\}$.

Let now $\psi(t)$ be a scalar differentiable function on $[0,T]$ such that $\psi(T)=0$. Multiplying (17) by $\psi(t)$ and integrating by t we get

$$\int_0^T (u_k(t),h_j)_+\psi(t)\,dt + \int_0^T \Big(\Big(\sum_{l=1}^m \tilde a_l(x)\mathfrak{D}_l + \delta a_0\Big)u_k(t),h_i\Big)_{L^2(\Omega)}\psi(t)\,dt+ \\ + \int_0^T \Big(\frac{\partial u_k(t)}{\partial t},h_i\Big)_{L^2(\Omega)}\psi(t)\,dt = \int_0^T < f(t),h_j > \psi(t)dt. \tag{34}$$

However

$$\int_0^T \Big(\frac{\partial u_k(t)}{\partial t},h_i\Big)_{L^2(\Omega)}\psi(t)\,dt \;=\; -\int_0^T (u_k(t),\psi'(t)h_j)_{L^2(\Omega)}\,dt - (u_k(0),h_j\psi(0))_{L^2(\Omega)}, \tag{35}$$

and it follows that

$$\int_0^T (u_k(t),h_j\psi(t))_+\,dt + \int_0^T \Big(\Big(\sum_{l=1}^m \tilde a_l(x)\mathfrak{D}_l + \delta a_0\Big)u_k(t),h_i\Big)_{L^2(\Omega)}\psi(t)\,dt- \\ - \int_0^T (u_k(t),\psi'(t)h_j)_{L^2(\Omega)}\,dt = (u_k(0),h_j\psi(0))_{L^2(\Omega)} + \int_0^T < f(t),h_j > \psi(t)dt. \tag{36}$$

Now we want to go to the limit in (36) with $k\to\infty$. It follows from 9, that

$$\int_0^T \Big(\Big(\sum_{l=1}^m \tilde a_l(x)\mathfrak{D}_l + \delta a_0\Big)u_k(t),h_i\Big)_{L^2(\Omega)}\psi(t)\,dt$$

is continuous linear functional on $L^2(0,T;H^+(\Omega))$. Since $u_k(t)\to u(t)$ with $k\to\infty$ in the weak topology of $L^2(0,T;H^+(\Omega))$, we have

$$\lim_{k\to\infty}\int_0^T \Big(\Big(\sum_{l=1}^m \tilde a_l(x)\mathfrak{D}_l + \delta a_0\Big)(u_k(t)-u(t)),h_i\Big)_{L^2(\Omega)}\psi(t)\,dt = 0.$$

From (31), (29), (33) and the fact that $u_k(0)\to u_0$ strongly in $L^2(\Omega)$ with $k\to\infty$ we get

$$\int_0^T (u(t),h_j\psi(t))_+\,dt + \int_0^T \Big(\Big(\sum_{l=1}^m \tilde a_l(x)\mathfrak{D}_l + \delta a_0\Big)u(t),h_i\psi(t)\Big)_{L^2(\Omega)}dt- \\ - \int_0^T (u(t),\psi'(t)h_j)_{L^2(\Omega)}\,dt = (u_0,h_j\psi(0))_{L^2(\Omega)} + \int_0^T < f(t),h_j > \psi(t)dt. \tag{37}$$

As the system $\{h_j\}_{j=1,2,\dots}$ is dense in $H^+(\Omega)$ and $L^2(\Omega)$, equality (37) holds by linearity and continuity for all $v\in H^+(\Omega)$,

$$\int_0^T (u(t),v)_+\psi(t)\,dt + \int_0^T \Big(\Big(\sum_{l=1}^m \tilde a_l(x)\mathfrak{D}_l + \delta a_0\Big)u(t),v\Big)_{L^2(\Omega)}\psi(t)\,dt- \\ - \int_0^T (u(t),v)_{L^2(\Omega)}\,\psi'(t)\,dt = (u_0,v)_{L^2(\Omega)}\psi(0) + \int_0^T < f(t),v > \psi(t)dt. \tag{38}$$

In particular, if we take by $\psi(t)$ differentiable functions with compact support in $(0, T)$, we get

$$(u(t), v)_+ + \Big(\Big(\sum_{l=1}^{m} \tilde{a}_l(x)\mathfrak{D}_l + \delta a_0\Big)u(t), v\Big)_{L^2(\Omega)} + \frac{d}{dt}\,(u(t), v)_{L^2(\Omega)} =< f(t), v > \tag{39}$$

in the sense of distributions. Now we have to show that $u(0) = u_0$. Indeed, multiplying (39) by $\psi(t)$ and integrating by parts we get

$$\int_0^T (u(t), v)_+\psi(t)\,dt + \int_0^T \Big(\Big(\sum_{l=1}^{m} \tilde{a}_l(x)\mathfrak{D}_l + \delta a_0\Big)u(t), v\Big)_{L^2(\Omega)} \psi(t)\,dt-$$
$$- \int_0^T (u(t), v)_{L^2(\Omega)}\, \psi'(t)\,dt = (u(0), v)_{L^2(\Omega)}\psi(0) + \int_0^T < f(t), v > \psi(t)dt.$$

Comparing it with (38) we get

$$(u(0) - u_0, v)_{L^2(\Omega)}\psi(0) = 0$$

for all $v \in H^+(\Omega)$. Taking $\psi(0) \neq 0$ we receive $u(0) = u_0$.

The continuity follows from (15). □

Corollary 5. *Under the hypothesis of Theorem 2, the Problem 3 has one and only one solution $u(t) \in C(0, T; L^2(\Omega))$, if*

$$\mathfrak{Re}\left(\Big(\Big(\sum_{l=1}^{m} \tilde{a}_l(x)\mathfrak{D}_l + \delta a_0\Big)v, v\Big)_{L^2(\Omega)}\right) \geqslant 0 \tag{40}$$

for all $v \in L^2(0, T; H^+(\Omega))$.

Proof. The existence of the solution follows from the Theorem 4. Let us now show, that the solution is unique, if the condition (40) and the hypothesis of Theorem 4 are fulfilled. Indeed, let $v \in L^2(0, T; H^+(\Omega))$ is another solution of Problem 3. Denote by $w = u - v$. Then w satisfies conditions of Problem 3 and

$$(w, v)_+ + \Big(\Big(\sum_{l=1}^{m} \tilde{a}_l(x)\mathfrak{D}_l + \delta a_0\Big)w, v\Big)_{L^2(\Omega)} + \frac{d}{dt}\,(w, v)_{L^2(\Omega)} = 0$$

for all $v \in H^+(\Omega)$, and $w(0) = 0$. It follows from (13), that

$$\frac{\partial w}{\partial t} + Lw = 0.$$

Multiplying scalar it by w we have

$$\|w\|_+^2 + \Big(\Big(\sum_{l=1}^{m} \tilde{a}_l(x)\mathfrak{D}_l + \delta a_0\Big)w, w\Big)_{L^2(\Omega)} + \left(\frac{\partial w}{\partial t}, w\right)_{L^2(\Omega)} = 0.$$

As the $\|w(t)\|_+^2$ is a real-valued function, therefore

$$\|w\|_+^2 + \mathfrak{Re}\left(\left(\frac{\partial w}{\partial t}, w\right)_{L^2(\Omega)}\right) + \mathfrak{Re}\left(\Big(\Big(\sum_{l=1}^{m} \tilde{a}_l(x)\mathfrak{D}_l + \delta a_0\Big)w, w\Big)_{L^2(\Omega)}\right) = 0.$$

On the other hand,

$$\mathfrak{Re}\left(\left(\frac{\partial w}{\partial t}, w\right)_{L^2(\Omega)}\right) = \frac{1}{2}\frac{d}{dt}\|w\|^2_{L^2(\Omega)}.$$

It follows from (40), that

$$2\mathfrak{Re}\left(\left(\frac{\partial w}{\partial t}, w\right)_{L^2(\Omega)}\right) = \frac{d}{dt}\|w\|^2_{L^2(\Omega)} \leqslant 0$$

and

$$\|w(t)\|^2_{L^2(\Omega)} \leqslant \|w(0)\|^2_{L^2(\Omega)} = 0,$$

hence $w(t) = 0$ for almost all $t \in [0, T]$, that completes the proof. □

As we already mentioned, the embedding $H^+(\Omega)$ into $H^{1/2-\varepsilon}(\Omega)$ is rather sharp. Let us show, that the space $L^2(0,T;H^+(\Omega))$ can not be continuously embedded into $L^2(0,T;H^s(\Omega))$ for all $s > 1/2$.

Eexample 6. Let Ω be a unit sphere in $\mathbb{C}$, matrix $\mathfrak{A}(x)$ has a form

$$\mathfrak{A}(x) = (a_{ij}(x))_{\substack{i=1,2\\ j=1,2}} = \begin{pmatrix} 1 & \sqrt{-1} \\ -\sqrt{-1} & 1 \end{pmatrix},$$

$S = \emptyset$, $a_l = 0$ for $l = 0, 1, \dots, m$, and $b_1 = b_0 = 1$. Then the series

$$u_\varepsilon(z,t) = \sum_{k=0}^{\infty} \frac{z^k t^{k/2}}{T^{(k+1)/2}(k+1)^{\varepsilon/2}},$$

$\varepsilon > 0$, converges in $L^2(0,T;H^+(\Omega))$ and

$$\|u_\varepsilon\|^2_{L^2(0,T;H^+(\Omega))} = \|u_\varepsilon\|^2_{L^2(0,T;L^2(\mathbb{S}))} = 2\pi \sum_{k=0}^{\infty} \frac{1}{(k+1)^{1+\varepsilon}}.$$

According to [20, Lemma 1.4]

$$\left\|u_\varepsilon\right\|^2_{L^2(0,T;H^s(\mathbb{B}))} \geqslant \pi \sum_{k=0}^{\infty} \frac{k^{2s-1}}{(k+1)^{1+\varepsilon}},\ 0 < s \leqslant 1.$$

It means, that for each $s \in (1/2, 1)$ there exist $\varepsilon > 0$ such that $u_\varepsilon \notin L^2(0,T;H^s(\mathbb{B}))$. Hence, the space $L^2(0,T;H^+(\mathbb{B}))$ can not be continuously embedded into $L^2(0,T;H^s(\mathbb{B}))$ for all $s > 1/2$.

The work was supported by the Foundation for the Advancement of Theoretical Physics and Mathematics "BASIS".

References

[1] O.A.Ladyzhenskaya, V.A.Solonnikov, N.N.Ural'tseva, Linear and Quasilinear Equations of Parabolic Type, Amer. Math. Soc., Providence, R.I., 1968

[2] J.L.Lions, E.Magenes, Non-Homogeneous Boundary Value Problems and Applications, Springer-Verlag, Berlin et al., 1972.

[3] V.P.Mikhailov, Partial Differential Equations, Moscow, Mir, 1978.

[4] R.Temam, Navier-Stokes equations: Theory and Numerical Analysis, Studies in Math. and its Appl. 2, 1979.

[5] S.Agmon, A.Douglis, L.Nirenberg, Estimates near the boundary for solutions of elliptic partial differential equations satisfying general boundary conditions, Part 1. *Comm. Pure Appl. Math.*, **12**(1959), 623–727.

[6] S.Campanato, Sui problemi al contorno per sistemi di equazioni differenziale lineari del tipo dell'elasticitá, *Ann. della Scuola Norm. Superiore, Cl. di Sci, Ser. III*, **13**(1959), no. 2, 223–258.

[7] S.Campanato, Proprietá di taluni spazi di distribuzioni e loro applicazione, *Ann. della Scuola Norm. Superiore, Cl. di Sci*, Ser. III, **14**(1960), no. 4, 363–376.

[8] A.S.Peicheva, A.A.Shlapunov, On the completeness of root functions of Sturm-Liouville problems for the Lame system in weighted spaces. *Z. Angew. Math. Mech.*, **95**(2015), no. 11, 1202–1214.

[9] A.N.Polkovnikov, A.A.Shlapunov, On spectral properties of a non-coercive mixed problem associated with the $\overline{\partial}$-operator, *Journal of Siberian Federal University. Math. and Phys.*, **6**(2013), no. 2, 247–261.

[10] A.Polkovnikov, A.Shapunov, On non-coercive mixed problems for parameter-dependent elliptic operators, *Math. Commun.*, **15**(2015), 1–20,

[11] A.Polkovnikov, A.A.Shlapunov, *Sib. Math. J.*, **58**(2017), no. 4, 676–686.
DOI: 10.1134/S0037446617040140

[12] A.Shlapunov, N. Tarkhanov, On completeness of root functions of Sturm-Liouville problems with discontinuous boundary operators, *J. of Differential Equations*, **10**(2013), 3305–3337.

[13] J.J.Kohn, Subellipticity of the $\overline{\partial}$-Neumann problem on pseudoconvex domains: sufficient conditions, *Acta Math.*, **142**(1979), no. 1-2, 79–12.

[14] A.Shlapunov, N.Tarkhanov, *Sib. Adv. Math.*, **26**(2016), no. 1, 30–76.
DOI: 10.3103/S105513441601003X

[15] M.Schechter, Negative norms and boundary problems, *Ann. Math.*, **72**(1960), no. 3, 581–593.
DOI: 10.1090/S0002-9904-1960-10516-2

[16] H.Gaevsky, K.Greger, K.Zaharias, Nonlinear Operator Equations and Operator Differential Equations, Mir, Moscow, 1978.

[17] A.F.Filippov, Differential equations with discontinuous right-hand side, Nauka, Moscow, 1985 (in Russian).

[18] T.H.Gronwall, Note on the derivatives with respect to a parameter of the solutions of a system of differential equations, *Ann. of Math.*, **20**(1919), no. 2, 292–296.

[19] D.S.Mitrinović, J.E.Pečarić, A.M.Fink, Inequalities Involving Functions and Their Integrals and Derivatives, Mathematics and its Applications (East European Series), Vol. 53, Kluwer Ac. Publ., Dordrecht, 1991.

[20] A.A.Shlapunov, Spectral decomposition of Green's integrals and existence of $W^{s,2}$-solutions of matrix factorizations of the Laplace operator in a ball, *Rend. Sem. Mat. Univ. Padova*, **96**(1996), 237–256.

О начально-краевой задаче для параболического дифференциального оператора с некоэрцитивными граничными условиями

Александр Н. Полковников
Сибирский федеральный университет
Красноярск, Российская Федерация

Аннотация. Мы рассматриваем начально-краевую задачу для равномерно 2-параболического дифференциального оператора второго порядка в цилиндрической области в $\mathbb{R}^n$ с некоэрцитивными граничными условиями. В отличие от коэрцитивного случая в данной ситуации происходит потеря гладкости решения в пространствах соболевского типа. Пользуясь методом Галеркина, мы доказываем, что проблема имеет единственное решение в специальных пространствах Бохнера.

Ключевые слова: некоэрцитивная задача, параболическая задача, метод Галеркина.

DOI: 10.17516/1997-1397-2020-13-5-559-567
УДК 517.55

Relationship Between the Bergman and Cauchy-Szegö Kernels in the Domains $\tau^+(n-1)$ and $\Re_{IV}^n$

Gulmirza Kh. Khudayberganov*
Jonibek Sh. Abdullayev†
National University of Uzbekistan
Tashkent, Uzbekistan

Received 02.05.2020, received in revised form 30.05.2020, accepted 08.07.2020

Abstract. In this paper, a connection has been established between the Bergman and Cauchy-Szegö kernels using the biholomorphic equivalence of the domains $\tau^+(n-1)$ and the Lie ball $\Re_{IV}^n$. Moreover, integral representations of holomorphic functions in these domains are obtained.

Keywords: classical domains, Lie ball, future tube, Shilov's boundary, Jacobian, Bergman's kernel, Cauchy-Szegö's kernel, Poisson's kernel.

Citation: G.Kh. Khudayberganov, J.Sh. Abdullayev, Relationship Between the Bergman and Cauchy-Szegö Kernels in the Domains $\tau^+(n-1)$ and $\Re_{IV}^n$, J. Sib. Fed. Univ. Math. Phys., 2020, 13(5), 559–567. DOI: 10.17516/1997-1397-2020-13-5-559-567.

1. Introduction and preliminaries

The selection of classes biholomorphically equivalent domains has great importance in multidimensional analysis and its applications. It is well known that all simply connected proper open subsets of the plane $\mathbb{C}$ are conformally equivalent (Rieman mapping theorem). The situation is completely different in the multidimensional case. For instance, an open unit ball and an open unit polydisc are not biholomorphically equivalent. In fact, there does not exist any holomorphic mapping from one to the other. Therefore, it is very important to have stocks of domains that are biholomorphically equivalent to each other.

Finding the kernels of representations of holomorphic functions in domains $\mathbb{C}^n$ and in the matrix domains from $\mathbb{C}^n[m \times m]$ is a rather difficult task (see [1–4]). Usually, in classical theory, kernels of such kind are constructed in bounded symmetric domains (see [5]). One of such domain is the matrix ball. One considers the following problems for it (see [4, 6]): finding the transitive group of automorphisms of a matrix ball; computing the Bergman and Cauchy-Szegö kernels for this domain; finding Carleman's formula, recovering values of a holomorphic function in a matrix ball by its values on some boundary (uniqueness) sets (see [7–9]).

By writing down explicitly the transitive group of automorphisms of the matrix ball, by direct calculation, we can find the Bergman and Cauchy-Szegö kernels for this domain. And then (using the properties of the Poisson kernel) we can find Carleman's formula, which recovers values of a holomorphic function in whole domain by its values on some boundary set of uniqueness

*gkhudaiberg@mail.ru
†jonibek-abdullayev@mail.ru

(see [9–11]). Here we use the scheme from ([5, 12, 13]) for finding the Bergman and Cauchy-Szegö's kernels. In [14], the volumes of the third type matrix ball and the generalized Lie ball are calculated. The full volumes of these domains are necessary for finding kernels of integral formulas for these domains (the Bergman, Cauchy-Szegö's, Poisson kernels etc.). It is also used for the integral representation of functions holomorphic in these domains as well as in the mean value theorem and other important concepts.

Bergman spaces of bounded symmetric domains are fundamental objects in the analysis. They are equipped with natural projection, i.e. with the Bergman projection, which is defined by the property of the reproducing kernel. On the other hand, the Bergman weighted spaces are very important in harmonic analysis also. For any transitive circular domain, the Bergman kernel is equal to the ratio of the volume density to the Euclidean volume of the domain. In the book of Hua Lo-ken (see [5]) the Bergman kernels are constructed for four types of classical domains, guided only by this consideration and without referring to complete orthonormal systems. In [15], holomorphic and pluriharmonic functions for classical domains of the first Cartan type were defined, and the Laplace and Hua Lo-Ken operators were studied. Moreover, the relationship was stated between these operators.

In homogeneous domains,the groups of automorphisms can be used for finding integral formulas ([2, 3]). Domains with rich automorphism groups are often realized as matrix domains ([5, 16]). They are very useful in solving various problems in theory of functions.

In this paper, we continue to develop the analysis in the future tube and move on to the study of the Lie ball. In [17, 18] it was noted that the Lie ball can be realized as a future tube. These realizations are subject of our research. We will be interested in integral formulas with holomorphic kernels in the future tube. There are two main types of formulas for restoring holomorphic functions: the Bergman formulas where integration is carried out over the entire domain and the Cauchy-Szegö formulas where integration is carried out over some set on the boundary of the domain (usually along its skeleton). This implementation turns out to be convenient for calculating the Bergman and Cauchy-Szegö kernels.

1. Realization of the Lie ball

We consider an n dimensional complex space $\mathbb{C}^n$, the set of all ordered n tuples of complex numbers $z = (z_1, z_2, \ldots, z_n)$. The domain $\Re^n_{IV}$ (the Lie ball (see [5])) consists of all n dimensional complex vectors z satisfying the conditions

$$\Re^n_{IV} = \left\{ z \in \mathbb{C}^n \colon |zz'|^2 - 2\overline{z}z' + 1 > 0,\, |zz'| < 1 \right\},$$

where z' is the transpose of a vector $z = (z_1, z_2, \ldots, z_n)$.

This domain is called the classical fourth type domain (according to E. Cartan's classification (see [19–21])) or the Lie ball. The Shilov boundary (the skeleton) $\Gamma_{\Re^n_{IV}}$ for the domain $\Re^n_{IV}$ is defined as follows:

$$\Gamma_{\Re^n_{IV}} = \left\{ z \in \mathbb{C}^n \colon \overline{z}z' = 1,\, |zz'| = 1 \right\}.$$

An unbounded domain of the form

$$\tau^+(n) = \left\{ w \in \mathbb{C}^{n+1} \colon (Imw_{n+1})^2 > (Imw_1)^2 + \cdots + (Imw_n)^2,\, Imw_{n+1} > 0 \right\}$$

is called the future tube in $\mathbb{C}^{n+1}$. The boundary $\partial\tau^+(n)$ of the domain $\tau^+(n)$ is defined as

$$\partial\tau^+(n) = \left\{ w \in \mathbb{C}^{n+1} \colon (Imw_{n+1})^2 = (Imw_1)^2 + \cdots + (Imw_n)^2,\, Imw_{n+1} > 0 \right\}$$

and the skeleton

$$\Gamma_{\tau^+(n)} = \left\{ w \in \mathbb{C}^{n+1} \colon\ Imw_1 = \dots = Imw_n = Imw_{n+1} = 0 \right\} = \mathbb{R}^{n+1},$$

on which the boundary degenerates.

The following statement is true.

Lemma 1. *The map* $\Phi : \mathbb{C}^n_z \to \mathbb{C}^n_w$ *defined by the equalities*

$$w_k = \frac{-2iz_k}{\sum\limits_{j=1}^{n-1} z_j^2 + (z_n - i)^2}, \ \ k = 1, \dots, (n-1), \quad w_n = \frac{2(z_n - i)}{\sum\limits_{j=1}^{n-1} z_j^2 + (z_n - i)^2} - i, \tag{1}$$

maps biholomorphically the domain $\Re^n_{IV}$ *onto* $\tau^+(n-1)$*, while* $\Gamma_{\Re^n_{IV}}$ *goes over to* $\Gamma_{\tau^+(n-1)}$.

We call the transformation (1) "*the generalized Cayley transform*". Then, from (1) we can find the inverse map $\Psi = \Phi^{-1} \colon \mathbb{C}^n_w \to \mathbb{C}^n_z$, which is defined as

$$z_k = \frac{-2iw_k}{\sum\limits_{j=1}^{n-1} w_j^2 - (w_n + i)^2}, \ \ k = 1, \dots, (n-1), \quad z_n = i - \frac{2(w_n + i)}{\sum\limits_{j=1}^{n-1} w_j^2 - (w_n + i)^2}. \tag{2}$$

Now we calculate the Jacobians of the transformation (1) and (2). For this purpose we denote

$$W = \sum_{k=1}^{n-1} w_k^2 - (w_n + i)^2 \ \text{ and } \ Z = \sum_{k=1}^{n-1} z_k^2 + (z_n - i)^2.$$

Lemma 2. *The Jacobians of the transformation* Φ *of the form (1) and* Φ^{-1} *of the form (2) are given by the next formulas respectively*

$$J_{\mathbb{C}}\Phi(z) = 2^n (-i)^{n+1} Z^{-n}$$

and

$$J_{\mathbb{C}}\Phi^{-1}(z) = -2^n (-i)^{n+1} W^{-n}.$$

2. Integral representation in the domain $\tau^+(n-1)$

We denote by dV the normalized Lebesgue measure in the domain $D \subset \mathbb{C}^n$ and define the Bergman space

$$A^2(D) = \left\{ f \in \mathcal{O}(D) : \int_D |f(z)|^2 \, dV(z) < \infty \right\}.$$

The inner product in the Bergman space is defined as:

$$\langle f, g \rangle = \int_D f(z) \overline{g(z)} dV(z).$$

Let the Bergman kernel $K_{\tau^+(n)}(w, \xi)$ of the domain $\tau^+(n)$ has the form [17]:

$$K_{\tau^+(n)}(w, \xi) = \frac{2^n (n+1)!}{\pi^{n+1} \left[\left(\frac{w - \bar{\xi}}{i} \right)^2 \right]^{n+1}}, \ \ w, \xi \in \tau^+(n), \tag{3}$$

where $\left(\frac{w-\bar{\xi}}{i}\right)^2 = \left[\left(w_1-\bar{\xi}_1\right)^2 + \dots + \left(w_{n-1}-\bar{\xi}_{n-1}\right)^2 - \left(w_n-\bar{\xi}_n\right)^2\right]$. If we denote

$$\Delta(y) := y_n^2 - (y')^2 = y_n^2 - \left(y_2^2 + y_3^2 + \ldots + y_{n-1}^2\right),$$

then the relation (3) can be written as

$$K_{\tau^+(n)}(w,\xi) = \frac{2^n (n+1)!}{\pi^{n+1}\Delta^{n+1}\left(\frac{w-\bar{\xi}}{i}\right)}, \quad w,\xi \in \tau^+(n).$$

It is known that the Bergman kernel for the Lie ball $\Re_{IV}^n$ has the form

$$K_{\Re_{IV}^n}(z,\zeta) = \frac{1}{V\left(\Re_{IV}^n\right)} \cdot \frac{1}{\left(1-2z\bar{\zeta}'+zz'\overline{\zeta\zeta'}\right)^n}, \tag{4}$$

where $V\left(\Re_{IV}^n\right) = \dfrac{\pi^n}{2^{n-1}n!}$ is the volume of the Lie ball $\Re_{IV}^n$ (see [5]).

We denote by $d\mu$, $d\nu$ and $d\eta$, $d\sigma$ the normalized Lebesgue measures in the domains $\tau^+(n-1)$, the Lie ball $\Re_{IV}^n$ and on the skeletons $\Gamma_{\tau^+(n)}$, $\Gamma_{\Re_{IV}^n}$, respectively.

Lemma 3. *Let $w = \Phi(z)$, $\xi = \Phi(\zeta)$. Then by the mapping (1) the Bergman kernel $K_{\tau^+(n-1)}(w,\xi)$ transforms as follows*

$$K_{\tau^+(n-1)}(\Phi(z),\Phi(\zeta)) = \frac{1}{4^n}\left[Z\overline{\Upsilon}\right]^n \cdot K_{\Re_{IV}^n}(z,\zeta), \tag{5}$$

where

$$Z = \sum_{k=1}^{n-1} z_k^2 + (z_n - i)^2, \quad \Upsilon = \sum_{k=1}^{n-1} \zeta_k^2 + (\zeta_n - i)^2.$$

Proof. Let $\Phi : \Re_{IV}^n \mapsto \tau^+(n-1)$ be biholomorphic and $\varphi \in A^2\left(\Re_{IV}^n\right)$. Then by replacing the variable $\zeta = \Phi^{-1}(\xi)$, we have

$$\begin{aligned}
&\int\limits_{\Re_{IV}^n} J_{\mathbb{C}}\Phi(z)\, K_{\tau^+(n-1)}(\Phi(z),\Phi(\zeta))\,\overline{J_{\mathbb{C}}\Phi(\zeta)}\varphi(\zeta)\,d\nu(\zeta) = \\
&= \int\limits_{\tau^+(n-1)} J_{\mathbb{C}}\Phi(z)\, K_{\tau^+(n-1)}(\Phi(z),\xi)\,\overline{J_{\mathbb{C}}\Phi\left(\Phi^{-1}(\xi)\right)}\varphi\left(\Phi^{-1}(\xi)\right) J_{\mathbb{R}}\Phi^{-1}(\xi)\,d\mu(\xi) = \\
&= J_{\mathbb{C}}\Phi(z) \int\limits_{\tau^+(n-1)} K_{\tau^+(n-1)}(\Phi(z),\xi)\,\overline{J_{\mathbb{C}}\Phi\left(\Phi^{-1}(\xi)\right)}\varphi\left(\Phi^{-1}(\xi)\right) \frac{1}{J_{\mathbb{R}}\Phi\left(\Phi^{-1}(\xi)\right)}d\mu(\xi).
\end{aligned} \tag{6}$$

By the Jacobian property ($J_{\mathbb{R}}\Phi = |J_{\mathbb{C}}\Phi|^2$) the last integral in (6) has the form:

$$\begin{aligned}
&J_{\mathbb{C}}\Phi(z) \int\limits_{\tau^+(n-1)} K_{\tau^+(n-1)}(\Phi(z),\xi)\,\overline{J_{\mathbb{C}}\Phi\left(\Phi^{-1}(\xi)\right)}\varphi\left(\Phi^{-1}(\xi)\right) \frac{1}{\left|J_{\mathbb{C}}\Phi\left(\Phi^{-1}(\xi)\right)\right|^2}d\mu(\xi) = \\
&= J_{\mathbb{C}}\Phi(z) \int\limits_{\tau^+(n-1)} K_{\tau^+(n-1)}(\Phi(z),\xi)\left[\left(J_{\mathbb{C}}\Phi\left(\Phi^{-1}(\xi)\right)\right)^{-1}\varphi\left(\Phi^{-1}(\xi)\right)\right]d\mu(\xi).
\end{aligned}$$

After changing variables, we can see that the expression $\left(J_{\mathbb{C}}\Phi\left(\Phi^{-1}(\xi)\right)\right)^{-1}\varphi\left(\Phi^{-1}(\xi)\right)$ in square brackets in the last integrand is an element of the space $A^2\left(\tau^+(n-1)\right)$. Applying the reproducing property of $K_{\tau^+(n-1)}$, we have

$$J_{\mathbb{C}}\Phi(z)\left(J_{\mathbb{C}}\Phi(z)\right)^{-1}\varphi\left(\Phi^{-1}(\Phi(z))\right) = \varphi(z).$$

From here it follows "*variable replacement formula*" for the Bergman kernels:

$$J_{\mathbb{C}}\Phi\left(z\right)K_{\tau^+(n-1)}\left(\Phi\left(z\right),\Phi\left(\zeta\right)\right)\overline{J_{\mathbb{C}}\Phi\left(\zeta\right)} = K_{\Re^n_{IV}}\left(z,\zeta\right).$$

Then

$$K_{\tau^+(n-1)}(\Phi(z),\Phi(\zeta)) = \left[J_{\mathbb{C}}\Phi t(z)\right]^{-1}\cdot K_{\Re^n_{IV}}(z,\zeta)\cdot\left[\overline{J_{\mathbb{C}}\Phi(\zeta)}\right]^{-1} =$$

$$= 2^{-n}(-i)^{-n-1}Z^n\cdot 2^{-n}(i)^{-n-1}\overline{\Upsilon}^n K_{\Re^n_{IV}}(z,\zeta) = \frac{1}{4^n}\left[Z\overline{\Upsilon}\right]^n\cdot K_{\Re^n_{IV}}(z,\zeta).$$

The lemma is proved □

In particular, when $n=1$, from formulas (4) and (5), we have

$$K_{\tau^+(0)}(w,\xi) = \frac{1}{\pi\Delta\left(\frac{w-\bar{\xi}}{i}\right)} = -\frac{1}{\pi\left(w-\bar{\xi}\right)^2} - \frac{1}{\pi\left(-i\frac{z+i}{z-i}-i\frac{\bar{\zeta}-i}{\bar{\zeta}+i}\right)^2} = \frac{(z-i)^2\left(\bar{\zeta}+i\right)^2}{4\pi\left(1-z\bar{\zeta}\right)^2}.$$

On the other hand

$$K_{\tau^+(0)}(w,\xi) = K_{\Re^1_{IV}}\left(z,\zeta\right)\frac{Z}{2}\frac{\overline{\Upsilon}}{2} = \frac{1}{\pi\left(1-z\bar{\zeta}\right)^2}\,\frac{(z-i)^2\left(\overline{\zeta-i}\right)^2}{4} = \frac{(z-i)^2\left(\bar{\zeta}+i\right)^2}{4\pi\left(1-z\bar{\zeta}\right)^2}.$$

Let $w=\Phi(z)$, $\xi=\Phi(\zeta)$. We have the following theorem

Theorem 1. *For any function $f\in A^2\left(\tau^+\left(n-1\right)\right)$ the formula holds:*

$$f(w) = \int\limits_{\tau^+(n-1)} f(\xi)K_{\tau^+(n-1)}(w,\xi)d\mu(\xi),\quad w\in\tau^+\left(n-1\right).$$

The integral in this formula defines an orthogonal projector of the space $L^2(\tau^+\left(n-1\right))$ into the space $A^2\left(\tau^+\left(n-1\right)\right)$.

Proof. Using the change of variables, according to Lemma 3 and the Jacobian properties we have:

$$\int\limits_{\tau^+(n-1)} f(\xi)K_{\tau^+(n-1)}(w,\xi)d\mu(\xi) = \frac{1}{4^n}\int\limits_{\Re^n_{IV}} f(\Phi(\zeta))K_{\Re^n_{IV}}(z,\zeta)Z^n\overline{\Upsilon}^n J_{\mathbb{R}}\Phi\left(\zeta\right)d\nu(\zeta) =$$

$$= \frac{Z^n}{2^n}\int\limits_{\Re^n_{IV}} f(\Phi(\zeta))K_{\Re^n_{IV}}(z,\zeta)\frac{\overline{\Upsilon}^n}{2^n}\left|J_{\mathbb{C}}\Phi\left(\zeta\right)\right|^2 d\nu(\zeta) =$$

$$= Z^n\int\limits_{\Re^n_{IV}}\frac{f(\Phi(\zeta))}{\Upsilon^n}K_{\Re^n_{IV}}(z,\zeta)d\nu(\zeta).$$

The last integral is the Bergman integral in the Lie ball $\Re^n_{IV}$ of the function $\dfrac{f(\Phi(\zeta))}{\Upsilon^n}$ and it is equal to $\dfrac{f(\Phi(z))}{Z^n}$. So, we obtain the first statement of the theorem.

Any function $g\in L^2(\tau^+\left(n-1\right))$ can be represented as $g=f+h$, where $f\in A^2(\tau^+\left(n-1\right))$ and $h\in A^{2^\perp}(\tau^+\left(n-1\right))$ are orthogonal function:

$$\int\limits_{\tau^+(n-1)} f(\xi)\overline{h(\xi)}d\mu(\xi) = 0.$$

We must show that

$$\int\limits_{\tau^+(n-1)} h(\xi)K_{\tau^+(n-1)}(w,\xi)d\mu(\xi)=0.$$

So,

$$\int\limits_{\tau^+(n-1)} f(\xi)\overline{h(\xi)}d\mu(\xi)=\int\limits_{\Re^n_{IV}} f(\Phi(\zeta))\overline{h(\Phi(\zeta))}\left|2^n\left(-i\right)^{n+1}\Upsilon^{-n}\right|^2 d\nu(\zeta)=$$

$$=2^n\int\limits_{\Re^n_{IV}} f(\Phi(\zeta))\Upsilon^{-n}\overline{h(\Phi(\zeta))\Upsilon^{-n}}d\nu(\zeta)=0.$$

Hence, $h(\Phi(\zeta))\Upsilon^{-n}\in A^{2^{\perp}}(\Re^n_{IV})$, i.e.

$$\int\limits_{\Re^n_{IV}} h(\Phi(\zeta))K_{\Re^n_{IV}}(z,\zeta)\Upsilon^{-n}d\nu(\zeta)=0.$$

Then

$$\int\limits_{\tau^+(n-1)} h(\xi)K_{\tau^+(n-1)}(w,\xi)d\mu(\xi)=\frac{Z^n}{4^n}\int\limits_{\Re^n_{IV}} h(\Phi(\zeta))\overline{\Upsilon}^n K_{\Re^n_{IV}}(z,\zeta)\left|2^n\left(-i\right)^{n+1}\Upsilon^{-n}\right|^2 d\nu(\zeta)=$$

$$=Z^n\int\limits_{\Re^n_{IV}} h(\Phi(\zeta))\Upsilon^n K_{\Re^n_{IV}}(z,\zeta)d\nu(\zeta)=0.$$

The theorem is completely proved. □

We define the Cauchy-Szegö kernel $C_{\tau^+(n-1)}(w,\xi)$ as follows (see [22])

$$C_{\tau^+(n-1)}(w,\xi)=\frac{2^n\pi^{\frac{n-1}{2}}\Gamma\left(\frac{n+1}{2}\right)}{\Delta^{\frac{n+1}{2}}\left(w-\xi\right)},$$

for $w\in\tau^+\left(n\right),\ \xi\in\Gamma_{\tau^+(n)}$.

The kernel $C_{\tau^+(n-1)}(w,\xi)$ is a holomorphic function in w and antiholomorphic in ξ.

The proof of the following lemma is similar to the proof of Lemma 1.

Lemma 4. *Let $w=\Phi(z),\ \ \xi=\Phi(\zeta)$. By mapping (1), the Cauchy-Szegö kernel $C_{\tau^+(n-1)}(w,\xi)$ transforms in the following way*

$$C_{\tau^+(n-1)}\left(\Phi\left(z\right),\ \Phi\left(\zeta\right)\right)=\frac{1}{2^n}Z^{\frac{n}{2}}\bar{\Upsilon}^{\frac{n}{2}}C_{\Re^n_{IV}}\left(z,\zeta\right),$$

where $C_{\Re^n_{IV}}(z,\zeta)$ is the Cauchy-Szegö kernel for the Lie ball $\Re^n_{IV}$ (see [19]).

$$C_{\Re^n_{IV}}(z,\zeta)=\frac{1}{V\left(\Gamma_{\Re^n_{IV}}\right)\left[\left(x-e^{-i\varphi}z\right)\left(x-e^{-i\varphi}z\right)'\right]^{\frac{n}{2}}},$$

$$\zeta=e^{i\varphi}x,\ \ x\in\mathbb{R}^n,\ \ xx'=1,\ \ \varphi\in[0;\,2\pi]\,.$$

Proof. According to Jacobian property (see [5]) for the Cauchy-Szegö kernel we have

$$\left[C_{\Re^n_{IV}}\left(z,\zeta\right)\right]^2=\left[C_{\tau^+(n-1)}\left(\Phi\left(z\right),\Phi\left(\zeta\right)\right)\right]^2 J_{\mathbb{C}}\Phi\left(z\right)\overline{J_{\mathbb{C}}\Phi\left(\zeta\right)},$$

it follows that

$$C_{\tau^+(n-1)}\left(\Phi\left(z\right),\Phi\left(\zeta\right)\right)=\frac{1}{2^n}Z^{\frac{n}{2}}\bar{\Upsilon}^{\frac{n}{2}}C_{\Re^n_{IV}}\left(z,\zeta\right)=\frac{Z^{\frac{n}{2}}\bar{\Upsilon}^{\frac{n}{2}}\Gamma\left(\frac{n}{2}\right)}{2^{n-1}\pi^{\frac{n+2}{2}}\left[\left(x-e^{-i\varphi}z\right)\left(x-e^{-i\varphi}z\right)'\right]^{\frac{n}{2}}}.$$

Theorem 2. *For any function* $f\in H^1(\tau^+(n-1))$ *the formula holds*‡

$$f(w)=\int\limits_{\Gamma_{\tau^+(n-1)}}f(\xi)C_{\tau^+(n-1)}(w,\xi)d\eta(\xi),\ \ w\in\tau^+\left(n-1\right).$$

Proof. It is known that the Poisson kernel for the domain $\tau^+(n)$ (see [17]) can be written in the form

$$P_{\tau^+(n)}\left(w,\xi\right)=\frac{2^n\Gamma\left(\frac{n+1}{2}\right)}{\pi^{\frac{n+3}{2}}}\frac{\Delta^{\frac{n}{2}}\left(Imw\right)}{\left|\Delta^n\left(w-\xi\right)\right|},\ \ w\in\tau^+\left(n\right),\ \ \xi\in\Gamma_{\tau^+(n)}.$$

By Lemma 3.4 from [23] we can get

$$P_{\tau^+(n-1)}\left(\Phi(z),\Phi(\zeta)\right)d\eta(\Phi(\zeta))=P_{\Re^n_{IV}}(z,\zeta)d\sigma(\zeta), \tag{7}$$

where

$$P_{\Re^n_{IV}}\left(z,\zeta\right)=\frac{\left(1+\left|(z,z)\right|^2-2\left|z\right|^2\right)^{\frac{n}{2}}}{\left|(z-\zeta,z-\zeta)\right|^n},\ z\in\Re^n_{IV},\ \zeta\in\Gamma_{\Re^n_{IV}},$$

is the Poisson kernel for the Lie ball.

On the other hand, due to the relation between the Cauchy-Szegö and Poisson kernels (see [23]) we have

$$P(w,\xi)=\frac{C(w,\xi)C(\xi,w)}{C(w,w)}=\frac{\left|C(w,\xi)\right|^2}{C(w,w)},$$

and by Lemma 2 we get that

$$P_{\tau^+(n-1)}\left(\Phi\left(z\right),\Phi\left(\zeta\right)\right)=\frac{\left|C_{\tau^+(n-1)}(w,\xi)\right|^2}{C_{\tau^+(n-1)}(w,w)}=\frac{\frac{1}{4^n}\left|Z\right|^n\left|\Upsilon^n\right|}{\frac{1}{2^n}\left|Z\right|^n}\frac{\left|C_{\Re^n_{IV}}(w,\xi)\right|^2}{C_{\Re^n_{IV}}(w,w)}=$$

$$=\frac{1}{2^n}\left|\Upsilon^n\right|P_{\Re^n_{IV}}\left(z,\,\zeta\right).$$

From that, we get

$$P_{\tau^+(n-1)}\left(\Phi\left(z\right),\Phi\left(\zeta\right)\right)=\frac{1}{2^n}\left|\Upsilon^n\right|P_{\Re^n_{IV}}\left(z,\,\zeta\right). \tag{8}$$

Now dividing the relation (7) by (8), we obtain

$$d\eta\left(\Phi\left(\zeta\right)\right)=2^n\left|\Upsilon^{-n}\right|d\sigma\left(\zeta\right).$$

Further on, after changing variable $\xi=\Phi(\zeta)$ and taking into account Lemma 2, we have

$$\int\limits_{\Gamma_{\tau^+(n-1)}}f(\xi)C_{\tau^+(n-1)}(w,\xi)d\eta(\xi)=Z^{\frac{n}{2}}\int\limits_{\Gamma_{\Re^n_{IV}}}\frac{f(\Phi(\zeta))}{\left|\Upsilon^n\right|}\overline{\Upsilon}^{\frac{n}{2}}C_{\Re^n_{IV}}(z,\zeta)d\sigma(\zeta)=$$

$$=Z^{\frac{n}{2}}\int\limits_{\Gamma_{\Re^n_{IV}}}\frac{f(\Phi(\zeta))}{\Upsilon^{\frac{n}{2}}}C_{\Re^n_{IV}}(z,\zeta)d\sigma(\zeta).$$

‡The Hardy class $H^1(D)$ is defined as follows: a function f holomorphic in D belongs to $H^1(D)$, if $\sup\limits_{0\leqslant r<1}\int\limits_{S(D)}\left|f\left(r\xi\right)\right|d\eta<\infty$, where η is the Lebesgue measure on the skeleton $S\left(D\right)$.

The last integral is the Cauchy-Szegö integral in the Lie ball of functions $\dfrac{f(\Phi(\zeta))}{\Upsilon^{\frac{n}{2}}}$, and it is equal to $\dfrac{f(\Phi(z))}{Z^{\frac{n}{2}}}$. It gives us the statement of the theorem.

Acknowledgement. OT-F4-(37+29) Functional properties of A-analytical functions and their applications. Some problems of complex analysis in matrix domains (2017–2021 y.) Ministry of Innovative Development of the Republic of Uzbekistan.

References

[1] B.V.Shabat, Introduction to Complex Analysis Part II Functions of Several Variables, Moscow, Nauka, Fiz. Mat. Lit., 1985 (in Russian).

[2] L.A.Aizenberg, A.P.Yuzhakov, Integral Representations and Residues in Multidimensional Complex Analysis, Transl. Math. Monogr., Vol. 58, Am. Math. Soc., Providence, 1983.

[3] L.A.Aizenberg, Carleman Formulas in complex analysis, Novosibirsk, Nauka, 1990 (in Russian).

[4] G.Khudayberganov, A.M.Kytmanov, B.A.Shaimkulov, Analysis in matrix domains, Monograph. Krasnoyarsk, Siberian Federal University, 2017 (in Russian).

[5] L.K.Hua, Harmonic analysis of functions of several complex variables, in classical domains, Moscow, IL, 1963 (in Russian).

[6] G.Khudayberganov, B.Kurbanov, Analysis in matrix domains, *Uzbek Mathematical Journal*, 2012, no. 3, 159–166

[7] G.Khudayberganov, B.P.Otemuratov, U.S.Rakhmonov, Boundary Morera theorem for the matrix ball of the third type, *Journal of Siberian Federal University. Mathematics* & *Physics*, **11**(2018), no. 1, 40–45.

[8] G.Khudayberganov, U.S.Rakhmonov, Z.Q.Matyakubov, Integral formulas for some matrix domains, *Contemporary Mathematics, AMS*, **662**(2016), 89–95.

[9] G.Khudayberganov, U.S.Rakhmonov, The Bergman and Cauchy-Szegő kernels for matrix ball of the second type, *Journal of Siberian Federal University. Mathematics* & *Physics*, **7**(2014), no. 3, 305–310.

[10] G.Khudayberganov, U.S.Rakhmonov, Carleman Formula for Matrix Ball of the Third Type, Algebra, Complex Analysis, and Pluripotential Theory. USUZCAMP 2017, Springer Proceedings in Mathematics & Statistics, vol. 264, Springer, Cham, 2017, 101–108.

[11] S.G.Myslivets, *Journal of Siberian Federal University. Mathematics* & *Physics*, **11**(2018), no. 6, 792–795. DOI: 10.17516/1997-1397-2018-11-6-792-795

[12] W.Rudin, Function Theory in the Unit Ball of $\mathbb{C}^n$, New York, Berlin, Heidelberg, Springer-Verlag, 1980.

[13] S.G.Myslivets, *Russ Math.*, **63**(2019), 35–41. DOI: 10.3103/S1066369X19010043

[14] U.S.Rakhmonov, J.Sh.Abdullayev, On volumes of matrix ball of third type and generalized Lie balls, *The Bulletin of Udmurt University. Mathematics. Mechanics. Computer Science*, **29**(2019), no. 4, 548–557.

[15] G.Khudayberganov, A.M.Khalknazarov, J.Sh.Abdullayev, *Russian Mathematics*, **64**(2020), no. 3, 66–71. DOI: 10.3103/S1066369X20030068

[16] G.Khudaiberganov, B.B.Khidirov, U.S.Rakhmonov, Automorphisms of matrix balls, *Acta NUUz*, 2010, no. 3, 205–210.

[17] V.S.Vladimirov, A.G.Sergeev, Complex analysis in the future tube, Several complex variables. II: Function theory in classical domains. Complex potential theory, Encycl. Math. Sci., Vol. 8, 1994, 179–253.

[18] G.Khudayberganov, J.Sh.Abdullayev, About one realization Lie ball, *Contemporary Mathematics. Fundamental Directions* (to appear).

[19] É.Cartan, Sur les domaines bornes homogenes de l 'espace de n variables complexes, *Abh. Math. Sem. Univ. Hamburg*, **11**(1935), 116–162.

[20] C.L.Siegel, Automorphic functions of several complex variables, Moscow, Inostrannaya Literatura, 1954 (in Russian).

[21] I.I.Pjateckiĭ-Šapiro, Geometry of classical domains and the theory of automorphic functions, Moscow, Fizmatgiz, 1961 (in Russian).

[22] V.S.Vladimirov, Generalized Functions in Mathematical Physics, Translated from 2nd. Russian ed., Mir, Moscow, 1979.

[23] A.Koranyi, The Poisson integral for generalized half-planes and bounded symmetric domains, *Ann. Math.*, **82**(1965), no. 2, 332–350.

Связь между ядрами Бергмана и Коши-Сеге в областях $\tau^+(n-1)$ и $\Re_{IV}^n$

Гулмирза Х. Худайберганов
Жонибек Ш. Абдуллаев
Национальный университет Узбекистана
Ташкент, Узбекистан

Аннотация. В работе с использованием биголоморфной эквивалентности областей $\tau^+(n-1)$ и шара Ли $\Re_{IV}^n$ найдена связь между ядрами Бергмана и Коши-Сеге. Получены интегральные представления голоморфных функций в этих областях.

Ключевые слова: классические области, шар Ли, труба будущего, граница Шилова, Якобиан, ядро Бергмана, ядро Коши-Сеге, ядро Пуассона.

DOI: 10.17516/1997-1397-2020-13-5-568-573
УДК 517.9

On a Problem that Does not Have Basis Property of Root Vectors, Associated with a Perturbed Regular Operator of Multiple Differentiation

Nurlan S. Imanbaev*
South Kazakhstan State Pedagogical University
Shymkent, Kazakhstan
Institute of Mathematics and Mathematical Modeling
Almaty, Kazakhstan

Received 18.05.2020, received in revised form 30.06.2020, accepted 04.07.2020

Abstract. A spectral problem for a multiple differentiation operator with integral perturbation of boundary value conditions which are regular but not strongly regular is considered in the paper. The feature of the problem is the absence of the basis property of the system of root vectors. A characteristic determinant of the spectral problem is constructed. It is shown that absence of the basis property of the system of root functions of the problem is unstable with respect to the integral perturbation of the boundary value condition.

Keywords: multiple differentiation operator, integral perturbation of boundary value conditions, basis property, root vectors, system of eigenfunctions and associated functions, eigenvalue, characteristic determinant.

Citation: N.S. Imanbaev, On a Problem that Does not Have Basis Property of Root Vectors, Associated with a Perturbed Regular Operator of Multiple Differentiation, J. Sib. Fed. Univ. Math. Phys., 2020, 13(5), 568–573. DOI: 10.17516/1997-1397-2020-13-5-568-573.

1. Introduction and formulation of the problem

It is well known that a system of eigenfunctions of an operator defined by a formally self-adjoint differential expression with arbitrary self-adjoint boundary value conditions, that provide a discrete spectrum, forms an orthonormal basis in the space L_2. The problem of preservation of basis properties with respect to some weak (in a certain sense) perturbation of the initial operator was investigated in many papers. For example, the problem was studied in the case of a self-adjoint initial operator [1–3], and in the case of a non self-adjoint initial operator [4–6]. The following spectral problem is considered in this paper (it is close to the problems considered in [1, 4, 7])

$$l(u) \equiv -u''(x) = \lambda u(x), \quad 0 < x < 1 \tag{1}$$

$$U_1(u) \equiv u'(0) - u'(1) - \alpha u(1) = 0, \quad \alpha > 0, \tag{2}$$

$$U_2(u) \equiv u(0) = \int_0^1 \overline{p(x)} u(x) dx, \quad p(x) \in L_2(0,1). \tag{3}$$

*imanbaevnur@mail.ru https://orcid.org/0000-0002-5220-9899

Where λ is an arbitrary complex number. Stability of the basis property of root vectors of the spectral problem has was studied in the case $\alpha = 0$ [8]. It is common knowledge that system of root functions of an ordinary differential operator with arbitrary strongly regular boundary value conditions forms Riesz basis in the space $L_2(0,1)$. In the case when boundary value conditions are regular but not strongly regular, the basis property of systems of root functions, in contrast to the completeness property, is not even determined by the specific type of boundary value conditions. This effect was first established by V. A. Il'in [5], and the corresponding example was constructed for some second-order differential operator. In this case it was shown that not only boundary value conditions but also the values of coefficients of the differential operator affect the existence of the basis property. Moreover, this property can be changed with an arbitrarily small change in values of the coefficients in the metric of those classes in which these coefficients are given. Let L_1 be an operator in $L_2(0,1)$. It is defined by (1) and "perturbed" boundary value conditions

$$U_1(u) \equiv u'(0) - u'(1) - \alpha u(1) = 0, \quad \alpha > 0, \quad U_2(u) \equiv u(0) = \int_0^1 \overline{p(x)} u(x) dx, \quad p(x) \in L_2(0,1).$$

Unperturbed operator (the case $p(x) = 0$) is denoted by L_0. Various variants of the integral perturbation of boundary value conditions were considered [9–13]. It was assumed that unperturbed operator L_0 has a system of eigenfunctions and associated functions (E&AF) that forms the Riesz basis in $L_2(0,1)$. Characteristic determinant of the spectral problem for operator L_1 was also constructed. Then inferences were drawn about stability or instability of the Riesz basis property of the E&AF problem with an integral perturbation of the boundary value condition.

The distinctive feature of this paper is that system of eigenfunctions of unperturbed problem (1)– (3) is complete but it does not form a basis in $L_2(0,1)$ [14]. Therefore, the method that we use in the previous papers cannot be applied in this case.

2. Characteristic determinant of the spectral problem

Boundary value conditions in problem (1)–(3) are regular but not strongly regular [15]. System of root functions of operator l is a complete system but does not even form a usual basis in $L_2(0,1)$ [14]. However, as it was shown in [16], a basis can be constructed with the use of these eigenfunctions. This makes it possible to apply the method of separation of variables to solve the initial-boundary value problem with boundary value condition (2). The problem that does not have the basis property of root vectors in $L_2(0,1)$ defined by expression (1) with integral perturbation of boundary value condition (2) is studied [17]. The general solution of equation (1) has the form

$$u(x,\lambda) = C_1 \cos\sqrt{\lambda}x + C_2 \sin\sqrt{\lambda}x.$$

Using boundary value conditions (2)–(3), we obtain the following linear system with respect to coefficients C_k :

$$\begin{cases} C_1\left(\sqrt{\lambda}\sin\sqrt{\lambda} - \alpha\cos\sqrt{\lambda}\right) + \left(\sqrt{\lambda}\left(1-\cos\sqrt{\lambda}\right) - \alpha\sin\sqrt{\lambda}\right) = 0, \\ C_1\left(1 - \int_0^1 \overline{p(x)}\cos\sqrt{\lambda}x dx\right) + C_2\int_0^1 \overline{p(x)}\sin\sqrt{\lambda}x dx = 0. \end{cases} \tag{4}$$

Therefore, characteristic determinant of problem (1)–(3) has the form

$$\Delta_1(\lambda) = \left(\sqrt{\lambda}\sin\sqrt{\lambda} - \alpha\cos\sqrt{\lambda}\right)\cdot\int_0^1 \overline{p(x)}\sin\sqrt{\lambda}x dx -$$
$$- \left(\sqrt{\lambda}\left(1-\cos\sqrt{\lambda}\right) - \alpha\sin\sqrt{\lambda}\right)\left(1 - \int_0^1 \overline{p(x)}\cos\sqrt{\lambda}x dx\right). \quad (5)$$

If $p(x) = 0$ then the characteristic determinant of unperturbed problem (1)–(3) is obtained. It is denoted by

$$\Delta_0(\lambda) = \sqrt{\lambda}\left(1-\cos\sqrt{\lambda}\right) - \alpha\sin\sqrt{\lambda}.$$

Solving the equation $\Delta_0(\lambda) = 0$, we have two series of eigenvalues $\lambda_k^{(1)} = (2\pi k)^2$, $k = 1, 2, \ldots$, $\lambda_k^{(2)} = (2\beta_k)^2$, $k = 1, 2, \ldots$, of unperturbed problem (1)–(3). Here β_k are roots of the equation

$$\operatorname{tg}\beta = \frac{\alpha}{2\beta}, \quad \beta > 0. \quad (6)$$

They are positive and satisfy the following inequalities:
$\pi k < \beta_k < \pi k + \frac{\pi}{2}$, $k = 0, 1, 2, \ldots$.

For the difference $\delta_k = \beta_k - \pi k$ at large enough k the following two-sided estimates are true

$$\frac{\alpha}{2\pi k}\left(1 - \frac{1}{2\pi k}\right) < \delta_k < \frac{\alpha}{2\pi k}\left(1 + \frac{1}{2\pi k}\right). \quad (7)$$

When $p(x) = 0$ eigenfunctions of the unperturbed problem have the following form

$$u_k^{(1)}(x) = \sin(2\pi k x), \quad k = 1, 2, \ldots, \quad u_k^{(2)}(x) = \sin(2\beta_k x), \quad k = 1, 2, \ldots .$$

When $p(x) = 0$ let us consider the problem adjoint to unperturbed problem (1)–(3)

$$l*(\nu) = \overline{\lambda}\nu, v'(1) + \alpha v(0) = 0; \quad v(0) - v(1) = 0 \quad (8)$$

It has eigenfunctions

$$v_k^{(1)}(x) = C_k^{(1)}\left(\cos(2\pi k x) - \frac{\alpha}{2\pi k}\sin(2\pi k x)\right), \quad k = 1, 2, \ldots,$$
$$v_k^{(2)}(x) = C_k^{(2)}\left(\cos(2\beta_k x) + \frac{\alpha}{2\beta_k}\sin(2\beta_k x)\right), \quad k = 1, 2, \ldots,$$

where $C_k^{(1)}, C_k^{(2)}$ are selected from the ratio of biorthogonality $\left(u_k^{(1)}, v_k^{(1)}\right) = 1$, $\left(u_k^{(2)}, v_k^{(2)}\right) = 1$. Hence,

$$C_k^{(1)} = -\frac{4\pi}{\alpha}, \quad C_k^{(2)} = \frac{4\pi}{\alpha} + O\left(\frac{1}{k}\right). \quad (9)$$

An auxiliary system was constructed that forms the basis in $L_2(0,1)$ [16]

$$u_0(x) = u_0^{(2)}(x)\cdot(2\beta_0)^{-1}, \quad u_{2k}(x) = u_k^{(1)}(x),$$
$$u_{2k-1}(x) = \left(u_k^{(2)}(x) - u_k^{(1)}(x)\right)\cdot(2\delta_k)^{-1}, \quad k = 1, 2, \ldots .$$

The system that is biorthogonal to auxiliary system is

$$v_0(x) = 2\beta_0 v_0^{(2)}(x), \quad v_{2k}(x) = v_k^{(2)}(x) + v_k^{(1)}(x), \quad v_{2k-1}(x) = 2\delta_k v_k^{(2)}(x), \quad k = 1, 2, \ldots .$$

It is constructed from the eigenfunctions of problem (8). The function $p(x)$ can be represented in terms of Fourier series based on the auxiliary system $\{v_k(x)\}$:

$$p(x) = a_0 v_0(x) + \sum_{k=1}^{\infty} (a_k v_{2k}(x) + b_k v_{2k-1}(x)). \tag{10}$$

We integrate expression (5) and obtain

$$\int_0^1 \overline{p(x)} \sin\sqrt{\lambda} x dx = \Delta_0(\lambda) \cdot \left\{ \bar{a}_0 \cdot C_0^{(2)} \frac{2\beta_0}{\lambda - (2\beta_0)^2} + \right.$$
$$\left. + \sum_{k=1}^{\infty} \bar{a}_k \cdot C_k^{(1)} \frac{1}{\lambda - (2k\pi)^2} + \sum_{k=1}^{\infty} \frac{C_k^{(2)}}{\lambda - (2\beta_k)^2} \left(\bar{a}_k + \bar{b}_k (2\delta_k) \right) \right\},$$

where

$$\Delta_0(\lambda) = \sqrt{\lambda}\left(1 - \cos\sqrt{\lambda}\right) - \alpha \sin\sqrt{\lambda};$$

$$\int_0^1 \overline{p(x)} \cos\sqrt{\lambda} x dx = \bar{a}_0 \frac{(2\beta_0) \cdot C_0^{(2)}}{\lambda - (2\beta_0)^2} \left(\sqrt{\lambda}\sin\sqrt{\lambda} - \alpha\cos\sqrt{\lambda} \right) +$$
$$+ \sum_{k=1}^{\infty} \bar{a}_k \frac{C_k^{(1)}}{\lambda - (2k\pi)^2} \left(\alpha \left(1 - \cos\sqrt{\lambda}\right) + \sqrt{\lambda}\sin\sqrt{\lambda} \right) +$$
$$+ \sum_{k=1}^{\infty} \left(\bar{a}_k + \bar{b}_k (2\delta_k) \right) \frac{C_k^{(2)}}{\lambda - (2\beta_k)^2} \left(\sqrt{\lambda}\sin\sqrt{\lambda} - \alpha\cos\sqrt{\lambda} - \alpha \right).$$

Using obtained results, determinant (5) is reduced to the form

$$\Delta_1(\lambda) = \Delta_0(\lambda) \left[-1 + \alpha \sum_{k=1}^{\infty} \left(a_k \left(C_k^{(1)} \frac{1}{\lambda - (2k\pi)^2} - C_k^{(2)} \frac{1}{\lambda - (2\beta_k)^2} \right) - \right. \right.$$
$$- b_k \frac{2\delta_k C_k^{(2)}}{\lambda - (2\beta_k)^2} \Bigg) + 2\left(\sqrt{\lambda}\sin\sqrt{\lambda} - \alpha\cos\sqrt{\lambda} \right) \cdot \left(a_0 \frac{(2\beta_0) \cdot C_0^{(2)}}{\lambda - (2\beta_0)^2} + \right.$$
$$\left. \left. + \sum_{k=1}^{\infty} \left(a_k \left(C_k^{(1)} \frac{1}{\lambda - (2k\pi)^2} + C_k^{(2)} \frac{1}{\lambda - (2\beta_k)^2} \right) + b_k \frac{(2\delta_k) C_k^{(2)}}{\lambda - (2\beta_k)^2} \right) \right) \right], \tag{11}$$

where $\Delta_0(\lambda) = \sqrt{\lambda}\left(1 - \cos\sqrt{\lambda}\right) - \alpha \sin\sqrt{\lambda}$.

Expression in square brackets is denoted by $A(\lambda)$. The obtained above result is formulated as the following theorem.

Theorem 2.1 *Characteristic determinant of spectral problem* (1)–(3) *with perturbed boundary value conditions can be represented in form* (11)*, where* $\Delta_0(\lambda)$ *is the characteristic determinant of the unperturbed problem, and* a_k, b_k *are the Fourier coefficients of expansion* (10) *of the function* $p(x)$ *in the biorthogonal system of eigenfunctions of adjoint problem* (8).

Function $A(\lambda)$ with $\lambda = (2\beta_k)^2$ and $\lambda = (2\pi k)^2$ has first order poles. Therefore, $\Delta_1(\lambda) = \Delta_0(\lambda) \cdot A(\lambda)$ is an entire analytic function of the variable λ. Characteristic determinant (11) has the simple form when $p(x)$ is represented as a finite sum in (10). When there exists a number N such that $a_k = b_k = 0$ for all $k > N$ then function $A(\lambda)$ has the form

$$A(\lambda) = -1 + \alpha \sum_{k=1}^{N} \left[a_k \left(C_k^{(1)} \frac{1}{\lambda - (2k\pi)^2} - C_k^{(2)} \frac{1}{\lambda - (2\beta_k)^2} \right) - \right.$$
$$\left. - b_k \frac{2\delta_k C_k^{(2)}}{\lambda - (2\beta_k)^2} \right] + 2\left(\sqrt{\lambda}\sin\sqrt{\lambda} - \alpha\cos\sqrt{\lambda}\right) \cdot \left(a_0 \frac{(2\beta_0)\cdot C_0^{(2)}}{\lambda - (2\beta_0)^2} + \right.$$
$$\left. + \sum_{k=1}^{N} \left[a_k \left(C_k^{(1)} \frac{1}{\lambda - (2k\pi)^2} + C_k^{(2)} \frac{1}{\lambda - (2\beta_k)^2} \right) + b_k \frac{(2\delta_k) C_k^{(2)}}{\lambda - (2\beta_k)^2} \right] \right. . \quad (12)$$

It follows from (12) that $\Delta_1(\lambda_k^1) = \Delta_1(\lambda_k^2) = 0$, for all $k > N$. Consequently, all eigenvalues $\lambda_k^1,\ \lambda_k^2,\ k > N$ of unperturbed problem (1)–(3) ($p(x) = 0$) are eigenvalues of perturbed problem (1)–(3). Multiplicity of eigenvalues $\lambda_k^1,\ \lambda_k^2,\ k > N$ is preserved. It follows from the condition of biorthogonality of the system of eigenfunctions $\left\{u_k^{(1)}(x),\ u_k^{(2)}(x)\right\}$ and $\left\{v_k^{(1)}(x),\ v_k^{(2)}(x)\right\}$ that $\int\limits_0^1 \overline{p\,(x)} u_k^{(1)}(x)dx = 0,\ \int\limits_0^1 \overline{p\,(x)} u_k^{(2)}(x)dx = 0,\ k > N$. Thus, eigenfunctions $\left\{u_k^{(1)}(x),\ u_k^{(2)}(x)\right\}$, $k > N$ of unperturbed problem ($p(x) = 0$) (1)–(3) satisfy the boundary value conditions of perturbed problem (1)–(3). Therefore, the system of eigenfunctions of perturbed problem (1)–(3) and the system of eigenfunctions of the unperturbed problem ($p(x) = 0$) coincide which do not form a basis with the exception of a finite number of the first terms. Then the system of eigenfunctions of perturbed problem (1)–(3) is also not a basis in $L_2(0, 1)$ in this particular case.

References

[1] A.S.Makin, *Differential Equations*, **42**(2006), no. 4, 599–602. DOI: 10.1134/S0012266106040185

[2] A.S.Markus, On decomposition of a weakly perturbedself-adjoint operator by root vectors, *Reports of the AS USSR*, **142**(1962), no. 3, 538–541 (in Russian).

[3] N.B.Kerimov, Kh.R.Mamedov, On the Riesz basis property of root functions of some regular boundary value problems, *Math. Notes.*, **64**(1998), 483–487.

[4] A.A.Shkalikov, On basis property of eigenfunctions of ordinary differential operators with integral boundary value conditions, *Vestnik MGU*, (1982), no. 6, 12-21 (in Russian).

[5] V.A.Ilyin, L.V.Kritskov, *Journal of Mathematical Sciences*, **116**(2003), 3489–3550. DOI: 10.1023/A:1024180807502

[6] N.S.Imanbaev, M.A.Sadybekov, Basis properties of root functions of loaded differential operators of the second order, *Reports of the NAS RK*, (2010), no. 2, 11–13 (in Russian).

[7] N.S.Imanbaev, M.A. Sadybekov, Stability of basis property of a type of problems with nonlocal perturbation of boundary conditions, AIP Conf. Proc., Vol. 1759, ICAAM, 2016, 020034–020039.

[8] N.S.Imanbaev, *International Journal of Differential Equations*, **2015**(2015), Article ID 641481. DOI: 10.1155/2015/641481

[9] N.S.Imanbaev, M.A. Sadybekov, On instability of basis property of root vectors system of the double differentiation operator with an integral perturbation of periodic type conditions, *AIP Publishing: Advancements in Mathematical Sciences*, **1676**(2015), 020083–020088.

[10] M.A.Sadybekov, N.S.Imanbaev, *Differential Equations*, **48**(2012), no. 6, 896–900. DOI: 10.1134/S0012266112060146

[11] M.A. Sadybekov, N.S.Imanbaev, On spectral properties of a periodic problem with an integral perturbation of the boundary condition, *Eurasian Mathematical Journal*, **4**(2013), no. 3, 53–62.

[12] N.S.Imanbaev, Stability of the basis property of eigenvalue systems of Sturm-Liouville operators with integral boundary condition, *ElectronicJournal of Differential Equations*, **2016**(2016), no. 87, 1–8.

[13] M.A.Sadybekov, N.S.Imanbaev, Regular differential operator with a perturbed boundary condition, *Mathematical Notes*, **101**(2017), no. 5, 878–887.

[14] P.Lang, J.Locker, Spectral theory of two-point differential operators determined by -D2, *Journal Math. Anal.Appl.*, **146**(1990), no. 1, 148–191.

[15] M.A.Naimark, Linear Differential Operators, George G. Harrap & Company, London, 1967.

[16] A.Yu.Mokin, *Differential equations*, **45**(2009), no. 1, 123–137. DOI: 10.1134/S0012266109010133

[17] M.A.Sadybekov, N.S.Imanbaev, Characteristic determinant of a boundary value problem, which does not have the basis property, *Eurasian Math. J.*, **8**(2017), no. 2, 40–46.

Об одной задаче, не обладающей свойством базисности корневых векторов, связанной с возмущенным регулярным оператором кратного дифференцирования

Нурлан С. Иманбаев
Южно-Казахстанский государственный педагогический университет
Чимкент, Казахстан
Институт математики и математического моделирования
Алма-Ата, Казахстан

Аннотация. В работе рассматривается спектральная задача для оператора кратного дифференцирования при интегральном возмущении краевых условий одного типа, являющихся регулярными, но не усиленно регулярными, где особенностью задачи считают отсутствие свойства базисности у системы корневых векторов. Построен характеристический определитель спектральной задачи. Показано, что отсутствие базисности у системы корневых функций задачи неустойчиво относительно интегрального возмущения краевого условия.

Ключевые слова: оператор кратного дифференцирования, интегральное возмущение краевых условий, базисность, корневые векторы, система собственных и присоединенных функций, собственное значение, характеристический определитель.

Journal of Siberian Federal University. Mathematics & Physics 2020, 13(5), 574–582

DOI: 10.17516/1997-1397-2020-13-5-574-582
УДК 517.9

Layered Motion of Two Immiscible Liquids with a Free Boundary

Elena N. Lemeshkova*
Institute of computational modelling SB RAS
Krasnoyarsk, Russian Federation

Received 26.05.2020, received in revised form 06.06.2020, accepted 17.07.2020

Abstract. The unidirectional motion of two viscous immiscible incompressible liquids in a flat channel is studied. An unsteady temperature gradient is set on the bottom solid wall, and the upper wall is a free boundary. Liquids contact on a flat interface. The motion is caused by the combined action of thermogravitational and thermocapillary forces and a given total unsteady flow rate in the layers. The corresponding initial boundary value problem is conjugate and inverse, since the pressure gradient along the channel is determined together with the velocity and temperature field. An exact stationary solution was found for it. In Laplace images, the solution of the non-stationary problem is found in the quadrature forms. It was established that if the temperature on the bottom wall and the flow rate stabilize with time, then the motion goes to a stationary state with time. This fact indicates the stability of the stationary solution with respect to unidirectional unsteady perturbations. The calculation results showing various methods of controlling motion by setting the temperature on the wall are given.

Keywords: thermocapillary, interface, Oberbek–Boussinesq equations.

Citation: E.N. Lemeshkova, Layered Motion of Two Immiscible Liquids with a Free Boundary, J. Sib. Fed. Univ. Math. Phys., 2020, 13(5), 574–582. DOI: 10.17516/1997-1397-2020-13-5-574-582.

It is well known that in a nonuniformly heated pure liquid medium there occurs a motion which is called convection. Under the condition close to zero gravity, the inhomogeneity of the temperature field affects the flow from the region of little to the region of large surface tension due to a spontaneous decrease of the interface free energy between two liquid media or the free boundary (Marangoni effect, see the detailed review in [1]). Observation of thermocapillary motion due to interfacial tension gradients in terrestrial conditions is very difficult, since gravitational convection becomes the dominant form of motion. However, in recent years in connection with the development of modern technologies, new problems have arisen in which it is necessary to take into account the thermocapillary effect in terrestrial conditions. For example, in laser annealing of semiconductors or in laser processing of materials with fusion, which is used to alloy the surface layer of metal [2]. In this case, relatively long thin layers of melt (of the order of several micrometers) appear on the surface of materials, in which, according to [3, 4], thermocapillary forces dominate over gravitational forces. Knowledge of the laws of thermal convection in liquid layered systems is of interest for understanding the hydrodynamics and heat and mass transfer processes when applying multilayer coatings, in thermal stabilization systems of power plants or cooling electronic devices, in the processes of growing single crystals and films, etc. [5]. It is known that the solution of R.V. Birikh [6] describes a stationary convective flow in a strip. For the first time, its generalization to the case when the longitudinal temperature

*elena_cher@icm.krasn.ru https://orcid.org/0000-0002-9059-2876

gradient depends on time was proposed by V. V. Pukhnachev in [7], as well as [8]. A similar generalization the solution of the problem of the motion of two immiscible liquids with a flat interface allows [9]. In the monograph [10], the problems of unidirectional motion of a viscous heat-conducting fluid in a flat horizontal channel in a field of gravity acting across the channel are studied. At the same time, an unsteady temperature field is set on a fixed solid wall, and the upper boundary can be a solid wall, either insulated or not, or a free boundary. The present work is devoted to the study of the unidirectional thermogravitational motion of two immiscible viscous liquids in a flat horizontal channel. In this case, an unsteady temperature field, which is linear in the longitudinal coordinate, is set on the lower solid wall and the upper wall is a free boundary. The system of Oberbek–Boussinesq equations is taken as a mathematical model. A detailed conclusion and analysis of the main assumptions that lead to this system are available in many works, for example, in [7].

1. Statement of problem

We consider a system of two incompressible immiscible liquids with an interface $y = 0$. The parameters of the fluid that moves in the strip $-h_1 < y < 0$, $-\infty < x < \infty$ are denoted by index "1", and the parameters of the fluid moving in the strip $0 < y < h_2$ will be denoted by index "2"; ρ_j, ν_j, μ_j, χ_j, β_j are densities, kinematic viscosities, dynamic viscosities, thermal diffusivity and volume expansion coefficients, respectively. Further on, it is assumed that these parameters are positive constants. Let x and y denote the horizontal and vertical coordinates, gravity with acceleration g acts in the negative direction of the y axis. Substitution of a solution of the form [8]

$$u_j = w_j(y,t), \quad \theta_j = -a_j(y,t)x + T_j(y,t), \quad p_j = -b_j(y,t)x + P_j(y,t), \quad j = 1,2. \tag{1}$$

into the Oberbek–Boussinesq system leads to the equations

$$\begin{aligned} &a_{jt} = \chi_j a_{jyy}, \\ &b_{jy} = \rho_j g \beta_j a_j, \\ &w_{jt} = \nu_j w_{jyy} + \frac{1}{\rho_j} b_j, \\ &T_{jt} = \chi_j T_{jyy} + a_j w_j, \\ &P_{jy} = \rho_j g \beta_j T_j. \end{aligned} \tag{2}$$

In (1) u_j is the projection of the velocity vector on the x axis, θ_j is temperature, p_j is pressure deviation from hydrostatic one. Next, we consider the problem only for determining the velocity field, that is, the problem for w_j and a_j. The functions T_j will be the solution of the conjugate problem similar to the problem for a_j. Using the known functions a_j and T_j, the functions $b_j(y,t)$ and $P_j(y,t)$ are reconstructed by quadratures from the second and fifth equations of (2)

$$\begin{aligned} b_j(y,t) &= \rho_j g \beta_j \int_0^y a_j(z,t)\,dz + C_j(t), \\ P_j(y,t) &= \rho_j g \beta_j \int_0^y T_j(z,t)\,dz + P_{0j}(t). \end{aligned} \tag{3}$$

For the system (2), we can pose various initial–boundary value problems that describe the motion in a flat channel: one liquid in a channel with solid impermeable walls, a solid wall and a free boundary, one or more interfaces, an interface and a free boundary (a part of the statements of such problems was considered in [8, 10]). In this paper, we consider the last case for a specific problem, namely, when the temperature is set on the solid lower wall $y = -h_1$, and the upper wall $y = h_2$ is free boundary. Since the interface $y = 0$ at the initial moment of time is assumed to be horizontal, at any moment of time it will be such [10, p. 41]. The large capillary pressure (Weber number We $\gg 1$) allows us to assume that the free surface $y = h_2$ also remains flat [11].

We introduce the characteristic length scales for the first and second layer, these are h_1 and h_2, for time this is $h_1^2\chi_1^{-1}$ and functions w_j, a_j, C_j these are $\chi_j h_j^{-1}$, $\tilde{a} = \max\limits_{t\geqslant 0} |a_1(-h_1, t)| > 0$, $\mu_1\chi_1 h_1^{-3}$ and write out the conjugate initial boundary value problem for the functions $a_j(y, t)$ in the dimensionless form

$$\begin{gathered} a_{1t}(y,t) = a_{1yy}(y,t), \quad -1 < y < 0, \\ a_{2t}(y,t) = h^2\chi^{-1}a_{2yy}(y,t), \quad 0 < y < 1, \end{gathered} \tag{4}$$

$$a_j(y,0) = a_{j0}(y), \tag{5}$$

$$a_1(-1,t) = a(t), \quad a_{2y}(1,t) + \mathrm{Bi}a_2(1,t) = 0, \tag{6}$$

$$a_1(0,t) = a_2(0,t), \quad ka_{1y}(0,t) = ha_{2y}(0,t). \tag{7}$$

In (4)–(7) $h = h_1/h_2$, $\chi = \chi_1/\chi_2$, $k = k_1/k_2$, $\mathrm{Bi} = \gamma h_2 k_2^{-1}$ is the Bio number, $\gamma \geqslant 0$ is the interfacial heat transfer coefficient, and functions $a_{j0}(y)$ and $a(t)$ are known. Relations (7) follow from the equality of temperatures and heat fluxes at the interface $y = 0$.

Let us pass to the formulation of the problem for dimensionless velocities $w_j(y, t)$

$$\mathrm{P}_1^{-1}w_{1t}(y,t) = w_{1yy}(y,t) + \mathrm{Ra}_1\int_0^y a_1(z,t)\,dz + C_1(t), \quad -1 < y < 0, \tag{8}$$

$$\chi h^{-2}\mathrm{P}_2^{-1}w_{2t}(y,t) = w_{2yy}(y,t) + \mathrm{Ra}_2\int_0^y a_2(z,t)\,dz + \mu\chi h^{-3}C_2(t), \quad 0 < y < 1, \tag{9}$$

$$w_j(y,0) = w_{j0}(y), \tag{10}$$

$$w_1(-1,t) = 0, \quad \chi h^{-1}w_1(0,t) = w_2(0,t). \tag{11}$$

The first condition in (11) is the sticking condition on the lower solid fixed wall $y = -1$, and the second is a consequence of the continuity of velocities at the interface $y = 0$. The dynamic condition on the interface $y = 0$ is reduced to two:

$$w_{2y}(0,t) - \mu\chi h^{-2}w_{1y}(0,t) = -\mathrm{Ma}_1 a_1(0,t), \quad p_1(0,t) = p_2(0,t). \tag{12}$$

The last condition, together with representations (1), (3), implies the equalities

$$C_1(t) = C_2(t) \equiv C(t), \quad P_1(0,t) = P_2(0,t).$$

Besides, on the free boundary $y = 1$, from the dynamic condition for tangential stresses it follows

$$w_{2y}(1,t) = \mathrm{Ma}_2 a_2(1,t). \tag{13}$$

Assuming We $\gg 1$, from the condition for normal stresses we obtain that the free surface remains flat. In equations (8), (9) and in conditions (12), (13), dimensionless parameters arise $\mu = \mu_1/\mu_2$,

$\mathrm{P}_j = \nu_j/\chi_j$ are the Prandtl numbers, $\mathrm{Ra}_j = g\beta_j \tilde{a} h_j^4/\nu_j\chi_j$ are the Rayleigh numbers, $\mathrm{Ma}_j = \text{æ}_j \tilde{a} h_2^2/\mu_2\chi_2$ are the Marangoni numbers.

If $C(t)$ is given, then the statement of the problem for velocities is completed. The aim of this work is to study the inverse problem, therefore, it is necessary to put another condition is the total consumption in layers

$$\chi \int_{-1}^{0} w_1(z,t)\,dz + \int_{0}^{1} w_2(z,t)\,dz = q(t). \tag{14}$$

2. Stationary solution

Suppose that $a(t) = a_0 = \text{const}$, $q(t) = q_0 = \text{const}$. We give the form of a stationary solution $a_j^0(y)$, $w_j^0(y)$ and $C^0 = \text{const}$

$$\begin{aligned} a_1^0(y) &= -\frac{\mathrm{Bi}a_0}{\delta}(y+1) + a_0, \quad -1 \leqslant y \leqslant 0, \\ a_2^0(y) &= -\frac{\mathrm{Bi}a_0}{\delta}\left(\frac{k}{h}y+1\right) + a_0, \quad 0 \leqslant y \leqslant 1, \end{aligned} \tag{15}$$

$$\begin{aligned} w_1^0(y) &= -\frac{a_0\mathrm{Ra}_1}{24}\left(-\frac{\mathrm{Bi}}{\delta}y^4 + 4\left(1-\frac{\mathrm{Bi}}{\delta}\right)y^3\right) - \frac{C^s}{2}y^2 + m_1 y + m_2, \quad -1 \leqslant y \leqslant 0, \\ w_2^0(y) &= -\frac{a_0\mathrm{Ra}_2}{24}\left(-\frac{k\mathrm{Bi}}{h\delta}y^4 + 4\left(1-\frac{\mathrm{Bi}}{\delta}\right)y^3\right) - \frac{\mu\chi C^s}{2h^3}y^2 + m_3 y + m_4, \quad 0 \leqslant y \leqslant 1, \\ C^s &= \delta_1 + h m_1. \end{aligned} \tag{16}$$

In (15), (16) we introduced the notation

$$\delta = \frac{(1+\mathrm{Bi})k}{h} + \mathrm{Bi},$$

$$\delta_1 = -\frac{a_0 h^3}{\mu\chi}\left[\frac{\mathrm{Ra}_2}{6}\left(3 - \frac{\mathrm{Bi}}{\delta}(3+\frac{k}{h})\right) + \mathrm{Ma}_1\left(1-\frac{\mathrm{Bi}}{\delta}\right) + \mathrm{Ma}_2\left(1 - \frac{\mathrm{Bi}}{\delta}(1+\frac{k}{h})\right)\right],$$

$$\delta_2 = \frac{a_0\mathrm{Ra}_1}{24}\left(\frac{3\mathrm{Bi}}{\delta} - 4\right) + \frac{\delta_1}{2}, \qquad \delta_3 = \frac{\chi a_0 \mathrm{Ra}_1}{120}\left(5 - \frac{4\mathrm{Bi}}{\delta}\right) - \frac{a_0\mathrm{Ra}_2}{120}\left(5 - \frac{\mathrm{Bi}}{\delta}(5+\frac{k}{h})\right),$$

$$m_1 = \frac{-6\delta_3 + \chi\delta_1(1+\mu h^{-3}) + 3a_0\mathrm{Ma}_1(1-\mathrm{Bi}\delta^{-1}) - 6\delta_2\chi(1+h^{-1}) + 6q_0}{-h\chi(1+\mu h^{-3}) - 3\chi(1-\mu h^{-2}) + 3h^{-1}\chi(h+1)(h+2)}.$$

$$m_2 = \delta_2 + \frac{m_1}{2}(h+2), \quad m_3 = \frac{\mu\chi}{h^2}m_1 - a_0\mathrm{Ma}_1\left(1-\frac{\mathrm{Bi}}{\delta}\right), \quad m_4 = \frac{\chi}{h}m_2,$$

Fig. 1 shows stationary velocity profiles $w_j^0(y)$ depending on the dimensionless flow rate q_0. Hereinafter, the function $w^0(y)$ coincides with the functions $w_j^0(y)$, $j = 1, 2$ on their domains of definition. The calculations show that with an increase in q_0, the velocity profile in the first layer becomes linear, i.e., a Couette flow arises. In the second layer, a Poiseuille flow arises (parabolic profile). With a decrease in the layers q_0, zones of the return flow arise.

Fig. 2 shows stationary velocity profiles for different values of the Bio number. It can be seen that in the second layer for $\mathrm{Bi} = 0$ (a thermally insulated free surface) the velocity profile is linear, and with an increase the Bio number it becomes parabolic. The dependence of the velocity profile in the layers on the dimensionless temperature gradient a_0 is shown in Fig. 3.

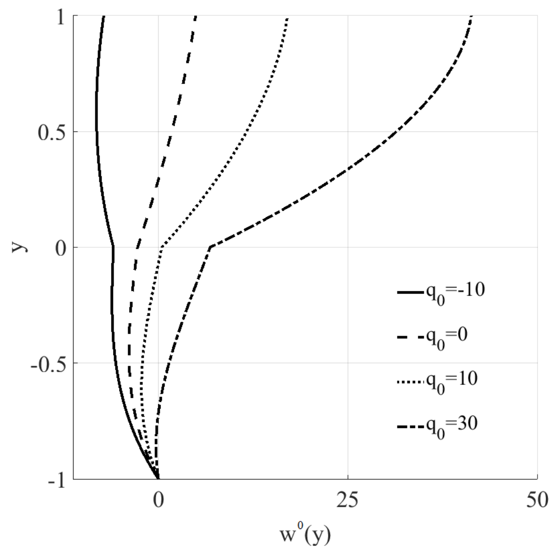

Fig. 1. Stationary profile of velocities $w_j^0(y)$ depending on the dimensionless flow rate

The case when $a_0 < 0$ means that the bottom wall is cooling. So, for $a_0 = 2$ a return flow arises near the interface $y = 0$, and for $a_0 = -2$ the direction of the current changes to the opposite. At $a_0 = 0$ in both layers the velocities have a parabolic profile and there are no zones of return flow.

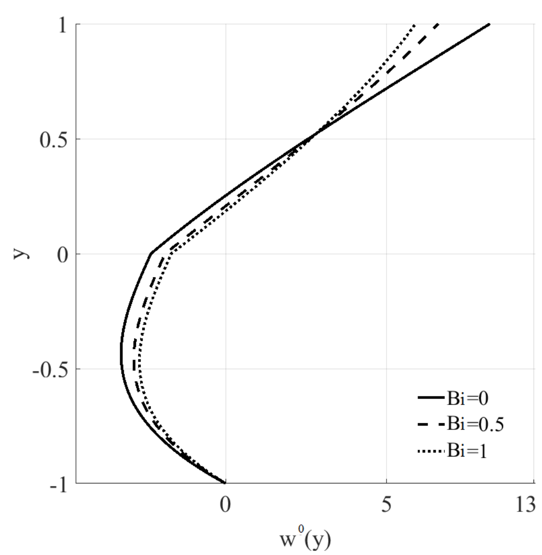

Fig. 2. Stationary profile of velocities $w_j^0(y)$ depending on the Bio number

Changes of the Marangoni numbers also affect the nature and intensity of the arising currents. So, a change in the number Ma_1 affects the direction and intensity of the flow near the interface $y = 0$ (see Fig. 4 a), and the number Ma_2 only affects the intensity near the free boundary (see Fig. 4 b).

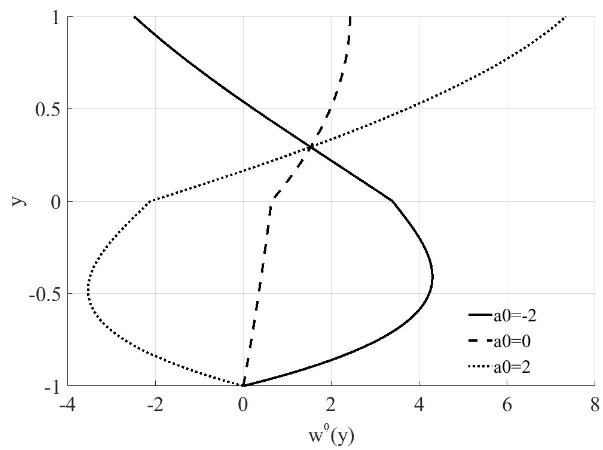

Fig. 3. Stationary profile of velocities $w_j^0(y)$ depending on the dimensionless temperature gradient a_0

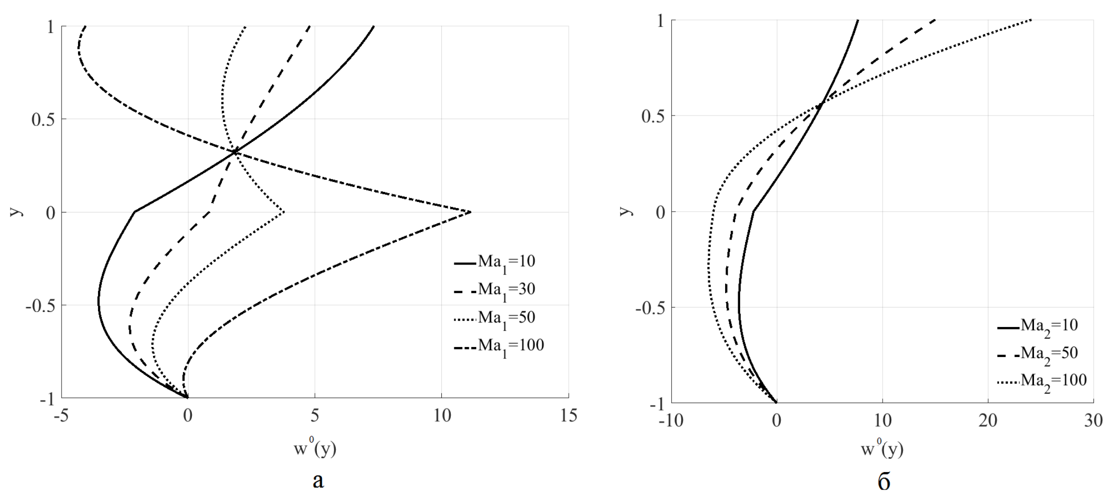

Fig. 4. Stationary profile of velocities $w_j^0(y)$ depending on the Marangoni numbers Ma_1 (a) and Ma_2 (b)

3. Solution of conjugate problem by the Laplace transformation method. Analysis of numerical results

To solve non-stationary problems (4)–(7) and (8)–(14) we apply the Laplace transform. As a result, we arrive to boundary value problems for images $A_j(y,s)$ of functions $a_j(y,t)$

$$\begin{gathered} A_{1yy}(y,s) - sA_1(y,s) = -a_{10}(y), \quad -1 < y < 0, \\ A_{2yy}(y,s) - s\chi h^{-2}A_2(y,s) = -\chi h^{-2}a_{20}(y), \quad 0 < y < 1, \end{gathered} \tag{17}$$

$$\begin{gathered} A_1(-1,s) = A(s), \quad A_{2y}(1,s) + \mathrm{Bi}A_2(1,s) = 0, \\ A_1(0,s) = A_2(0,s), \quad kA_{1y}(0,s) = hA_{2y}(0,s) \end{gathered} \tag{18}$$

and images $W_j(y,s)$ of functions $w_j(y,t)$

$$
\begin{aligned}
&W_{1yy}(y,s) - s\,\mathrm{P}_1^{-1}W_1(y,s) = -\left(\mathrm{P}_1^{-1}w_{10}(y) + \mathrm{Ra}_1\int_0^y A_1(z,s)\,dz + \hat{C}(s)\right), \quad -1<y<0,\\
&W_{2yy}(y,s) - \frac{s\chi}{h^2\,\mathrm{P}_2}W_2(y,p) = -\left(\frac{\chi}{h^2\mathrm{P}_2}w_{20}(y) + \mathrm{Ra}_2\int_0^y A_2(z,s)dz + \frac{\mu\chi}{h^3}\hat{C}(s)\right), \ 0<y<1,
\end{aligned} \tag{19}
$$

$$
\begin{aligned}
&W_1(-1,s)=0, \quad \chi h^{-1}W_1(0,s) = W_2(0,s), \quad W_{2y}(1,s) = \mathrm{Ma}_1 A_2(1,s),\\
&W_{2y}(0,s) - \mu\chi h^{-2}W_{1y}(0,s) = -\mathrm{Ma}_1A_1(0,s), \ \ \chi\int_{-1}^0 W_2(y,s)dy + \int_0^1 W_2(y,s)dy = Q(s).
\end{aligned} \tag{20}
$$

When deriving equations (17), (19), the initial data (5), (10) were used. In (18), (20) the $A(s)$ and $Q(s)$ are images of the given functions $a(t)$ and $q(t)$ respectively (see conditions (6), (14)). The general solution of equations (17), (19) has the form

$$
\begin{aligned}
A_1(y,s) &= b_1\,\mathrm{sh}\,\sqrt{s}y + b_2\,\mathrm{ch}\,\sqrt{s}y - \frac{1}{\sqrt{s}}\int_0^y a_{10}(z)\,\mathrm{sh}\,\sqrt{s}(y-z)\,dz,\\
A_2(y,s) &= b_3\,\mathrm{sh}\,\sqrt{s}y + b_4\,\mathrm{ch}\,\sqrt{s}y - \frac{\sqrt{\chi}}{h\sqrt{s}}\int_0^y a_{20}(z)\,\mathrm{sh}\,\sqrt{s}(y-z)\,dz,\\
W_1(y,s) &= d_1\,\mathrm{sh}\,\alpha_1 y + d_2\,\mathrm{ch}\,\alpha_1 y - \frac{1}{\alpha_1}\int_0^y f_1(z,s)\,\mathrm{sh}\,\alpha_1(y-z)dz - \frac{\hat{C}(s)}{\alpha_1}\left(\mathrm{ch}\,\alpha_1 y - \frac{1}{\alpha_1}\right),\\
W_2(y,s) &= d_3\,\mathrm{sh}\,\alpha_2 y + d_4\,\mathrm{ch}\,\alpha_2 y - \frac{1}{\alpha_2}\int_0^y f_2(z,p)\,\mathrm{sh}\,\alpha_2(y-z)dz - \frac{\mu\mathrm{P}_2\hat{C}(s)}{sh}(\mathrm{ch}\,\alpha_2 y - 1),
\end{aligned} \tag{21}
$$

$$
f_1(y,p) = \mathrm{P}_1^{-1}w_{10}(y) + \mathrm{Ra}_1\int_0^y A_1(z,s)\,dz, \quad f_2(y,s) = \frac{\chi}{h^2\mathrm{P}_2}w_{20}(y) + \mathrm{Ra}_2\int_0^y A_2(z,s)\,dz,
$$

$$
\alpha_1 = \sqrt{s\mathrm{P}_1^{-1}}, \quad \alpha_2 = \sqrt{s\chi\mathrm{P}_2^{-1}}h^{-1}.
$$

The values m_k, d_k, $k=\overline{1,4}$, appearing in (21), and function $\hat{C}(s)$ are determined from the boundary conditions (18), (20). The type of these values is not presented here because of its complexity.

Suppose, that $\lim\limits_{t\to\infty} a(t) = a_0$ and $\lim\limits_{t\to\infty} q(t) = q_0$. Using the obtained representations for $A_j(y,s)$, $W_j(y,s)$ and $\hat{C}(p)$, we can prove the limit equalities

$$
\begin{aligned}
&\lim_{t\to\infty} a_j(y,t) = \lim_{s\to 0} sA_j(y,s) = a_j^s(y), \quad \lim_{t\to\infty} w_j(y,t) = \lim_{s\to 0} sW_j(y,s) = w_j^s(y),\\
&\lim_{t\to\infty} C(t) = \lim_{s\to 0} s\hat{C}(s) = C^s,
\end{aligned} \tag{22}
$$

where $a_j^s(y), w^s(y)$ and C^s are given by formulas (15), (16).

Using the method of numerical inversion of the Laplace transform, we obtain some results for the velocities. The case when $q(t)=0$ (the flow rate is zero and the movement occurs only due to thermogravitational forces) is considered, and the longitudinal temperature gradient on the bottom wall is distributed according to the law $a_(t) = a_0 + \gamma_1 e^{-\gamma_2 t}\sin(\gamma_3 t)$, where the coefficients γ_1, γ_2 are responsible for the amplitude and frequency of the oscillations, respectively. In the case when $a \neq 0, \gamma_2 > 0$, then, according to equalities (22), the solution converges to the stationary state (see Fig. 5 a), and for $\gamma_2 \leqslant 0$, the limits of the functions $a_(t)$ at $t\to\infty$ do not exist and the solution does not tend to the stationary state (see Fig. 5 b). As $a(t)$ discontinuous functions can also be specified, thereby also influencing the nature of the flow.

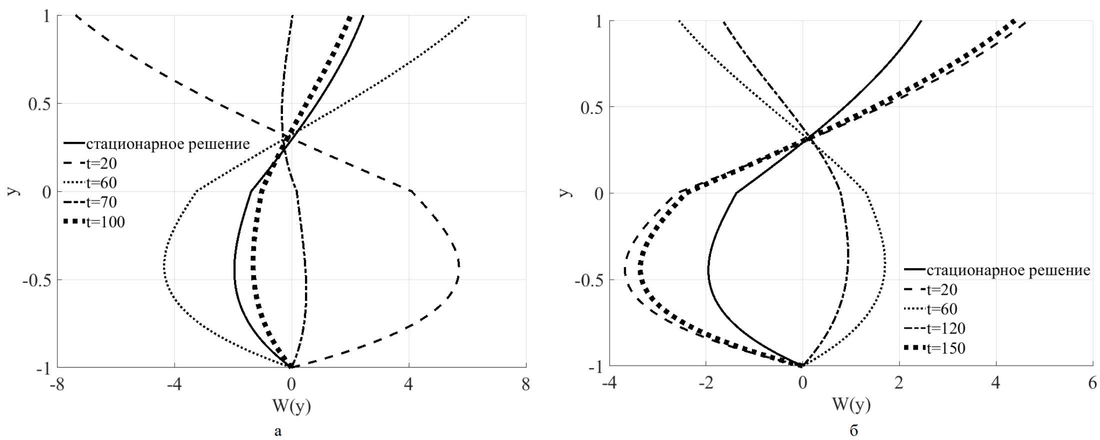

Fig. 5. Profile of dimensionless velocities $W_j(y)$ at $a(t) = 1 - 5e^{-0.01t}\sin(0.1t)$ (a) and $a(t) = 2\sin(0.1t)$ (b)

This work is supported by the Krasnoyarsk Mathematical Center and financed by the Ministry of Science and Higher Education of the Russian Federation in the framework of the establishment and development of regional Centers for Mathematics Research and Education (Agreement no. 075-02-2020-1631).

References

[1] M.K.Smith, S.H.Davis, Instabilities of dynamic thermocapillary liquid layers. Part 2. Surface wave instabillities, *J. Fluid Mech.*, **132**(1983), no. 7, 145–162.

[2] A.A.Vedenov, G.G.Gladush, Physical processes during laser processing of metals, Moscow, Nauka, 1985 (in Russian).

[3] E.A.Ryabitskii, Thermocapillarity instability of liquid layer with internal heat generation, *Microgravity Science and Technology*, **7**(1994), no. 1, 20–23.

[4] H.J.Palmer, J.C.Berg, Hydrodynamic stability of surfactant solutions heated from below, *J. Fluid Mech.*, **51**(1972), no. 2, 385–402.

[5] R.Kh.Zeytovnian, Convection in Fluids, Dordrecht, Springer, 2009.

[6] R.V.Birikh, Thermocapillary convection in a horizontal layer of liquid, *J. Appl. Mech. Tech. Phys.*, **5**(1966), 69–72.

[7] V.K.Andreev, Yu.A.Gaponenko, O.N.Goncharova, V.V.Pukhnachev, Mathematical Models of Convection, Berlin/Boston, Walter de Gruyter GmbH and CO KG, 2012.

[8] V.V.Pukhnachov, Non-stationary Analogues of the Birikh Solution, *Nauchnyi zhurnal teoreticheskih i prikladnyh issledovanii. Novosti Altaiskogo Gos. Universiteta*, **69**(2011), no. 1-2, 62–69 (in Russian).

[9] V.K.Andreev, Birikh solution of convection equations and some their generalizations, Preprint of ICM SB RAS, Krasnoyarsk, 2010 (in Russian).

[10] V.K.Andreev, E.N.Lemeshkova, Linear problems of motions with interfaces, Krasnoyarsk, Sib. Feder. Univ., 2018 (in Russian).

[11] R.Kh.Zeytounian, The Benard-Marangoni thermocapillary-instability problem, *Phys. Usp.*, **41**(1998), 241–267.

Слоистое движение двух несмешивающихся жидкостей со свободной границей

Елена Н. Лемешкова
Институт вычислительного моделирования СО РАН
Красноярск, Российская Федерация

Аннотация. Изучено однонаправленное движение двух вязких несжимаемых жидкостей в плоском канале. На нижней твердой стенке задан нестационарный градиент температуры, а верхняя стенка — свободная граница. Жидкости контактируют по плоской поверхности раздела. Движение вызвано совместным действием термогравитационных и термокапиллярных сил и заданного общего нестационарного расхода в слоях. Соответствующая начально-краевая задача является сопряжённой и обратной, поскольку градиент давления вдоль канала должен находиться вместе с полем скоростей и температур. Для нее найдено точное стационарное решение. В изображениях по Лапласу решение нестационарной задачи находится в виде квадратур. Установлено, что если температура на нижней стенке и расход стабилизируются со временем, то движение выходит на стационарный режим с ростом времени, что говорит об устойчивости стационарного решения относительно однонаправленных нестационарных возмущений. Приведены результаты расчетов, показывающие различные способы управления движением с помощью задания температуры на стенке.

Ключевые слова: термокапиллярность, поверхность раздела, уравнения Обербека–Буссинеска.

Journal of Siberian Federal University. Mathematics & Physics 2020, 13(5), 583–595

DOI: 10.17516/1997-1397-2020-13-5-583-595
УДК 512.5

On the Equationally Artinian Groups

Mohammad Shahryari*
Department of Mathematics
College of Science
Sultan Qaboos University
Muscat, Oman

Javad Tayyebi†
Department of Pure Mathematics
Faculty of Mathematical Sciences
University of Tabriz
Tabriz, Iran

Received 26.05.2020, received in revised form 02.07.2020, accepted 16.08.2020

Abstract. In this article, we study the property of being equationally Artinian in groups. We define the radical topology corresponding to such groups and investigate the structure of irreducible closed sets of these topologies. We prove that a finite extension of an equationally Artinian group is again equationally Artinian. We also show that a quotient of an equationally Artinian group of the form $G[t]$ by a normal subgroup which is a finite union of radicals, is again equationally Artnian. A necessary and sufficient condition for an Abelian group to be equationally Artinian will be given as the last result. This will provide a large class of examples of equationally Artinian groups.

Keywords: algebraic geometry over groups, systems of group equations, radicals, Zariski topology, radical topology, equationally Noetherian groups, equationally Artinian groups.

Citation: M. Shahryari, J. Tayyebi, On the Equationally Artinian Groups, J. Sib. Fed. Univ. Math. Phys., 2020, 13(5), 583–595. DOI: 10.17516/1997-1397-2020-13-5-583-595.

In the mid-twentieth century, Alfred Tarski asked whether two arbitrary non-abelian free groups are elementary equivalent. To answer this question, it was necessary to investigate systems of equations over groups. Makanin and Razborov proved that the existence of solutions for systems of equations over free groups is a decidable problem and an algorithm to solve such systems of equation is discovered (Makanin-Razborov diagrams, see [10] and [14]). The work of Makanin and Razborov as well as many other mathematicians was the beginning of *algebraic geometry over groups*. Since then, this new area of algebra was the subject of importent studies in group theory. The work of Baumslag, Myasnikov and Remeslennikov provides a complete account of this new subject, [1]. Positive solution to the problem of Tarski is discovered by Kharlampovich, Myasnikov and Sela at the the beginning of the recent century (see [7–9] and [15]). After that, many mathematicians investigated the algebraic geometry over general algebraic systems and this new area of algebra is now known as *universal algebraic geometry*. The reader can see the works of Daniyarova, Myasnikov, and Remeslennikov as well as the lecture notes of Plotkin as introduction to this branch, [3–6], and [13].

One of the very important notions in the algebraic geometry of groups (as well as other algebraic structures) is the property of being *equationally Noetherian*. Note that if S is a system of equations over a group A, then we say that the system S implies an equation $w \approx 1$, if every

*m.ghalehlar@squ.edu.om
†j.tayyebi@tabrizu.ac.ir

solution of S in A is also a solution of $w \approx 1$. This gives us an equational logic over the group A which is not in general similar to the first order logic. For example, the compactness theorem may fails in this equational logic. There are examples of groups such that the compactness for the systems of equations fails (see [1] and [5] for some examples). In some groups, every system of equations is equivalent to a finite subsystem, such groups are called equationally Noetherian. Free groups, Abelian groups, linear groups over Noetherian rings and torsion-free hyperbolic groups are equationally Noetherian. To see interesting properties of this types of groups, the reader can consult [11] and [16]. This kind of groups have very important roles in algebraic geometry of groups. There are many equivalent conditions for the property of being equationally Noetherian, for example, it is known that a group A has this property, if and only if, for any natural number n, every descending chain of algebraic sets in A^n is finite. According to this equivalent condition, in [11] and [12], the dual property of being *equationally Artinian* is defined. A group A is equationally Artinian, if and only if, for any natural number n, every ascending chain of algebraic sets in A^n is finite. In [12], many equivalent conditions to this property is given.

In 1997, Baumslag, Myasnikov, and Romankov proved two important theorems about equationally Noetherian groups: first, they showed that a virtually equationally Noetherian group is equationally Noetherian. They also showed that quotient of an equationally Noetherian group by a normal subgroup which is a finite union of algebraic sets, is again equationally Noetherian (see [2]). In this Article we prove similar results for the case of equationally Artinian groups. These results will provide a large class of examples for equationally Artinian groups. Also, we study irreducible closed subsets of the radical topology in the case of equationally Artinian groups and we obtain a necessary and sufficient condition for an Abelian group to be equationally Artinian.

1. Preliminaries

Let G be an arbitrary group and suppose that $X = \{x_1, \ldots, x_n\}$ is a finite set of variables. Consider the free product $G[X] = G * F[X]$, where $F[X]$ is the free group over X. Every element $w \in G[X]$ corresponds to an equation $w \approx 1$, which is called a group equation with coefficients from G. If $w = w(x_1, \ldots, x_n, g_1, \ldots, g_m) \in G[X]$, then the expression $w \approx 1$ is a G-equation with coefficients $g_1, \ldots, g_m \in G$. Suppose H is a group which contains G as a distinguished subgroup. Then we say that H is a G-group. A tuple $\overline{h} = (h_1, \ldots, h_n) \in H^n$ is called a root of the equation $w \approx 1$, if

$$w(h_1, \ldots, h_n, g_1, \ldots, g_m) = 1.$$

An arbitrary set of G-equations is called *a system of equation with coefficients from G*. The set of all common roots of the elements of S in H is called the corresponding *algebraic set* of S and denoted by $V_H(S)$. Clearly, the intersection of a non-empty family of algebraic sets is again an algebraic set but the same is not true for unions of algebraic sets. If we define a closed subset of H^n to be an arbitrary intersection of finite unions of algebraic sets, then we get a topology on H^n, which is known as *Zariski topology*.

For a subset $E \subseteq H^n$, we define the corresponding radical $\mathrm{Rad}(E)$ to be the set of all elements $w \in G[X]$ such that every element of E is a solution of $w \approx 1$. This is a normal subgroup of $G[X]$ which is called the *radical* of E and the quotient group $\Gamma(E) = G[X]/\mathrm{Rad}(E)$ is called the *coordinate group* of E. Similarly, for a system S, we define its radical to be $\mathrm{Rad}_H(S) = \mathrm{Rad}(V_H(S))$. This is the largest system of G-equations equivalent to S over H. The corresponding coordinate group is $\Gamma_H(S) = G[X]/\mathrm{Rad}_H(S)$. It is proved that the study of coordinate groups is equivalent to the study of Zariski topology, i.e. algebraic geometry of H reduces to the study coordinate groups, [1].

A G-group H is called G-equationally Noetherian, if for every system S, there exists a finite subsystem S_0, such that $V_H(S) = V_H(S_0)$. Such G-groups have important role in the study of algebraic geometry over G-groups. There are two extremal cases: if $G = 1$, we say that H is 1-equationally Noetherian or equationally Noetherian without coefficients, and if $G = H$, then we say that H is equationally Noetherian (or equationally Noetherian in Diophantine sense). It is proved that a 1-equationally Noetherian finitely generated group is equationally Noetherian as well, [1]. The class of equationally Noetherian groups is very large, containing all Free groups, Abelian groups, linear groups over Noetherian rings and torsion-free hyperbolic groups are equationally Noetherian. It is not hard to see that the following statements are equivalent for a G-group H:

i- H is G-equationally Noetherian.

ii- the Zariski topology on H^n is Noetherian for all n.

iii- every chain of coordinate groups and proper epimorphisms

$$\Gamma(E_1) \to \Gamma(E_2) \to \Gamma(E_3) \to \cdots$$

is finite.

The authors of [2] proved two important theorems about equationally Noetherian groups. The first theorem shows that a finite extension of an equationally Noetherian group is again equationally Noetherian. The second theorem says that if G is equationally Noetherian and N is a normal subgroup which is a finite union of algebraic sets (in Diophantine case), then G/N is also equationally Noetherian. In this article, we are dealing with the dual notion, the property of being equationally Artinian and we prove the similar statements for this type of groups.

2. Equationally Artinian groups

Equationally Artinian algebras are introduced in [11] and [12]. In this section, we review this notion for the case of G-groups. We say that a G-group H is G-equationally Artinian, if for any n, every ascending chain of algebraic sets in H^n terminates. This is not equivalent to the property of being Artinian for the Zariski topology, instead we define a new topological space which becomes Noetherian if H is equationally Artinian. Suppose

$$T = \{u\mathrm{Rad}(E) : \ E \subseteq H^n, \ u \in G[X]\}.$$

Note that for arbitrary cosets $u\mathrm{Rad}(E)$ and $v\mathrm{Rad}(F)$, if their intersection is non-empty, then for an arbitrary element $w \in u\mathrm{Rad}(E) \cap v\mathrm{Rad}(F)$, we have $w\mathrm{Rad}(E) = u\mathrm{Rad}(E)$ and $w\mathrm{Rad}(F) = v\mathrm{Rad}(F)$. Hence $u\mathrm{Rad}(E) \cap v\mathrm{Rad}(F) = w(\mathrm{Rad}(E) \cap \mathrm{Rad}(F)) = w\mathrm{Rad}(E \cup F)$. This shows that the intersection of two cosets of radicals, is again a coset of a radical subgroup (or it is empty). The set T is a subbasis of closed sets of a topology on the set $G[X]$ which is called *the radical topology* on $G[X]$ corresponding to H (this topology is finer than the previous one defined in [12], in fact the subbasis introduced in [12] is a fundamental system of closed sets containing the identity of $G[X]$). Every closed set in $G[X]$ is an arbitrary intersection of finite unions of cosets of the form $u\mathrm{Rad}(E)$, with $E \subseteq H^n$ and $u \in G[X]$. In [12], it is proved that the following statements are equivalent for a G-group H:

i- H is G-equationally Artinian.

ii- for any n and any subset $E \subseteq H^n$, there exists a finite subset $E_0 \subseteq E$, such that $\mathrm{Rad}(E) = \mathrm{Rad}(E_0)$.

iii- the corresponding radical topology over $G[X]$ is Noetherian.

Remark 1. The proof is essentially the same as in [12], but since we used here our enhanced definition of finer radical topology, so we show that why the proof remains unchanged. We only need to show that for a G-group H, being G-equationally Artinian is equivalent to the property of being Noetherian for the corresponding radical topology on $G[X]$. So, let H be G-equationally Artinian and

$$T = \{u\mathrm{Rad}(E) :\ E \subseteq H^n,\ u \in G[X]\}.$$

We first prove that T satisfies the descending chain condition. Suppose

$$u_1\mathrm{Rad}(E_1) \supseteq u_2\mathrm{Rad}(E_2) \supseteq u_3\mathrm{Rad}(E_3) \supseteq \cdots$$

is a descending chain of elements of T. Then we have also the following chain

$$\mathrm{Rad}(E_1) \supseteq \mathrm{Rad}(E_2) \supseteq \mathrm{Rad}(E_3) \supseteq \cdots .$$

Therefore,

$$V_H(\mathrm{Rad}(E_1)) \subseteq V_H(\mathrm{Rad}(E_2)) \subseteq V_H(\mathrm{Rad}(E_3)) \subseteq \cdots ,$$

and this later chain terminates, as H is G-equationally Artinian. So, for some k, we have

$$V_H(\mathrm{Rad}(E_k)) = V_H(\mathrm{Rad}(E_{k+1})) = V_H(\mathrm{Rad}(E_{k+2})) = \cdots .$$

Taking one more radical, we get

$$\mathrm{Rad}(E_k) = \mathrm{Rad}(E_{k+1}) = \mathrm{Rad}(E_{k+2}) = \cdots .$$

This shows that

$$u_k\mathrm{Rad}(E_k) = u_{k+1}\mathrm{Rad}(E_{k+1}) = u_{k+2}\mathrm{Rad}(E_{k+2}) = \cdots ,$$

and hence T satisfies the descending chain condition. Now, let T_1 be the set of all finite unions of elements of T and T_2 be the set of all arbitrary intersection of elements of T_1. Note that T_2 is the set of all closed subsets of $G[X]$ with respect to the radical topology. We show that T_1 also satisfies the descending chain condition. Suppose that

$$M_1 = u_1\mathrm{Rad}(E_1) \cup \ldots \cup u_m\mathrm{Rad}(E_m),\ M_2 = v_1\mathrm{Rad}(F_1) \cup \ldots \cup v_k\mathrm{Rad}(F_k)$$

are sets in T_1 and $M_2 \subset M_1$. For every $i \leqslant m$ and $j \leqslant k$, we have $u_i\mathrm{Rad}(E_i) \cap v_j\mathrm{Rad}(F_j) \subseteq u_i\mathrm{Rad}(E_i)$. Hence we can gain a tree with root vertex $u_i\mathrm{Rad}(E_i)$ and with a unique edge from the root to every proper subset $u_i\mathrm{Rad}(E_i) \cap v_j\mathrm{Rad}(F_j) \subset u_i\mathrm{Rad}(E_i)$. Suppose there exists a strictly descending chain of subsets in T_1:

$$M_1 \supset M_2 \supset M_3 \supset \cdots .$$

As we mentioned, we obtain a tree for any inclusion $M_i \supset M_{i+1}$, such that each vertex is a finite intersection of sets in T, hence each vertex is in T itself, since as we saw above, the non-empty intersections of a finite number of elements from T are again belong to T. Since each vertex is connected to only finite number of other vertices, so each vertex has finite degree. So, every path corresponds to a strictly descending chain of radicals and since H is G-equationally Artinian, so the path is finite. By the well-known König's lemma of graph theory, this implies that the graph is finite. Therefore the above chain is also finite. So T_1 satisfies the descending chain condition and is closed under finite intersection.

Now, we prove that T_2 satisfies the descending chain condition too. Suppose $\bigcap_{i=1}^{\infty} R_i$ is an infinite intersection of elements of T_1. Then we have the following chain:

$$R_1 \supseteq R_1 \cap R_2 \supseteq R_1 \cap R_2 \cap R_3 \supseteq \cdots .$$

Since T_1 satisfies descending chain condition and is closed under finite intersection, so the chain terminates. Therefore

$$\exists k : R_1 \cap R_2 \cap \ldots \cap R_k = \bigcap_{i=1}^{\infty} R_i.$$

Hence, every infinite intersection of subsets of T_1 is in fact a finite intersection in T_1 and so it belongs to T_1. Consequently, we have $T_2 = T_1$ and hence it satisfies the descending chain condition. This shows that the radical topology on $G[X]$ is Noetherian. The proof of the converse statement is trivial.

By $(EA)_G$, we denote the class of all G-equationally Artinian G-groups, by $(EA)_1$, the class of 1-equationally Artinian groups and EA will be used for the class of Equationally Artinian groups (Diophantine case where $G = H$). In this article, we first prove the following theorem.

Theorem 1. *Let $G \in EA$ be torsion-free and $E \subseteq G^n$ be an algebraic set. Then the set $\mathrm{Rad}(E)$ is irreducible and all irreducible closed subset of $G[X]$ is a coset of some radical.*

Our main tool to prove this result is a well-known theorem of B. Neumann which says that if a group covered by a finite set of cosets of subgroups, then at least one of those subgroups has finite index. This result of Neumann also will be used to prove the following result.

Theorem 2. *Let $G \in EA$ be torsion-free and $E \subseteq G^n$ be a non-empty algebraic set with $\mathrm{Rad}(E) \neq G[X]$. Then the interior of $\mathrm{Rad}(E)$ is empty.*

Note that every Noetherian topological space has finite number of irreducible components. In the case of a torsion-free equationally Artinian group G, the space $G[X]$ has a unique irreducible component, say $G[X]$ itself. Theorem 2, also shows that if $G \in EA$ is torsion-free, then $G[X]$ is connected. We will prove the converse for coefficient-free case.

Theorem 3. *Let $G \in (EA)_1$. Then G is torsion-free if and only if, $F[X]$ is connected.*

Note that there are many equationally Noetherian groups which are not equationally Artinian, for example, the additive group $\mathbb{Q}/\mathbb{Z}$, the multiplicative group of complex numbers, the quasi-cyclic groups $\mathbb{Z}_{p^\infty}$ (see also Theorem 8). Many other groups like non-Abelian free groups and torsion-free hyperbolic groups are failed to be equationally Artinian (as they are domains and every equationally Artinian domain is finite). It must be said that, at the time of writing this paper, we don't know if there is equationally Artinian group which is not equationally Noetherian. But, both classes are included in a larger class of groups which we call *equationally semi-Noetherian*. A group G has this property, if for every system of equations $S \subseteq G[X]$, almost every finite subset $T \subseteq S$ can be omitted solving the system over G, i.e. there exists a finite subset $S_0 \subseteq S$ such that for all other finite subset $T \subseteq S \setminus S_0$, we have $V_G(S) = V_G(S \setminus T)$. Clearly, every equationally Noetherian group has this property. We will prove,

Theorem 4. *If $G \in EA$, then G is equationally semi-Noetherian.*

Our next theorem concerns about an important relation between the classes $(EA)_1$ and $(EA)_G$. We prove,

Theorem 5. *Let G be a finitely generated group and let H be a G-group. If $H \in (EA)_1$, then $H \in (EA)_G$, and as a result, any finitely generated element of $(EA)_1$ is equationally Artinian.*

In our sixth theorem, we deal with finite extensions of equationally Artinian groups. We prove,

Theorem 6. *Let a group A contains a finite index subgroup H which is equationally Artinian. Then A is also equationally Artinian.*

This theorem enables us to conclude that any virtually finitely generated Abelian group is equationally Artinian as well as any finite extension of the additive group of any field. This gives us a large class of examples of such groups. This theorem is EA-version of the similar theorem in [2].

Note that the quotient of an equationally Artinian group is not necessarily equationally Artinian (for example the group $\mathbb{Q}/\mathbb{Z}$), but, there exists an important situation, the quotient in which, has this property. Our nest result concerns with these situations. Note that in this theorem, we use the group $G[t] = G * \langle t \rangle$.

Theorem 7. *Let G be an arbitrary group such that $G[t]$ is equationally Artinian. Let R be a normal subgroup of $G[t]$ which is closed in the radical topology of $G[t]$. Then $G[t]/R$ is also equationally Artinian.*

Finally, we will show that an Abelian group G is equationally Artinian, if and only if, it has finite number of periods: let $p(G)$ be the set of orders of torsion elements of G. We will prove,

Theorem 8. *An Abelian group G is equationally Artinian, if and only if, $p(G)$ is finite.*

3. The proofs

Proof. (**Theorem 1 and 2**) Suppose G is equationally Artinian. Let Y be an irreducible closed subset of $G[X]$. Since Y is a finite union of cosets of the form $u\mathrm{Rad}(E)$, so $W = u\mathrm{Rad}(E)$, for some algebraic set $E \subseteq G^n$ and $u \in G[X]$. Now, for an algebraic set E, we show that $\mathrm{Rad}(E)$ is irreducible. Note that every closed subset of $\mathrm{Rad}(E)$ has the form $v_1\mathrm{Rad}(L_1) \cup \ldots \cup v_p\mathrm{Rad}(L_p)$, where $v_i \in G[X]$ and $E \subseteq L_i$. Now, if $\mathrm{Rad}(E)$ can be written as a union of two closed subsets, then we have

$$\mathrm{Rad}(E) = \bigcup_{i=1}^{m} u_i\mathrm{Rad}(K_i),$$

for some elements $u_i \in G[X]$ and algebraic sets K_i with $E \subseteq K_i$. It is a well-known theorem of B. Neumann which says that if a group is covered by a finite number of cosets of subgroups, then at least one of those subgroups has finite index. So, we have for example $[\mathrm{Rad}(E) : \mathrm{Rad}(K_1)] < \infty$. Suppose now that G is torsion-free and $\mathrm{Rad}(K_1) \neq \mathrm{Rad}(E)$. Choose an element $w \in \mathrm{Rad}(E)$, such that for some $\overline{a} \in K_1$, we have $w(\overline{a}) \neq 1$. Then, for all non-zero integers k we have also $w^k(\overline{a}) \neq 1$ and hence all cosets $w^j\mathrm{Rad}(K_1)$, $(1 \leqslant j)$, are distinct. This shows that $\mathrm{Rad}(E) = \mathrm{Rad}(K_1)$ and so $\mathrm{Rad}(E)$ is irreducible. Note that in any Noetherian space, there is a finite number of maximal irreducible sets (irreducible components) and in the case of $G[X]$, $\mathrm{Rad}(\varnothing) = G[X]$ is the only irreducible component.

Now, we show that the interior of $\mathrm{Rad}(E)$ is empty for any $E \neq \varnothing$. Let an open set $G[X] \setminus \bigcup_{j=1}^{m} w_j\mathrm{Rad}(E_j)$ be contained in $\mathrm{Rad}(E)$. Then we have

$$G[X] = \mathrm{Rad}(E) \cup \bigcup_{j=1}^{m} w_j\mathrm{Rad}(E_j),$$

and again using the theorem of Neumann, some of these subgroups has finite index, which is shows that $G[X] \setminus \bigcup_{j=1}^{m} w_j\mathrm{Rad}(E_j) = \varnothing$. Hence $\mathrm{Rad}(E)$ has empty interior. □

Proof. (**Theorem 3**) In this proof, we denote the coefficient-free radical of a subset $E \in G^n$ by $\mathrm{Rad}^0(E)$, i.e.

$$\mathrm{Rad}^0(E) = \{w \in F[X] : \ \forall \overline{a} \in E \ w(\overline{a}) = 1\}.$$

Suppose first that $F[X]$ is not connected. Then we have

$$F[X] = \bigcup_{i=1}^{m} u_i \mathrm{Rad}^0(E_i),$$

for some elements $u_i \in F[X]$ and subsets $E_i \subseteq G^n$, where $m \geqslant 2$ is minimal. Again by the theorem of Neumann, there is an index i such that $[F[X] : \mathrm{Rad}^0(E_i)]$ is finite and is not equal to 1 by the minimality of m. This shows that the coefficient-free coordinate group $\Gamma(E_i)$ is finite and non-trivial. But, we know that this coordinate group embeds inside a direct power of G. So, G is not torsion-free, a contradiction.

Conversely, assume that G is not torsion-free. Let $a \in G$ be a non-trivial element of finite order m and put $\overline{a} = (a, 1, \ldots, 1) \in G^n$. Let $w = x_1$. Then clearly, $w^m \in \mathrm{Rad}^0(\overline{a})$. Consider the subgroup $\langle w \rangle \mathrm{Rad}^0(\overline{a})$. This subgroup contains all elements $x_1, \ldots, x_n$, and so we have $F[X] = \langle w \rangle \mathrm{Rad}^0(\overline{a})$. Now, we have

$$F[X] = \bigcup_{i=0}^{m-1} w^i \mathrm{Rad}^0(\overline{a}),$$

and hence, $F[X]$ is not connected. □

We now prove Theorem 4. Note that the proof can be applied for arbitrary algebraic structures as well.

Proof. (**Theorem 4**) Suppose G is equationally Artinian and $S \subseteq G[X]$ is an infinite system. For simplicity, assume $S = \{v_1, v_2, v_3, \ldots\}$. We have the ascending chain

$$V_G(S) = V_G(v_1, v_2, v_3, \ldots) \subseteq V_G(v_2, v_3, v_4, \ldots) \subseteq V_G(v_3, v_4, v_5, \ldots) \subseteq \cdots .$$

This chain terminates as G is equationally Artinian, so there exists k such that

$$V_G(v_k, v_{k+1}, v_{k+2}, \ldots) = V_G(v_{k+1}, v_{k+2}, v_{k+3}, \ldots) = V_G(v_{k+2}, v_{k+3}, v_{k+4}, \ldots) \cdots .$$

This shows that

$$\bigcap_{j \geqslant k} V_G(v_j) = \bigcap_{j \geqslant k+1} V_G(v_j) = \bigcap_{j \geqslant k+2} V_G(v_j) = \cdots ,$$

and hence

$$\begin{aligned} V_G(S) &= V_G(v_1, \ldots, v_{k-1}) \cap \bigcap_{j \geqslant k} V_G(v_j) \\ &= V_G(v_1, \ldots, v_{k-1}) \cap \bigcap_{j \geqslant k+1} V_G(v_j) \\ &= V_G(v_1, \ldots, v_{k-1}) \cap \bigcap_{j \geqslant k+2} V_G(v_j). \end{aligned}$$

In other words, this argument shows that the algebraic sets $V_G(v_j)$ can be drop in the intersection for $j \geqslant k$. Let $S_0 = \{v_1, \ldots, v_{k-1}\}$. Then by this argument, for any finite subset $T \subseteq S \setminus S_0$, we have $V_G(S) = V_G(S \setminus T)$. □

Proof. (**Theorem 5**) Let $a_1, \ldots, a_k$ be a finite set of generators for the group G. Suppose $E \subseteq H^n$. We prove that there exists a finite subset $E_0 \subseteq E$, such that $\mathrm{Rad}_G(E) = \mathrm{Rad}_G(E_0)$ (note that, here Rad_G denotes the radical with coefficient in G). Let $S = \mathrm{Rad}_G(E) \subseteq G[X]$. Every element of S has the form

$$w = w(x_1, \ldots, x_n, a_1, \ldots, a_k).$$

We replace every coefficient a_i by a new variable y_i, and then a coefficient-free system of equations $S(\overline{x},\overline{y})$ appears. Let $T = E\times\{(a_1,\dots,a_k)\}\subseteq H^{n+k}$. Now, since $H\in(EA)_1$, so there is a finite subset $T_0\subseteq T$, such that $\mathrm{Rad}_1(T)=\mathrm{Rad}_1(T_0)$. Clearly, we have $T_0 = E_0\times\{(a_1,\dots,a_k)\}$, for some finite subset $E_0\subseteq E$. Obviously, $S(\overline{x},\overline{y})\subseteq \mathrm{Rad}_1(T)$. Let $u(\overline{x},\overline{y})\in\mathrm{Rad}_1(T)$. Then for all $\overline{e}\in E$, we have $u(\overline{e},\overline{a})=1$, so $u(\overline{x},\overline{a})\in\mathrm{Rad}_G(E)$, and therefore $u\in S(\overline{x},\overline{y})$. This proves that $S(\overline{x},\overline{y})=\mathrm{Rad}_1(T)$, and hence $S(\overline{x},\overline{y})=\mathrm{Rad}_1(T_0)$.

Now, we show that $S(\overline{x},\overline{a})=\mathrm{Rad}_G(E_0)$. Suppose $w(\overline{x},\overline{a})\in S(\overline{x},\overline{a})$. For any $\overline{e}\in E_0$, we have $w(\overline{e},\overline{a})=1$, so $w(\overline{x},\overline{a})\in\mathrm{Rad}_G(E_0)$. Conversely, if $w(\overline{x},\overline{a})\in\mathrm{Rad}_G(E_0)$, then for $w(\overline{x},\overline{y})\in\mathrm{Rad}_1(T_0)=\mathrm{Rad}_1(T)$, and this shows that $w(\overline{x},\overline{a})\in\mathrm{Rad}_G(E)=S(\overline{x},\overline{a})$. This proves that $\mathrm{Rad}_G(E)=\mathrm{Rad}_G(E_0)$ and hence $H\in(EA)_G$. □

Theorem 5, enables us to prove that every finitely generated Abelian group belongs to the class EA (we also can deduce this from Theorem 8). Here we give an elementary proof which shows the infinite cyclic group is equationally Artinian.

Lemma 1. *Let $H=\langle a\rangle$ be infinite cyclic group. Then H is equationally Artinian.*

Proof. We first show that $H\in(EA)_1$. Let $E\subseteq H^n$. Every element of E has the form $\overline{e} = = (a^{j_1},\dots,a^{j_n})$ for some integers $j_1,\dots,j_n$. Let $w=x_1^{\alpha_1}x_2^{\alpha_2}\dots x_n^{\alpha_n}\in\mathrm{Rad}_1(E)$. Then $w(\overline{e})=1$ and hence $a^{j_1\alpha_1+\dots+j_n\alpha_n}=1$. This shows that

$$\mathrm{Rad}_1(E)=\bigcap_{j_1,\dots,j_n}\{x_1^{\alpha_1}x_2^{\alpha_2}\dots x_n^{\alpha_n} : (a^{j_1},\dots,a^{j_n})\in E, j_1\alpha_1+\dots+j_n\alpha_n=0\}.$$

Suppose

$$E=\{(a^{j_1^{(1)}},\dots,a^{j_n^{(1)}}),(a^{j_1^{(2)}},\dots,a^{j_n^{(2)}}),(a^{j_1^{(3)}},\dots,a^{j_n^{(3)}}),\dots\}.$$

Suppose S is the following set of equations

$$j_1^{(t)}\alpha_1+\dots+j_n^{(t)}\alpha_n=0,\quad (t=1,2,3,\dots).$$

Since the additive group $\mathbb{Z}$ is equationally Noetherian, so there exists a finite subset $S_0\subseteq S$, such that $V_{\mathbb{Z}}(S)=V_{\mathbb{Z}}(S_0)$. Suppose S_0 consists of the equations

$$j_1^{(t)}\alpha_1+\dots+j_n^{(t)}\alpha_n=0\quad (t=1,2,\dots,m).$$

Let $E_0=\{(a^{j_1^{(1)}},\dots,a^{j_n^{(1)}}),(a^{j_1^{(2)}},\dots,a^{j_n^{(2)}}),\dots,(a^{j_1^{(m)}},\dots,a^{j_n^{(m)}})\}$. Then we have obviously, $\mathrm{Rad}_1(E)=\mathrm{Rad}_1(E_0)$. This shows that H is 1-equationally Artinian and hence by Theorem 2, it belongs to EA. □

Now, we show that any direct product of finitely many element of $(EA)_1$ is again in $(EA)_1$. This will prove that every finitely generated Abelian group belongs to $(EA)_1$ and hence to EA.

Lemma 2. *Suppose A and B are equationally Artinian (1-equationally Artinian). Then so is $A\times B$.*

Proof. For a number n and a subset $E\subseteq(A\times B)^n$, suppose that

$$E=\{c_i=(u_1^i,u_2^i,\dots,u_n^i):\ i\in I\},$$

where I is an index set. We have $u_i^j=(a_i^j,b_i^j)$, for some $a_i^j\in A$ and $b_i^j\in B$. Now, let

$$T=\{t_i=(a_1^i,a_2^i,\dots,a_n^i):\ i\in I\},$$

and

$$S=\{s_i=(b_1^i,b_2^i,\dots,b_n^i):\ i\in I\}.$$

Since A and B are equationally Artinian, so there are two finite subsets $T_0 \subseteq T$ and $S_0 \subseteq S$, such that

$$\mathrm{Rad}_A(T) = \mathrm{Rad}_A(T_0),\ \mathrm{Rad}_B(S) = \mathrm{Rad}_B(S_0).$$

Suppose for example $T_0 = \{t_1, \ldots, t_l\}$ and $S_0 = \{s_1, \ldots, s_k\}$ and $k \geqslant l$. Suppose $t_i = (a_1^i, \ldots, a_n^i)$ and $s_i = (b_1^i, \ldots, b_n^i)$. Using these elements, we can define a finite subset

$$E_0 = \{c_i = ((a_1^i, b_1^i), \ldots, (a_n^i, b_n^i) : 1 \leqslant i \leqslant l\},$$

such that $\mathrm{Rad}(E) = \mathrm{Rad}(E_0)$. This shows that $A \times B$ is equationally Artinian. □

Summarizing, we have

Corollary 1. *Every finitely generated Abelian group is equationally Artinian.*

There are also infinitely generated Abelian groups which are equationally Artinian: let K be a field and consider its additive group $H = (K, +)$. Every equation with coefficient in K has the form $a_1x_1 + \cdots + a_nx_n = b$ for some elements $a_1, \ldots, a_n \in \mathbb{Z}, b \in K$, so the corresponding algebraic set is an affine subspace of K^n. This shows that every ascending chain of algebraic sets terminates and hence H is equationally Artinian. However, some Abelian groups are not equationally Artinian. For example, consider the additive group $H = \mathbb{Q}/\mathbb{Z}$. Let

$$E = \left\{\frac{1}{p} + \mathbb{Z} :\ p = prime\right\} \subseteq H^1.$$

If $w(x) = mx + \left(\frac{a}{b} + \mathbb{Z}\right) \in \mathrm{Rad}(E)$, then for any prime p, we have $w\left(\frac{1}{p} + \mathbb{Z}\right) = \mathbb{Z}$, and this means that for any prime p, $\frac{m}{p} + \frac{a}{b} \in \mathbb{Z}$, which is not true. Another example is the quasi-cyclic groups $G = \mathbb{Z}_{p^\infty}$, for prime numbers p. This is because, the ascending chain of algebraic sets $V_G(x^{p^n} \approx 1)$, $(n \geqslant 1)$ does not terminate (this fact will be used in the proof of Theorem 8).

Before proving Theorem 6, we introduce some notations from [2]. Let a group A be the semidirect product of a finite subgroup T and a normal subgroup H. Assume that $T = \{t_1 = 1, t_2, \ldots, t_k\}$. Let $w(x_1, \ldots, x_n, g_1, \ldots, g_m)$ be a group word with coefficients in A and $v \in A^n$. We can express v uniquely in the form $v = (s_1h_1, \ldots, s_nh_n)$ with $s_i \in T$ and $h_i \in H$. We also have $g_i = r_ib_i$ for unique elements $r_i \in T$ and $b_i \in H$. Define the map $\lambda : A^n \to T^n$ by $\lambda(v) = (s_1, \ldots, s_n)$ and

$$\overline{w}(x_1, \ldots, x_n) = w(x_1, \ldots, x_n, r_1, \ldots, r_m).$$

Note that $\overline{w}$ is an element of $T[X]$ which depends only on w. For any $1 \leqslant i \leqslant n$ and $1 \leqslant j \leqslant k$, define $h_{ij} = t_j^{-1}h_it_j \in H$. Denote the tuple

$$(h_{11}, \ldots, h_{1k}, \ldots, h_{n1}, \ldots, h_{nk})$$

by v'. Consider the new variables y_{ij} for $1 \leqslant i \leqslant n$ and $1 \leqslant j \leqslant k$. In [2], it is proved that there exists a unique element

$$w'_v \in H[y_{11} \ldots, y_{1k}, \ldots, y_{n1}, \ldots, y_{nk}],$$

such that $w(v) = \overline{w}(\lambda(v))w'_v(v')$, and w'_v depends only on the value of $\lambda(v)$. As a result, it is shown that $v \in A^n$ is a root of $w \approx 1$, if and only if, $\lambda(v)$ is a root of $\overline{w} \approx 1$ and v' is a root of $w'_v \approx 1$. We are now ready to prove Theorem 6.

Proof. (**Theorem 6**) Replacing H by its core, we can suppose that H is a normal subgroup of A with finite index. Let $T = A/H$. Then A embeds into the wreath product $H \wr T$. Recall that this wreath product is the semidirect product of T and $H^{|T|}$. We know that (Lemma 2), $H^{|T|}$ is equationally Artinian and any subgroup of an equationally Artinian group is again equationally Artinian. So, it is enough to prove our theorem using the further assumption $A = TH$, with T finite, H normal and $T \cap H = 1$. We will use all the above notations.

Suppose $E \subseteq A^n$ is an algebraic set and $S = \mathrm{Rad}_A(E)$. We must show that there exists a finite subset $E_0 \subseteq E$, such that $\mathrm{Rad}_A(E_0) = S$. Let $\overline{S} = \{\overline{w} : \ w \in S\}$ (see the above discussion). Suppose

$$V_T(\overline{S}) = \{v_1, \ldots, v_d\} \subseteq T^n.$$

For any $1 \leqslant i \leqslant d$, put $L_i = V_H(S'_{v_i}) \subseteq H^{nk}$. Here S'_{v_i} denotes the set of all w'_{v_i}, such that $w \in S$. Define also

$$K_i = \{\overline{h} \in H^n : \ (\overline{h})' \in L_i\} \subseteq H^n.$$

We have $(K_i)' \subseteq H^{nk}$ and since H is equationally Artinian, there exists a finite subset $K_i^0 \subseteq K_i$, such that

$$\mathrm{Rad}_H((K_i^0)') = \mathrm{Rad}_H((K_i)').$$

Assume that $E_0 = \cup_{i=1}^d v_i K_i^0 \subseteq A^n$. We show that $E_0 \subseteq E$. Let $v_i\overline{h} \in E_0$. Then $\overline{h} \in K_i$ and hence

$$\overline{S}(\lambda(v_i\overline{h})) = \overline{S}(v_i) = 1,$$

and

$$S'_{v_i}((v_i\overline{h})') \in S'_{v_i}(L_i) = 1.$$

This means that $v_i\overline{h} \in V_A(S) = E$. Therefore $E_0 \subseteq E$.

Now, we claim that $\mathrm{Rad}_A(v_i K_i^0) = \mathrm{Rad}_A(v_i K_i)$. To prove this claim, assume that w belongs to the left hand side. Then $w(v_i K_i^0) = 1$ and hence $w'((v_i K_i^0)') = 1$. This shows that $w'_{v_i} \in \mathrm{Rad}_H((v_i K_i^0)')$. Recall that, by the definition of the map $v \mapsto v'$, we have $(v_i K_i^0)') = (K_i^0)'$ and hence $w'_{v_i} \in \mathrm{Rad}_H((K_i^0)') = \mathrm{Rad}_H((K_i)') = \mathrm{Rad}_H((v_i K_i)')$. Therefore, for any $\overline{h} \in K_i$, we have $w'_{v_i}((v_i\overline{h})') = 1$, and since in the same time $\overline{w}(\lambda(v_i\overline{h})) = 1$, we have $w(v_i K_i) = 1$. This proves the claim.

We now, prove that $\mathrm{Rad}_A(E_0) = \mathrm{Rad}_A(E)$. Let w be an element of the left hand side and $v \in E$. We have $S(v) = 1$ and

$$w \in \bigcap_{i=1}^{d} \mathrm{Rad}_A(v_i K_i^0).$$

Note that $v = \lambda(v)\overline{h}$, for some $\overline{h} \in H^n$. We have $\overline{S}(\lambda(v)) = 1$, so there is an index i such that $\lambda(v) = v_i$. Therefore, $v = v_i\overline{h}$. On the other side, since $S'_v(V') = 1$, so

$$1 = S'_v(v') = S'_{v_i}((v_i\overline{h})').$$

Hence, $(v_i\overline{h})' \in L_i$, and therefore $\overline{h} \in K_i$. Now, by the above claim, we have

$$w \in \mathrm{Rad}_A(v_i K_i^0) = \mathrm{Rad}_A(v_i K_i),$$

and hence $w(v) = 1$. This shows that $w \in \mathrm{Rad}_A(E)$. □

Theorem 6 shows that any virtually finitely generated Abelian group is equationally Artinian as well as any finite extension of the additive group of any field. This gives us a large class of examples of such groups. We now come to Theorem 7. Note that the similar theorem ([2]) for the equationally Noetherian case deals with the Zarizki topology of G^1 and its closed normal subgroups. The dual case here deals with the radical topology of $G[t]$ and its closed normal subgroups.

Proof. Assume that

$$R = \bigcup_{i=1}^{m} \mathrm{Rad}_G(K_i),$$

where $K_i \subseteq G$. Note that G is equationally Artinian as $G[t]$ is so. Hence every K_i can be chosen finite. Let $H = G[t]/R$ be not equationally Artinian. Hence there exists a number n and a subset $E \in H^n$ such that $\mathrm{Rad}_H(E) \neq \mathrm{Rad}_H(E_0)$, for any finite subset $E_0 \subseteq E$. Assume that $e_0 \in E$ is an arbitrary element. As $\mathrm{Rad}_H(E) \neq \mathrm{Rad}_H(\{e_0\})$, there exist elements $f_1 \in \mathrm{Rad}_H(\{e_0\})$ and $e_1 \in E$, such that $f_1(e_1) \neq 1$. Similarly, we have $\mathrm{Rad}_H(E) \neq \mathrm{Rad}_H(\{e_0, e_1\})$, so there exist elements $f_2 \in \mathrm{Rad}_H(\{e_0, e_1\})$ and $e_2 \in E$, such that $f_2(e_2) \neq 1$. Repeating this argument, we obtain two infinite sequences

$$f_1, f_2, f_3, \ldots \in H[X],$$

$$e_0, e_1, e_2, \ldots \in E,$$

such that for any i, $f_i(e_0) = f_i(e_1) = \cdots = f_i(e_{i-1}) = 1$, but $f_i(e_i) \neq 1$. Note that, here $X = \{x_1, \ldots, x_n\}$ and so every element of $H[X]$ is a word in t and elements of X with coefficients in G. Suppose $q : G[t, X] \to H[X]$ is the canonical map sending elements of G to their cosets, and fixing elements of X and the element t. Suppose also that $\psi : (G[t])^n \to H^n$ is the map

$$\psi(u_1, \ldots, u_n) = (u_1 R, \ldots, u_n R).$$

Choose a pre-image $\overline{f}_i$ for f_i under q and a pre-image $\overline{e}_i$ for e_i under ψ. Hence, we have $\overline{f}_i \in G[t, X]$ and $\overline{e}_i \in (G[t])^n$. For any i, we have $f_i(e_0) = 1$, so $\overline{f}_i(\overline{e}_0) \in R$. This shows that, there exists an infinite sequence of numbers

$$i_1(0) < i_2(0) < i_3(0) < \cdots,$$

and a number $1 \leqslant p_0 \leqslant m$, such that

$$\overline{f}_{i_1(0)}(\overline{e}_0), \overline{f}_{i_2(0)}(\overline{e}_0), \overline{f}_{i_3(0)}(\overline{e}_0), \ldots \in \mathrm{Rad}_G(K_{p_0}).$$

Equivalently, this shows that for all s, we have

$$\overline{f}_{i_s(0)} \in \mathrm{Rad}_{G[t]}(\overline{e}_0(K_{p_0})).$$

By a similar argument, we obtain an infinite subsequence of $\{i_s(0)\}$ of the form

$$i_1(1) < i_2(1) < i_3(1) < \cdots,$$

and a number $1 \leqslant p_1 \leqslant m$, such that for all s, we have

$$\overline{f}_{i_s(1)} \in \mathrm{Rad}_{G[t]}(\overline{e}_1(K_{p_1})).$$

We continue this process to find an infinite subsequence

$$i_1(k) < i_2(k) < i_3(k) < \cdots,$$

of the previous sequence, and a number $1 \leqslant p_k \leqslant m$, such that

$$\overline{f}_{i_s(k)} \in \mathrm{Rad}_{G[t]}(\overline{e}_k(K_{p_k})),$$

for all s. Note that all sets $\overline{e}_i(K_{p_i})$ are finite as K_i's are finite. Let

$$K = \bigcup_{i=0}^{\infty} \overline{e}_i(K_{p_i}) \subseteq (G[t])^n.$$

By assumption, $G[t]$ is equationally Artinian, so there exists an index l, such that

$$\mathrm{Rad}_{G[t]}(K) = \mathrm{Rad}_{G[t]}(\bigcup_{i=0}^{l} \overline{e}_i(K_{p_i})).$$

Assume that $j > l$. Then for any s, we have

$$\overline{f}_{i_s(j)} \in \bigcap_{i=1}^{l} \mathrm{Rad}_{G[t]}\overline{e}_i(K_{p_i}) = \mathrm{Rad}_{G[t]}(K).$$

Suppose $k = i_1(j)$. Then $\overline{f}_k \in \mathrm{Rad}_{G[t]}(\overline{e}_k(K_{p_k}))$, and hence $\overline{f}_k(\overline{e}_k) \in \mathrm{Rad}_G(K_{p_k}) \subseteq R$. This shows that $f_k(e_k) = 1$, a contradiction. Hence H is equationally Artinian. □

Finally, we give a proof for Theorem 8.

Proof. (**Theorem 8**) We first, show that a divisible Abelian group G is equationally Artinian, if and only if, it is torsion-free. Recall that a divisible Abelian group has the form $G = \mathbb{Q}^I \oplus \sum_{p\in J} \mathbb{Z}_{p^\infty}$, for an index set I and a set J of prime numbers. If G is torsion-free then $G = \mathbb{Q}^I$, and since the additive group of rationales is equationally Artinian, so is G. Now, suppose that G is equationally Artinian but is not torsion-free. Then for some prime p, we have $\mathbb{Z}_{p^\infty} \leqslant G$, and this implies that $\mathbb{Z}_{p^\infty}$ is equationally Artinian, a contradiction.

Now, suppose that G is an arbitrary Abelian group. Assume that $p(G)$ is finite. We know that $G = \mathrm{Tor}(G) \oplus G_1$, where $\mathrm{Tor}(G)$ is the torsion part of G and G_1 is a torsion-free subgroup. We know that G_1 can be embedded in some divisible Abelian group and hence it is equationally Artinian. The torsion part has finite exponent and hence can be written in the form $\mathrm{Tor}(G) = \bigoplus_{m\in p(G)} \mathbb{Z}_m^{I_m}$, where for all $m \in p(G)$, an index set I_m is associated. Clearly, every component $\mathbb{Z}_m^{I_m}$ is equationally Artinian and since $p(G)$ is finite, so the direct sum is also so. This shows that $G \in EA$.

Finally, suppose that in a group G, the set $p(G)$ is infinite. Let $m_1 < m_2 < m_3 < \cdots$ be elements of $p(G)$ such that for all i the integer $m_1 m_2 \dots m_{i-1}$ is not divisible by m_i. For any i, assume that a_i is an element of order m_i. Consider the ascending chain

$$V_G(x^{m_1} \approx 1) \subseteq V_G(x^{m_1 m_2} \approx 1) \subseteq V_G(x^{m_1 m_2 m_3} \approx 1) \subseteq \cdots .$$

This chain does not terminate, because for any i, we have

$$a_1, \dots, a_i \in V_G(x^{m_1 \dots m_i} \approx 1),$$

but a_{i+1} does not belong to it. Therefore G is not equationally Artinian. □

References

[1] G.Baumslag, A.Myasnikov, V.Remeslennikov, Algebraic geometry over groups, I. Algebraic sets and ideal theory, *J. Algebra*, **219**1999, 16–79.

[2] G.Baumslag, A.Myasnikov, V.Romankov, Two theorems about equationally Noetherian groups, *J. Algebra*, 1997, **194**, 654–64.

[3] E.Daniyarova, A.Myasnikov, V.Remeslennikov, *Algebra and Discrete Mathamatics*, **1**(2008), 80–112. DOI: 10.1142/9789812793416_0007

[4] E.Daniyarova, A.Myasnikov, V.Remeslennikov, Algebraic geometry over algebraic structures, II: Fundations, *J. Math. Sci.*, **185**(2012), no. 3, 389–416.

[5] E.Daniyarova, A.Myasnikov, V.Remeslennikov, Algebraic geometry over algebraic structures, III: Equationally noetherian property and compactness, *South. Asian Bull. Math.*, **35**(2011), no. 1, 35–68.

[6] E.Daniyarova, A.Myasnikov, V.Remeslennikov, Algebraic geometry over algebraic structures, IV: Equatinal domains and co-domains, *Algebra and Logic*, **49**(2011), no. 6, 483–508.

[7] O.Kharlampovich, A.Myasnikov, Tarski's problem about the elementary theory of free groups has a psitive solution, *E.R.A. of AMS*, **4**(1998), 101–108.

[8] O.Kharlampovich, A.Myasnikov, Irreducible affine varieties over a free group. I: irreducibility of quadratic equations and Nullstellensatz, *J. Algebra*, **200**(1998), no. 2, 472–516.

[9] O.Kharlampovich, A.Myasnikov, The elemntary theory of free non-abelian groups, *J. Algebra*, **302**(2006), 451–552.

[10] G.Makanin, Equations in free groups, *Math. USSR-Izv.*, **21**(1982), no. 3, 483–546.

[11] P.Modabberi, M.Shahryari, Compactness conditions in universal algebraic geometry, *Algebra and Logic*, **55**(2016), no. 2, 146–172.

[12] P.Modabberi, M.Shahryari, On the equational Artinian algebras, *Siberian Electronic Mathematical Reports*, **13**(2016), 875–881.

[13] B.Plotkin, Seven lectures in universal algebraic geometry, preprints, Arxiv.

[14] A.Razborov, On systems of equations in free groups, *Math. USSR-Izv.*, 1982, **25**, no. 1, 115–162.

[15] Z.Sela, Diophantine geometry over groups: I-X, preprints, Arxiv.

[16] M.Shahryari, A.Shevliyakov, Direct products, varieties, and compacness conditions, *Groups Complex, Cryptol*, **9**(2017), no. 2, 159–166.

Об эквивалентно артиновых группах

Мохаммед Шахриари
Математический факультет, Колледж Науки
Университет Султана Кабуса
Мускат, Оман
Джавад Тайеби
Кафедра чистой математики, Факультет математических наук
Университет Тебриза
Тебриз, Иран

Аннотация. В этой статье мы изучаем свойство быть артиновым в группах. Определяем радикальную топологию, соответствующую таким группам, и исследуем структуру неприводимых замкнутых множеств этих топологий. Докажем, что конечное расширение уравновешенно артиновой группы снова уравновешенно артиново. Мы также показываем, что частное от артиново-уравновешенной группы вида $G[t]$ по нормальной подгруппе, являющейся конечным объединением радикалов, опять-таки уравновешенно артново. В качестве последнего результата будет дано необходимое и достаточное условие, чтобы абелева группа была эквивалентно артиновой. Это обеспечит большой класс примеров уравновешенно артиновых групп.

Ключевые слова: алгебраическая геометрия над группами, системы групповых уравнений, радикалы, топология Зариского, радикальная топология, нетеровы группы, эквационально артиновы группы.

DOI: 10.17516/1997-1397-2020-13-5-596-607
УДК 517.95

On a Transmission Problem Related to Models of Electrocardiology

Yulia L. Shefer*
Siberian Federal University
Krasnoyarsk, Russian Federation

Received 18.05.2020, received in revised form 09.07.2020, accepted 15.08.2020

Abstract. We consider a generalization of a transmission problem for matrix elliptic operators related to mathematical models of cardiology. We find sufficient conditions when the approach developed for scalar elliptic operators is still valid in this much more general situation.

Keywords: transmission problems for elliptic systems, models of electrocardiology.

Citation: Y.L. Shefer, On a Transmission Problem Related to Models of Electrocardiology, J. Sib. Fed. Univ. Math. Phys., 2020, 13(5), 596–607. DOI: 10.17516/1997-1397-2020-13-5-596-607.

Introduction

In this paper we consider a family of transmission problems for elliptic operators with constant coefficients related to models of electrocardiology. More precisely, for many years for satisfactory models of heart activity one uses Cauchy, Dirichlet, and Neumann problems for scalar strongly elliptic operators, see, for example, [1, 2]. A modification of such a model involving boundary problems for the Laplace operator has been recently studied in [3].

We consider similar problems for more general matrix linear elliptic operators and find sufficient conditions under which the scheme for solving the problems suggested in [3] allows to construct their solutions. Our approach is essentially based on the general theory of Fredholm problems for strongly elliptic (matrix) linear operators, see, e.g., [4], and the theory of regularization of an ill-posed Cauchy problem for operators with an injective principal symbol, see [3].

1. A model example

To begin with, we consider a basic example related to models of electrocardiology. As known from clinical practice, see, e.g., [1, 2], electrical activity of cardiac cells is crucial for pumping function of heart, which is the result of rhythmical cycles of contraction-relaxation of the cardiac tissue. Anomalies of electrical activity often cause heart diseases, which makes these investigations, in particular, development of adequate mathematical models, very relevant nowadays.

Let us illustrate this by one model of electrocardiology [1,2,5]. Denote by Ω_B and Ω_H three-dimensional domains with piecewise smooth boundaries with $\partial\Omega_B$ and $\partial\Omega_H$ corresponding to a body and a heart (see Fig. 1). Then the domain $\Omega = \Omega_B \backslash \Omega_H$ with the boundary $\partial\Omega = \partial\Omega_B \cup \partial\Omega_H$ corresponds to the body without heart.

*yushefer@mail.ru

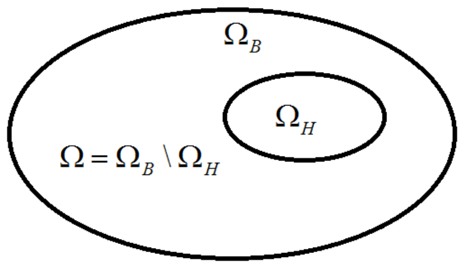

Fig. 1. Geometry of the model

Usually, in standard models one assumes that the cardiac tissue can be divided into two parts – intracellular and extracellular parts separated by a membrane – to which the electric potential u_i and u_e, respectively, is assigned. Regarding the cardiac tissue as a continuous medium we think of the potentials as defined in each point of Ω_H and satisfying the equation

$$\nabla^* M_i \nabla u_i + \nabla^* M_e \nabla u_e = 0, \tag{1}$$

where M_i and M_e are known tensor matrices that characterize intracellular and extracellular parts, and ∇ is the gradient operator in $\mathbb{R}^3$.

One often considers the case when M_i and M_e are positively defined matrices with constant coefficients with entry values defined by conductivity of the cardiac tissue. For simplicity of the further analysis one assumes that these matrices are proportional

$$M_i = \lambda M_e, \qquad \lambda > 0.$$

Based on equation (1) one considers two models of heart activity. In one model it is assumed that the heart is isolated and one considers the problem

$$\begin{gathered} \nabla^* M_i \nabla u_i + \nabla^* M_e \nabla u_e = 0 \text{ in } \Omega_H, \\ (\nu_1, \nu_2, \nu_3) M_i \nabla u_i = 0 \text{ on } \partial\Omega_H, \\ (\nu_1, \nu_2, \nu_3) M_e \nabla u_e = -(\nu_1, \nu_2, \nu_3) M_b \nabla u_b \text{ on } \partial\Omega_H, \\ u_b = u_e \text{ on } \partial\Omega_H, \end{gathered} \tag{2}$$

where M_b is the tensor matrix characterizing conductivity of the body, ν is the vector field of unit outward normal vectors to the boundary of the domain under the consideration and u_b is the electric potential of the body.

In the second model one takes the body into account, and from the electrodynamics of stationary currents it follows that the electric potential of the body u_b in the domain Ω is defined by the equations

$$\begin{gathered} \nabla^* M_b \nabla u_b = 0 \text{ in } \Omega, \\ (\nu_1, \nu_2, \nu_3) M_b \nabla u_b = 0 \text{ on } \partial\Omega_B. \end{gathered} \tag{3}$$

A feature of the model is the fact that one is more interested not in potentials u_i and u_e separately but in their difference $v = u_i - u_e$ in Ω_H or at least on its boundary.

Since matrices M_i and M_e are positively defined and not degenerate, the problems (2), (3) can be studied in the framework of the theory of boundary (maybe ill-posed) problems for elliptic formally self-adjoint equations, see [1, 2, 5]. Moreover, notice that the problems above may be regarded as transmission problems for elliptic equations with discontinuous coefficients describing solutions in different domains of a continuum with the help of additional conditions on separating surfaces, see, for example, [6, 7].

Until now we have not used any functional spaces in the problems description, in the next section we give a precise formulation of a more general problem and specify functional classes for its solution.

2. Formulation of a problem

Let θ be a measurable set in $\mathbb{R}^n$, $n \geqslant 2$. Denote by $L^2(\theta)$ a Lebesgue space of complex-valued functions on θ with the scalar product

$$(u, v)_{L^2(\theta)} = \int_\theta \overline{v}(x) u(x)\, dx.$$

If D is a domain in $\mathbb{R}^n$ with a piecewise smooth boundary ∂D, then for $s \in \mathbb{N}$ we denote by $H^s(D)$ the standard Sobolev space with the scalar product

$$(u, v)_{H^s(D)} = \int_D \sum_{|\alpha| \leqslant s} \overline{(\partial^\alpha v)} (\partial^\alpha u) dx.$$

It is well-known that this scale extends for all $s > 0$. Let now $H^s(D)$ for $s \in \mathbb{R}_+ \setminus \mathbb{Z}_+$ be the standard Sobolev-Slobodeckij spaces. Denote by $H^s_0(D)$ the closure of the subspace $C^\infty_{\mathrm{comp}}(D)$ in $H^s(D)$, where $C^\infty_{\mathrm{comp}}(D)$ is the linear space of functions with compact supports in D.

The space of k-vectors $u = (u_1, \ldots, u_k)$ whose components lie in $H^s(D)$ equipped with the scalar product

$$(u, v)_{[H^s(D)]^k} = \sum_{j=1}^k \int_D \sum_{|\alpha| \leqslant s} \overline{(\partial^\alpha v_j)} (\partial^\alpha u_j) dx = \int_D \sum_{|\alpha| \leqslant s} (\partial^\alpha v)^* (\partial^\alpha u) dx$$

we shall denote by $[H^s(D)]^k$.

Further on, we shall consider linear matrix operators

$$A = \sum_{|\alpha| \leqslant p} A_\alpha \partial^\alpha, \; x \in D,$$

where $p \in \mathbb{N}$ is the order of operator A, $\alpha \in \mathbb{Z}^n_+$, and A_α are $(l \times k)$-matrices with constant coefficients. By a formal adjoint of A we call the differential operator

$$A^* = \sum_{|\alpha| \leqslant p} A^*_\alpha \partial^\alpha,$$

where A^*_α is the adjoint matrix for A_α or, equivalently,

$$(Au, v)_{[L^2(D)]^l} = (u, A^* v)_{[L^2(D)]^k} \;\; \text{для всех } u \in [C^\infty_0(D)]^k, \; v \in [C^\infty_0(D)]^l.$$

As usual, the principal symbol of an operator A is the matrix

$$\sigma(A)(x, \zeta) = \sum_{|\alpha| = p} A_\alpha \zeta^\alpha, \; x \in D, \; \zeta \in \mathbb{C}^n.$$

We say that the principal symbol of A is injective if $l \geqslant k$ and

$$\operatorname{rang} \sigma(A)(x,\zeta) = k, \text{ для всех } \zeta \in \mathbb{R}^n \setminus \{0\} \text{ и всех } x \in \overline{D}.$$

If $l = k$ operators with injective principal symbols are called elliptic.

Let now A_e, A_i, and A_b be linear differential operators of the first order with constant coefficients on $\overline{D}_m$, i.e.

$$A_m = \sum_{j=1}^{n} a_j^{(m)} \frac{\partial}{\partial x_j} + a_0^{(m)},$$

where $m \in \{e, i, b\}$, $D_e \equiv D_i \equiv \Omega_H$, $D_b \equiv \Omega$.

Further on, we assume that principal symbols of operators A_m are injective in the corresponding domains.

Denote by A_m^* a formal adjoint of A_m and consider a generalized Laplacian $A_m^* A_m$.

Under assumptions made above, the operator $A_m^* A_m$ is a strongly elliptic $(k \times k)$-matrix second order operator, i.e. it is elliptic and there exists a positive constant c such that

$$\Re\Big(- w^* \sigma(A_m^* A_m)(x,\zeta) w \Big) \geqslant c\,|w|^2\,|\zeta|^2 \text{ for all } \zeta \in \mathbb{R}^n \setminus \{0\}\,,\ w \in \mathbb{C}^k \setminus \{0\}\,,\ x \in \overline{D}_m.$$

The operator $A_m^* A_m$ is also formally self-adjoint, i.e.

$$(A_m^* A_m u, v)_{[L^2(D_m)]^k} = (u, A_m^* A_m v)_{[L^2(D_m)]^k} = (A_m u, A_m v)_{[L^2(D_m)]^l} \text{ for all } u, v \in [C_0^\infty(D_m)]^k;$$

in particular, the operator $A_m^* A_m$ is (formally) positively defined

$$(A_m^* A_m u, u)_{[L^2(D_m)]^k} \geqslant 0 \text{ for all } u \in [C_0^\infty(D_m)]^k.$$

Let, as before, ν be the outward normal vector operator on the boundary of the domain of the operator A_m. Introduce the conormal derivatives

$$\nu_{A_m} = \sigma^*(A_m)(\nu) A_m,$$

associated with these operators via Green's formula:

$$\int_{\partial\Omega} v \nu_{A_m} u ds = \int_{\Omega} (v^* (A_m^* A_m u) - (A_m v)^* A_m u) dx \text{ for all } u, v \in \left[H^2(\overline{D_m})\right]^k. \tag{4}$$

Assume that bounded domains Ω_H, Ω, and Ω_b have twice smooth boundaries and consider the following problem (5), (6): find vector-functions u_i, u_e from $\left[H^2(\Omega_H)\right]^k$ and a vector-function u_b from $\left[H^2(\Omega)\right]^k$ such that

$$\begin{cases} A_i^* A_i u_i + A_e^* A_e u_e = 0 \text{ in } \Omega_H, \\ \nu_{A_i} u_i = 0 \text{ on } \partial\Omega_H, \\ \nu_{A_e} u_e = -\nu_{A_b} u_b \text{ on } \partial\Omega_H, \\ u_e = u_b \text{ on } \partial\Omega_H, \end{cases} \tag{5}$$

$$\begin{cases} A_b^* A_b u_b = 0 \text{ in } \Omega, \\ \nu_{A_b} u_b = 0 \text{ on } \partial\Omega_B, \end{cases} \tag{6}$$

where the equality on the boundary is in the sense of traces, and the equality in the domains is in the sense of distributions. In this case we can assume that traces of functions and their conormal derivatives are well-defined.

It is obvious that the problem (5), (6) is a generalization of the problem (2), (3). Note also that it incorporates several classical boundary problems.

Example 2.1. Consider first the classical case $A_b = \nabla$ ($k = 1$, $l = n$), then $\nu_{A_b} = \dfrac{\partial}{\partial \nu}$ is a directional derivative along the outward normal vector to $\partial\Omega_B$. If we assume that u_e is known on $\partial\Omega_H$ and equal to a function $v_0 \in H^{3/2}(\partial\Omega_H)$, then (5), (6) gives the following problem: find a function $u_b \in H^2(\Omega)$ satisfying

$$\begin{cases} -\Delta u_b = 0 \text{ in } \Omega, \\ \dfrac{\partial u_b}{\partial \nu} = 0 \text{ on } \partial\Omega_B, \\ u_b = v_0 \text{ on } \partial\Omega_H. \end{cases} \tag{7}$$

This is a classical mixed problem that is often called a Zaremba problem, see, e.g. [4, 8]. This problem can be studied by standard methods in Sobolev and Hölder spaces. It is well-known that this problem has a unique solution in these classes that can be written with the help of the Green function $\mathcal{Z}_\Omega(x, y)$ having the standard properties

$$u_b(x) = \int_{\partial\Omega} \mathcal{Z}_\Omega(x, y) v_0(y) dS(y), \ x \in \Omega_H,$$

where $dS(y)$ is the volume form on the surface $\partial\Omega$, see [4, 8].

Analogously, if we assume that $A_e = \nabla$ ($k = 1$, $l = n$), then $\nu_{A_e} = \dfrac{\partial}{\partial \nu}$ is a directional derivative along the outward normal vector to $\partial\Omega_H$. If the conormal derivative $\nu_{A_e} u_e$ is known on $\partial\Omega_H$ and equal to a function $v_1 \in H^{1/2}(\partial\Omega_H)$, then (5), (6) gives a special case of a classical Neumann problem for a Laplace operator: find a function $u_b \in H^2(\Omega)$ satisfying

$$\begin{cases} -\Delta u_b = 0 \text{ in } \Omega, \\ \dfrac{\partial u_b}{\partial \nu} = 0 \text{ on } \partial\Omega_B, \\ \dfrac{\partial u_b}{\partial \nu} = v_1 \text{ on } \partial\Omega_H, \end{cases} \tag{8}$$

see [4, 9]. It is known that this problem is Fredholm in Sobolev and Hölder spaces, its solution is defined up to an additive constant, and the necessary and sufficient condition for solvability is the following

$$\int_{\partial\Omega_H} v_1(y) dS(y) = 0. \tag{9}$$

If this condition is satisfied the problem has a unique solution u_b in these classes that satisfies, for example,

$$\int_{\partial\Omega_H} u_b(y) dS(y) = 0. \tag{10}$$

It can be written with the help of an appropriate parametrix $\mathcal{N}_\Omega(x, y)$ that has the standard properties

$$u_b(x) = \int_{\partial\Omega} \mathcal{N}_\Omega(x, y) v_0(y) dS(y), \ x \in \Omega_H,$$

However, the general theory of boundary problems suggests that knowledge of u_e or $\nu_{A_e} u_e$ on $\partial\Omega_H$ does not allow to recover the potential u_i uniquely from the remaining data and equations without additional conditions (see also Uniqueness Theorem 3.1 for the problem (5), (6) proved under additional assumptions below).

Besides that, cardiology models are special in the sense that additional conditions necessary for recovering of unknown potentials u_i, u_e, u_b in the problem (5), (6) should preferably be set on the boundary of 'the body' Ω, since all measurements must be less traumatic for a patient and not invasive.

3. Application of an ill-posed Cauchy problem

On of the simplest additional conditions mentioned above leads to using of an ill-posed Cauchy problem. More precisely, it implies measuring the potential u_b on the boundary of 'the body':

$$u_b = f \text{ on } \partial\Omega_B, \tag{11}$$

where f is a given vector-function from $\left[H^{3/2}(\Omega)\right]^k$.

Unfortunately, as known very well, the problem (6), (11) is nothing else but an ill-posed problem for an elliptic operator $A_b^* A_b$. Let us see what the addition of the property (11) gives in a more general problem than those in cardiology.

Denote by $N(\Omega)$ the set of solutions to the problem (5), (6), (11) under the condition $f = 0$. Let $S_{A_e}(\Omega_H)$ be the space of generalized solutions of the equation $A_e h = 0$ в Ω_H. Since the operator A_e has an injective symbol and its coefficients are real analytic, the Petrovsky theorem yields that the elements of the space $S_{A_e}(\Omega_H)$ are real analytic vector-functions in Ω_H.

Theorem 3.1. *Let bounded domains Ω_H, Ω, and Ω_b have twice smooth boundaries and let for some constant $\lambda > 0$,*

$$A_i = \lambda A_e. \tag{12}$$

Then the set $N(\Omega)$ consists of triples $(u_i, u_e, u_b) \subset \left[H^2(\Omega_H)\right]^k \times \left[H^2(\Omega_H)\right]^k \times \left[H^2(\Omega)\right]^k$ such that

$$u_i = \frac{h - w}{\lambda^2}, \quad u_e = w, \quad u_b = 0, \tag{13}$$

where h is an arbitrary function from the space $S_{A_e}(\Omega_H) \cap \left[H^2(\Omega_H)\right]^k$, and w is an arbitrary function from $\left[H_0^2(\Omega_H)\right]^k$.

Proof. Let a vector h belong to $S_{A_e}(\Omega_H) \cap \left[H^2(\Omega_H)\right]^k$ and a vector w belong to $\left[H_0^2(\Omega_H)\right]^k$. Then w satisfies the following conditions

$$w = 0 \text{ on } \partial\Omega_H, \quad \nu_{A_i}(w) = 0 \text{ on } \partial\Omega_H, \tag{14}$$

and $A_i^* A_i = \lambda^2 A_e^* A_e$. Therefore the vector functions from (13) give a solution to the problem (5), (6), (11) for $f = 0$.

Let $u_i, u_e \in \left[H^2(\Omega_H)\right]^k$, and $u_b \in \left[H^2(\Omega)\right]^k$ is a triple of functions from $N(\Omega)$. Then from (5), (6) it follows that u_b is a solution to the Cauchy problem for the operator $A_b^* A_b$:

$$A_b^* A_b u_b = 0 \text{ in } \Omega, \quad \nu_{A_b}(u_b) = 0 \text{ on } \partial\Omega_B, \quad u_b = 0 \text{ on } \partial\Omega_B.$$

Since the operators A_m have injective symbols, we have

$$\operatorname{rang}(\nu_{A_m})(x, \nu(x)) = \sigma^*(A_m)(x, \nu(x))\sigma(A_m)(x, \nu(x)) = k$$

for any $m = e, i, b$ and all $x \in \partial\Omega_H$ or $\partial\Omega$, respectively. In particular, the systems of boundary operators $\{I, \nu_{A_e}\}$, $\{I, \nu_{A_i}\}$ are first order Dirichlet systems on $\partial\Omega_H$, while the system of boundary operators $\{I, \nu_{A_b}\}$ is a first order Dirichlet system on $\partial\Omega$ (see, for example, [3]). Then by the uniqueness theorem for a Cauchy problem for elliptic operators (see, for example, [3, Theorem 10.3.5]), $u_b \equiv 0$ in Ω. Now by the trace theorem for Sobolev spaces and by equations from (5) we see that $u_e \equiv 0$ нон $\partial\Omega_H$ and $\nu_{A_e}(u_e) \equiv 0$ on $\partial\Omega_H$. However, since the system of boundary operators $\{I, \nu_{A_e}\}$ is a first order Dirichlet system on $\partial\Omega_H$, it follows from the theorem on spectral synthesis (see [10]) that $u_e \in [H_0^2(\Omega_H)]^k$.

To complete the proof of the theorem we need the following lemma.

Lemma 3.1. *Let Ω_H be a bounded domain in $\mathbb{R}^n$ with a twice smooth boundary and* (12). *If the functions u_e, $u_i \in \left[H^2(\Omega_H)\right]^k$ satisfy the equations (5) then they are related in Ω_H by*

$$u_e + \lambda^2 u_i = h, \tag{15}$$

where h as a function from the space $S_{A_e^* A_e}(\Omega_H) \cap \left[H^2(\Omega_H)\right]^k$.

Moreover, if $u_b \equiv 0$ on $\partial\Omega_H$, then the functions u_e, u_i are related in Ω_H by 15, where h is a function from the space $S_{A_e}(\Omega_H) \cap \left[H^2(\Omega_H)\right]^k$.

Proof. Since $A_i = \lambda A_e$, the first equation in (5) can be rewritten in the form

$$A_e{}^* A_e h = 0 \text{ in } \Omega_H, \tag{16}$$

with $h = u_e + \lambda^2 u_i$, and clearly $h \in S_{A_e^* A_e}(\Omega_H) \cap \left[H^2(\Omega_H)\right]^k$.

If we additionally know that $u_b \equiv 0$ on $\partial\Omega_H$ then, as noticed above, $u_b \equiv 0$ in Ω. Therefore $\nu_{A_e}(u_e) = 0$ on $\partial\Omega_H$, and $\nu_{A_i}(u_i) = 0$ and $\partial\Omega_H$, which implies that

$$\nu_{A_e}(h) = 0 \text{ on } \partial\Omega_H. \tag{17}$$

From this, by the Green formula (4) we obtain

$$0 = (A_e^* A_e h, h)_{[L^2(\Omega_H)]^k} = \int_{\Omega_H} h^* (A_e^* A_e h) dx =$$
$$= \int_{\Omega_H} (A_e h)^* (A_e h) dx + \int_{\partial\Omega_H} h^* \nu_{A_e}(h) ds = \|A_e h\|^2_{[L^2(\Omega_H)]^l}.$$

Therefore, the vector function h defined by the equality (15) belongs to $S_{A_e}(\Omega_H) \cap \left[H^2(\Omega_H)\right]^k$. □

Thus, the functions $u_i, u_e \in \left[H^2(\Omega_H)\right]^k$ satisfy (5), and by Lemma 3.1 we get $u_i = \dfrac{h - v}{\lambda^2}$, where $v \in [H_0^2(\Omega_H)]^k$ and $h \in S_{A_e}(\Omega_H) \cap \left[H^2(\Omega_H)\right]^k$. □

In particular, it follows from Lemma 3.1 that the zero space of the problem (5) coincides with the space $S_{A_e}(\Omega_H) \cap \left[H^2(\Omega_H)\right]^k$.

Denote by $\ker A_e$ the kernel of a continuous linear operator $A_e : [H^2(\Omega_H)]^k \to [H^1(\Omega_H)]^l$ and consider several examples. In fact, $\ker A_e = S_{A_e}(\Omega_H) \cap \left[H^2(\Omega_H)\right]^k$.

Example 3.1. Let $A_e = \begin{pmatrix} \nabla \\ 1 \end{pmatrix}$, $(k = 1,\ l = n + 1)$. Then $A_e^* = \big(-\mathrm{div}, 1\big)$, $\nu_{A_e} = \dfrac{\partial}{\partial \nu}$, $A_e^* A_e = -\Delta + 1$, and the problem (16)–(17) becomes a Neumann problem for the Helmholtz operator

$$\begin{cases} -\Delta h + h = 0 \text{ in } \Omega_H, \\ \quad \dfrac{\partial h}{\partial \nu} = 0 \text{ on } \partial\Omega_H, \end{cases} \tag{18}$$

and the equation $A_e h = 0$ takes the form

$$\begin{cases} \nabla h = 0 \text{ in } \Omega_H, \\ h = 0 \text{ in } \Omega_H. \end{cases}$$

Consequently, $\ker A_e = \{0\}$ and coincides with the space of solutions of the homogeneous problem (18).

Example 3.2. Let $A_e = \nabla$ then $A_e^* = -\text{div}$. In this case ($k = 1$, $l = n$), $A_e^* = -\text{div}$, $\nu_{A_e} = \dfrac{\partial}{\partial \nu}$, $A_e^* A_e = -\Delta$, and the problem (16), (17) becomes a Neumann problem for the Laplace operator

$$\begin{cases} \Delta h = 0 \text{ in } \Omega_H, \\ \dfrac{\partial h}{\partial \nu} = 0 \text{ on } \partial\Omega_H, \end{cases} \tag{19}$$

and the equation $A_e h = 0$ takes the form

$$\nabla h = 0 \text{ in } \Omega_H.$$

Therefore, $\ker A_e = \mathbb{R}$ and coincides with the space of solutions of the problem (19).

Example 3.3. Consider the case where $A_e = \overline{\partial} = \partial_x - i\partial_y$ is the Cauchy–Riemann operator in $\mathbb{R}^2 \cong \mathbb{C}$ where i stands for imaginary unit. Then $A_e^* = -\partial = -\partial_x - i\partial_y$, and the kernel of A_e is holomorphic functions. The problem (16)–(17) defines then the zero space of a non- coercive $\overline{\partial}$-Neumann problem, see, for example, [11, 12].

It is clear that the operator A_e should be chosen in a way that its kernel is at least finite dimensional.

Under assumptions of Theorem 3.1 the rest of the scheme of solving the problem (5), (6), (11) differs little from the standard one, see [5]. Namely, first we introduce a function $h(x)$ such that $h(x) = \lambda^2 u_i + u_e$, where $x \in \Omega_H$. From the conditions on the boundaries in (5) and the fact that $\nu_{A_i} = \lambda \nu_{A_e}$ we get that

$$\nu_{A_e} h = -\nu_{A_b} u_b \text{ on } \partial\Omega_H.$$

Thus, we can rewrite the original problem (5), (6), (11) in new notation: knowing a vector $f \in [H^{3/2}(\partial\Omega_H)]^k$, find vectors $h \in [H^2(\partial\Omega_H)]^k$ and $u_b \in [H^2(\partial\Omega)]^k$ such that

$$\begin{cases} A_e^* A_e h = 0 \text{ in } \Omega_H, \\ \nu_{A_e} h = -\nu_{A_b} u_b \text{ on } \partial\Omega_H, \end{cases} \tag{20}$$

$$\begin{cases} A_b^* A_b u_b = 0 \text{ in } \Omega, \\ \nu_{A_b} u_b = 0 \text{ on } \partial\Omega_B, \\ u_b = f \text{ on } \partial\Omega_B. \end{cases} \tag{21}$$

The original problem splits into two — (20) and (21). The problem (21), as noticed above, is an ill-posed Cauchy problem for an elliptic operator $A_b^* A_b$. It is known that if a solution to this problem exists it is unique. The problem (20) is a Neumann problem for an elliptic operator $A_e^* A_e$. Unfortunately, in general the Neumann problem may also be ill-posed. For it to be Fredholm, the so called Shapiro-Lopatinsky conditions must be placed [13, Chapter 1, Sec. 3, condition II for $q = 0$], [14] on the pair $(A_e^* A_e, \nu_{A_e})$. In particular, they guarantee that the space $S_{A_e}(\Omega_H) \cap [H^2(\Omega_H)]^k$ is finite dimensional.

More precisely, let us consider the following Neumann problem: for a given vector $h_0 \in [H^{1/2}(\partial\Omega_H)]^k$ find a vector $h \in [H^2(\partial\Omega_H)]^k$ such that

$$\begin{cases} A_e^* A_e h = 0 \text{ in } \Omega_H, \\ \nu_{A_e} h = h_0 \text{ on } \partial\Omega_H, \end{cases} \tag{22}$$

and formulate conditions for solvability.

Theorem 3.2. *If for a pair of operators $(A_e^* A_e, \nu_{A_e})$ the Shapiro-Lopatinsky conditions are fulfilled then the problem (22) is Fredholm. To be precise,*

1) *the zero space of the problem coincides with the finite-dimensional space $S_{A_e}(\Omega_H) \cap [H^2(\Omega_H)]^k$;*
2) *the problem is solvable if and only if*

$$(h_0, \varphi)_{[L^2(\partial\Omega_H)]^k} = 0 \text{ for all } \varphi \in S_{A_e}(\Omega_H) \cap [H^2(\Omega_H)]^k; \tag{23}$$

3) *under (23) there exists a unique solution h_1 of the problem (22) satisfying*

$$(h_1, \varphi)_{[L^2(\partial\Omega_H)]^k} = 0 \text{ for all } \varphi \in S_{A_e}(\Omega_H) \cap [H^2(\Omega_H)]^k. \tag{24}$$

Proof. See [4]. □

Thus, under hypothesis of Theorem 3.2 for solvability of the Neumann problem (20) it is necessary and sufficient that for the vector $h_0 = -\nu_{A_b} u_b$ the condition (23) is fulfilled. This can be achieved if we place additional conditions on relations between the operators A_e and A_b. Namely, as we have seen above, it is quite natural to assume that

$$A_e = \tilde{\lambda} A_b \text{ for some constant } \tilde{\lambda} > 0. \tag{25}$$

Denote by $S_{A_b}(\Omega)$ the zero space of solutions to the problem (21) in the domain Ω.

Corollary 3.1. *Let for the pair of operators $(A_e^* A_e, \nu_{A_e})$ the Shapiro–Lopatinsky conditions be fulfilled. Besides that assume that the identity (25) holds and the spaces $S_{A_e}(\Omega_H) \cap [H^2(\Omega_H)]^k$ and $S_{A_b}(\Omega) \cap [H^2(\Omega)]^k$ coincide. Then for any vector $u_b \in [H^2(\Omega)]^k$ satisfying (21) there exists a unique vector $h_1 \in [H^2(\Omega_H)]^k$ that satisfies (20) and (24).*

Proof. By Theorem 3.2 for solvability of the problem (20) it is necessary and sufficient that

$$(\nu_{A_e} u_b, \varphi)_{[L^2(\partial\Omega_H)]^k} = 0 \text{ for all } \varphi \in S_{A_e}(\Omega_H) \cap [H^2(\Omega_H)]^k. \tag{26}$$

If the vector $u_b \in [H^2(\Omega)]^k$ satisfies (21), then by the Green formula (4) for the operator A_b

$$-\int_{\partial\Omega_H} \nu_{A_b} u_b \psi ds = \int_{\partial\Omega} \nu_{A_b} u_b \psi ds = (\psi, A_b^* A_b u)_{[L^2(\Omega)]^k} - (A_b \psi, A_b u)_{[L^2(\Omega)]^k} = 0,$$

for any $\psi \in S_{A_b}(\Omega_H) \cap [H^2(\Omega_H)]^k$.

On the other hand, the relation (25) guarantees that $(\tilde{\lambda})^2 \nu_{A_b} = -\nu_{A_e}$, and therefore

$$(\nu_{A_e} u_b, \varphi)_{[L^2(\partial\Omega_H)]^k} = -(\tilde{\lambda})^2 (\nu_{A_b} u_b, \varphi)_{[L^2(\partial\Omega_H)]^k} = -(\tilde{\lambda})^2 \int_{\partial\Omega_H} \nu_{A_b} u_b \psi ds$$

for any $\varphi \in S_{A_e}(\Omega_H) \cap [H^2(\Omega_H)]^k$. Due to the fact that the spaces $S_{A_e}(\Omega_H) \cap [H^2(\Omega_H)]^k$ and $S_{A_b}(\Omega) \cap [H^2(\Omega)]^k$ coincide, (26) holds. Then by statement 3 of Theorem 3.2 for any vector u_b there exists a unique vector $h_1 \in [H^2(\Omega_H)]^k$ satisfying (20) and (24). □

The condition that the spaces $S_{A_e}(\Omega_H)\cap[H^2(\Omega_H)]^k$ and $S_{A_b}(\Omega)\cap[H^2(\Omega)]^k$ coincide seems to be rather strong, especially since these are spaces of solutions to different differential equations in different domains. Nevertheless, provided (25) holds, such a coincidence is possible if the operator A_e is so much overdetermined that the space of its solutions in any domain is finite dimensional and coincides with the space of solutions in $\mathbb{R}^n$; the typical examples are the so-called stationary holonomic systems. Let us illustrate this by the following examples.

Example 3.4. Let $A_e = \nabla$ and $A_b = \tilde{\lambda}\nabla$ $(k = 1,\, l = n)$. The function $u = const$ is a solution to the equation $\nabla u = 0$ in Ω_H and extends to Ω, where it is a solution to $\tilde{\lambda}\nabla u = 0$. Thus we get that the spaces $S_{A_e}(\Omega_H)$ and $S_{A_b}(\Omega)$ coincide.

Example 3.5. Let $A_e = \begin{pmatrix}\nabla\\ 1\end{pmatrix}$ and $A_b = \tilde{\lambda}\begin{pmatrix}\nabla\\ 1\end{pmatrix}$, $(k = 1,\, l = n+1)$. A solution to $A_e u = 0$ in Ω_H is $u \equiv 0$ and it extends to Ω, where it is a solution $A_b u = 0$. Thus, the spaces $S_{A_e}(\Omega_H)$ and $S_{A_b}(\Omega)$ coincide.

Example 3.6. Consider the following operators A_i, A_e и A_b:

$$A_e = \begin{pmatrix}\partial_x & 0 & 0\\ \partial_y & 0 & 0\\ 0 & \partial_x & 0\\ 0 & \partial_y & 0\\ -1 & 0 & \partial_x\\ 0 & -1 & \partial_y\end{pmatrix},\quad A_i = \lambda\begin{pmatrix}\partial_x & 0 & 0\\ \partial_y & 0 & 0\\ 0 & \partial_x & 0\\ 0 & \partial_y & 0\\ -1 & 0 & \partial_x\\ 0 & -1 & \partial_y\end{pmatrix},\quad A_b = \tilde{\lambda}\begin{pmatrix}\partial_x & 0 & 0\\ \partial_y & 0 & 0\\ 0 & \partial_x & 0\\ 0 & \partial_y & 0\\ -1 & 0 & \partial_x\\ 0 & -1 & \partial_y\end{pmatrix}.$$

These operators have injective principal symbols and are equivalent to second order operators

$$\tilde{A}_e = \begin{pmatrix}\partial_{xx}\\ \partial_{yy}\\ \partial_{xy}\end{pmatrix},\quad \tilde{A}_i = \lambda\begin{pmatrix}\partial_{xx}\\ \partial_{yy}\\ \partial_{xy}\end{pmatrix},\quad \tilde{A}_b = \tilde{\lambda}\begin{pmatrix}\partial_{xx}\\ \partial_{yy}\\ \partial_{xy},\end{pmatrix}.$$

Therefore the space of solutions of the system $A_e u = 0$ in Ω_H coincides with the set of all linear functions $u = c_1 x + c_2 y + c_3$, and any function of this form extends to Ω, where it is a solution to the equation $A_b u = 0$. Therefore, the spaces $S_{A_e}(\Omega_H)$ and $S_{A_b}(\Omega)$ coincide.

As noticed above, if a solution to the Neumann problem (20) exists, it is 'unique' up to an element of the space $S_{A_e}(\Omega_H)\cap[H^2(\Omega_H)]^k$ (additively).

Recall that the aim of solving the original problem (5), (6), (11) is to find the transmembrane potential v on the surface $\partial\Omega_H$. Let us write down the algorithm for solving the problem (20), (21):

1. Find a function u_b and its conormal derivative $\nu_{A_b}(u_b)$ on the surface $\partial\Omega_H$ by solving an ill-posed Cauchy problem (21) for an elliptic operator $A_b^* A_b$.

2. Compute values of $h(x)$ on the surface $\partial\Omega_H$ by solving a Neumann problem (20) for an elliptic operator $A_e^* A_e$ with the data $\nu_{A_b} u_b$ on $\partial\Omega_H$ obtained in Step 1. The possibility of this depends on whether the restrictions on operators A_e and A_b described above hold.

3. Find the transmembrane potential v on the surface $\partial\Omega_H$ by using the relation (15) together with u_b and h on $\partial\Omega_H$, found in Steps 1 and 2, respectively

$$v = u_i - u_e = \frac{h - u_b}{\lambda^2} - u_b \text{ on } \partial\Omega_H. \tag{27}$$

In conclusion we note that solvability conditions for an ill-posed Cauchy problem in Sobolev spaces for a rather wide class of operators with real analytic coefficients are well known, see, for example, [3]. Moreover, in [3, 15] on can find constructive procedures for its regularization, i.e. for construction exact and approximate solutions (the so called Carleman formulas). Regarding areas with models with the geometry corresponding to that in cardiology and operators that are first order matrix factorizations of the Laplace operator, or more generally, of a Lamé-type operator such Carleman formulas were obtained in [17].

The work was supported by the Foundation for the Advancement of Theoretical Physics and Mathematics "BASIS".

References

[1] M.Burger, K.A.Mardal, B.F.Nielsen, Stability analysis of the inverse transmembrane potential problem in electrocardiography, *Inverse Problems*, **26**(2010), no. 10, 105012

[2] J.Sundnes, G.T.Lines, X.Cai, B.F.Nielsen, K.A.Mardal, A.Tveito, Computing the Electrical Activity in the Heart, Springer-Verlag, 2006.

[3] N.N.Tarkhanov, The Cauchy Problem for Solutions of Elliptic Equations, Akademie-Verlag, Berlin, 1995.

[4] S.Simanca, Mixed Elliptic Boundary Value Problems, *Comm. in PDE*, **12**(1987), 123–200.

[5] V.Kalinin, A.Kalinin, W.H.W.Schulze, D.Potyagaylo, A.Shlapunov, On the correctness of the transmembrane potential based inverse problem of ECG, *Computing in Cardiology*, 2017, 1–4.

[6] M.Schechter, A generalization of the transmission problem, *Ann. SNS di Pisa, Cl. Sci.*, **14**(1960), no. 3, 207–236.

[7] M.Borsuk, Transmission problem for Elliptic second-order Equations in non-smooth domains, Birkhäuser, Berlin, 2010.

[8] S.Zaremba, Sur un problème mixte relatif à l'équation de Laplace, *Bull. Acad. Sci. Cracovie*, **2**(1910), 314–344.

[9] V.P.Mikhailov, Partial differential equations, Nauka, Moscow, 1976 (in Russian).

[10] L.I.Hedberg, T.H.Wolff, Thin sets in nonlinear potential theory, *Ann. Inst. Fourier*, **33**(1983), no. 4, 161–187.

[11] G.B.Folland, J.J.Kohn, The Neumann Problem for the Cauchy-Riemann Complex, Princeton University Press, 1973.

[12] A.A.Shlapunov, N.N.Tarkhanov, On completeness of root functions of Sturm-Liouville problems with discontinuous boundary operators, *J. of Differential Equations*, **255**(2013), 3305–3337.

[13] M.S.Agranovich, M.I.Vishik, Elliptic problems with a parameter and parabolic problems of general type, *Rus. Math. Surv.*, **19**(1964), 3, 53–157.

[14] Z.Ya.Shapiro, On general boundary problems for equations of elliptic type, *Izv. Akad. Nauk SSSR Ser. Mat.*, **17**(1953), 6, 539–562 (in Russian).

[15] A.A.Shlapunov, N.N.Tarkhanov, Bases with double orthogonality in the Cauchy problem for systems with injective symbols, *Proceedings LMS*, **3**(1995), 1, 1–52.

[16] D.Gilbarg, N.Trudinger, Elliptic Partial Differential Equations of Second Order, Springer–Verlag, Berlin, 1977.

[17] A.A.Shlapunov, On the Cauchy problem for some elliptic systems in a shell in $\mathbb{R}^n$, *Zeitschrift für Angewandte Mathematik und Mechanik*, **77**(1997), 11, 849–859.

Об одной задаче трансмиссии, связанной с моделями электрокардиологии

Юлия Л. Шефер
Сибирский федеральный университет
Красноярск, Российская Федерация

Аннотация. В настоящей работе рассмотрено одно обобщение задачи трансмиссии для матричных эллиптических операторов, связанной с математическими моделями кардиологии. Указаны достаточные условия, при которых подход, разработанный для скалярных операторов, все еще работает в новой, гораздо более общей ситуации.

Ключевые слова: задачи трансмиссии для эллиптических операторов, модели электрокардиологии.

DOI: 10.17516/1997-1397-2020-13-5-608-621
УДК 517.9

Baranchick-type Estimators of a Multivariate Normal Mean Under the General Quadratic Loss Function

Abdenour Hamdaoui*
Department of Mathematics
University of Sciences and Technology, Mohamed Boudiaf, Oran
Laboratory of Statistics and Random Modelisations (LSMA), Tlemcen
Algeria

Abdelkader Benkhaled†
Department of Biology
Mascara University Mustapha Stambouli
Laboratory of Geomatics, Ecology and Environment (LGEO2E)
Mascara, Algeria

Mekki Terbeche‡
Department of Mathematics
University of Sciences and Technology, Mohamed Boudiaf, Oran
Laboratory of Analysis and Application of Radiation (LAAR), USTO-MB
Oran, Algeria

Received 08.04.2020, received in revised form 01.06.2020, accepted 16.07.2020

Abstract. The problem of estimating the mean of a multivariate normal distribution by different types of shrinkage estimators is investigated. We established the minimaxity of Baranchick-type estimators for identity covariance matrix and the matrix associated to the loss function is diagonal. In particular the class of James-Stein estimator is presented. The general situation for both matrices cited above is discussed.

Keywords: covariance matrix, James-Stein estimator, loss function, multivariate gaussian random variable, non-central chi-square distribution, shrinkage estimator.

Citation: A. Hamdaoui, A. Benkhaled, M. Terbeche, Baranchick-type Estimators of a Multivariate Normal Mean Under the General Quadratic Loss Function, J. Sib. Fed. Univ. Math. Phys., 2020, 13(5), 608–621. DOI: 10.17516/1997-1397-2020-13-5-608-621.

1. Introduction and Preliminaries

The field of estimation of a multivariate normal mean using shrinkage estimators was introduced in [10]. The author showed that the maximum likelihood estimator (MLE) of the mean θ of a multivariate gaussian distribution $N_p\left(\theta, \sigma^2 I_p\right)$ is inadmissible in mean squared sense when the dimension of the parameters space $p \geqslant 3$. In particular, he proved the existence of an estimator which always achieves the smaller total mean squared error regardless of the true θ. Perhaps the best known estimator of such kind is James-Stein's estimator introduced in [7]. This

*abdenour.hamdaoui@yahoo.fr, abdenour.hamdaoui@univ-usto.dz
†benkhaled08@yahoo.fr
‡mekki.terbeche@gmail.com

one is a special case of a larger class of estimators known as shrinkage estimators which is a combination of a model with low bias and high variance, and a model with high bias but low variance. In this context we can cite for example Baranchik [2] for his work on the minimaxity of the estimators of the form $\delta_r(X,S) = (1 - r(F)/F)X$ where $F = \|X\|^2/S$, the statistics $S \sim \sigma^2\chi_n^2$ is the estimator of the unknown parameter σ^2 and $r(.)$ is a real mesurable function. Strawderman [12] was interested to study the estimation of the mean vector of a scale mixture of multivariate distribution under squared error loss. He showed the analogous results obtained by Baranchik [2]. Xie et al [13] have introduced a class of semiparametric/parametric shrinkage estimators and established their asymptotic optimality properties. Selahattin et al [9], provided several alternative methods for derivation of the restricted ridge regression estimator (RRRE). The optimal extended balanced loss function (EBLF) estimators and predictors are introduced and derived from [8] and discussed their performances. In [6], the authors considered the model $X \sim N_p\left(\theta, \sigma^2 I_p\right)$ where σ^2 is unknown and estimated by S^2 ($S^2 \sim \sigma^2\chi_n^2$). They studied the following class of shrinkage estimators $\delta_\psi = \delta^{JS} + l(S^2\psi(S^2, \|X\|^2)/\|X\|^2)X$ with l is real parameter. Benkhaled and Hamdaoui [3], have considered the model $X \sim N_p\left(\theta, \sigma^2 I_p\right)$ where σ^2 is unknown. They studied the minimaxity of two different forms of shrinkage estimators of θ: estimators of the form $\delta^\psi = (1 - \psi(S^2, \|X\|^2)S^2/\|X\|^2)X$, and estimators of Lindley-type given by $\delta^\varphi = (1 - \varphi(S^2, T^2)S^2/T^2)(X - \overline{X}) + \overline{X}$.

In this work, we deal with the model $X \sim N_p(\theta, \Sigma)$ and the loss matrix Q where the covariance matrix Σ is known. Our aims is to estimate the unknown parameter θ by shrinkage estimators deduced by the MLE. The paper is organized as follows. In Section 2, we study the standard case $\Sigma = I_p$ and $Q = D = \mathrm{diag}(d_1, d_2, \ldots, d_p)$, we find the explicit formula of the risk function of considered estimators and we treat there minimax property. As a special case, the James-Stein estimator and its risk are also found. In Section 3, we study the considered problem with the generalized matrices Σ and Q. In Section 4, we graphically illustrate risks ratios of the James-Stein estimator and the estimators of Baranchick-type to the MLE for various values of p. We end the manuscript by giving an Appendix which contains technical lemmas used in the proofs of our results.

We recall that if $X \sim N_p\left(\theta, \sigma^2 I_p\right)$, then $\|X\|^2/\sigma^2 \sim \chi_p^2(\lambda)$ where $\chi_p^2(\lambda)$ denotes the non-central chi-square distribution with p degrees of freedom and non-centrality parameter $\lambda = \|\theta\|^2/2\sigma^2$. We also recall the following results that are useful in our proofs.

Definition 1. *For any measurable function $f : \mathbb{R}_+ \longrightarrow \mathbb{R}$, $\chi_p^2(\lambda)$ integrable, we have*

$$E\left[f(\chi_p^2(\lambda))\right] = E_{\chi_p^2(\lambda)}\left[f(U)\right] = \sum_{k=0}^{+\infty}\left[\int_{\mathbb{R}_+} f(u)\chi_{p+2k}^2 du\right] P\left(\frac{\lambda}{2}; dk\right),$$

where $P(\lambda/2)$ being the Poisson's distribution of parameter $\lambda/2$ and χ_{p+2k}^2 is the central chi-square distribution with $p+2k$ degrees of freedom.

Lemma 1. *(Stein [11]). Let X be a $N\left(\upsilon, \sigma^2\right)$ real random variable and let $f : \mathbb{R} \longrightarrow \mathbb{R}$ be an indefinite integral of the Lebesgue measurable function, f' essentially the derivative of f. Suppose also that $E\left|f'(X)\right| < +\infty$, Then*

$$E\left[\left(\frac{X-\upsilon}{\sigma^2}\right) f(X)\right] = E\left(f'(X)\right).$$

For the next, if $X \sim N_p(\theta, \Sigma)$, we assume that the loss incurred in estimating θ by δ is the function $L_Q(\delta, \theta) = (\delta - \theta)^t Q(\delta - \theta)$ and the risk function associated to this loss is $R_Q(\delta, \theta) = E_\theta(L_Q(\delta, \theta))$.

2. Results for standard case

Let $X \sim N_p(\theta, I_p)$ be a multivariate gaussian random variable in $\mathbb{R}^p$ and for any estimator δ we take the loss function $L_Q(\delta, \theta) = (\delta - \theta)^t Q(\delta - \theta)$ where $Q = D = \text{diag}\,(d_1, d_2, \dots, d_p)$. It is well known that the MLE of the parameter θ is X and its risk function associated to the loss function L_D is $\sum\limits_{i=1}^{p} d_i = \text{Tr}\,(D)$. Endeed

$$R_D\,(X, \theta) = E\,(L_D\,(X, \theta)) = E\left(\sum_{i=1}^{p} d_i\,(X_i - \theta_i)^2\right) = \sum_{i=1}^{p} d_i E\,(X_i - \theta_i)^2 = \text{Tr}\,(D)\,,$$

because for any i $(i = 1, \dots, p)$, $(X_i - \theta_i)^2 \sim \chi_1^2$ where χ_1^2 is the chi-square distribution with 1 degrees of freedom, then $E_\theta(X_i - \theta_i)^2 = 1$. It is easy to check that the MLE X is minimax, thus any estimator dominates it, is also minimax.

Next, we suppose that $\underline{K} = (K_1, \dots, K_p)$ where K_i $(i = 1, \dots, p)$ are independent Poisson $P\left(\theta_i^2/2\right)$ and $K = \sum\limits_{i=1}^{p} K_i \left(K \sim P\left(\left\|\theta^2\right\|/2\right)\right)$. We give the following Lemma, that can be used in our proofs and its proof is postponed to the Appendix.

Lemma 2. *Let $X \sim N_p\,(\theta, I_p)$ where $X = (X_1, \dots, X_p)^t$ and $\theta = (\theta_1, \dots, \theta_p)^t$. If $p \geqslant 3$, we have*

i) $$E\left(\frac{X_i^2}{\|X\|^2}\right) = E\left(\frac{1 + \dfrac{2\theta_i^2}{\|\theta\|^2}K}{p + 2K}\right);$$

ii) $$E\left(\frac{X_i^2}{\|X\|^4}\right) = E\left(\frac{1 + \dfrac{2\theta_i^2}{\|\theta\|^2}K}{(p - 2 + 2K)\,(p + 2K)}\right).$$

2.1. Baranchick-type estimators

In this part, we study the minimaxity of Baranchick-type estimator, which is given by

$$\delta_\psi = \left(1 - \frac{\psi\left(\|X\|^2\right)}{\|X\|^2}\right) X. \tag{1}$$

Proposition 1. *The risk function of the estimator defined in* (1) *under the loss function L_D is*

$$R_D(\delta_\psi, \theta) = \text{Tr}(D) + E\left\{\frac{\sum\limits_{i=1}^{p} d_i\left(1 + \dfrac{2\theta_i^2}{\|\theta\|^2}K\right)}{p + 2K}\left[\frac{\psi^2(\chi_{p+2K}^2)}{\chi_{p+2K}^2} - 4\psi'(\chi_{p+2K}^2) + 4\frac{\psi(\chi_{p+2K}^2)}{\chi_{p+2K}^2}\right]\right\} -$$
$$- 2\text{Tr}\,(D)\,E\left(\frac{\psi\left(\chi_{p+2K}^2\right)}{\chi_{p+2K}^2}\right). \tag{2}$$

Proof. We have

$$R_D(\delta_\psi,\theta) = E\left[L_D\left(\delta_\psi,\theta\right)\right] = E\left\{\left(X-\theta-\frac{\psi\left(\|X\|^2\right)}{\|X\|^2}X\right)^t D\left(X-\theta-\frac{\psi\left(\|X\|^2\right)}{\|X\|^2}X\right)\right\} =$$

$$= E\left\{(X-\theta)^t D\,(X-\theta)\right\} + E\left\{\left(\frac{\psi\left(\|X\|^2\right)}{\|X\|^2}X\right)^t D\left(\frac{\psi\left(\|X\|^2\right)}{\|X\|^2}X\right)\right\} -$$

$$- 2E\left\{(X-\theta)^t D\left(\frac{\psi\left(\|X\|^2\right)}{\|X\|^2}X\right)\right\} =$$

$$= \mathrm{Tr}\,(D) + E\left(\frac{\psi^2\left(\|X\|^2\right)\sum\limits_{i=1}^{p} d_i X_i^2}{\|X\|^4}\right) - 2\sum_{i=1}^{p} d_i E\left[(X_i-\theta_i)\left[\frac{\psi\left(\|X\|^2\right)X_i}{\|X\|^2}\right]\right].$$

Using Lemma 1, we obtain

$$R_D(\delta_\psi,\theta) = \mathrm{Tr}\,(D) + E\left(\frac{\psi^2\left(\|X\|^2\right)\sum\limits_{i=1}^{p} d_i X_i^2}{\|X\|^4}\right) - 2\sum_{i=1}^{p} d_i E\left[\frac{\partial}{\partial X_i}\left[\frac{\psi\left(\|X\|^2\right)X_i}{\|X\|^2}\right]\right] =$$

$$= \mathrm{Tr}\,(D) + E\left(\frac{\psi^2\left(\|X\|^2\right)\sum\limits_{i=1}^{p} d_i X_i^2}{\|X\|^4}\right) - 4E\left(\frac{\psi'\left(\|X\|^2\right)\sum\limits_{i=1}^{p} d_i X_i^2}{\|X\|^2}\right) -$$

$$- 2\left(\sum_{i=1}^{p} d_i\right) E\left(\frac{\psi\left(\|X\|^2\right)}{\|X\|^2}\right) + 4E\left(\frac{\psi\left(\|X\|^2\right)\sum\limits_{i=1}^{p} d_i X_i^2}{\|X\|^4}\right) =$$

$$= \mathrm{Tr}\,(D) + E\left\{\frac{\sum\limits_{i=1}^{p} d_i X_i^2}{\|X\|^2}\left[\frac{\psi^2\left(\|X\|^2\right)}{\|X\|^2} - 4\psi'\left(\|X\|^2\right) + 4\frac{\psi\left(\|X\|^2\right)}{\|X\|^2}\right]\right\} -$$

$$- 2\mathrm{Tr}\,(D)\, E\left(\frac{\psi\left(\|X\|^2\right)}{\|X\|^2}\right).$$

From the independence given $\underline{K}$ between $X_i^2/\|X\|^2$ and $\|X\|^2$ for $i = 1,\ldots,p$, we get

$$R_D(\delta_\psi,\theta) = \mathrm{Tr}\,(D) +$$

$$+ E\left\{\sum_{i=1}^{p} d_i E\left(\frac{X_i^2}{\|X\|^2}|\underline{K}\right) E\left[\left(\frac{\psi^2\left(\|X\|^2\right)}{\|X\|^2} - 4\psi'\left(\|X\|^2\right) + 4\frac{\psi\left(\|X\|^2\right)}{\|X\|^2}\right)|\underline{K}\right]\right\} -$$

$$- 2\left(\sum_{i=1}^{p} d_i\right) E\left(\frac{\psi\left(\|X\|^2\right)}{\|X\|^2}\right).$$

Using the Lemma 2, we have

$$R_D(\delta_\psi,\theta) = \mathrm{Tr}\,(D) + \\ + E\left\{ \frac{\sum_{i=1}^{p} d_i\left(1+\frac{2\theta_i^2}{\|\theta\|^2}K\right)}{p+2K} E\left[\frac{\psi^2\left(\|X\|^2\right)}{\|X\|^2} - 4\psi'\left(\|X\|^2\right) + 4\frac{\psi\left(\|X\|^2\right)}{\|X\|^2}|\underline{K}\right]\right\} - \\ - 2\mathrm{Tr}\,(D)\,E\left(\frac{\psi\left(\|X\|^2\right)}{\|X\|^2}\right) = \\ = \mathrm{Tr}\,(D) - 2\mathrm{Tr}\,(D)\,E\left(\frac{\psi\left(\|X\|^2\right)}{\|X\|^2}\right) + \\ + E\left\{E\left\{\frac{\sum_{i=1}^{p} d_i\left(1+\frac{2\theta_i^2}{\|\theta\|^2}K\right)}{p+2K}\left[\frac{\psi^2\left(\|X\|^2\right)}{\|X\|^2} - 4\psi'\left(\|X\|^2\right) + 4\frac{\psi\left(\|X\|^2\right)}{\|X\|^2}|\underline{K}\right]\right\}\right\}.$$

From Definition 1 and using properties of conditional expectation we have, for any two measurable functions G and H, $E\left[G\left(\|X\|^2\right)\right] = E\left[G\left(\chi^2_{p+2K}\right)\right]$ and $E\left\{E\left[H\left(K\right)|\underline{K}\right]\right\} = E\left[H\left(K\right)\right]$, where $K \sim P\left(\|\theta\|^2/2\right)$, thus we get the desired result. □

Note that the classical result of minimaxity of Baranchick-type estimators which is obtained for the loss function $L\left(\delta,\theta\right) = \sum_{i=1}^{p}\left(\delta_i-\theta_i\right)^2$ (i.e. $d_i = 1$ for any $i = 1,\ldots,p$), is also available and it is established in the following Theorem.

Theorem 1. *Assume that δ_ψ is given in (1) with $p \geqslant 3$. Under the loss function L_D with $\frac{\mathrm{Tr}\,(D)}{\max_{1\leqslant i\leqslant p}\,(d_i)} \geqslant 2$, if*

i) $\psi\,(.)$ is monotone non-decreasing function;

ii) $0 \leqslant \psi\,(.) \leqslant 2\left(\frac{\mathrm{Tr}\,(D)}{\max_{1\leqslant i\leqslant p}\,(d_i)} - 2\right)$,

then δ_ψ is minimax.

Proof. From formula (2), we have

$$R_D(\delta_\psi,\theta) \leqslant \mathrm{Tr}\,(D) + \\ + E\left\{\frac{-4\psi'\left(\chi^2_{p+2K}\right)\sum_{i=1}^{p} d_i\left(1+\frac{2\theta_i^2}{\|\theta\|^2}K\right)}{p+2K} + \max_{1\leqslant i\leqslant p}\,(d_i)\left[\frac{\psi^2\left(\chi^2_{p+2K}\right)+4\psi\left(\chi^2_{p+2K}\right)}{\chi^2_{p+2K}}\right]\right\} - \\ - 2\mathrm{Tr}\,(D)\,E\left(\frac{\psi\left(\chi^2_{p+2K}\right)}{\chi^2_{p+2K}}\right) \leqslant$$

$$\leqslant \mathrm{Tr}\,(D) + E\left[\frac{-4\psi'\left(\chi^2_{p+2K}\right)\sum_{i=1}^{p} d_i\left(1+\frac{2\theta_i^2}{\|\theta\|^2}K\right)}{p+2K}\right] +$$

$$+ E\left\{\frac{\psi\left(\chi^2_{p+2K}\right)}{\chi^2_{p+2K}}\left[\max_{1\leqslant i\leqslant p}(d_i)\left[\psi\left(\chi^2_{p+2K}\right)+4\right]-2\mathrm{Tr}\,(D)\right]\right\}.$$

Then, a sufficient condition for that δ_ψ is minimax is that $\psi\,(.)$ is a positive monotone non-decreasing function and $\max_{1\leqslant i\leqslant p}(d_i)\left[\psi\left(\chi^2_{p+2K}\right)+4\right]-2\mathrm{Tr}\,(D)\leqslant 0$. Which are equivalent to

$$0\leqslant \psi\,(.)\leqslant 2\left(\frac{\mathrm{Tr}\,(D)}{\max_{1\leqslant i\leqslant p}(d_i)}-2\right)$$

and $\psi\,(.)$ is monotone non-decreasing. □

Example 1. Let the shrinkage functions $\psi^{(1)}\left(\|X\|^2\right) = \|X\|^2/\left(\|X\|^2+1\right)$, $\psi^{(2)}\left(\|X\|^2\right) = = 1-exp(-\|X\|^2)$ and the matrices $D^{(1)} = \mathrm{diag}\,(d_1 = 1, d_2 = 1/2, \ldots, d_p = 1/p)$ with $p \geqslant 7$ and $D^{(2)} = \mathrm{diag}\,(d_1 = 1/2, d_2 = 2/3, \ldots, d_p = p/p+1)$ with $p \geqslant 4$. It is clear that the functions $\psi^{(1)}\,(.)$ and $\psi^{(2)}\,(.)$ satisfie conditions of Theorem 1. Then the estimators $\delta_{\psi^{(1)}}$ and $\delta_{\psi^{(2)}}$ are minimax for $p \geqslant 7$ under the loss function $L_D^{(1)}$ and are minimax for $p \geqslant 4$ under the loss function $L_D^{(2)}$.

Now, we discuss the special case where $\psi\,(.) = a$ with a is a positive constant.

2.2. James-Stein estimator

Consider the estimator $\delta_a = \left(1-a/\|X\|^2\right)X = X-\left(a/\|X\|^2\right)X$, where a is a real parameter that can depend on p. Using the Proposition 1, the risk function of the estimator δ_a is

$$R_D(\delta_a,\theta) = \mathrm{Tr}\,(D) + a\,(a+4)\,E\left(\frac{\sum_{i=1}^{p} d_i\left(1+\frac{2\theta_i^2}{\|\theta\|^2}K\right)}{(p-2+2K)\,(p+2K)}\right) - 2a\mathrm{Tr}\,(D)\,E\left(\frac{1}{p-2+2K}\right). \quad (3)$$

Proposition 2. *Under the loss function L_D with $p \geqslant 3$ and $\frac{\mathrm{Tr}\,(D)}{\max_{1\leqslant i\leqslant p}(d_i)} \geqslant 2$, we have*

i) a sufficient condition for that δ_a dominates the MLE X is

$$0\leqslant a\leqslant 2\left(\frac{\mathrm{Tr}\,(D)}{\max_{1\leqslant i\leqslant p}(d_i)}-2\right);$$

ii) the optimal value of a that minimizes the risk function $R_D(\delta_a,\theta)$ is

$$\widehat{a} = \frac{\mathrm{Tr}\,(D)\,E\left(\frac{1}{(p-2+2K)}\right)}{\alpha} - 2,$$

where $\alpha = E\left(\sum\limits_{i=1}^{p} d_i\left(1+\left(2\theta_i^2/\|\theta\|^2\right)K\right)/(p-2+2K)\,(p+2K)\right)$.

Proof. i) From formula (3), a sufficient condition so that δ_a dominating the MLE X is

$$a\,(a+4)\,E\left(\frac{\sum\limits_{i=1}^{p} d_i\left(1+\dfrac{2\theta_i^2}{\|\theta\|^2}K\right)}{(p-2+2K)\,(p+2K)}\right)-2a\mathrm{Tr}\,(D)\,E\left(\frac{1}{p-2+2K}\right)\leqslant 0.$$

As

$$E\left(\frac{\sum\limits_{i=1}^{p} d_i\left(1+\dfrac{2\theta_i^2}{\|\theta\|^2}K\right)}{(p-2+2K)\,(p+2K)}\right)\leqslant \max_{1\leqslant i\leqslant p}(d_i)\,E\left(\frac{1}{p-2+2K}\right),$$

thus, a sufficient condition so that δ_a dominates the MLE X is

$$a\left[(a+4)\max_{1\leqslant i\leqslant p}(d_i)-2\mathrm{Tr}\,(D)\right]E\left(\frac{1}{p-2+2K}\right)\leqslant 0,$$

which is equivalent to the desired result.

ii) Using the convexity of the risk function $R_D(\delta_a,\theta)$ on a, one can easily show that the optimal value of a that minimizes the risk function $R_D(\delta_a,\theta)$ is $\widehat{a}=(\mathrm{Tr}\,(D)\,E\,(1/(p-2+2K))/\alpha)-2$, where $\alpha = E\left(\sum\limits_{i=1}^{p} d_i\left(1+\left(2\theta_i^2/\|\theta\|^2\right)K\right)/(p-2+2K)\,(p+2K)\right)$. □

For $a=\widehat{a}$ we obtain the James-Stein estimator $\delta_{JS}\left(=\delta_{\widehat{a}}=\left(1-\widehat{a}/\|X\|^2\right)X\right)$ which minimizes the risk function of estimators δ_a, so that from formula (3), the risk function of the James-Stein estimator δ_{JS} under the loss function L_D is

$$R_D(\delta_{JS},\theta)=\mathrm{Tr}\,(D)-\frac{\left[\mathrm{Tr}\,(D)\,E\left(\dfrac{1}{p-2+2K}\right)-2\alpha\right]^2}{\alpha}, \tag{4}$$

As the constant α is non-negative and using the formula (4), it is clear that the James-Stein estimator δ_{JS}, has a risk less than $\mathrm{Tr}\,(D)$, then δ_{JS} is minimax.

3. The case of generalized Σ and Q

Let $X\sim N_p(\theta,\Sigma)$ and the loss function $L_Q(\delta,\theta)=(\delta-\theta)^t\,Q\,(\delta-\theta)$ where the covariance matrix Σ is known and $\Sigma^{1/2}Q\Sigma^{1/2}$ is diagonalizable matrix. Take the change of variables $Y=P\Sigma^{-1/2}X$ where P is an orthogonal matrix ($PP^t=I_p$) that diagonalizes the matrix $\Sigma^{1/2}Q\Sigma^{1/2}$ such as $P\Sigma^{1/2}Q\Sigma^{1/2}P^t=D^*=\mathrm{diag}\,(a_1,\ldots,a_p)$. Then we have $Y\sim N_p(\nu,I_p)$ with $\nu=P\Sigma^{-1/2}\theta$. Thus the risk function of the MLE X associated to the loss function L_Q is $\sum\limits_{i=1}^{p} a_i=\mathrm{Tr}\,(D^*)$. Endeed

$$R_Q(X,\theta)=E\left[(X-\theta)^t\,Q\,(X-\theta)\right]=E\left\{\left[\Sigma^{1/2}P^{-1}(Y-\nu)\right]^t Q\left[\Sigma^{1/2}P^{-1}(Y-\nu)\right]\right\}=$$
$$=E\left\{(Y-\nu)^t\,P\Sigma^{1/2}Q\Sigma^{1/2}P^t\,(Y-\nu)\right\}=E\left\{(Y-\nu)^t\,D^*\,(Y-\nu)\right\}=$$
$$=\sum_{i=1}^{p}a_iE\left[(Y_i-\nu_i)^2\right]=\mathrm{Tr}\,(D^*),$$

because for any i $(i = 1, \dots, p)$ $(Y_i - \nu_i)^2 \sim \chi_1^2$ where χ_1^2 is the chi-square distribution with 1 degrees of freedom, thus $E\,(Y_i - \nu_i)^2 = 1$. As the MLE X is minimax, then any estimator dominates it, is also minimax.

3.1. Baranchik-type estimators

Now, consider the estimator given by

$$\delta_\phi = \left(1 - \frac{\phi\left(X^t\Sigma^{-1}X\right)}{X^t\Sigma^{-1}X}\right)X. \tag{5}$$

Proposition 3. *Under the loss function L_Q the risk function of the estimator $\delta_\phi$is*

$$R_Q(\delta_\phi, \theta) = \mathrm{Tr}\left(D^*\right) + E\left\{\frac{\sum\limits_{i=1}^{p} a_i\left(1 + \dfrac{2\theta_i^2}{\|\theta\|^2}K\right)}{p + 2K}\left[\frac{\phi^2\left(\chi^2_{p+2K}\right)}{\chi^2_{p+2K}} - 4\phi'\left(\chi^2_{p+2K}\right) + 4\frac{\phi\left(\chi^2_{p+2K}\right)}{\chi^2_{p+2K}}\right]\right\} -$$
$$- 2\mathrm{Tr}\left(D^*\right)E\left(\frac{\phi\left(\chi^2_{p+2K}\right)}{\chi^2_{p+2K}}\right),$$

where $K \sim P\left(\|\nu\|^2/2\right)$.

Proof.

$$R_Q(\delta_\phi, \theta) = E\left\{\left[\left(1 - \frac{\phi\left(X^t\Sigma^{-1}X\right)}{X^t\Sigma^{-1}X}\right)X - \theta\right]^t Q\left[\left(1 - \frac{\phi\left(X^t\Sigma^{-1}X\right)}{X^t\Sigma^{-1}X}\right)X - \theta\right]\right\} =$$
$$= E\left\{\left[(X - \theta) - \frac{\phi\left(X^t\Sigma^{-1}X\right)}{X^t\Sigma^{-1}X}X\right]^t Q\left[(X - \theta) - \frac{\phi\left(X^t\Sigma^{-1}X\right)}{X^t\Sigma^{-1}X}X\right]\right\}.$$

Using the change variable $Y = P\Sigma^{-1/2}X$ where P is an orthogonal matrix and P diagonalizes the matrix $\Sigma^{1/2}Q\Sigma^{1/2}$, then $Y \sim N_p\left(\nu, I_p\right)$ with $\nu = P\Sigma^{-1/2}\theta$ and

$$R_Q(\delta_\phi, \theta) = E\left\{\left[\Sigma^{1/2}P^{-1}\left[(Y - \nu) - \frac{\phi\left(Y^tY\right)}{Y^tY}Y\right]\right]^t Q\left[\Sigma^{1/2}P^{-1}\left[(Y - \nu) - \frac{\phi\left(Y^tY\right)}{Y^tY}Y\right]\right]\right\} =$$
$$= E\left\{\left[(Y - \nu) - \frac{\phi\left(Y^tY\right)}{Y^tY}Y\right]^t \left(P^{-1}\right)^t\Sigma^{1/2}Q\Sigma^{1/2}P^{-1}\left[(Y - \nu) - \frac{\phi\left(Y^tY\right)}{Y^tY}Y\right]\right\} =$$
$$= E\left\{\left[(Y - \nu) - \frac{\phi\left(\|Y\|^2\right)}{\|Y\|^2}Y\right]^t P\Sigma^{1/2}Q\Sigma^{1/2}P^t\left[(Y - \nu) - \frac{\phi\left(\|Y\|^2\right)}{\|Y\|^2}Y\right]\right\} =$$
$$= E\left\{\left[(Y - \nu) - \frac{\phi\left(\|Y\|^2\right)}{\|Y\|^2}Y\right]^t D^*\left[(Y - \nu) - \frac{\phi\left(\|Y\|^2\right)}{\|Y\|^2}Y\right]\right\} = R_{D^*}(\delta^*_\phi, \theta),$$

where $\|.\|$ is the usual euclidean norm in $\mathbb{R}^p$, $P\Sigma^{1/2}Q\Sigma^{1/2}P^t = D^* = \mathrm{diag}\left(a_1, \dots, a_p\right)$ and $\delta^*_\phi = \left(1 - (\phi(\|Y\|^2/\|Y\|^2)\right)Y$. From Proposition 1, we obtain the desired result. □

Theorem 2. *Assume that δ_ϕ is given by (5) where $p \geqslant 3$. Under the loss function L_Q with $\dfrac{\mathrm{Tr}(D^*)}{\max\limits_{1\leqslant i\leqslant p}(a_i)} \geqslant 2$, if*

i) $\phi(.)$ *is monotone non-decreasing;*

ii) $0 \leqslant \phi(.) \leqslant 2\left(\dfrac{\mathrm{Tr}(D^*)}{\max\limits_{1\leqslant i\leqslant p}(a_i)} - 2\right)$,

then δ_ϕ is minimax.

The proof is the same given for the Theorem 1.

3.2. James-Stein estimator

Consider the estimator $\delta_b = \left(1 - b/\left(X^t\Sigma^{-1}X\right)\right)X$. Using the Proposition 3, one can show easily that the risk function of the estimator δ_b under the loss function L_Q is.

$$R_Q(\delta_b, \theta) = \mathrm{Tr}(D^*) + b(b+4)E\left(\frac{\sum\limits_{i=1}^{p} a_i\left(1 + \dfrac{2\theta_i^2}{\|\theta\|^2}K\right)}{(p-2+2K)(p+2K)}\right) - 2b\mathrm{Tr}(D^*)E\left(\frac{1}{(p-2+2K)}\right),$$

where $K \sim P\left(\|\nu\|^2/2\right)$. From the last formula, we deduce immediately that, a sufficient condition for that δ_b dominating the MLE X is $0 \leqslant b \leqslant 2\left((\mathrm{Tr}(D^*)/\max\limits_{1\leqslant i\leqslant p} a_i) - 2\right)$, and the optimal value of b that minimizes the risk function $R_Q(\delta_b, \theta)$ is

$$\widehat{b} = \frac{\mathrm{Tr}(D^*)E\left(\dfrac{1}{(p-2+2K)}\right)}{\beta} - 2,$$

where $\beta = E\left(\sum\limits_{i=1}^{p} a_i\left(1 + \left(2\theta_i^2/\|\theta\|^2\right)K\right)/(p-2+2K)(p+2K)\right)$.

For $b = \widehat{b}$ we obtain the James-Stein estimator $\delta_{JS}^* = \delta_{\widehat{b}} = \left(1 - \widehat{b}/\left(X^t\Sigma^{-1}X\right)\right)X$ which minimizes the risk function of δ_b. Its risk function associated to the loss function L_Q is

$$R_Q(\delta_{JS}^*, \theta) = \mathrm{Tr}(D^*) - \frac{\left[\mathrm{Tr}(D^*)E\left(\dfrac{1}{p-2+2K}\right) - 2\beta\right]^2}{\beta}. \tag{6}$$

From formula (6), we note that δ_{JS}^* dominates the MLE X, thus δ_{JS}^* is minimax.

4. The simulation results

In this section we take the model $X \sim N_p(\theta, I_p)$ where $\theta = (\theta_1, \theta_1, \ldots, \theta_1)^t$ and we recall the estimators of Baranchick-type and the matrices $D^{(1)}$ and $D^{(2)}$ given in Example 1, i.e., $\delta_{\psi^{(1)}} = \left(1 - \psi^{(1)}\left(\|X\|^2\right)/\|X\|^2\right)X$, $\delta_{\psi^{(2)}} = \left(1 - \psi^{(2)}\left(\|X\|^2\right)/\|X\|^2\right)X$ with $\psi^{(1)}\left(\|X\|^2\right) = \|X\|^2/\left(\|X\|^2 + 1\right)$, $\psi^{(2)}\left(\|X\|^2\right) = 1 - exp\left(-\|X\|^2\right)$, $D^{(1)} = \mathrm{diag}(d_1 = 1, d_2 = 1/2, \ldots, d_p = 1/p)$

and $D^{(2)} = \mathrm{diag}\,(d_1 = 1/2, d_2 = 2/3, \ldots, d_p = p/p+1)$. We also recall the form of the James-Stein estimator $\delta_{JS}\left(= \delta_{\widehat{a}} = \left(1 - \widehat{a}\,/\,\|X\|^2\right) X\right)$, where $\widehat{a} = (\mathrm{Tr}\,(D)\,E\,(1/\,(p-2+2K))/\alpha) - 2$ and $\alpha = E\left(\sum_{i=1}^{p} d_i \left(1 + \left(2\theta_i^2/\|\theta\|^2\right) K\right)/(p-2+2K)\,(p+2K)\right)$. We graph the risks ratios of estimators cited above, to the MLE associated the the losses functions $L_{D^{(1)}}$ and $L_{D^{(2)}}$ denoted respectively: $R(\delta_{JS}, \theta)/R(X, \theta)$, $R(\delta_{\psi^{(1)}}, \theta)/R(X, \theta)$ and $R(\delta_{\psi^{(2)}}, \theta)/R(X, \theta)$ as function of $\lambda = \theta_1^2$ for various values of p.

In Figs. 1–4, we note that the risks ratios $R(\delta_{JS}, \theta)/R(X, \theta)$, $R(\delta_{\psi^{(1)}}, \theta)/R(X, \theta)$ and $R(\delta_{\psi^{(2)}}, \theta)/R(X, \theta)$ are less than 1, thus the estimators δ_{JS}, $\delta_{\psi^{(1)}}$ and $\delta_{\psi^{(2)}}$ are minimax for $p = 8$ and $p = 12$ under the loss function $L_{D^{(1)}}$, and also minimax for $p = 4$ and $p = 6$ under the loss function $L_{D^{(2)}}$.

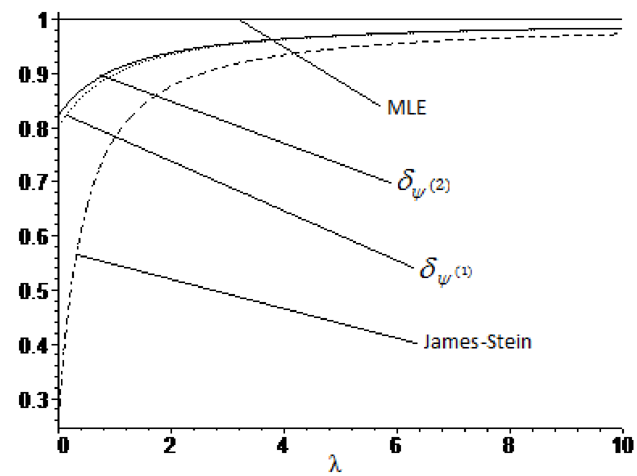

Fig. 1. Graph of risks ratios $R(\delta_{JS}, \theta)/R(X, \theta)$, $R(\delta_{\psi^{(1)}}, \theta)/R(X, \theta)$ and $R(\delta_{\psi^{(2)}}, \theta)/R(X, \theta)$ as function of $\lambda = \theta_1^2$ for $p = 8$ under the loss function $L_{D^{(1)}}$

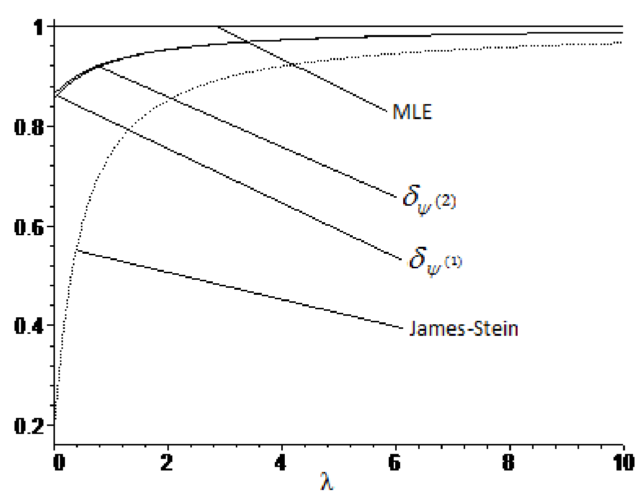

Fig. 2. Graph of risks ratios $R(\delta_{JS}, \theta)/R(X, \theta)$, $R(\delta_{\psi^{(1)}}, \theta)/R(X, \theta)$ and $R(\delta_{\psi^{(2)}}, \theta)/R(X, \theta)$ as function of $\lambda = \theta_1^2$ for $p = 12$ under the loss function $L_{D^{(1)}}$

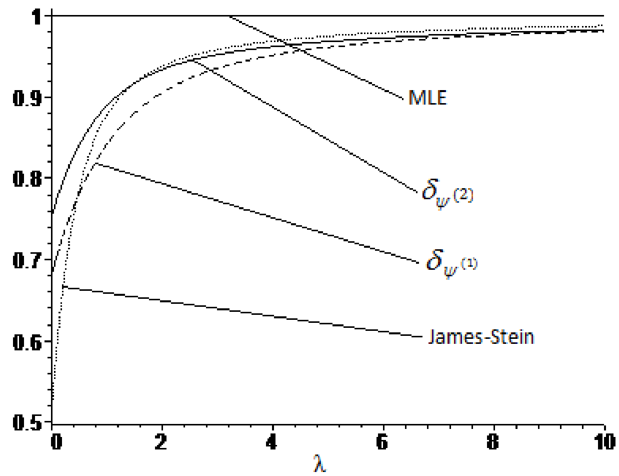

Fig. 3. Graph of risks ratios $R(\delta_{JS}, \theta)/R(X, \theta)$, $R(\delta_{\psi^{(1)}}, \theta)/R(X, \theta)$ and $R(\delta_{\psi^{(2)}}, \theta)/R(X, \theta)$ as function of $\lambda = \theta_1^2$ for $p = 4$ under the loss function $L_{D^{(2)}}$

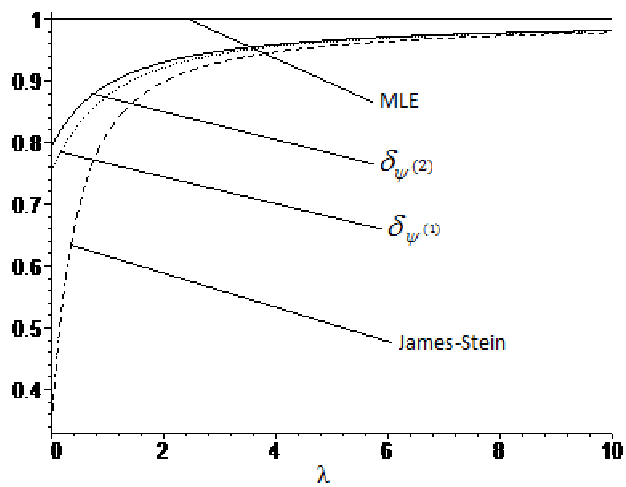

Fig. 4. Graph of risks ratios $R(\delta_{JS}, \theta)/R(X, \theta)$, $R(\delta_{\psi^{(1)}}, \theta)/R(X, \theta)$ and $R(\delta_{\psi^{(2)}}, \theta)/R(X, \theta)$ as function of $\lambda = \theta_1^2$ for $p = 6$ under the loss function $L_{D^{(2)}}$

5. Appendix

Lemma 3 (Bock [5]). *Let $X \sim N_p(\theta, I_p)$ where $X = (X_1, \ldots, X_p)^t$ and $\theta = (\theta_1, \ldots, \theta_p)^t$, then, For any measurable function $h : [0, +\infty[\to \mathbb{R}$, we have*

$$E\left(h\left(\|X\|^2\right) X_i^2\right) = E\left[h\left(\chi^2_{p+2}\left(\|\theta\|^2\right)\right)\right] + \theta_i^2 E\left[h\left(\chi^2_{p+4}\left(\|\theta\|^2\right)\right)\right].$$

where $K \sim P\left(\|\theta\|^2/2\sigma^2\right)$ being the Poisson's distribution of parameter $\|\theta\|^2/2\sigma^2$.

Lemma 4 (Bock [5])**.** *Let f be a real-valued mesurable function defined on the integer. Let $K \sim P(\lambda/2)$ being the Poison's distribution of parameter $\lambda/2$. Then*

$$\lambda E[f(K)] = E[2Kf(K-1)],$$

if both sides exist.

Proof Lemma 2. i) Using Lemma 3 and the Definition 1, we obtain

$$E\left(\frac{X_i^2}{\|X\|^2}\right) = E_{\chi^2_{p+2}(\|\theta\|^2)}\left(\frac{1}{u}\right) + \theta_i^2 E_{\chi^2_{p+4}(\|\theta\|^2)}\left(\frac{1}{u}\right) = E\left(\frac{1}{p+2K}\right) + \theta_i^2 E\left(\frac{1}{p+2+2K}\right),$$

where $K \sim P\left(\|\theta\|^2/2\right)$ being the Poisson's distribution of parameter $\|\theta\|^2/2$.
From Lemma 4, we have

$$E\left(\frac{X_i^2}{\|X\|^2}\right) = E\left(\frac{1}{p+2K}\right) + \frac{\theta_i^2}{\|\theta\|^2}E\left(\frac{2K}{p+2K}\right) = E\left(\frac{1+2\dfrac{\theta_i^2}{\|\theta\|^2}K}{p+2K}\right).$$

ii) Using Lemma 3 and the Definition 1, we obtain

$$\begin{aligned} E\left(\frac{X_i^2}{\|X\|^4}\right) &= E_{\chi^2_{p+2}(\|\theta\|^2)}\left(\frac{1}{u^2}\right) + \theta_i^2 E_{\chi^2_{p+4}(\|\theta\|^2)}\left(\frac{1}{u^2}\right) = \\ &= E\left(\frac{1}{(p-2+2K)(p+2K)}\right) + \theta_i^2 E\left(\frac{1}{(p+2K)(p+2+2K)}\right), \end{aligned}$$

where $K \sim P\left(\|\theta\|^2/2\right)$. From Lemma 4, we have

$$\begin{aligned} E\left(\frac{X_i^2}{\|X\|^2}\right) &= E\left(\frac{1}{(p-2+2K)(p+2K)}\right) + \frac{\theta_i^2}{\|\theta\|^2}E\left(\frac{2K}{(p-2+2K)(p+2K)}\right) = \\ &= E\left(\frac{1+2\dfrac{\theta_i^2}{\|\theta\|^2}K}{(p-2+2K)(p+2K)}\right). \end{aligned}$$

□

Conclusion

Stein [10], has started to study the estimation of the mean θ of a multivariate gaussian random $N_p\left(\theta, \sigma^2 I_p\right)$ in $\mathbb{R}^p$, by the shrinkage estimators deduced from the usual estimator. Many authors continued to work in this field. The majority among them have studied the minimaxity of these estimators under the usual quadratic risk function, we cite for example [5,7]. Other authors research the stability of the minimaxity property in the case where the dimension of the parameter space and the sample size are large, we refer to [3,6]. In this work we studied the minimaxity of Baranchick-type estimators, relatively to the general loss function. We showed similar results to those found in the classical case. An idea would be to see whether one can obtain similar results of the minimaxity and the asymptotic behavior of risk ratios in the general case of the symmetrical spherical models.

The authors are extremely grateful to the editor and the referees for carefully reading the paper and for valuable comments and helpful remarks and suggestions which greatly improved the presentation of this paper. This research is supported by the Thematic Research Agency in Science and Technology (ATRST-Algeria).

References

[1] A.J.Baranchick, Multiple Regression and estimation of the mean of a multivariate normal distribution, Stanford Univ., Technical Report, no. 51, 1964.

[2] A.J.Baranchick, A family of minimax estimators of the mean of a multivariate normal distribution, *The Annals of Mathematical Statistics*, **41**(1970), no. 2, 642–645.

[3] A.Benkhaled, A.Hamdaoui, General classes of shrinkage estimators for the multivariate normal mean with unknown variance: minimaxity and limit of risks ratios, *Kragujevac J. Math.*, **46**(2019), no. 2, 193–213.

[4] D.Benmansour, A.Hamdaoui, Limit of the ratio of risks of James-Stein estimators with unknown variance, *Far East J. Theo. Stat.*, **36**(2011), no. 1, 31–53.

[5] M.E.Bock, Minimax estimators of the mean of a multivariate normal distribution, *Ann. Statist.*, **3**(1975), no. 1, 209–218.

[6] A.Hamdaoui, D.Benmansour, Asymptotic properties of risks ratios of shrinkage estimators, *Hacet. J. Math. Stat.*, **44**(2015), no. 5, 1181–1195.

[7] W.James, C.Stein, Estimation with Quadratic Loss, Proc. 4th Berkeley Symp, Math. Statist. Prob, Univ of California Press, Berkeley, Vol. 1, 1961, 361–379.

[8] K.Selahattin, D.Issam, The optimal extended balanced loss function estimators, *J. Comput. Appl. Math.*, **345**(2019), 86–98.

[9] K.Selahattin, S.Sadullah, M.RÖzkale, H.Güler, More on the restricted ridge regression estimation, *J. Stat. Comput. Simul.*, **81**(2011), 1433–1448.

[10] C.Stein, Inadmissibilty of the usual estimator for the mean of a multivariate normal distribution, Proc. 3th Berkeley Symp, Math. Statist. Prob. Univ. of California Press, Berkeley, Vol. 1, 1956, 197–206.

[11] C.Stein, Estimation of the mean of a multivariate normal distribution, *Ann. Statis.*, **9**(1981), no. 6, 1135–1151.

[12] W.E.Strawderman, Minimax estimation of location parameters for certain spherically symmetric distribution, *J. Multivariate Anal.*, **4**(1974), no. 1, 255–264.

[13] X.Xie, S.C.Kou, L.Brown, Optimal shrinkage estimators of mean parameters in family of distribution with quadratic variance, *Ann. Statis.*, **44**(2016), no. 2, 564–597.

Об оценках решений задачи расщепления для некоторых многомерных дифференциальных уравнений в частных производных

Абденур Хамдауи
Университет науки и технологии Оран Мохамед-Будиаф
Оран, Алжир
Университет Тлемсена (LSMA)
Тлемсен, Алжир
Абделькадер Бенхалед
Университет Мюстафа Стамбули (LGEO2E)
Маскара, Алжир
Мекки Тербече
Университет науки и технологии Оран Мохамед-Будиаф (LAAR), USTO-MB
Оран, Алжир

Аннотация. Исследована проблема оценки среднего многомерного нормального распределения различными типами оценок усадки. Мы установили минимаксность оценок типа Баранчика для единичной ковариационной матрицы, а матрица, связанная с функцией потерь, является диагональной. В частности, представлен класс оценки Джеймса-Стейна. Обсуждается общая ситуация для обеих упомянутых выше матриц.

Ключевые слова: ковариационная матрица, оценка Джеймса-Стейна, функция потерь, многомерная гауссовская случайная величина, нецентральное распределение хи-квадрат, оценка усадки.

Journal of Siberian Federal University. Mathematics & Physics 2020, 13(5), 622–630

DOI: 10.17516/1997-1397-2020-13-5-622-630
УДК 517.53

On the Differentiation in the Privalov Classes

Eugenia G. Rodikova*
Bryansk State University
Bryansk, Russian Federation
Faizo A. Shamoyan†
Saratov State University
Saratov, Russian Federation

Received 22.09.2019, received in revised form 25.04.2020, accepted 26.07.2020

Abstract. The invariance of the Privalov classes with respect to the differentiation operator is studied.

Keywords: Privalov spaces, the Bloch-Nevanlinna conjecture, differentiation operator.

Citation: E.G. Rodikova, F.A. Shamoyan, On the Differentiation in the Privalov Classes, J. Sib. Fed. Univ. Math. Phys., 2020, 13(5), 622–630.
DOI: 10.17516/1997-1397-2020-13-5-622-630.

Introduction

Let $\mathbb{C}$ be the complex plane, D be the unit disk on $\mathbb{C}$, $H(D)$ be the set of all functions, holomorphic in D. For all $0 < q < +\infty$ we define the Privalov class of function Π_q as follows (see [11]):

$$\Pi_q = \left\{ f \in H(D) : \sup_{0 \leqslant r < 1} \frac{1}{2\pi} \int_{-\pi}^{\pi} \left(\ln^+ |f(re^{i\theta})|\right)^q d\theta < +\infty \right\}.$$

$\ln^+ |a| = \max(\ln |a|, 0)$, $\forall a \in \mathbb{C}$.

The classes Π_q were first considered by I. I. Privalov in [11]. If $q = 1$ the Privalov class coincides with the Nevanlinna class N of analytic functions in D with bounded characteristic $T(r, f) = \frac{1}{2\pi} \int\limits_{-\pi}^{\pi} \ln^+ |f(re^{i\theta})| d\theta$, $0 \leqslant r < 1$. This is well-known in scientific literature (see [9]). Using Hölder's inequality, it is easy to prove the inclusion chain:

$$\Pi_q \,(q > 1) \subset N \subset \Pi_q \,(0 < q < 1).$$

Since for all $0 < q < q'$

$$(\ln^+ |f|)^q < (\ln^+ |f| + 1)^q < (\ln^+ |f| + 1)^{q'} < 2^{q'} \cdot \left((\ln^+ |f|)^{q'} + 1 \right),$$

we have

$$\Pi_{q'} \subset \Pi_q.$$

In the case of $1 \leqslant q < +\infty$ the Privalov spaces were studied by M. Stoll, V. I. Gavrilov, A. V. Subbotin, D. A. Efimov, R. Mestrovic, Z. Pavicevic, etc. The monograph [6] contains a brief overview of their results. Certain results were extended to the case $0 < q < 1$ by the first author of this paper (see [13]). Notice that the case $0 < q < 1$ was little studied. The questions

*evheny@yandex.ru
†shamoyanfa@yandex.ru

of interpolation in the Privalov classes, as well as properties of root sets of analytic functions from these classes were investigated in recent works by the authors (see [14–16, 20]).

In this paper we study a question of the invariance of the classes Π_q with respect to the differentiation operator. In other words, we verify the validity of the Bloch-Nevanlinna conjecture in the Privalov spaces.

The assumption, known as the Bloch-Nevanlinna conjecture, was clearly formulated by Nevanlinna in 1929 (see [9]) as follows: a derivative of any analytic function in the unit disk with bounded characteristic is a function of bounded characteristic.

The famous result refuting this hypothesis belongs to O. Frostman (see [5]). He proved that there is a Blaschke product whose derivative is not a function with a bounded characteristic.

Subsequently, many counterexamples that refute the Bloch-Nevanlinna conjecture were constructed in the works of others such as H. Fried (1946), W. Rudin (1955), W. Hayman (1964), P. Duren (1969), J. Anderson (1971), L.-Sh. Khan (1972), et. al. D. Campbell and G. Weeks [1] provide a brief overview of these results, as well as a general approach to the construction of such examples.

The invariance with respect to the integro-differential operators of other classes of analytic functions have been studied by many mathematicians. A brief overview of their results is contained in the work of S. V. Shvedenko [22]. In particular, a closure of the classes of analytic functions in a disk with the restrictions on Nevanlinna's characteristic function regarding the operations of differentiation and integration was studied by F. A. Shamoyan, I. S. Kursina, V. A. Bednazh (see [19]).

We state the Bloch-Nevanlinna conjecture in the Privalov spaces: for whatever $q > 0$, the derivative of a function from the class Π_q belongs to the class Π_q.

The paper is organized as follows. In the first part of the article we refute the Bloch-Nevanlinna conjecture in the Privalov spaces for all $0 < q < +\infty$. In the second part of the article we indicate the class to which the derivative of any function from the Privalov space belongs.

1. The Bloch-Nevanlinna conjecture for the Privalov spaces

The following statement is true.

Theorem 1.1. *The Bloch-Nevanlinna conjecture fails in the spaces* Π_q, $0 < q < +\infty$.

In other words, the Privalov spaces Π_q are not invariant under the differentiation operator for all $0 < q < +\infty$, not only for $q = 1$.

In the sequel, unless otherwise noted, we denote by $c, c_1, \dots, c_n(\alpha, \beta, \dots)$ some arbitrary positive constants depending on $\alpha, \beta, \dots$, whose specific values are immaterial.

Proof of this statement reproduces the arguments from [21], the method goes back to the work of Hayman [8].

Let λ be a sufficiently large positive integer, $0 < \alpha < 1$, H^∞ be the class of bounded analytic functions in D. We define a function f_λ as follows:

$$f_\lambda = \sum_{k=0}^{+\infty} \lambda^{-k(1-\alpha)} z^{\lambda^k}.$$

It is obvious that $f_\lambda \in H(D)$, and $|f_\lambda| \leqslant \sum\limits_{k=0}^{+\infty} \lambda^{-k(1-\alpha)} = \dfrac{\lambda^{1-\alpha}}{\lambda^{1-\alpha} - 1}$, that is $f_\lambda \in H^\infty$. Since $H^\infty \subset \Pi_q$, we have $f_\lambda \in \Pi_q$ for all $0 < q < +\infty$.

In the same time we have

$$f'_\lambda = \sum_{k=0}^{+\infty} \lambda^{\alpha k} z^{\lambda^k - 1}. \qquad (1)$$

Show that $f'_\lambda \notin \Pi_q$. We fix $n \in \mathbb{N}$ and denote $r_n = \exp(-\alpha/\lambda^n)$, $r_n \to 1-0$, $n \to +\infty$. Let $u_n(z)$ be the n-th term of the series (1):

$$u_n(z) = \lambda^{\alpha n} z^{\lambda^n - 1}.$$

By $S_n(z)$ we denote the n-th partial sum of the series (1):

$$S_n(z) = \sum_{k=0}^{n-1} \lambda^{\alpha k} z^{\lambda^k - 1},$$

and by $R_n(z)$ we denote the n-th remainder of the series (1):

$$R_n(z) = \sum_{k=n+1}^{+\infty} \lambda^{\alpha k} z^{\lambda^k - 1}.$$

We estimate these sums on the circle $|z| = r_n$.

$$|S_n(z)| \leqslant \sum_{k=0}^{n-1} \lambda^{\alpha k} r_n^{\lambda^k - 1} = \sum_{k=0}^{n-1} \lambda^{\alpha k} \exp\left(-\frac{\alpha}{\lambda^n} \cdot (\lambda^k - 1)\right) = \exp\left(\frac{\alpha}{\lambda^n}\right) \sum_{k=0}^{n-1} \lambda^{\alpha k} \exp\left(-\alpha \cdot \lambda^{-(n-k)}\right) \leqslant$$

$$\leqslant \exp\left(\frac{\alpha}{\lambda^n}\right) \sum_{k=0}^{n-1} \lambda^{\alpha k} = \exp\left(\frac{\alpha}{\lambda^n}\right) \cdot \frac{\lambda^{n\alpha} - 1}{\lambda^\alpha - 1} = \lambda^{n\alpha} \exp(-\alpha - 1) \cdot A(\lambda, \alpha),$$

where $A(\lambda,\alpha) = \exp\left[\alpha(1 + \frac{1}{\lambda^n}) \cdot \frac{(1 - \lambda^{-n\alpha} \cdot e)}{\lambda^\alpha - 1}\right] < \frac{1}{4}$ for $\lambda > \lambda_0$.

Therefore we have $|S_n(z)| \leqslant \frac{1}{4} |u_n(z)|$.

Now we estimate $R_n(z)$ on the circle $|z| = r_n$.

$$|R_n(z)| \leqslant \sum_{k=n+1}^{+\infty} \exp\left(\frac{\alpha}{\lambda^n}\right) \lambda^{\alpha k} \sum_{m=1}^{+\infty} \frac{\lambda^{\alpha m}}{\exp(\alpha \lambda^m)}.$$

Since $\exp(\alpha\lambda^m) \geqslant \exp(m\alpha\lambda)$ for $m \geqslant 1$ and sufficient large λ,

$$\sum_{m=1}^{+\infty} \frac{\lambda^{\alpha m}}{\exp(\alpha \lambda^m)} \leqslant \frac{\lambda^\alpha}{e^{\alpha\lambda} - \lambda^\alpha},$$

so we have

$$|R_n(z)| \leqslant \exp(2\alpha + 1)|u_n(z)| \frac{\lambda^{\alpha m}}{\exp(\alpha\lambda^m)} \leqslant \frac{\lambda^\alpha}{e^{\alpha\lambda} - \lambda^\alpha} \leqslant \frac{1}{4}|u_n(z)|$$

for $\lambda > \lambda_1$.

As a result, we obtain:

$$|f'_\lambda(z)| \geqslant \frac{1}{2}|u(z)|,\ |z| = r_n,$$

for $\lambda > \max(\lambda_0, \lambda_1)$.

But

$$\ln|u_n(z)| \geqslant c_\alpha \ln \frac{1}{1 - r_n},\ \ n = 1, 2, \ldots$$

Thus, we have

$$\int_{-\pi}^{\pi} \left(\ln^+ |f'_\lambda(r_n e^{i\theta})|\right)^q d\theta \geqslant c_\alpha^q \ln^q \frac{1}{1 - r_n},$$

this means that $f'_\lambda \notin \Pi_q$. Theorem 1.1 is proved. $\square$

2. On the differentiation in the Privalov spaces

An important place in the theory of analytic functions belongs to the Nevanlinna N-class of analytic functions in D with bounded characteristic $T(r,f)$. It was introduced by A. Ostrovsky and brothers R. Nevanlinna and F. Nevanlinna (see [10]). As noted above, $N=\Pi_1$. Unlike the class N, the area Nevanlinna class is defined as follows (see ibid.):

$$\mathbf{N}=\left\{f\in H(D):\iint\limits_D \ln^+|f(z)|dxdy<+\infty\right\},\ z=x+iy,$$

or equivalent to this

$$\mathbf{N}=\left\{f\in H(D):\int\limits_0^1\int\limits_{-\pi}^{\pi}\ln^+|f(re^{i\theta})|d\theta dr<+\infty\right\}.$$

The area Nevanlinna classes are a natural generalization of the classes N. As it was established in the works [2, 17], these classes are close with respect to the properties of root sets and the factorization of functions. The class $\mathbf{N}$ is included in the scale of the Nevanlinna-Djrbashian classes N_α (see ibid.):

$$N_\alpha=\left\{f\in H(D):\int_0^1(1-r)^\alpha T(r,f)dr<+\infty\right\},\ \alpha>-1,$$

and in the scale of S^q_α-classes of F.A. Shamoyan (see [18]):

$$S^q_\alpha=\left\{f\in H(D):\int_0^1(1-r)^\alpha T^q(r,f)dr<+\infty\right\},\quad \alpha>-1,\ 0<q<+\infty.$$

Similar to the definition of the area Nevanlinna class, for all $0<q<+\infty$ we introduce the area Privalov class:

$$\tilde{\Pi}_q=\left\{f\in H(D):\int\limits_0^1\int\limits_{-\pi}^{\pi}\left(\ln^+|f(re^{i\theta})|\right)^q d\theta dr<+\infty\right\}.$$

It is clear that $\tilde{\Pi}_1=\mathbf{N}$. Using Hölder's inequality, it is easy to prove that $\tilde{\Pi}_q\subset S^q_0$ for $q>1$ and $\tilde{\Pi}_q\supset S^q_0$ for $0<q<1$.

The main result of the second part of this paper is the following theorem.

Theorem 2.1. *If $f\in\Pi_q$ $(0<q<+\infty)$ and function f has no zeros, then $f'\in\tilde{\Pi}_q$.*

To prove this statement, we need auxiliary statements.

Theorem 2.2 (see [13])**.** *If $f\in\Pi_q,(0<q<1)$, then*

$$\ln^+M(r,f)=o((1-r)^{-1/q}),\ r\to 1-0, \tag{2}$$

where $M(r,f)=\max\limits_{|z|=r}|f(z)|$, and the estimate is exact.

Lemma 2.3 (The Minkowski inequality, see [7], p. 178)**.** *Let $\{f_k\}_{k=1}^{+\infty}$ be the sequence of non-negative functions. For all $0<p<1$ the following inequality is valid:*

$$\left[\int\left\{\sum_k f_k(x)\right\}^p dx\right]^{1/p}\geqslant\sum_k\left\{\int f_k^p(x)dx\right\}^{1/p}.$$

Lemma 2.4 (see [6], p. 144). *Let $P(r,\theta)$ denote the Poisson kernel in D, i.e.*

$$P(r,\theta)=\frac{1-r^2}{1+r^2-2r\cos\theta}.$$

For each real number q there exist finite positive constants c_q, d_q, such that

$$c_q\phi_q(r)\leqslant\frac{1}{2\pi}\int_{-\pi}^{\pi}P^q(r,\theta)d\theta\leqslant d_q\phi_q(r),$$

where

$$\phi_q(r)=\begin{cases}(1-r)^q,\ q<\dfrac{1}{2},\\ \sqrt{1-r}\ln\left(1+\dfrac{1}{1-r}\right),\ q=\dfrac{1}{2},\\ (1-r)^{1-q},\ q>\dfrac{1}{2}.\end{cases}$$

Proof of Theorem 2.1. Let $z=re^{i\theta}$, $t=Re^{i\varphi}$, $0<r<R<1$. Since $f\in H(D)$ and function f has no zeros, we have, by the Schwarz formula, that:

$$\ln f(z)=\frac{1}{2\pi}\int_0^{2\pi}\ln|f(t)|\cdot\frac{t+z}{t-z}d\varphi+iC, \tag{3}$$

where the main branch of the logarithm is chosen.

Differentiate (3) by z:

$$\frac{f'(z)}{f(z)}=\frac{1}{\pi}\int_0^{2\pi}\ln|f(t)|\cdot\frac{t}{(t-z)^2}d\varphi,$$

$$f'(z)=\frac{f(z)}{\pi}\int_0^{2\pi}\ln|f(Re^{i\varphi})|\cdot\frac{Re^{i\varphi}}{(Re^{i\varphi}-re^{i\theta})^2}d\varphi,$$

whence

$$|f'(z)|\leqslant\frac{|f(z)|}{\pi}\int_0^{2\pi}\ln^+|f(Re^{i\varphi})|\cdot\frac{R}{R^2-2Rr\cos(\varphi-\theta)+r^2}d\varphi,$$

$$|f'(z)|\leqslant\frac{|f(z)|}{\pi R}\int_0^{2\pi}\ln^+|f(Re^{i\varphi})|\cdot\frac{1}{1-2\frac{r}{R}\cos(\varphi-\theta)+\frac{r^2}{R^2}}d\varphi. \tag{4}$$

Let us consider 3 cases.

Case 1. We assume that $0<q<1$.

Rewrite the last inequality in the form:

$$|f'(z)|\leqslant\frac{|f(z)|}{\pi R}\int_0^{2\pi}(\ln^+|f(Re^{i\varphi})|)^{q^2}\cdot(\ln^+|f(Re^{i\varphi})|)^{1-q^2}\cdot\frac{1}{1-2\frac{r}{R}\cos(\varphi-\theta)+\frac{r^2}{R^2}}d\varphi.$$

Applying Hölder's inequality with exponents $\dfrac{1}{q}$ and $\dfrac{1}{1-q}$, we have:

$$|f'(z)|\leqslant\frac{|f(z)|}{\pi R}\left[\int_0^{2\pi}(\ln^+|f(Re^{i\varphi})|)^q\right]^q\cdot\left[\int_0^{2\pi}\frac{(\ln^+|f(Re^{i\varphi})|)^{1+q}}{\left(1-2\frac{r}{R}\cos(\varphi-\theta)+\frac{r^2}{R^2}\right)^{1/(1-q)}}d\varphi\right]^{1-q}.$$

Since the function f belongs to the class Π_q, we have by Theorem 2.2:

$$|f'(z)| \leqslant \frac{|f(z)|}{\pi R} \cdot \frac{c_q \varepsilon_q}{(1-R)^{(1-q^2)/q}\left(1-\frac{r^2}{R^2}\right)} \left[\int_0^{2\pi} \left(P\left(\frac{r}{R}, \varphi-\theta\right)\right)^{1/(1-q)} d\varphi\right]^{1-q},$$

where $P\left(\frac{r}{R}, \varphi-\theta\right)$ is the Poisson kernel. We use the Poisson kernel estimate for $\frac{1}{1-q} > \frac{1}{2}$ from Lemma 2.4:

$$|f'(z)| \leqslant \frac{|f(z)|}{\pi R} \cdot \frac{c_q}{(1-R)^{(1-q^2)/q}} \cdot \frac{\varepsilon_q}{\left(1-\frac{r^2}{R^2}\right)} \cdot \frac{D_q}{\left(1-\frac{r}{R}\right)^q}.$$

Suppose $R = \frac{1+r}{2}$. After elementary transformations we obtain:

$$|f'(re^{i\theta})| \leqslant A_q \cdot |f(re^{i\theta})| \cdot \frac{1}{(1-r)^{(1+q)/q}}.$$

We proceed with the logarithm of the last inequality and take into account that $\ln^+|ab| \leqslant \ln^+|a| + \ln^+|b|$, $a>0$, $b>0$:

$$\ln^+|f'(re^{i\theta})| \leqslant \ln^+|f(re^{i\theta})| + \ln\left(\frac{A_q}{(1-r)^{(1+q)/q}}\right).$$

Next, raise both sides to the power q, and take into account $(a+b)^q \leqslant a^q + b^q$ for all $a>0$, $b>0$, $0<q<1$, after integration over $\theta \in [-\pi, \pi]$ we have:

$$\int_{-\pi}^{\pi} \left(\ln^+|f'(re^{i\theta})|\right)^q d\theta \leqslant \int_{-\pi}^{\pi} \left(\ln^+|f(re^{i\theta})|\right)^q d\theta + B_q + \left(\ln\frac{1}{(1-r)^{(1+q)/q}}\right)^q.$$

Since $f \in \Pi_q$ we have:

$$\int_{-\pi}^{\pi} \left(\ln^+|f'(re^{i\theta})|\right)^q d\theta \leqslant \tilde{B}_q + 2\pi\left(\ln\frac{1}{(1-r)^{(1+q)/q}}\right)^q.$$

Integrate over $r \in [0,1]$. In view of the convergence of the integrals on the right-hand side of the inequality, we conclude that $f' \in \tilde{\Pi}_q$.

Case 2. Now we suppose that $q>1$.

Applying Hölder's inequality with exponents q and $1+\frac{1}{q-1}$ in (4), we obtain

$$|f'(z)| \leqslant \frac{|f(z)|}{\pi R\left(1-\frac{r^2}{R^2}\right)} \left[\int_0^{2\pi} (\ln^+|f(Re^{i\varphi})|)^q d\varphi\right]^{1/q} \cdot \left[\int_0^{2\pi} P\left(\frac{r}{R}, \varphi-\theta\right)^{1+\frac{1}{q-1}} d\varphi\right]^{1-1/q}.$$

Since the function f belongs to the class Π_q, we have

$$|f'(z)| \leqslant \frac{|f(z)|}{\pi R\left(1-\frac{r^2}{R^2}\right)} c_q \cdot \left[\int_0^{2\pi} P\left(\frac{r}{R}, \varphi-\theta\right)^{1+\frac{1}{q-1}} d\varphi\right]^{1-1/q}.$$

We use the Poisson kernel estimate for $1+\frac{1}{q-1} > \frac{1}{2}$ from Lemma 2.4:

$$|f'(z)| \leqslant \tilde{c}_q \frac{|f(z)|}{\pi R\left(1-\frac{r^2}{R^2}\right)} \cdot \frac{1}{\left(1-\frac{r}{R}\right)^{1/q}}.$$

Suppose that $R = \dfrac{1+r}{2}$, then we have:

$$|f'(z)| \leqslant C_q \frac{|f(z)|}{(1-r)^{\frac{q}{q-1}}}.$$

We proceed with the logarithm of the last inequality and take into account that $\ln^+ |ab| \leqslant \ln^+ |a| + \ln^+ |b|$, $a > 0$, $b > 0$:

$$\ln^+ |f'(z)| \leqslant \ln^+ |f(z)| + \ln \frac{C_q}{(1-r)^{\frac{q}{q-1}}}.$$

Further, raise both sides to the power q, and take into account $(a+b)^q \leqslant a^q + b^q$ for all $a > 0$, $b > 0$, $0 < q < 1$. After integration in $\theta \in [-\pi, \pi]$ we obtain:

$$\int_{-\pi}^{\pi} \left(\ln^+ |f'(re^{i\theta})|\right)^q d\theta \leqslant \int_{-\pi}^{\pi} \left(\ln^+ |f(re^{i\theta})|\right)^q d\theta + \ln \frac{\tilde{C}_q}{(1-r)^{\frac{q}{q-1}}}.$$

Since $f \in \Pi_q$, we see that:

$$\int_{-\pi}^{\pi} \left(\ln^+ |f'(re^{i\theta})|\right)^q d\theta \leqslant a_q + \ln \frac{\tilde{C}_q}{(1-r)^{\frac{q}{q-1}}}.$$

Integrate over $r \in [0, 1]$. In view of the convergence of the integrals on the right-hand side of the inequality, we conclude that $f' \in \tilde{\Pi}_q$.

Case 3. We assume $q = 1$. Using the estimate of S. N. Mergelyan for a function of the Nevanlinna class (see [12, c. 84]), we get from (4):

$$|f'(z)| \leqslant C \frac{|f(z)|}{\pi R(1-R)\left(1 - \frac{r^2}{R^2}\right)} \int_0^{2\pi} P\left(\frac{r}{R}, \varphi - \theta\right) d\varphi,$$

whence by the property of the Poisson integral

$$|f'(z)| \leqslant C \frac{|f(z)|}{\pi R(1-R)\left(1 - \frac{r^2}{R^2}\right)}.$$

Further, the proof repeats the argument for Case 2. Theorem 2.1 is completely proved. □

Remark 2.1. *Note that W. Hayman indicates the invariance of the class* Π_q, $(1 < q < +\infty)$ *with respect to the integration operator [8].*

First author was financially supported by Russian Foundation for Fundamental Research, project number 18-31-00180.

Second author was financially supported by Russian Foundation for Fundamental Research, project number 17-51-15005.

References

[1] D.Campbell, Wickes, The Bloch–Nevanlinna conjecture revisited, *Bull. Austral. Math. Soc.*, **18**(1978), 447–453.

[2] M.M.Djrbashian, On the problem of the representation of analytic functions, *Soobshch. inst. matem. i mehan. Acad. Nauk Arm. SSR.*, **2**(1948), 3–40 (in Russian).

[3] P.L.Duren, On the Bloch-Nevanlinna conjecture, *Colloq. Math.*, **20**(1969), 295–297.

[4] P.L.Duren, Theory of H^p spaces, *Pure and Appl. Math.*, NY: Academic Press., Vov. 38, 1970.

[5] O.Frostman, Sur les produits des Blaschke, *Kungl. Fysiografiska Sallskapets i Lund Forhandlingar*, [Proa. Roy. Physiog. Soa. Lund], **12**(1942), no. 15, 169–182.

[6] V.I.Gavrilov, A.V.Subbotin, D.A.Efimov, Boundary properties of analytic functions (further contribution), Moscow, Publishing House of the Moscow University, 2012 (in Russian).

[7] G.Hardy, Inequalities, G. Hardy, J. Littlewood, G. Polia, Translate from English: S. B. Stechkin (eds.); V. I. Levin (transl.), Moscow, GITTL, 1948.

[8] W.K.Hayman, On the characteristics of functions meromorphic in the unit disk and of their integrals, *Acta. math.*, **112**(1964), no. 3–4, 181–214.

[9] R.Nevanlinna, Le theoreme de Picard-Borel et la theorie des fonctions meromorphes, Paris, Gauthiers-Villars, 1929.

[10] R.Nevanlinna, Eindeutige analytische Funktionen, 2nd ed., Berlin, Springer-Verlag, 1953.

[11] I.I.Privalov, Boundary properties of single-valued analytic functions, Moscow, Izd. Moscow State University, 1941 (in Russian).

[12] I.I.Privalov, Boundary properties of analytic functions, M.-L.: GITTL, 1950 (in Russian).

[13] E.G.Rodikova, Coefficient multipliers for the Privalov classes in a disk, *J. Sib. Fed. Univ. Math. Phys.*, **11**(2018), no. 6, 723–732.

[14] E.G.Rodikova, V.A.Bednazh, On interpolation in the Privalov classes in a disk, *Sib. electronic matem. reports.*, **16**(2019), 1762—1775 (in Russian).

[15] E.G.Rodikova, On properties of zeros of functions from the Privalov class in a disk, *Scientific notes of Bryansk State Univ.*, (2019), no. 4, 19–22 (in Russian).

[16] E.G.Rodikova, On interpolation sequences in the Privalov spaces, Complex analysis, mathematical physics and nonlinear equations: collection of abstracts of the International Scientific Conference, Ufa, 2020, 52–53 (in Russian).

[17] F.A.Shamoyan, Factorization theorem M. M. Djrbashian and characterization of zeros of analytic functions with a majorant of finite growth, *Izv. Acad. nauk Arm. SSR, Matem.*, **13**(1978), no. 5-6, 405–422 (in Russian).

[18] F.A.Shamoyan, Parametric representation and description of the root sets of weighted classes of functions holomorphic in the disk, *Siberian Math. J.*, 40(1999), no. 6, 1211–1229.

[19] F.A.Shamoyan, Weighted spaces of analytic functions with a mixed norm, Bryansk, Bryansk St. Univ., 2014 (in Russian).

[20] F.A.Shamoyan, On some properties of zero sets of the Privalov class in a disk, *Zap. nauch. semin. POMI*, **480**(2019), 199–205 (in Russian).

[21] F.A.Shamoyan, I.S.Kursina, On the invariance of some classes of holomorphic functions under integral and differential operators, *J. Math. Sci. (New York).*, **107**(2001), no. 4, 4097–4107.

[22] S.V.Shvedenko, Hardy classes and the spaces of analytic functions associated with them in the unit disc, polydisc, and ball, *Itogi Nauki i Tekhniki. Ser. Mat. Anal.*, **23**(1985), 3–124 (in Russian).

О дифференцировании в классах И. И. Привалова

Евгения Г. Родикова
Брянский государственный университет
Брянск, Российская Федерация
Файзо А. Шамоян
Саратовский государственный университет
Саратов, Российская Федерация

Аннотация. В статье исследуется инвариантность классов И. И. Привалова относительно оператора дифференцирования.

Ключевые слова: класс Привалова, гипотеза Блоха-Неванлинны, оператор дифференцирования.

Journal of Siberian Federal University. Mathematics & Physics 2020, 13(5), 631–643

DOI: 10.17516/1997-1397-2020-13-5-631-643
УДК 532.526

Experimental and Numerical Study of Free Convection Heat Transfer Around the Junction of Circular Cylinder and Heated Vertical Plate

Hamid Malah*
Yurii S. Chumakov†
Peter the Great St. Petersburg Polytechnic University
St. Petersburg, Russian Federation

Received 10.06.2020, received in revised form 21.07.20, accepted 23.08.2020

Abstract. The present study investigates the effects of a circular cylinder on the three-dimensional characteristics of free convective heat transfer. The circular cylinder is mounted horizontally on a heated vertical plate and is categorized as high aspect ratio obstacle, which means the height of cylinder is comparable to its diameter. The obtained results are provided for the laminar flow regime. In addition, during numerical study the governing differential equations are solved around the Grashof number equals to 3×10^8. In order to illustrated the regions of high gradients of temperature, the flow temperature is shown in terms of non-dimensional contours and diagrams. At the near junction region in upstream of cylinder, by description of heat transfer coefficients represented to the temperature gradients at intended points, the effects of cylinder emplacement on the heat transfer rate is surveyed. As expected, the value of the buoyancy-induced heat transfer coefficient increases at the cylinder junction in the upstream side. The maximum value of heat transfer coefficient is seen at the symmetry plane of study domain, which is corresponded to the location of horseshoe vortex system core. Finally, by deviation calculating between numerical and experimental results also by analysis of the experimental method uncertainty the validity and reliability of numerical and experimental approaches are proved.

Keywords: juncture flow, free convection heat transfer, laminar boundary layer, heat transfer coefficient.

Citation: H. Malah, Y.S. Chumakov, Experimental and Numerical Study of Free Convection Heat Transfer Around the Junction of Circular Cylinder and Heated Vertical Plate, J. Sib. Fed. Univ. Math. Phys., 2020, 13(5), 631–643. DOI: 10.17516/1997-1397-2020-13-5-631-643.

Introduction

Free convection flows around surface-mounted isothermal circular cylinders are found in many industrial and environmental applications. Flow through bluff body located on a heated plate is complex and can contribute to significant non-uniformities of local heat transfer [1]. Such flows can attributed to the body junction of different heat exchanger units and cooling systems. In most industrial applications, the focus is on heat transfer and factor that governs the flow pattern. Typical geometry of such applications can be represented, as a circular cylinder mounted on heated vertical plate with a free convection boundary layer developing along it. A large

*hamid.malah@gmail.com
†chymakov@yahoo.com

adverse pressure gradient in the encounter region where the free convection boundary layer approaches the circular cylinder leads to separate the boundary layer in front of the stagnation point. Development of such flows depends both on intensity of the separation and on the arising buoyancy force acting on the heated flow pushed away from the heated plate. In addition, flow structure forming depends on bluff body shape and its aspect ratio. Compared to the flow past a short obstacle, the flow past a high aspect ratio bluff body placed on the heated vertical plate is more complex due to the effects of the plate and free convection boundary layer [2].

Experimental and computational studies of free convection flows over a bluff body have been reported in the several literature [3–5], that all confirming the anticipated increase in heat transfer in the stagnation region in front of the bluff body and its decrease in the wake region. Heat transfer from simple plate or around single cylinder intensively differs from a complex system consist of bluff body placed on the plate due to effects of boundary layer arising from the plate leading edge on bluff body in the encounter region. Despite unconfined cylinders or bluff body is placed on the horizontal plate [6, 7], the free convection heat transfer around a complex system included high aspect ratio circular cylinder mounted on heated vertical plate in laminar regime is not well understood. The better understanding of this complex system is critically important because of its practical applications. In the case of bluff body mounted on heated plate, there are some papers, which present the analysis of free convection heat transfer at the junction of low aspect ratio isothermal cylinder mounted on a heated wall [2, 8]. These papers [2, 8] show that the low aspect ratio bluff body which completely embedded in free convection boundary layer leads to heat transfer enhancement.

The present paper covers numerical results of time-based computations, which categorize the effects of high aspect ratio cylinder mounting on free convection heat transfer at the junction of the circular cylinder and heated vertical plate with approaching laminar free convection boundary layer. The axis of the circular cylinder directed normally to the plate. The study considers the case, which an adiabatic circular cylinder obviously crosses from thermal boundary layer arisen on the vertical plate, which means the cylinder height (H) is comparable to the diameter of cylinder (D). In order to confirm obtained computational results in upstream region of cylinder, the experimental data for the same problem are presented in parallel.

1. Problem definition

This study is performed by using three-dimensional Navier–Stokes and energy balance equations to solve the problem of laminar flow of a viscous incompressible fluid, which regarded to the Boussinesq approximation over a circular cylinder mounted on a rectangular heated vertical plate. In order to reduce the non-linearity of the governing equations meanwhile the numerical solution, the density variation was neglected in all terms of governing equations except the buoyancy term. Therefore, the following set of equations, which will applied to consider computational domain by finite volume method is in good accuracy with original set of governing equations:

$$\begin{gathered} \nabla . V = 0, \\ \rho \frac{\partial V}{\partial t} + \rho (V . \nabla) V = \rho \beta g (T - T_\infty) - \nabla P + \mu \nabla^2 V, \\ \rho C_p \frac{\partial T}{\partial t} + \rho C_P V . \nabla T = \lambda \nabla^2 T. \end{gathered} \tag{1}$$

The system of equations (1) include: the velocity of a parcel of fluid, V; density, ρ; time, t; volume thermal expansion coefficient, β; gravitational acceleration, g; the temperature of a

Table 1. Boundary conditions of problem

Index	Applied conditions
1	$V_X = V_Y = V_Z = 0,\ \partial P/\partial n = 0,\ \partial T/\partial n = 0$
2	$V_X = V_Y = V_Z = 0,\ \partial P/\partial Z = 0,\ T = T_W$
3	$V_X = V_Y = V_Z = 0,\ \partial P/\partial Z = 0,\ T = T_\infty$
4	$V_X = V_Y = V_Z = 0,\ \partial P/\partial Z = 0,\ T = T_\infty$
5	Symmetry
6	Symmetry
7	$P = P_\infty, T = T_\infty$

parcel of fluid, T; surrounding medium temperature, T_∞; the local pressure of a parcel of fluid, P; dynamic viscosity, μ; specific heat at constant pressure, C_P and thermal conductivity, λ.

2. Computational aspects

In this study, for different analyzed cases, the diameter of the cylinder (D) equals to $0.02\,[m]$ and the height of the cylinder (H) equals to $2D$. The height of cylinder is selected in this way to cross thermal boundary layer entirely. The emplacement position of cylinder is considered to be at the place, where the Grashof number along vertical plate (Gr_y) on simply vertical plate without cylinder (undisturbed upward flow), equals to desired value of the laminar Grashof number. The origin of the system of coordinate is placed $31.5D$ from the leading edge of the isothermal vertical plate and coincides with the center of circular cylinder cross section at junction region and consequently the approaching flow at the location of cylinder is fully developed laminar flow. The X-axis corresponds to the horizontal direction, the Y-axis aligns to the vertical direction and the Z-axis is normal to the vertical plate. The coordinate system and computational domain, which used in the present study, are shown in Fig. 1. In order to cautiously simulate the characteristics of the flow in the vicinity of the cylinder, the dimensions of computational domain are drawn out $117D$ in height and $7D$ width. For the normal direction, the domain size is set to $10D$, to provide a sufficient wide space for ultimate develop of convective boundary layer thickness (δ_0) arising from heated vertical plate. In order to avert the effects of the solid boundaries on the free convective flow pattern also on the heat transfer characteristics, the lower face of the computational domain is located at $40.5D$ upstream of the cylinder axis and the upper one is placed at $76.5D$ downstream, as shown in Fig. 1a.

By authors' knowledge, the free convective flow in laminar regime does not depend on initial condition of solution, especially in our case the final solution is independent from given initial condition [2]. In this paper, there is no initial velocity around leading edge of heated vertical plate and at the beginning of calculations (first time step). Therefore, velocity in all directions assumed as zero. By starting numerical procedure ($t = 0\,[s]$), the temperature at the vertical plate is increased suddenly from $T_\infty = 293.15\,[K]$ to $T_W = 333.15\,[K]$ and maintained at this value. Heat is transferred initially by pure conduction to the surrounding medium and this initial conductive phase is characterized by arisen thermal boundary layer on vertical plate, until a certain critical time is happened when buoyant flow arises and interacts with cylinder surfaces also heated vertical plates and finally leads to forming free convective layer around the solid surfaces.

The boundary conditions are formulated as corresponding mathematical equations as follows in Tab. 1 (see Fig. 1a for numbering of the computational domain boundaries).

Fig. 1b shows the blocked structure of grid for meshing system, which covers the computational domain by eight-node hexahedron elements. The computational domain is discretized using body-fitted mesh adapted to the geometry of the domain in order to carry out high-resolution numerical simulation. The final quantity of three-dimensional computational grid elements approximately equal to 6.2 million. The grid is clustered to heated vertical plate along Z-axis by considering that the elements size in the near-wall regions are about $0.02D$. In the normal direction to the vertical plate, the grid consists of 150 nodes, where 50 nodes are located inside the thickness of undisturbed boundary layer. The multi-block structured grid is originated from two-dimensional grid of the XY plane.

In Fig. 1b is shown the numbering of the blocked structure of two-dimensional grid, which is explained in the following. An O-type block (8) embedded in the clustered region close to the cylinder (9) to discretize this enclosing area. The diameter of this O-type block, which is centered at the center of the cylinder equals to $2.1D$. An additional square block (10) with edges dimension equal to $3D$, which enclose the O-type mesh area, is embedded to enhance fine grid computation around the cylinder. In the XY plane, the elements size of grid in the O-type block and square block respectively varied between $0.013 - 0.027D$ and $0.023 - 0.093D$. Four additional grid blocks (11) are selected between the refined square block around the cylinder and outer blocks (12) with coarse mesh, which served to align the skewness of the forming elements. Finally, six blocks (13) were constructed adjacent of leading and trailing edges of vertical plate, where the grid of these blocks were clustered to the heated vertical plate edges in order to correctly describe arising flow in these critical regions.

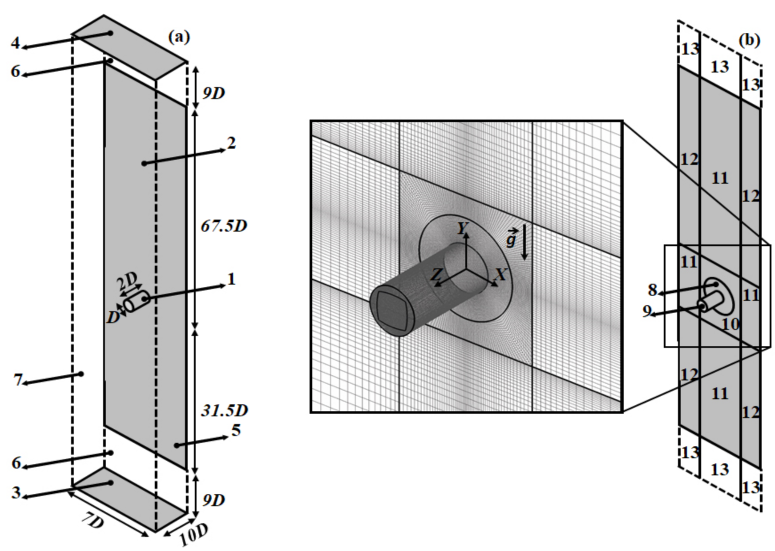

Fig. 1. The schematic representation of the problem configuration: (a) — computational domain, (b) — grid layout

In this work, the commercial package ANSYS-Fluent 16.2 [9] was used to simulate numerically three-dimensional disturbing effects around circular cylinder in the upstream region. To evaluate the effects of computational parameters include time step and mesh sizes during numerical solution, here, four different simulations were performed, which differs from each other in grid refinement and time step. These different configurations were applied to surface-mounted adiabatic circular cylinder on the heated vertical plate. In Tab. 2, two symmetry plane characteristics of the horseshoe vortex system include diagonal distance of necklace vortex axis coordinates relative to the circular cylinder axis (L_{PV}) and in normal direction to the vertical plate (C_{PV}) were provided to compare. Tab. 2 shows a high degree of fidelity between refined and coarse grids. In addition, Tab. 2 proves the ineffectiveness of reducing physical time step on characteristics of horseshoe vortex system. Furthermore, Tab. 2 was used to validate the numerical simulation method by previous numerical study [2]. The analysis of given values in Tab. 2 confirms the reliability of numerical solutions. As result of reliability analysis, which presented in Tab. 2, the effectiveness of applied numerical configuration in ANSYS-Fluent 16.2, is emphasized. Therefore, the time step used in the present computation was, in general, at 0.002 $[s]$ and flow parameters are stabilized with an accuracy equal to 10^{-6} at physical time approximately equal to 200 $[s]$.

Table 2. Computational parameters effects on the necklace vortex characteristics

Index	**Number of cells**	**Time step [s]**	$\mathbf{L_{PV}}$ **[m]**	$\mathbf{C_{PV}}$ **[m]**
Coarse	6 200 000	0.001	0.0128	0.0024
Coarse	6 200 000	0.002	0.0128	0.0024
Refined	20 500 000	0.001	0.0129	0.0024
Refined	20 500 000	0.002	0.0129	0.0024
[2]	3 650 000	Steady State	0.0135	0.0022

3. Experimental investigation

The difficulty in the experimental study of free convection flows compared with forced flows is due to the complexity of providing experimental apparatus, which can generate stable free convection flow with determined parameters. In this study, free convective boundary layer near a heated vertical plate is provided, by using an experimental setup described in experiments of other authors [10]. Here the characteristics of the used experimental apparatus allow simulating different flow regimes, up to the Grashof number equal 5×10^{11}.

The main part of this experimental apparatus is the junction of cylinder and heated vertical plate. A solid polymer circular cylinder is mounted on vertical plate using light glue. The diameter of cylinder fixed to 0.02 $[m]$ and its height is equal $2D$. The adiabatic condition on cylinder surfaces is provided by very low thermal conductivity of polymer. The vertical plate is made of duralumin sheet with dimensions of $250D \times 44D \times 0.02D$, and its working surface, along which free convection flow arises, is polished.

The temperature of studying domain is measured using a resistance thermometer sensor, which is controlled remotely by computer. During the experimental investigation, the surrounding medium temperature varies between 293.15 and 295.15$[K]$. Since it is not possible to measure the temperature of vertical plate (T_W) by resistance thermometer sensor precisely, the wall temperature is approximated by linearization of temperature profiles on desired points.

4. Results and discussion

In this part, the results of the numerical simulation in comparison with experimental data for a high aspect ratio circular cylinder ($H/D = 2$), which crosses the formed boundary layer entirely, are presented. It is noteworthy that the thermal boundary layer thickness (δ) assumed as the Z coordinate where the local temperature differs from the surrounding medium temperature less than two percent. Here, the computed thickness of developed thermal boundary layer along vertical plate at emplacement of cylinder ($Y/D = 0$) equals $0.5H$, which ensures that cylinder height entirely crosses the formed boundary layer. In this case, the free convection heat transfer is controlled by the intensity of buoyancy force, which depends on the local temperature difference between heated vertical plate and surrounding medium ($T_W - T_\infty$) [2]. Subsequently in order to study free convection heat transfer parameters at first, the temperature field is illustrated by using dimensionless temperature. In continue, by demonstrating dimensionless temperature profiles at different YZ sections ($X/D = 0, 0.6, 0.7, 0.85$), the free convection heat transfer coefficient (α) is investigated in both numerical and experimental study to realize the effects of cylinder emplacement on the flow configuration in the upstream region of the cylinder. In this paper, numerical and experimental results of temperature fields and heat transfer rate were obtained for laminar Grashof number equals to 3×10^8. The extremely high Grashof number leads to apply enormous time steps for numerical solution procedure also leads to convergence problems. Furthermore, the provided experimental study demonstrated that transient regime is occurred only after $45D$ from leading edge of the heated vertical plate. Therefore, the characteristics of laminar regime approximately extend to Grashof number equals 5×10^9. Thus, the Grashof number was selected in this way to satisfy the requirements of the practical applications by considering the numerical limitations. To sum up, by comparing the numerical results to experimental data the validity of the computational simulation and their underlying assumptions and simplifications are verified.

Since the temperature of the vertical plate and surrounding medium in numerical simulation are given as isothermal temperature, and in experimental investigation is measured by thermal sensors, there is dissimilarity among these values in numerical and experimental study. Thus, the obtained values from the experimental investigation cannot be directly compared to the computed values. To combat this problem, in present work, a dimensionless temperature (T^+) is defined as follows:

$$T^+ = \frac{T_W - T}{T_W - T_\infty}. \tag{2}$$

Which is determined as the ratio of the temperature difference between vertical plate and desired point of investigated ambient ($T_W - T$), to the temperature difference between vertical plate and surrounding medium ($T_W - T_\infty$). As stated in previous parts, the temperature of surrounding medium (T_∞) in numerical simulation considered as $293.15\,[K]$ and in experimental study varies between $293.15\,[K]$ to $295.15\,[K]$. In addition, the temperature of vertical plate in numerical solution fixed on $333.15\,[K]$ and in experimental investigation is obtained by linearization of temperature profile on desired point.

The dimensionless temperature field comparison between the numerical solution and experimental investigation is shown in Fig. 2 for different YZ planes around the junction of cylinder and heated vertical plate. The present experimental results qualitatively are in good agreement with their numerical results at different YZ planes. The obtained temperature field in upstream region of the cylinder is symmetric. It is especially noted here that the existence of step-like structure in Fig. 2 for experimental temperature field is due to spatial step among gathered experimental data. Subsequently when the contour of temperature is drawn by using the triangulation of data method in modeling software, there are depicted step-like structure. Furthermore, the white regions in Fig. 2a and Fig. 2b correspond to very closed cylinder surfaces regions, which could not be considered due to the limitation of experimental measurement.

As seen in Fig. 2, the buoyancy-induced flow develops on the vertical plate due to temperature gradient between the heated vertical plate and the cold ambient fluid. This gradient leads to

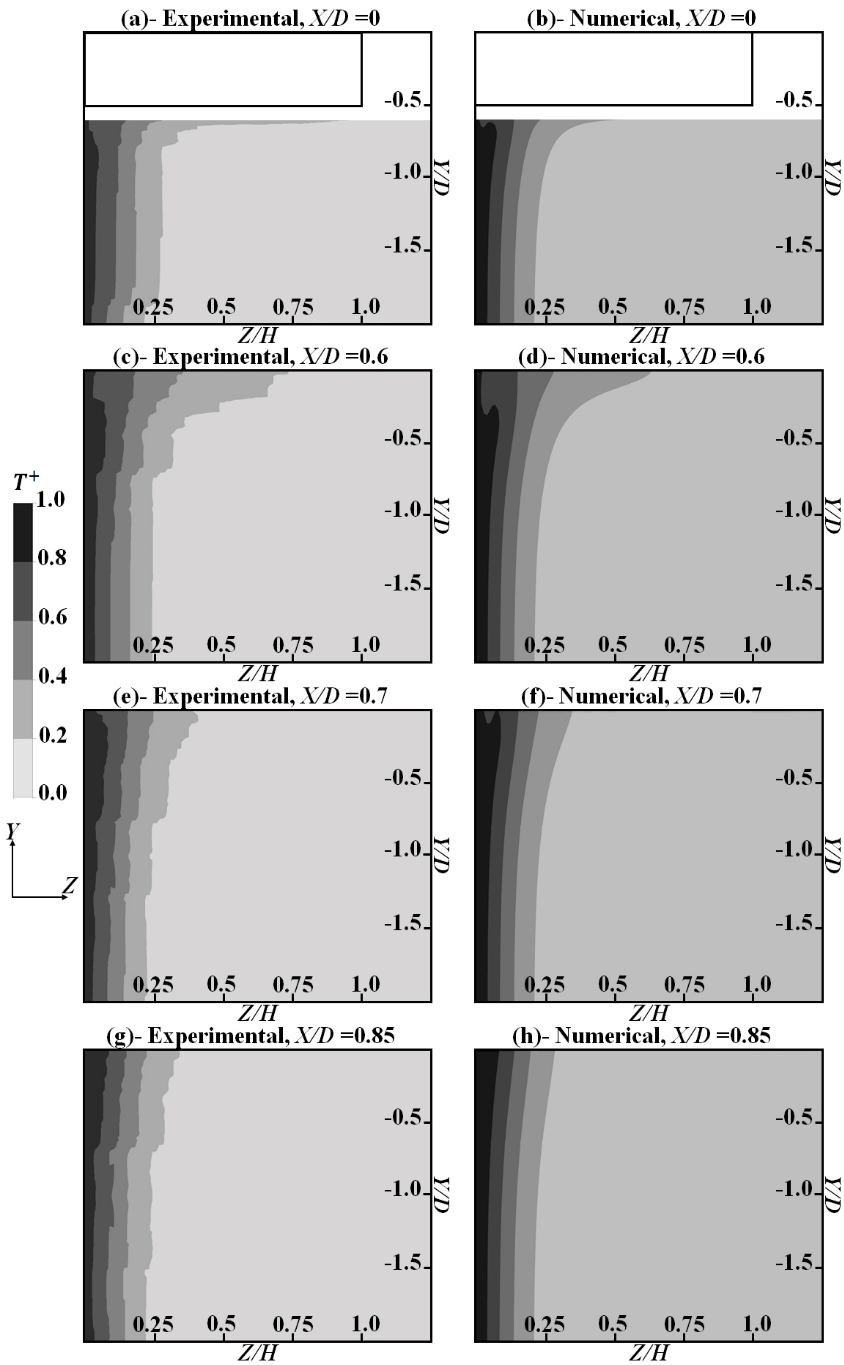

Fig. 2. The dimensionless temperature fields: (a) — experiment for $X/D = 0$, (b) — numeric for $X/D = 0$, (c) — experiment for $X/D = 0.6$, (d) — numeric for $X/D = 0.6$, (e) — experiment for $X/D = 0.7$, (f) — numeric for $X/D = 0.7$, (g) — experiment for $X/D = 0.85$, (h) — numeric for $X/D = 0.85$

form a non-uniform density field. The arisen buoyant force encounters leading edge of circular cylinder. In the upstream region of the cylinder, the temperature of approaching flow to the leading edge of cylinder is very small except in the formed boundary layer very closed to the cylinder surface. It is due to this fact that diffusion arisen away from cylinder surface is less than approaching advection toward the cylinder [11]. When heated fluid approaches the leading edge of cylinder, at stagnation point the temperature distribution is divided into hot and cold regions by an imaginary line because of higher pressure at this point compared with adjacent regions. The hot flow moves toward the vertical plate, recirculates at junction of cylinder and heated plate, and leads to form horseshoe vortex system. For symmetry YZ plane ($X/D = 0$), the core of horseshoe vortex structure, which is located near the cylinder junction, can be seen in Fig. 2. In this work, as result of time-based solution, vortex structure is formed near the cylinder junction with elapse of time, and then a steady state condition will prevail along solution time so the center of the vortex stays at a certain position in accordance with flow recirculation.

In order to have better understanding of the temperature field in Fig. 2, the dimensionless temperature profiles in the symmetry plane ($X/D = 0$) for both numerical and experimental studies have been plotted in Fig. 3a at the different Y coordinate below the leading edge of cylinder.

Fig. 3a displays the role of free convection flow approaching to the cylinder in modifying the flow field in the upstream region of the cylinder. The agreement between experimental and numerical results is quantitatively and qualitatively very good, especially in the thermal boundary layer region. The temperature distributions in the near wall domain and in the cylinder junction, where flow recirculates compare favorably. However, by moving away from cylinder in upstream region in Y direction the consistency of experimental and numerical results get much closer. This is due to cylinder junction effects in the experiment where different disturbances appears in computational ambient compare to numerical simulation which simplifies the ambient parameters to resolve problem.

In order to demonstrate the effects of cylinder emplacement on approaching free convection flow, in Fig. 3b have been drawn experimental and numerical results of dimensionless temperature profiles for different YZ plane ($X/D = 0.6$, 0.7, 0.85). These graphs show dimensionless temperature at the angular coordinate equal to 90 degree relative to the leading edge of cylinder ($Y/D = 0$). Here, the angular coordinate zero is measured from the lower symmetry line of the cylinder (leading edge of cylinder). In Fig. 3b, analyzing the dimensionless temperature profiles on the heated vertical plate reveals the thermal boundary layer thickness, which in the upstream region of cylinder (angular coordinate between zero to 90 degree) is fairly uniform and approximately equals to $0.45H$. As shown in Fig. 3b, the effect of cylinder emplacement disappears by moving away from the cylinder in X direction. In addition, the numerical results are in good compatibility with the experimental data.

As mentioned above, a part of flow, which separate at stagnation point on the leading edge of cylinder toward the vertical plate, recirculates at junction region, and leads to form horseshoe vortex system in the upstream of the cylinder. The arisen horseshoe vortex contributes intensive interaction of cold and hot flows, which undoubtedly have significant effect on free convective heat transfer in junction region [12]. By considering equation (1) for thermal energy, the heat transfer coefficient (α) is determined as ratio of the heat flux (q) and the temperature difference between heated vertical plate and surrounding medium ($T_W - T_\infty$):

$$\alpha = \frac{q}{T_W - T_\infty}. \tag{3}$$

For the numerical solution, the values of free convection heat transfer coefficient are computed directly from numerical simulation based on finite volume method and equation (3). In experimental investigation, based on the collected data, the values of the heat transfer coefficient

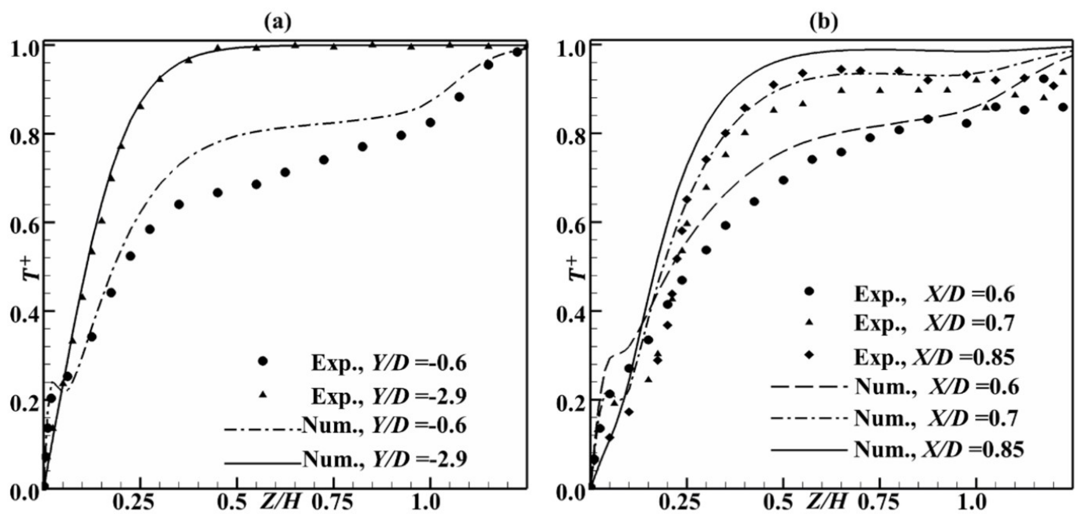

Fig. 3. The dimensionless temperature profiles; (a) — in the symmetry plane ($X/D = 0$), (b) — at the angular coordinate equal to 90 degree ($Y/D = 0$)

are calculated by using equation (3). To aim this purpose, a thermal layer on the vertical plate is considered, on which the heat transfer is only due to conduction. Thus, the rate of convection heat transfer of ambient in the arisen convective layer on the heated vertical plate assumed equal to the rate of conduction heat transfer in imaginary conductive sub-layer. Therefore, the convection heat transfer is defined as follows:

$$\alpha = \frac{\lambda}{T_W - T_\infty}\frac{\partial T}{\partial z}. \tag{4}$$

In equation (4), λ is the thermal conductivity of ambient flow, which in this case considered constant and equals to $0.0242\,[Wm^{-1}K^{-1}]$. The surrounding medium temperature (T_∞) was probed for each point during the measurement of computational ambient temperature. The value of the vertical plate temperature (T_W) was obtained by linearization of temperature profile for desired point. Lastly, the value of temperature derivative relatively to the Z coordinate (normal direction to the plate) was approximated by the line slope of temperature profile linearization in the vicinity of vertical plate with in conductive sub-layer. The existence of horseshoe vortex system in the junction region increase the convection heat transfer rate and subsequently leads to a thinner conductive sub-layer around cylinder, which may be considered during the temperature profile linearization.

Comparison of the numerical calculation of free convective heat transfer coefficient at different YZ plane with experimental results is shown in Fig. 4. For graphs within Fig. 4, the values of heat transfer coefficient meet a minimum at the junction of cylinder and vertical plate, which is described by existence of horseshoe vortex system and its configuration. As demonstrated in Fig. 4, the heat transfer rate near the cylinder surface is intensively more than other regions. This phenomenon can be explained by temperature differences between arisen thermal boundary layer and heated vertical plate. The bigger buoyancy force formed around the cylinder leads to increase in flow rate in this region and consequently the free convection heat transfer coefficient increases obviously.

As shown in Fig. 4, for spatial coordinates more than $0.4D$ relative to cylinder surface, the experimental values quantitatively agree very well with the numerical results. For spatial

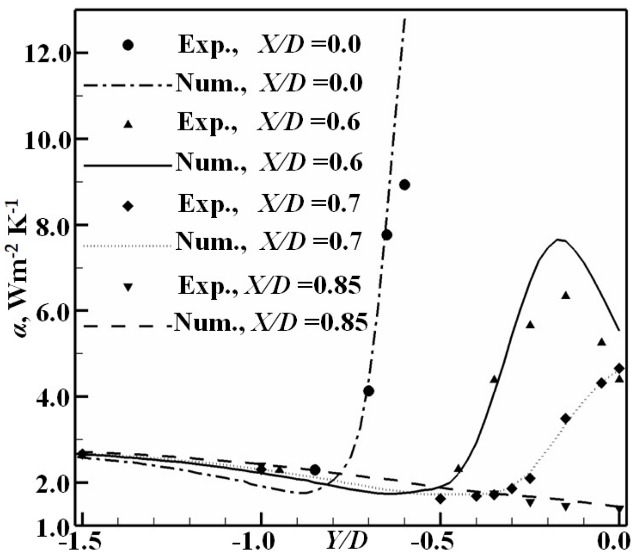

Fig. 4. The free convective heat transfer coefficient at different YZ plane

coordinates less than $0.4D$ and more than $0.2D$, the experimental results quantitatively and qualitatively are in acceptable agreement with numerical solution. Finally, for spatial coordinates less than $0.2D$ (near cylinder surface), the experimental data only match with numerical calculations in configurative view. The maximum deviation of experimental heat transfer coefficients and its numerical values, which is seen in Fig. 4 around cylinder surface, is about 20%. The appeared deviation between numerical and experimental results may be occurred due to disturbances existence near the junction of cylinder in experimental measurement, which were simplified to compute flow parameters in numerical solution. In addition, the accuracy of experimental procedure and devices, which can be affected by arisen errors during experiments, leads to accumulate deviation. The deviation can be categorized as researcher's errors and measurement errors, which consist of random errors and systematic errors. The random errors are due to real situation during the process of measurement and the systematic errors are because of specific experimental procedure. The statistical methods can foreshow the random errors and the calibration of devices can reduce the systematic errors. These errors must be analyzed to minimize the uncertainty of the obtained results. In this study, in order to analysis uncertainty of experimental data, we determine the parameters affecting the computation of heat transfer coefficient in equation (4). In this work, uncertainty of obtained values is due to measurement accuracy of surrounding medium temperature, vertical plate temperature and the rate of free convection heat transfer. These three parameters are considered as independent variables, so the uncertainty of heat transfer coefficient is defined as follows, where errors are neglected on length measurements [13]:

$$\omega_\alpha = \pm\left[\left(\frac{\partial\alpha}{\partial q}\omega_q\right)^2 + \left(\frac{\partial\alpha}{\partial T_W}\omega_{T_W}\right)^2 + \left(\frac{\partial\alpha}{\partial T_\infty}\omega_{T_\infty}\right)^2\right]^{\frac{1}{2}}. \quad (5)$$

It should be noted that the conduction and radiation heat transfer are neglected due to assuming pure buoyant flow on heated vertical plate. Therefore, the uncertainty of free convection heat transfer coefficient per unit area is determined as:

$$\omega_\alpha = \pm \left[\left(\frac{1}{T_W - T_\infty} \omega_{q_{conv}} \right)^2 + \left(\frac{-q_{conv}}{(T_W - T_\infty)^2} \omega_{T_W} \right)^2 + \left(\frac{q_{conv}}{(T_W - T_\infty)^2} \omega_{T_\infty} \right)^2 \right]^{\frac{1}{2}}. \qquad (6)$$

From equation (6), the uncertainty of heat transfer coefficient for different points, which considered in experimental study, varies between 2.96% and 7.41%. From uncertainty analysis, it is found out that the maximum value of uncertainty for heat transfer coefficient is associated to near cylinder junction zone in symmetry plane ($X/D = 0$), where the disturbances expect to be maximum due to forming the core of horseshoe vortex system in this region. However, the uncertainty in order less than 10% is common. The high range of uncertainty about 10% appears from variation of the flow physical properties affected by temperature. Thus, in this work, the accuracy of the gathered experimental data are acceptable and satisfactory.

Conclusion

In this paper, the results of computational fluid dynamics simulation of free convective flow around circular cylinder, mounted on heated vertical plate in comparison with experimental case measurements are presented to understand the effect of cylinder junction on heat transfer rate in the upstream region of the cylinder. As result of experimental data gathering and non-stationary numerical solution, the following conclusions can be drawn:

- In the case of numerical study, by performing time-based simulation using Boussinesq approximation to resolve governing equation, a steady state solution was obtained around the junction area.

- By comparing the numerical and experimental results of dimensionless temperature and heat transfer coefficient, the validity of computational approach, its configurations and applied simplifications are approved.

- The computed values of heat transfer coefficient, demonstrate the considerable effects of cylinder emplacement on the free convection heat transfer rate. The heat transfer coefficient experiences its maximum at the symmetry plane of computational domain ($X/D = 0$) around the cylinder junction.

- The significant role of horseshoe vortex system on heat transfer calculation [6, 8, 13] is illustrated partially and therefore is confirmed.

- By analyzing of heat transfer rate, this fact is emphasized that the high aspect ratio cylinder leads to enhance heat transfer coefficient in the upstream region of cylinder, as well as low aspect ratio cylinder [2,8]. Therefore, the increase of heat transfer coefficient in this region does not depend on cylinder height.

- In the vicinity region of the cylinder, flow is more complicated and the reliability of experimental study is decreased. In addition, associated uncertainty in this area is increased. It is occurred due to empirical measurement method sensitivity and arisen disturbances in reality.

- In the junction region of cylinder, for spatial coordinates less than $0.2D$, the results show that the compatibility of numerical solution and experimental data is reduced. Although the trend line of obtained results is matched each other favorably.

To sum up, by considering the limitations of experiment facility and numerical method deviations, it is appreciated to state that the accuracy of the obtained numerical and experimental results are satisfactory, acceptable and in good agreement. Furthermore, the used experimental and numerical approaches are valid to compute and analyze the characteristics of buoyancy-induced flow around the junction of horizontal cylinder, which located on the heated vertical plate.

References

[1] R.L.Simpson, Junction flows, *Annu. Rev. Fluid Mech.*, **33**(2001), 415–443.

[2] Y.S.Chumakov, A.M.Levchenya, H.Malah, The vortex structure formation around a circular cylinder placed on a vertical heated plate, *J. Phys. Math.: St. Petersburg State Polytechnical University*, **11(1)**(2018), 56–66. DOI: 10.18721/JPM.11108

[3] H.Bhowmik, A.Faisal, A.Al-Yaarubi, N.Al-Alawi, Analyses of Natural Convection Heat Transfer from a Heated Cylinder Mounted in Vertical Duct, *International Journal of Mechanical and Mechatronics Engineering*, **12(3)**(2018), 213–218.

[4] M.El-Amrani, M.Seaid, Numerical simulation of natural and mixed convection flows by Galerkin-characteristic method, *Int. J. Numer. Meth. Fluids*, **53**(2007), 1819–1845.

[5] S.M.Elsherbiny, M.A.Teamah, A.R.Moussa, Natural convection heat transfer from an isothermal horizontal square cylinder, *Alexandria Engineering Journal*, **56**(2017), 181–187.

[6] C.E.Clifford, M.L.Kimber, Optimizing laminar natural convection for a heat generating cylinder in a channel, *Journal of Heat Transfer*, **136**(2014), no. 112502.

[7] T.J.Praisner, C.R.Smith, The dynamics of the horseshoe vortex and associated endwall heat transfer—Part I: Temporal behavior, *Journal of Turbomachinery*, **128**(2006), 747–754.

[8] G.Polidori, J.Padet, *Journal of Flow Visualization & Image Processing*, **10**(2003), no. 1, 13–26. DOI: 10.1615/JFlowVisImageProc.v10.i12.20

[9] ANSYS Academic Research Mechanical, Release 16.2, Help System, Fluent Theory Guide, ANSYS, Inc. 2015.

[10] S.B.Koleshko, Y.S.Chumakov, Turbulent natural-convection heat transfer on a vertical surface under stepwise heating with inversion of heat flux, *High Temp.*, **48**(2010), 397–401 (in Russian).

[11] D.M.Christopher, B.X.Wang, Natural convection around a horizontal cylinder in a fluid-saturated porous media using Fourier series, *Transactions on Engineering Sciences*, **5**(1994), 4–11.

[12] H.Malah, Y.S.Chumakov, A.M.Levchenya, A study of the vortex structures around circular cylinder mounted on vertical heated plate, *AIP Conference Proceedings*, **1959**(2018), no. 1, 050018.

[13] S.O.Atayilmaz, Experimental and numerical study of natural convection heat transfer from horizontal concentric cylinders, *International Journal of Thermal Sciences*, **50**(2011), 1472–1483.

Экспериментальное и численное исследование свободно-конвективного теплообмена в окрестности стика кругового цилиндра с вертикальной нагретой поверхностью

Хамид Малах
Юрий С. Чумаков
Санкт-Петербургский политехнический университет Петра Великого
Санкт-Петербург, Российская Федерация

Аннотация. В данной работе исследуется влияние круглого цилиндра на трехмерные характеристики свободно-конвективного теплообмена. Круглый цилиндр установлен горизонтально на нагретой вертикальной поверхности и классифицируется как препятствие с высоким соотношением сторон, что означает, что высота цилиндра значительно больше чем его диаметр. Полученые результаты приведены для режима ламинарного потока. Кроме того, в ходе численного исследования определяющие дифференциальные уравнения решаются вокруг числа Грасгофа, равного 3×10^8. Чтобы определить области высоких градиентов температуры, температура потока показана в виде безразмерных контуров и диаграмм. В ближней области цилиндра вверх по потоку с помощью описания коэффициентов теплообмена, представленных градиентам температуры в заданных точках, изучается влияние расположения цилиндра на скорость теплопередачи. Как и ожидалось, значение коэффициента теплообмена, вызванного плавучестью, увеличивается в области соединения цилиндра и поверхности вверх по потоку. Максимальное значение коэффициента теплообмена появилось на плоскости симметрии исследуемой области, что соответствует расположению центра подковообразной структуры. В итоге обоснованность и достоверность численного и экспериментального подходов подтверждается путем вычисления девиации между числовыми и экспериментальными результатами, а также с помощью анализа неопределенности экспериментального метода.

Ключевые слова: поток, проходящий через стыки, свободно-конвективный теплообмен, ламинарный пограничный слой, коэффициент теплообмена.

Journal of Siberian Federal University. Mathematics & Physics 2020, 13(5), 644–654

DOI: 10.17516/1997-1397-2020-13-5-644-654
УДК 519.63

Finite Difference Schemes for Modelling the Propagation of Axisymmetric Elastic Longitudinal Waves

Vladimir M. Sadovskii*
Oxana V. Sadovskaya†
Evgenii A. Efimov‡
Institute of Computational Modeling SB RAS
Krasnoyarsk, Russian Federation

Received 16.06.2020, received in revised form 11.07.2020, accepted 16.08.2020

Abstract. An efficient finite difference shock-capturing scheme for the solution of direct seismic problems is constructed. Problem formulation is based on equations of the dynamics of elastic medium with axial symmetry. When implementating the scheme on multiprocessor computing systems, the two-cyclic splitting method with respect to spatial variables is used. One-dimensional systems of equations that arise in the context of splitting procedure are represented as subsystems for longitudinal, transverse and torsional waves. The case of longitudinal waves is considered in this paper. The results of simulations with the use of explicit grid-characteristic schemes and implicit schemes of the "predictor–corrector" type with controllable dissipation of energy are compared with exact solutions that describe propagation of monochromatic waves.

Keywords: elastic medium, cylindrical waves, splitting method, finite difference scheme, monotonicity, dissipativity, parallel computing.

Citation: V. M. Sadovskii, O.V. Sadovskaya, E. A. Efimov, Finite Difference Schemes for Modelling the Propagation of Axisymmetric Elastic Longitudinal Waves, J. Sib. Fed. Univ. Math. Phys., 2020, 13(5), 644–654. DOI: 10.17516/1997-1397-2020-13-5-644-654.

Introduction

The key problem in simulation of axially symmetric wave propagation consists in selection of appropriate approximation of lowest terms in equations of dynamic elasticity written in cylindrical system of coordinates. These terms cause the degeneracy of the equations at the axis of symmetry. Our aim is to find such appropriate approximations while remaining within the framework of conservative Godunov's scheme which is widely used in numerical solution of two- and three-dimensional problems [1]. The scheme can be modified to simulate wave propagation in granular and porous materials with different resistances to compression and tension [2–4], wave propagation and fracturing in blocky media [5–7] and other nonlinear processes.

Many methods were developed to solve axially symmetric equations of dynamic elasticity. The method of characteristics was used for the analysis of one-dimensional cylindrical wave [8, 9]. To solve two- and three-dimensional equations finite difference schemes based on the method of characteristics were proposed [10, 11]. These schemes allow one to compute the discontinuities of velocities and stresses. Such methods were applied for the analysis of wave processes in linear elastic, viscoelastic and elastic-plastic media [12, 13]. Numerical analysis of the dynamics of plates and shells of revolution was implemented with the methods based on the axially symmetric equations [14–16].

*sadov@icm.krasn.ru https://orcid.org/0000-0001-9695-0032
†o_sadov@icm.krasn.ru https://orcid.org/0000-0001-7222-8088
‡eugene6467@mail.ru

1. Equations of axially symmetric motion

Equations of the dynamic theory of elasticity with axially symmetric fields of stresses and velocities are written as

$$\begin{aligned}
&\rho\, r\,\frac{\partial v_r}{\partial t}=\frac{\partial(r\,\sigma_r)}{\partial r}+\frac{\partial(r\,\sigma_{rz})}{\partial z}-\sigma_\varphi\,, \qquad \rho\, r\,\frac{\partial v_\varphi}{\partial t}=\frac{\partial(r\,\sigma_{r\varphi})}{\partial r}+\frac{\partial(r\,\sigma_{\varphi z})}{\partial z}+\sigma_{r\varphi}\,,\\
&\rho\, r\,\frac{\partial v_z}{\partial t}=\frac{\partial(r\,\sigma_{rz})}{\partial r}+\frac{\partial(r\,\sigma_z)}{\partial z}\,, \qquad \frac{1}{E}\,\frac{\partial\sigma_r}{\partial t}-\frac{\nu}{E}\,\frac{\partial}{\partial t}\big(\sigma_\varphi+\sigma_z\big)=\frac{\partial v_r}{\partial r}\,,\\
&\frac{1}{E}\,\frac{\partial\sigma_\varphi}{\partial t}-\frac{\nu}{E}\,\frac{\partial}{\partial t}\big(\sigma_z+\sigma_r\big)=\frac{v_r}{r}\,, \qquad \frac{1}{E}\,\frac{\partial\sigma_z}{\partial t}-\frac{\nu}{E}\,\frac{\partial}{\partial t}\big(\sigma_r+\sigma_\varphi\big)=\frac{\partial v_z}{\partial z}\,,\\
&\frac{1}{\mu}\,\frac{\partial\sigma_{r\varphi}}{\partial t}=\frac{\partial v_\varphi}{\partial r}-\frac{v_\varphi}{r}\,, \qquad \frac{1}{\mu}\,\frac{\partial\sigma_{rz}}{\partial t}=\frac{\partial v_z}{\partial r}+\frac{\partial v_r}{\partial z}\,, \qquad \frac{1}{\mu}\,\frac{\partial\sigma_{\varphi z}}{\partial t}=\frac{\partial v_\varphi}{\partial z}\,.
\end{aligned} \tag{1}$$

Here $E=2\,\mu\,(1+\nu)$ is Young's modulus, μ is the shear modulus, ν is the Poisson ratio; the axes r and z of the cylindrical coordinate system are directed along the radius and the axis of symmetry, respectively. This form of equations is convenient for obtaining the equation of energy balance. Let us multiply the equations of motions by v_r, v_φ, v_z and the constitutive equations by $r\,\sigma_r$, $r\,\sigma_\varphi$, $r\,\sigma_z$, $r\,\sigma_{r\varphi}$, $r\,\sigma_{rz}$, $r\,\sigma_{\varphi z}$. Then the left-hand and right-hand sides of the obtained equations are summed up. The result is

$$\begin{aligned}
&\frac{\partial}{\partial t}\Big(\rho\, r\,\frac{v_r^2+v_\varphi^2+v_z^2}{2}+r\,W\Big) = \\
&=\frac{\partial}{\partial r}\Big(r\,v_r\,\sigma_r+r\,v_\varphi\,\sigma_{r\varphi}+r\,v_z\,\sigma_{rz}\Big)+\frac{\partial}{\partial z}\Big(r\,v_r\,\sigma_{rz}+r\,v_\varphi\,\sigma_{\varphi z}+r\,v_z\,\sigma_z\Big)\,,
\end{aligned} \tag{2}$$

where W is the elastic potential which is a quadratic form with respect to stresses:

$$W=\frac{1}{4\,\mu}\Big(\sigma_r^2+\sigma_\varphi^2+\sigma_z^2+2\,\sigma_{r\varphi}^2+2\,\sigma_{rz}^2+2\,\sigma_{\varphi z}^2-\frac{\nu}{1+\nu}\big(\sigma_r+\sigma_\varphi+\sigma_z\big)^2\Big)\,.$$

System (1) is a system of hyperbolic partial differential equations. It can be splitted into two independent subsystems. The first subsystem (equations 1, 3–6 and 8) describes motion in plane of symmetry, and the second subsystem (equations 2, 7 and 9) describes torsional motion. Motion in rz-plane is represented by superposition of longitudinal and transverse waves with velocities

$$c_p=\sqrt{\frac{2\,\mu}{\rho}\,\frac{1-\nu}{1-2\,\nu}}\,,\qquad c_s=\sqrt{\frac{\mu}{\rho}}\,,$$

respectively. The torsional wave has velocity c_s.

For seismic analysis, system of equations (1) is solved using the method of two-cyclic splitting with respect to spatial variables with solution of one-dimensional problems in parallel mode at different stages. Contrary to the conventional splitting, the method of two-cyclic splitting maintains the second-order accuracy if second-order finite difference schemes are used to solve one-dimensional systems [17]. Numerical implementation of splitting in z direction presents no difficulties. We have the system with constant coefficients after all derivatives with respect to r and terms containing r are canceled. Then obtained system is splitted into subsystems of plane P- and S-waves. These subsystems can be solved using Godunov's scheme [1] or grid-characteristic finite difference scheme with limiting reconstruction of Riemann invariants [18].

One-dimensional system of equations in the direction of radial axis r is splitted into three subsystems of longitudinal, transverse and torsional waves. The system contains terms without derivatives. Thus direct application of standard finite difference schemes for plane problem of the elasticity theory may lead to unwanted effects such as asymptotic instability, accumulation of rounding errors in simulations with many time steps or imbalance in momentum and energy. In consequence of these features the correctness of numerical results may be doubtful.

A common approach to suppress such effects is to use fully conservative finite difference schemes [19, 20] in combination with the method of artificial viscosity [21] that smoothes off oscillations of numerical solution in calculation of discontinuities due to artificial dissipation of energy. In the case of equations of the dynamic theory of elasticity, an approach to construct schemes with controllable dissipation of energy was developed [22, 23]. This approach is applied to the subsystem of equations for one-dimensional cylindrical longitudinal waves.

2. Finite difference schemes for propagation of cylindrical longitudinal wave

The equations for longitudinal waves can be written in an equivalent form in terms of the Lamé parameters $\lambda = 2\,\mu\,\nu/(1-2\,\nu)$ and μ as follows

$$\begin{aligned} &\rho\, r\,\frac{\partial v_r}{\partial t} = \frac{\partial (r\,\sigma_r)}{\partial r} - \sigma_\varphi\,, \qquad && \frac{\partial \sigma_r}{\partial t} = (\lambda + 2\,\mu)\,\frac{\partial v_r}{\partial r} + \lambda\,\frac{v_r}{r}\,,\\ &\frac{\partial \sigma_\varphi}{\partial t} = \lambda\,\frac{\partial v_r}{\partial r} + (\lambda+2\,\mu)\,\frac{v_r}{r}\,, \qquad && \frac{\partial \sigma_z}{\partial t} = \lambda\Big(\frac{\partial v_r}{\partial r} + \frac{v_r}{r}\Big)\,. \end{aligned} \tag{3}$$

Equation (2) of the energy balance for this subsystem takes the form

$$\frac{\partial}{\partial t}\Big(\rho\, r\,\frac{v_r^2}{2} + r\,W\Big) = \frac{\partial (r\, v_r\,\sigma_r)}{\partial r}\,.$$

Integrating equations (3) over the rectangular space–time grid leads to discrete equations of the "corrector" step

$$\begin{aligned} &\rho\, r^0\,\frac{\hat v_r - v_r}{\tau} = \frac{r^+\sigma_r^+ - r^-\sigma_r^-}{h} - \sigma_\varphi^0\,, \qquad && \frac{\hat\sigma_r - \sigma_r}{\tau} = (\lambda+2\,\mu)\,\frac{v_r^+ - v_r^-}{h} + \lambda\,\frac{v_r^0}{r^0}\,,\\ &\frac{\hat\sigma_\varphi - \sigma_\varphi}{\tau} = \lambda\,\frac{v_r^+ - v_r^-}{h} + (\lambda+2\,\mu)\,\frac{v_r^0}{r^0}\,, \qquad && \frac{\hat\sigma_z - \sigma_z}{\tau} = \lambda\,\frac{v_r^+ - v_r^-}{h} + \lambda\,\frac{v_r^0}{r^0}\,. \end{aligned} \tag{4}$$

In what follows, the values with circumflex belong to the upper time layer, and the values without circumflex belong to the lower time layer. The values with superscripts "$\pm$" belong to the right-hand and left-hand boundaries of a cell, respectively and $r^0 = (r^+ + r^-)/2$. The input velocity v_r^0 and stress σ_φ^0, alongside with $\sigma_r^\pm$ and $v_r^\pm$, are determined at the "predictor" step.

Multiplying equations (4) respectively by $(\hat v_r + v_r)/2$, $r^0(\hat\sigma_r + \sigma_r)/2$, $r^0(\hat\sigma_\varphi + \sigma_\varphi)/2$ and $r^0(\hat\sigma_z + \sigma_z)/2$, the difference analogue of the energy balance equation (2) for longitudinal waves is obtained. It has the form

$$\rho\, r^0\,\frac{\hat v_r^2 - v_r^2}{2\,\tau} + r^0\,\frac{\hat W - W}{\tau} = \frac{r^+ v_r^+\sigma_r^+ - r^- v_r^-\sigma_r^-}{h} - D\,,$$

$$D = \frac{r^+\sigma_r^+ - r^-\sigma_r^-}{h}\Big(\frac{v_r^+ + v_r^-}{2} - \frac{\hat v_r + v_r}{2}\Big) + \frac{v_r^+ - v_r^-}{h}\Big(\frac{r^+\sigma_r^+ + r^-\sigma_r^-}{2} - r^0\,\frac{\hat\sigma_r + \sigma_r}{2}\Big) + {}$$
$$+\ \sigma_\varphi^0\,\frac{\hat v_r + v_r}{2} - v_r^0\,\frac{\hat\sigma_\varphi + \sigma_\varphi}{2}\,.$$

The idea to control the dissipation of energy consists in setting expression for D explicitly in the form of a positive definite quadratic form. This form can be identically equal to zero. Then we have a dissipation-free (totally conservative) scheme.

Let the quadratic form be $D = \gamma\,(v_r^+ - v_r^-)^2/h^2$ with a free parameter $\gamma = O(h) \geqslant 0$, with an assumption that

$$v_r^0 = \frac{\hat v_r + v_r}{2} = \frac{v_r^+ + v_r^-}{2}\,, \quad \sigma_\varphi^0 = \frac{\hat\sigma_\varphi + \sigma_\varphi}{2} = \frac{\sigma_\varphi^+ + \sigma_\varphi^-}{2}\,,$$
$$\frac{r^+\sigma_r^+ + r^-\sigma_r^-}{2} - r^0\,\frac{\hat\sigma_r + \sigma_r}{2} = \gamma\,\frac{v_r^+ - v_r^-}{h}\,. \tag{5}$$

In this case, artificial dissipation of energy in the scheme is nonnegative. This automatically ensures stability of calculations. Further, it decreases with refinement of the grid and only depends on the rate of deformation of the medium. When $\gamma = 0$ the scheme is totally conservative, and the law of energy conservation is preserved on discrete level. However, in practice this is inapplicable in calculations of discontinuities and solutions with high gradients because it results in nonmonotonic solutions.

Considering equations (4), the closure equations in the scheme with controllable energy dissipation lead to the system

$$r^+\sigma_r^+ - r^-\sigma_r^- = a_{j-1/2}\,v_r^+ + b_{j-1/2}\,v_r^- + f_{j-1/2}\,,$$
$$r^+\sigma_r^+ + r^-\sigma_r^- = c_{j-1/2}\,v_r^+ + d_{j-1/2}\,v_r^- + g_{j-1/2}\,,$$

where coefficients $a_{j-1/2}$, $b_{j-1/2}$, $c_{j-1/2}$, $d_{j-1/2}$, $f_{j-1/2}$ and $g_{j-1/2}$ depend on the cell number $j = 1, 2, \ldots, n$ (fractional indices belong to the centers of cells). They are calculated as

$$a_{j-1/2} = \frac{\rho\,h\,r^0}{\tau} + (\lambda + 2\,\mu)\,\frac{\tau\,h}{4\,r^0} + \lambda\,\frac{\tau}{2}\,, \quad b_{j-1/2} = \frac{\rho\,h\,r^0}{\tau} + (\lambda + 2\,\mu)\,\frac{\tau\,h}{4\,r^0} - \lambda\,\frac{\tau}{2}\,,$$
$$c_{j-1/2} = \lambda\,\frac{\tau}{2} + (\lambda + 2\,\mu)\,\frac{\tau\,r^0}{h} + \frac{2\,\gamma}{h}\,, \quad d_{j-1/2} = \lambda\,\frac{\tau}{2} - (\lambda + 2\,\mu)\,\frac{\tau\,r^0}{h} - \frac{2\,\gamma}{h}\,,$$
$$f_{j-1/2} = h\,\sigma_\varphi - 2\,h\,r^0\,\frac{\rho\,v_r}{\tau}\,, \quad g_{j-1/2} = 2\,r^0\sigma_r\,.$$

Hence

$$2\,r^+\sigma_r^+ = (a_{j-1/2} + c_{j-1/2})\,v_r^+ + (b_{j-1/2} + d_{j-1/2})\,v_r^- + f_{j-1/2} + g_{j-1/2}\,,$$
$$2\,r^-\sigma_r^- = (c_{j-1/2} - a_{j-1/2})\,v_r^+ + (d_{j-1/2} - b_{j-1/2})\,v_r^- + g_{j-1/2} - f_{j-1/2}\,. \tag{6}$$

Equating these expressions and changing index j, we obtain system of linear equations with three-diagonal matrix to determine velocities $v_r^+ = v_r^j$ and $v_r^- = v_r^{j-1}$ at the cell boundaries:

$$A_j\,v_r^{j+1} + C_j\,v_r^j + B_j\,v_r^{j-1} = F_j \tag{7}$$

with coefficients

$$A_j = c_{j+1/2} - a_{j+1/2}\,, \quad C_j = d_{j+1/2} - b_{j+1/2} - a_{j-1/2} - c_{j-1/2}\,,$$
$$B_j = -\,b_{j-1/2} - d_{j-1/2}\,, \quad F_j = f_{j+1/2} + f_{j-1/2} - g_{j+1/2} + g_{j-1/2}\,.$$

This system is supplemented with the boundary condition $v_r^0 = 0$ at the axis of symmetry in the first cell of the grid and with the condition $v_r^n = v$ in the last cell of the grid if the particle velocity v is set at $r = R$ or with the condition

$$(a_{n-1/2} + c_{n-1/2})\,v_r^n + (b_{n-1/2} + d_{n-1/2})\,v_r^{n-1} + f_{n-1/2} + g_{n-1/2} = 2\,R\,\sigma\,,$$

which follows from (6) if the external stress σ is set at the boundary. In either case, the system of equations with boundary conditions is solved with the use of the three-point sweep method.

In this manner, the algorithm of transition to the next time level starts with computation of $v_r^\pm$ and $\sigma_r^\pm$ using equations (6), (7) at the "predictor" stage. Then, v_r^0 and σ_r^0 are determined from (5). The final computations of $\hat{v}_r$, $\hat{\sigma}_r$, $\hat{\sigma}_\varphi$ and $\hat{\sigma}_z$ are performed using equations (4) of the "corrector" stage.

For comparison, we consider three versions of explicit finite difference schemes based on the solution of the Riemann problem. The schemes are constructed by means of approximation of the first equation in (3) written in equivalent (nonconservative) form

$$\rho\,\frac{\partial v_r}{\partial t} = \frac{\partial \sigma_r}{\partial r} + \frac{\sigma_r - \sigma_\varphi}{r}\,.$$

The "predictor–corrector" scheme with explicit approximation of lowest terms

$$\rho\,\frac{\hat{v}_r - v_r}{\tau} = \frac{\sigma_r^+ - \sigma_r^-}{h} + \frac{\sigma_r - \sigma_\varphi}{r^0}\,, \qquad \frac{\hat{\sigma}_r - \sigma_r}{\tau} = (\lambda + 2\,\mu)\,\frac{v_r^+ - v_r^-}{h} + \lambda\,\frac{v_r}{r^0}\,,$$

$$\frac{\hat{\sigma}_\varphi - \sigma_\varphi}{\tau} = \lambda\,\frac{v_r^+ - v_r^-}{h} + (\lambda + 2\,\mu)\,\frac{v_r}{r^0}\,, \qquad \frac{\hat{\sigma}_z - \sigma_z}{\tau} = \lambda\Big(\frac{v_r^+ - v_r^-}{h} + \frac{v_r}{r^0}\Big),$$

$$v_r^+ - \frac{\sigma_r^+}{\rho\,c_p} = v_r - \frac{\sigma_r}{\rho\,c_p}\,, \qquad v_r^- + \frac{\sigma_r^-}{\rho\,c_p} = v_z + \frac{\sigma_r}{\rho\,c_p}\,,$$

is applicable if the Courant number $K_p = c_p\,\tau/h$ is in the range from 0 to 0.8. When the Courant number exceeds 0.8 the parasitic oscillations appear in the vicinity of the axis of symmetry. The amplitude of oscillations unlimitedly increases with increasing K_p from 0.9 to 1. This results in distortion of the solution. When the value of K_p is low, the viscosity of the scheme considerably smoothes off the solution. For these two reasons, it is inadvisable to use this scheme in simulations.

The scheme with implicit approximation of lowest terms is derived by replacing the stresses σ_r, σ_φ and the velocity v_r in the lowest terms with $\hat{\sigma}_r$, $\hat{\sigma}_\varphi$ and $\hat{v}_r$. This scheme is stable and monotone for $0 < K_p \leqslant 1$ but it also smoothes off the solution at sufficiently low values of the Courant number.

The scheme with implicit approximation of lowest terms by the Crank–Nicolson method includes

$$\frac{\hat{\sigma}_r + \sigma_r}{2}\,, \quad \frac{\hat{\sigma}_\varphi + \sigma_\varphi}{2}\,, \quad \frac{\hat{v}_r + v_r}{2}\,.$$

In regard to the accuracy of numerical solution, this scheme has certain advantages over the explicit and implicit schemes.

3. Results of computations

Computational schemes were verified by comparing the results of simulations with the exact solution obtained for the monochromatic wave with frequency ω with the use of the method of separation of variables. The exact solution has the form

$$v_r = \frac{\sigma_0}{\rho\,c_p}\,\sin\omega t\,J_1(\xi_p)\,, \quad \sigma_r = \frac{\sigma_0}{\lambda + 2\,\mu}\,\cos\omega t\Big((\lambda + 2\,\mu)\,J_2(\xi_p) - \frac{2\,(\lambda+\mu)}{\xi_p}\,J_1(\xi_p)\Big),$$

$$\sigma_\varphi = \frac{\sigma_0}{\lambda + 2\,\mu}\,\cos\omega t\Big(\lambda\,J_2(\xi_p) - \frac{2\,(\lambda+\mu)}{\xi_p}\,J_1(\xi_p)\Big), \quad \sigma_z = \frac{\lambda\,\sigma_0}{\lambda + 2\,\mu}\,\cos\omega t\Big(J_2(\xi_p) - \frac{2}{\xi}\,J_1(\xi_p)\Big),$$

where $\xi_p = \omega\,r/c_p$ is the dimensionless variable, $J_k(x)$ is the Bessel function of an integer order k.

Tabs. 1 − 4 present relative errors of the schemes for various frequencies as a function of the Courant number. The dimensionless frequency $\bar{\omega} = \omega\,R/c_s$, where R is the radius of the computational domain, was varied between 10 and 50. At such frequencies the number of half-waves in the computational domain varies between one and a half and seven and a half (Fig. 1).

Table 1. Relative errors for the dissipation-free scheme ($\gamma = 0$)

K_p \ $\bar{\omega}$	10	20	30	40	50
0.5	0.00023	0.00110	0.00837	0.01408	0.03500
0.75	0.00009	0.00045	0.00405	0.00701	0.01818
1	0.00019	0.00081	0.00201	0.00349	0.00563
1.25	0.00049	0.00210	0.00977	0.01621	0.03527
1.5	0.00085	0.00372	0.01920	0.03209	0.07125

Table 2. Relative errors for the scheme with explicit approximation of lowest terms

K_p \ $\bar{\omega}$	10	20	30	40	50
0.5	0.04334	0.09600	0.37513	0.40646	0.63148
0.75	0.02148	0.05053	0.19970	0.25226	0.39363

Table 3. Relative errors for the scheme with implicit approximation of lowest terms

K_p \ $\bar{\omega}$	10	20	30	40	50
0.5	0.04802	0.10547	0.39297	0.41076	0.65137
0.75	0.02860	0.06545	0.23341	0.26373	0.44424
1	0.01025	0.02666	0.03320	0.04613	0.05591

Table 4. Relative errors forn the scheme with Crank–Nicolson approximation of lowest terms

K_p \ $\bar{\omega}$	10	20	30	40	50
0.5	0.03824	0.09015	0.36576	0.40030	0.62509
0.75	0.01350	0.04288	0.18149	0.23903	0.37536
0.97	0.00912	0.02940	0.02788	0.05726	0.04975
1	0.01220	0.03379	0.05646	0.06663	0.12026

To calculate the error of numerical solution a discrete equivalent of the norm of the space $L_\infty\big(0,T;L_2(0,R)\big)$ was used:

$$\Big\|\big(v_r,\sigma_r,\sigma_\varphi,\sigma_z\big)\Big\| = \sup_{0<t<T}\sqrt{\pi\int_0^R\Big(\rho\,\frac{v_r^2}{2}+W\Big)dr^2}\,.$$

The time T was set so that the cylindrical longitudinal wave in the interval $(0,T)$ travels a distance $2\,R$ with single reflection from the axis of symmetry.

The finite difference grid has 200 cells. Analysis of the data in the tables shows that numerical solution obtained with implicit approximation of lowest terms and with approximation by

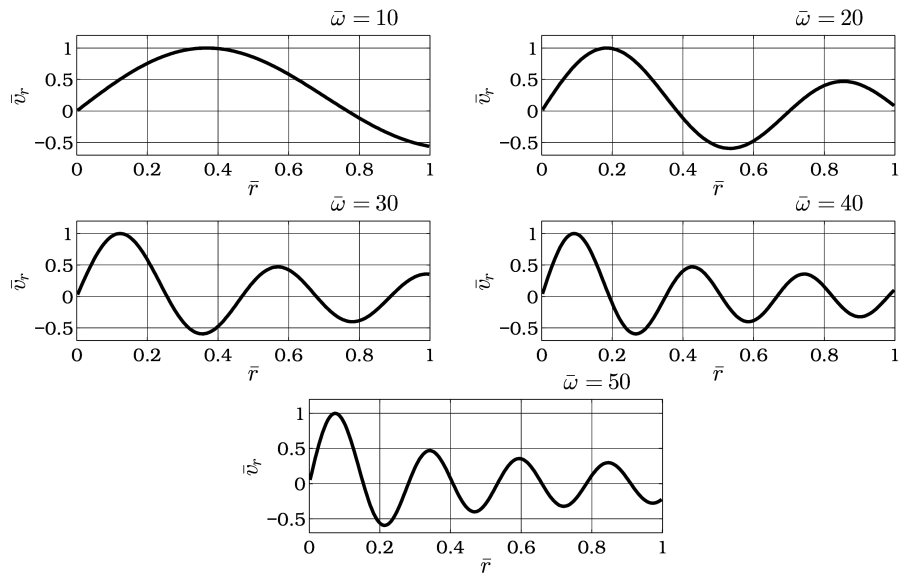

Fig. 1. Exact solution for cylindrical longitudinal waves; dimensionless velocity versus dimensionless distance

the Crank–Nicolson method is inaccurate if one half-wave contains less than $60-70$ cells. The dissipation-free scheme gives more accurate results at all frequencies within the specified range.

Figs. 2 and 3 show the velocity profiles behind the front of a strong discontinuity when sudden constant stress is applied at the boundary of the domain. These results are obtained using the scheme with Crank–Nicolson approximation (Fig. 2) and the dissipation-free scheme (Fig. 3).

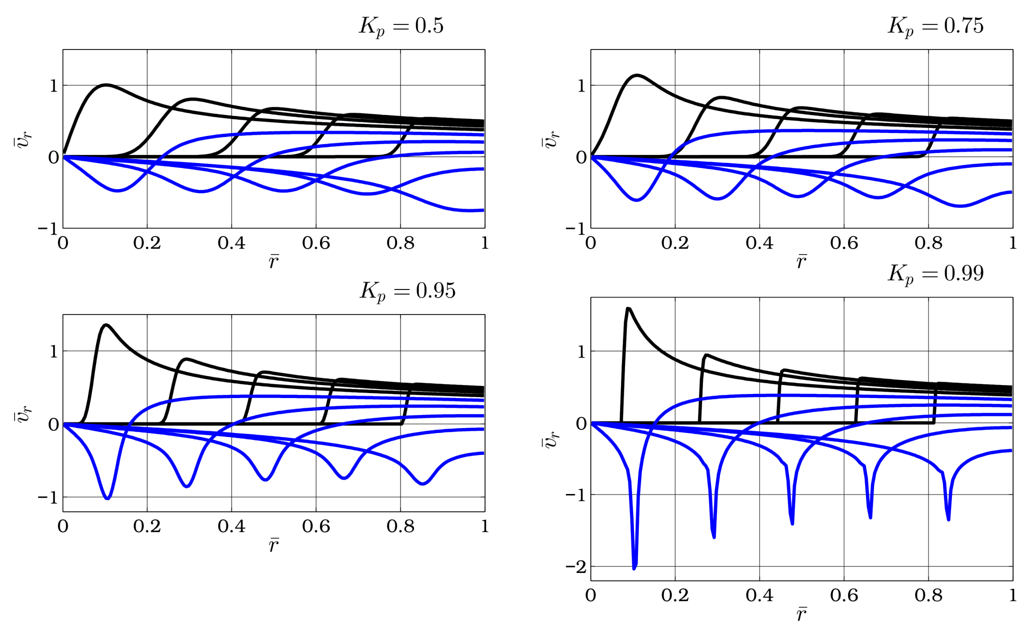

Fig. 2. Velocity profiles behind the front of discontinuity: scheme with approximation of lowest terms by the Crank–Nicolson method

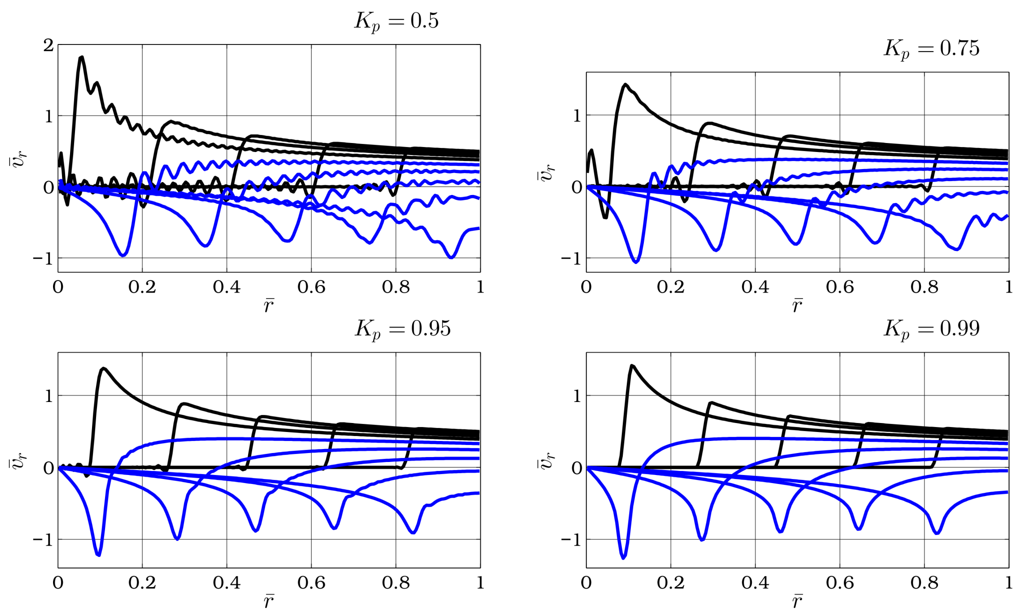

Fig. 3. Velocity profiles behind the front of discontinuity: dissipation-free scheme

In the case of the dissipation-free scheme, the profiles of velocities and stresses are only monotonous at $K_p = 1$, while at $K_p = 0.9$ parasitic oscillations appear ahead of the wave front, and they grow as the Courant number decreases. The oscillations are smoothed off setting $\gamma > 0$ in doing so the artificial dissipation of energy is introduced or replacing the sudden application of stress at the boundary by a monotone increasing stress that is changed from zero to a constant value during not less than 10 time steps.

To implement the scheme with controllable energy dissipation on computational clusters it is possible to use the iterative process which demonstrates rarely high rate of convergence of approximate solutions in umerical experiments. It appears that errors presented in Tab. 1 are already obtained with one or two iterations.

The problem consists in parallel computing at the "predictor" stage of the finite difference scheme. System of equations (7) is changed at the junction points of the neighbor processors and corresponding three-point equations of the system are replaced by the relations of the Godunov scheme:

$$v_r^j = \frac{v_{r\,j+1/2} + v_{r\,j-1/2}}{2} + \frac{\sigma_{r\,j+1/2} - \sigma_{r\,j-1/2}}{2\,\rho\,c_p},$$

where fractional indices mark the velocities and stresses that belong to the boundary of grid cell of the neighbor processors. This procedure allows one to implement the three-point sweep method in parallel mode on a cluster and obtain solution in the first approximation.

Referring to figures given above, the nonconservative finite difference scheme with approximation of lowest terms by the Crank–Nicolson method provides much more reliable results for solutions with discontinuities over the whole range of the Courant number $K_p \leqslant 1$ (this is stability condition of the scheme).

It is worth to mention that the analogous schemes with conservative equations (3) inadequately distort the pattern of wave reflection from the axis of symmetry even in the case of smooth solutions, and this may finally result in the total loss of accuracy.

Conclusions

The finite difference scheme with controllable dissipation of energy and typical grid-characteristic schemes of the "predictor–corrector" type were considered. The comparison of the results of computations shows that the scheme with controllable dissipation has undisputable advantages over the other approaches in the case of smooth solutions. As for solutions with discontinuities, the grid-characteristics schemes are preferable due to their monotonicity. In the case of solutions with discontinuities the scheme with controllable energy dissipation produces parasitic oscillations. To smooth off these oscillations one should introduce artificial energy dissipation.

This work is supported by the Krasnoyarsk Mathematical Center and financed by the Ministry of Science and Higher Education of the Russian Federation in the framework of the establishment and development of regional Centers for Mathematics Research and Education (Agreement No. 075-02-2020-1631). Computations were performed using computational cluster at the Institute of Computational Modeling SB RAS, Krasnoyarsk, and at the Joint Supercomputer Center of the Russian Academy of Sciences, Moscow.

References

[1] S.K.Godunov, A.V.Zabrodin, M.Ya.Ivanov, A.N.Kraiko, G.P.Prokopov, Numerical Solution of Multidimensional Problems of Gas Dynamics, Moscow, Nauka, 1976 (in Russian).

[2] O.Sadovskaya, V.Sadovskii, Mathematical Modeling in Mechanics of Granular Materials, Ser. Advanced Structured Materials, vol. 21, Heidelberg – New York – Dordrecht – London, Springer, 2012. DOI: 10.1007/978-3-642-29053-4

[3] O.V.Sadovskaya, V.M.Sadovskii, Numerical implementation of mathematical model of the dynamics of a porous medium on supercomputers of cluster architecture, AIP Conf. Proc., vol. 1684, 2015, 070005-1–070005-9. DOI: 10.1063/1.4934306

[4] V.M.Sadovskii, O.V.Sadovskaya, Analyzing the deformation of a porous medium with account for the collapse of pores, *J. Appl. Mech. Techn. Phys.*, **57**(2016), no. 5, 808–818. DOI: 10.1134/S0021894416050072

[5] V.M.Sadovskii, O.V.Sadovskaya, Modeling of elastic waves in a blocky medium based on equations of the Cosserat continuum, *Wave Motion*, **52**(2015), 138–150. DOI: 10.1016/j.wavemoti.2014.09.008

[6] V.M.Sadovskii, O.V.Sadovskaya, A.A.Lukyanov, Modeling of wave processes in blocky media with porous and fluid-saturated interlayers, *J. Comput. Phys.*, **345**(2017), 834–855. DOI: 10.1016/j.jcp.2017.06.001

[7] V.M.Sadovskii, O.V.Sadovskaya, Supercomputer Modeling of Wave Propagation in Blocky Media Accounting Fractures of Interlayers, In: Nonlinear Wave Dynamics of Materials and Structures (Eds. H. Altenbach, V.A. Eremeyev, I.S. Pavlov, A.V. Porubov), Ser. Advanced Structured Materials, vol. 122, chapt. 22, Cham, Springer, 2020, 379–398. DOI: 10.1007/978-3-030-38708-2_22

[8] P.F.Sabodash, R.A.Cherednichenko, Application of the method of spatial characteristics to the solution of axisymmetric problems on the propagation of elastic waves, *Prikl. Mekh. Tehn. Fiz.*, **12**(1971), no. 4, 101–109 (in Russian).

[9] K.M.Magomedov, A.S.Kholodov, Grid-Characteristic Numerical Methods, Moscow, Nauka, 1988 (in Russian).

[10] V.N.Kukudzhanov, Difference Methods for the Solution of Problems of Mechanics of Deformable Media, Moscow, MFTI, 1992 (in Russian).

[11] V.N. Kukudzhanov, Numerical Continuum Mechanics, Ser. De Gruyter Studies in Mathematical Physics, vol. 15, Berlin – Boston, De Gruyter, 2013. DOI: 10.1515/9783110273380

[12] N.G.Burago, A.B.Zhuravlev, I.S.Nikitin, Continuum model and method of calculating for dynamics of inelastic layered medium, *Math. Models Comput. Simul.*, **11**(2019), no. 3, 488–498. DOI: 10.1134/S2070048219030098

[13] I.S.Nikitin, Theory of Inelastic Layered and Blocky Media, Moscow, Fizmatlit, 2019 (in Russian).

[14] N.G.Burago, V.N.Kukudzhanov, Buckling and Supercritical Deformations of Elastic-Plastic Shells under Axial Symmetry, In: Collection of Numerical Methods in the Mechanics of Deformable Solids, Moscow, Computing Center of the USSR Academy of Sciences, 1978, 47–66 (in Russian).

[15] V.G.Bazhenov, E.V.Igonicheva, Nonlinear Processes of Shock Buckling of Elastic Structural Elements in the Form of Orthotropic Shells of Rotation, N. Novgorod, UNN, 1991 (in Russian).

[16] V.G.Bazhenov, D.T.Chekmarev, Solution of the Problems of Non-Stationary Dynamics of Plates and Shells by the Variational-Difference Method, N. Novgorod, UNN, 2000 (in Russian).

[17] G.I.Marchuk, Splitting and Alternating Direction Methods, In: Handbook of Numerical Analysis, vol. 1 (Eds. P.G. Ciarlet, J.-L. Lions), North-Holland, Elsevier, 1990, 197–462. DOI: 10.1016/S1570-8659(05)80035-3

[18] A.G.Kulikovskii, N.V.Pogorelov, A.Yu.Semenov, Mathematical Aspects of Numerical Solution of Hyperbolic Systems, Ser. Monographs and Surveys in Pure and Applied Mathematics, vol. 118, Boca Raton – London – New York – Wasington, Chapman & Hall/CRC, 2001. DOI: 10.1201/9781482273991

[19] Yu.P.Popov, A.A.Samarskii, Completely conservative difference schemes, *USSR Comput. Math. Math. Phys.*, **9**(1969), no. 4, 296–305. DOI: 10.1016/0041-5553(69)90049-4

[20] A.A.Samarskii, The Theory of Difference Schemes, Ser. Pure and Applied Mathematics, New York – Basel, CRC Press, 2001. DOI: 10.1201/9780203908518

[21] B.L.Roždestvenskii, N.N.Janenko, Systems of Quasilinear Equations and Their Applications to Gas Dynamics, Ser. Translations of Mathematical Monographs, vol. 55, Providence, American Mathematical Society, 1983.

[22] G.V.Ivanov, V.D.Kurguzov, Schemes for solving one-dimensional problems of the dynamics of inhomogeneous elastic bodies on the basis of approximation by linear polynomials, *Dynam. Contin. Medium*, **49**(1981), 27–44 (in Russian).

[23] G.V.Ivanov, Yu.M.Volchkov, I.O.Bogulskii, S.A.Anisimov, V.D.Kurguzov, Numerical Solution of Dynamic Elastic-Plastic Problems of Deformable Solids, Novosibirsk, Sib. Univ. Izd., 2002 (in Russian).

Разностные схемы для анализа продольных волн на основе осесимметричных уравнений динамической теории упругости

Владимир М. Садовский
Оксана В. Садовская
Евгений А. Ефимов
Институт вычислительного моделирования СО РАН
Красноярск, Российская Федерация

Аннотация. Цель исследования состоит в построении экономичной разностной схемы сквозного счета для решения прямых задач сейсмики на основе уравнений динамики упругой среды в осесимметричной постановке. При численной реализации схемы на многопроцессорных вычислительных системах применяется метод двуциклического расщепления по пространственным переменным. Одномерные системы уравнений на этапах расщепления распадаются на подсистемы продольных, поперечных и крутильных волн. В данной работе рассматривается случай продольных волн. Проводится сравнение явных сеточно-характеристических схем и неявных схем типа "предиктор–корректор" с контролируемой диссипацией энергии на точных решениях, описывающих бегущие монохроматические волны.

Ключевые слова: упругая среда, цилиндрические волны, метод расщепления, разностная схема, монотонность, диссипативность, параллельная реализация.

Retraction Note to: A Recursive Algorithm for Estimating the Correlation Matrix of the Interference Based on the QR Decomposition

The authors acknowledge that text of the paper [1] partly correlates with the material of the monograph [2] and sincerely apology to the reviewers, editorial staff and the readers of *Journal of Siberian Federal University. Mathematics* & *Physics*. The paper has been retracted.

References

[1] Valery N.Tyapkin, Dmitry D. Dmitriev, Andrey B. Gladyshev, Peter Yu. Zverev A Recursive Algorithm for Estimating the Correlation Matrix of the Interference Based on the QR Decomposition, Journal of Siberian Federal University. Mathematics & Physics, 2020, 13(2), 160-169.

[2] V. I. Dzhigan, Adaptive signal filtering: theory and algorithms, http://www.technosphera.ru/lib/book/333.