ISSN 1563 – 0277
eISSN 2617 – 4871

ӘЛ-ФАРАБИ атындағы ҚАЗАҚ ҰЛТТЫҚ УНИВЕРСИТЕТІ

ХАБАРШЫ

Математика, механика, информатика сериясы

КАЗАХСКИЙ НАЦИОНАЛЬНЫЙ УНИВЕРСИТЕТ имени АЛЬ-ФАРАБИ

ВЕСТНИК

Серия математика, механика, информатика

AL-FARABI KAZAKH NATIONAL UNIVERSITY

Journal of Mathematics, Mechanics and Computer Science

№3 (115)

Алматы
«Қазақ университеті»
2022

Зарегистрирован в Министерстве информации и коммуникаций Республики Казахстан, свидетельство №16508-Ж от 04.05.2017 г. (Время и номер первичной постановки на учет №766 от 22.04.1992 г.). Язык издания: казахский, русский, английский. Выходит 4 раза в год. Тематическая направленность: теоретическая и прикладная математика, механика, информатика.

Научное издание
Вестник КазНУ. Серия “Математика, механика, информатика”, № 3 (115) 2022.
Редактор – С.Е. Айтжанов. Компьютерная верстка – С.Е. Айтжанов

ИБ N 15[illegible]
Формат 60 × 84 1/8. Бумага офсетная. Печать цифровая. Объем 11,18 п.л.
Заказ N 13463. Издательский дом “Қазақ университеті”
Казахского национального университета им. аль-Фараби. 050040, г. Алматы, пр.аль-Фараби, 71, КазНУ.
Отпечатано в типографии издательского дома “Қазақ университеті”.

1-бөлім **Раздел 1** **Section 1**

Математика **Математика** **Mathematics**

IRSTI 27.31.21 DOI: https://doi.org/10.26577/JMMCS.2022.v115.i3.01

S.E. Aitzhanov[1*], **J. Ferreira**[2], **K.A. Zhalgassova**[3]
[1]Al-Farabi Kazakh National University, Kazakhstan, Almaty
[2]Federal University of Fluminense, Brazil, Volta Redonda
[3] M.Auezov South Kazakhstan University, Kazakhstan, Shymkent
*e-mail: aitzhanovserik81@gmail.com

SOLVABILITY OF THE INVERSE PROBLEM FOR THE PSEUDOHYPERBOLIC EQUATION

This paper investigates the solvability of the inverse problem of finding a solution and an unknown coefficient in a pseudohyperbolic equation known as the Klein-Gordon equation. A distinctive feature of the given problem is that the unknown coefficient is a function that depends only on the time variable. The problem is considered in the cylinder, the conditions of the usual initial-boundary value problem are set. The integral overdetermination condition is used as an additional condition. In this paper, the inverse problem is reduced to an equivalent problem for the loaded nonlinear pseudohyperbolic equation. Such equations belong to the class of partial differential equations, not resolved with respect to the highest time derivative, and they are also called composite type equations. The proof uses the Galerkin method and the compactness method (using the obtained a priori estimates). For the problem under study, the authors prove existence and uniqueness theorems for the solution in appropriate classes.
Key words: Pseudohyperbolic equation, inverse problem, Klein-Gordon equation, Galerkin method, compactness method, existence, uniqueness.

С.Е. Айтжанов[1*], Ж. Феррейра[2], К.А. Жалғасова[3]
[1]Әл-Фараби атындағы Қазақ ұлттық университеті, Қазақстан, Алматы қ.
[2]Флюминенс Федерал университеті, Бразилия, Вольта-Редонда қ.
[3]М.Әуезов атындағы Оңтүстік Қазақстан университеті, Қазақстан, Шымкент қ.
*e-mail: aitzhanovserik81@gmail.com

Псевдогиперболалық теңдеу үшін кері есептің шешімділігі

Мақалада Клейн-Гордон теңдеуі деген атпен белгілі псевдогиперболалық теңдеудің шешімін және оң жақ коэффициентін табу кері есебі зерттеледі. Бұл есеп ізделінді коэффициенттің тек уақыттан тәуелді функция болуымен ерекшеленеді. Есеп цилиндрлік аймақта қарастырылады, әдеттегідей бастапқы-шеттік есептің шарттары қойылады. Қосымша шарт ретінде интегралдық түрдегі артық анықталған шарт берілген. Бұл жұмыста кері есеп жүктелген сызықтық емес псевдогиперболалық теңдеу үшін қойылған эквивалентті есепке келтіріледі. Мұндай теңдеулер уақыт бойынша ең жоғары туындыға қатысты шешілмеген дербес туындылы дифференциалдық теңдеулер класына жатады және оларды құрама типті теңдеулер деп те атайды. Дәлелдеуде Галеркин әдісі және компакт әдісі (априорлық бағалаулар алу арқылы) қолданылады. Жұмыста зерттеліп отырған есептің сәйкес кластардағы шешімнің бар болу және жалғыздық теоремалары дәлелденеді.
Түйін сөздер: Псевдогиперболалық теңдеу, кері есеп, Клейн-Гордон теңдеуі, Галеркин әдісі, компакт әдісі, шешімнің бар болуы және жалғыздығы.

С.Е. Айтжанов[1*], Ж. Феррейра[2], К.А. Жалгасова[3]
[1]Казахский национальный университет имени аль-Фараби, Казахстан, г. Алматы
[2]Федеральный университет Флуминенсе, Бразилия, г. Вольта-Редонда
[3]Южно-Казахстанский университет имени М.Ауезова, Казахстан, г. Чимкент
*e-mail: aitzhanovserik81@gmail.com

Разрешимость обратной задачи для псевдогиперболического уравнения

Исследуется разрешимость обратной задачи нахождения решения и неизвестного коэффициента в псевдогиперболическом уравнении, известного как уравнение Клейна-Гордона. Отличительной особенностью изучаемой задачи является то, что неизвестный коэффициент является функцией, зависящей лишь от временной переменной. Задача рассматривается в цилиндрической области, задаются условия обычной начально-краевой задачи. В качестве дополнительного условия используется условие интегрального переопределения. В работе обратная задача сводится к эквивалентной задаче для нагруженного нелинейного псевдогиперболического уравнения. Подобные уравнения относятся к классу дифференциальных уравнений в частных производных, не разрешенные относительно старшей производной по времени и они также называются уравнениями составного типа. При доказательстве применяются метод Галеркина и метод компактности (с использованием полученных априорных оценок). Для изучаемой задачи авторы доказывают теоремы существования и единственности решения в рассматриваемых классах.
Ключевые слова: Псевдогиперболическое уравнение, обратная задача, уравнение Клейна-Гордона, метод Галеркина, метод компактности, существование, единственность.

1 Introduction

The work is devoted to the study of the solvability of the inverse problem of reCЃovering an external influence in the pseudohyperbolic equation known as the Klein-Gordon equation. Nowadays, inverse problems have become a powerful and rapidly developing field of knowledge, penetrating almost all areas of mathematics. Similar inverse problems arise naturally in the mathematical modeling of certain processes occurring in the media with unknown characteristics. Since it is the characteristics of the medium that determine the coefficients of the corresponding differential equation or the coefficients of the external influence. The Klein-Gordon equation plays an important role in mathematical physics. This equation is used in modeling various phenomena of relativistic quantum mechanics [1] and nonlinear optics, in studying the behavior of elementary particles and dislocation propagation in crystals, as well as in studying nonlinear wave equations [2]. For such equations, many problems have been investigated in different formulations by various methods [3]-[14].

In this paper, the inverse problem under study is reduced to an equivalent problem for the loaded nonlinear pseudohyperbolic equation. Pseudohyperbolic equations belong to the class of partial differential equations, not solved with respect to the highest time derivative, and they are also known as composite type equations. Initial-boundary value problems for linear and nonlinear pseudohyperbolic equations were studied in various works [15]-[20]. Moreover, it is necessary to note the works [21]-[25], where studied the qualitative properties of solutions of inverse problems for hyperbolic type equations.

In the cylinder $Q_T = \{(x,t) : \ x \in \Omega, \ 0 < t < T\}$ we consider the inverse problem of reCЃovering the right-hand side of the Klein-Gordon equation

$$u_{tt}-\chi\Delta u_t-\left(a_0+a_1\left\|\nabla u\right\|_{2,\Omega}^{2r}\right)\Delta u+|u_t|^{q-2}u_t=b(x,t)|u|^{p-2}u+f(t)h(x,t),\ \ (x,t)\in Q_T,\ (1)$$

with initial conditions

$$u(x,0)=u_0(x),\ \ u_t(x,0)=u_1(x),\ \ x\in\Omega, \qquad (2)$$

the boundary condition

$$u|_S=0, \qquad (3)$$

and the overdetermination condition

$$\int\limits_{\Omega}u(x,t)\omega(x)dx=\varphi(t),\ t\in(0,T). \qquad (4)$$

Here $\Omega\subset R^N,\ N\geq 1$ is bounded area, $\partial\Omega$ is sufficiently smooth boundary, $b(x,t),\ h(x,t)$, $u_0(x),\ u_1(x),\ \omega(x),\ \varphi(t)$ are the given functions, $\chi,\ a_0,\ a_1,\ p,\ q$ and r are positive constants. Let the given functions of the problem (1)-(4) satisfy the conditions

$$\begin{array}{l}\omega\in L_2(\Omega)\bigcap\overset{0}{W}{}_2^2(\Omega),\\ h(x,t)\in C^1(Q_T),\quad h_1(t)\equiv\int\limits_{\Omega}h(x,t)\omega(x)dx\neq 0,\ \ \forall t\in[0,T],\end{array} \qquad (5)$$

$$\begin{array}{l}\varphi(t)\in W_2^2(0,T),\\ \int\limits_{\Omega}u_0(x)\omega(x)dx=\varphi(0),\ \ \int\limits_{\Omega}u_1(x)\omega(x)dx=\varphi'(0),\\ u_0\in\overset{0}{W}{}_2^2(\Omega),\quad u_1\in\overset{0}{W}{}_2^1(\Omega).\end{array} \qquad (6)$$

2 Materials and methods

2.1 The Equivalent Problem

Lemma 1. The problem (1)-(4) is equivalent to the next problem for nonlinear pseudoparabolic equation containing nonlinear nonlocal operator from function $u(x,t)$

$$u_{tt}-\chi\Delta u_t-\left(a_0+a_1\left\|\nabla u\right\|_{2,\Omega}^{2r}\right)\Delta u+|u_t|^{q-2}u_t=b(x,t)|u|^{p-2}u+F(t,u)h(x,t),\ \ x\in\Omega,\ \ t>0, \qquad (7)$$

$$u(x,0)=u_0(x),\ \ x\in\Omega,\quad u|_S=0. \qquad (8)$$

Here

$$\begin{array}{l}F(t,u)=\frac{1}{h_1(t)}\left(\varphi''(t)+\chi\int\limits_{\Omega}\nabla u_t\nabla\omega dx+\left(a_0+a_1\left\|\nabla u\right\|_{2,\Omega}^{2r}\right)\int\limits_{\Omega}\nabla u\nabla\omega dx+\right.\\ \left.+\int\limits_{\Omega}|u_t|^{q-2}u_t\omega dx-\int\limits_{\Omega}b(x,t)|u|^{p-2}u\omega dx\right).\end{array} \qquad (9)$$

Proof. Indeed, it follows from equation (1) that

$$
\begin{aligned}
&\int_\Omega (u_{tt}-\chi\Delta u_t)\omega dx-\left(a_0+a_1\left\|\nabla u\right\|_{2,\Omega}^{2r}\right)\int_\Omega \Delta u\omega dx- \\
&-\int_\Omega b(x,t)|u|^{p-2}u\omega dx=\int_\Omega f(t)h(x,t)\omega dx,
\end{aligned}
\tag{10}
$$

next, if conditions (4) and (5) are performed, then

$$
\begin{aligned}
&F(t,u)=\tfrac{1}{h_1(t)}\left(\varphi''(t)+\chi\int_\Omega \nabla u_t\nabla\omega dx+\left(a_0+a_1\left\|\nabla u\right\|_{2,\Omega}^{2r}\right)\int_\Omega \nabla u\nabla\omega dx+\right. \\
&\left.+\int_\Omega |u_t|^{q-2}u_t\omega dx-\int_\Omega b(x,t)|u|^{p-2}u\omega dx\right).
\end{aligned}
\tag{11}
$$

Therefore, the relation (9) is satisfied.

Now let us consider the problem (7)-(8). If the relation (9) is satisfied, then equality (11) obviously follows from it. Then

$$
\begin{aligned}
&F(t,u)=\tfrac{1}{h_1(t)}\left(\varphi''(t)+\chi\int_\Omega \nabla u_t\nabla\omega dx+\left(a_0+a_1\left\|\nabla u\right\|_{2,\Omega}^{2r}\right)\int_\Omega \nabla u\nabla\omega dx+\right. \\
&\left.+\int_\Omega |u_t|^{q-2}u_t\omega dx-\int_\Omega b(x,t)|u|^{p-2}u\omega dx\right)= \\
&=\tfrac{1}{h_1(t)}\left(\varphi''(t)-\chi\int_\Omega \Delta u_t\omega dx-\left(a_0+a_1\left\|\nabla u\right\|_{2,\Omega}^{2r}\right)\int_\Omega \Delta u\omega dx+\right. \\
&\left.+\int_\Omega |u_t|^{q-2}u_t\omega dx-\int_\Omega b(x,t)|u|^{p-2}u\omega dx\right).
\end{aligned}
$$

By virtue of (10), we obtain that

$$
\begin{aligned}
&F(t,u)=\tfrac{1}{h_1(t)}\left(\varphi''(t)+\chi\int_\Omega \nabla u_t\nabla\omega dx+\left(a_0+a_1\left\|\nabla u\right\|_{2,\Omega}^{2r}\right)\int_\Omega \nabla u\nabla\omega dx-\int_\Omega b(x,t)|u|^{p-2}u\omega dx\right)= \\
&=\tfrac{1}{h_1(t)}\left(\varphi''(t)-\chi\int_\Omega \Delta u_t\omega dx-\left(a_0+a_1\left\|\nabla u\right\|_{2,\Omega}^{2r}\right)\int_\Omega \Delta u\omega dx-\int_\Omega b(x,t)|u|^{p-2}u\omega dx\right) = \\
&=\tfrac{1}{h_1(t)}\left(\varphi''(t)-\int_\Omega u_{tt}\omega dx+\int_\Omega b(x,t)|u|^{p-2}u\omega dx+\int_\Omega f(t)h(x,t)\omega dx-\int_\Omega b(x,t)|u|^{p-2}u\omega dx\right).
\end{aligned}
$$

$$
\varphi''(t)-\int_\Omega u_{tt}\omega dx=0.
$$

In this way, $\frac{d^2}{dt^2}\left(\varphi(t)-\int_\Omega u\omega dx\right)=0$. Denote by $v(t)=\varphi(t)-\int_\Omega u\omega dx$. Then the function $v(t)$ can be found as a solution of the Cauchy problem: $v''(t)=0,\ v(0)=0,\ v'(0)=0$. ($v(0)=0,\ v'(0)=0$ follows from the matching condition (5)). The unique solution of the problem is the function $v(t)=0$, consequently, $\int_\Omega u(x,t)\omega(x)dx=\varphi(t)$.

3 Existence of the solution. Galerkin approximations

Theorem 1. *Let the conditions* (5), (6) *and* $2 \le p < \frac{2n-2}{n-2}$, $n \ge 3$, $q \ge 2$, $r > 1$ *are performed. Then there exists the generalized solution* Δu, Δu_t, $u_{tt} \in L_2(Q_T)$ *of the problem* (7)-(8).

Proof. Let us choose in $\overset{0}{W}{}_2^1(\Omega)$ some system of functions $\{\Psi_j(x)\}$ forming a basis in the given space. As a basis, we can take the eigenfunctions of the Sturm-Liouville problem

$$\Delta\Psi + \lambda\Psi = 0, \quad \Psi|_{\partial\Omega} = 0.$$

We will look for an approximate solution of the problem (7)-(8) in the form

$$u_m(x,t) = \sum_{k=1}^{m} C_{mk}(t)\Psi_k(x) \tag{12}$$

where coefficients $C_{mk}(t)$ are searched out from the relations

$$\begin{gathered}\sum_{k=1}^{m} C''_{mk}(t)\int_\Omega \Psi_k\Psi_j dx + \chi\sum_{k=1}^{m} C'_{mk}(t)\int_\Omega \nabla\Psi_k\nabla\Psi_j dx + \left(a_0 + a_1\|\nabla u_m\|_{2,\Omega}^{2r}\right)\int_\Omega \nabla u_m\nabla\Psi_j dx + \\ + \sum_{k=1}^{m} C'_{mk}(t)\int_\Omega |\partial_t u_m|^{p-2}\Psi_k\Psi_j dx - \int_\Omega b(x,t)\,|u_m|^{p-2}u_m\Psi_j dx = \int_\Omega F(t,u_m)\Psi_j dx.\end{gathered} \tag{13}$$

$$\begin{aligned} u_{m0} = u_m(0) = \sum_{k=1}^{m} C_{mk}(0)\Psi_k = \sum_{k=1}^{m}\alpha_{0k}\Psi_k, \\ u_{m1} = u'_m(0) = \sum_{k=1}^{m} C'_{mk}(0)\Psi_k = \sum_{k=1}^{m}\alpha_{1k}\Psi_k \end{aligned} \tag{14}$$

and besides

$$\begin{aligned} &u_{m0} \to u_0 \text{ strongly in } \overset{0}{W}{}_2^2(\Omega) \text{ at } m\to\infty \\ &u_{m1} \to u_1 \text{ strongly in } \overset{0}{W}{}_2^1(\Omega) \text{ at } m\to\infty \end{aligned} \tag{15}$$

Let us introduce denotations

$$\vec{C}_m \equiv \{C_{1m}(t),...,C_{mm}(t)\}^T, \vec{\alpha} \equiv \{\alpha_1,...,\alpha_m\}^T, a_{kj} = \int_\Omega \Psi_k\Psi_j dx, b_{kj} = \chi\int_\Omega (\nabla\Psi_k,\nabla\Psi_j)\,dx,$$

$$\begin{gathered} f_{kj} = \chi\int_\Omega (\nabla\Psi_k,\nabla\Psi_j)\,dx + \left(a_0 + a_1\|\nabla u_m\|_{2,\Omega}^{2r}\right)\int_\Omega \nabla\Psi_k\nabla\Psi_j dx + \\ + \sum_{k=1}^{m} C'_{mk}(t)\int_\Omega |\partial_t u_m|^{p-2}\Psi_k\Psi_j dx - \int_\Omega b(x,t)\,|u_m|^{p-2}\Psi_k\Psi_j dx + \int_\Omega F(t,u_m)\Psi_j dx, \\ A_m\left(\vec{C}_m\right) \equiv \left\{a_{jk}\left(\vec{C}_m\right)\right\}, \vec{F}_m\left(\vec{C}_m,\vec{C}'_m\right) \equiv \left\{f_{jk}\left(\vec{C}_m,\vec{C}'_m\right)\right\}\vec{C}_m. \end{gathered}$$

Then the system of equations (13) and condition (14) take the matrix form

$$\begin{aligned} &A_m \vec{C}_m'' \equiv \vec{F}_m\left(\vec{C}_m,\, \vec{C}_m'\right), \\ &\vec{C}_m(0) = \vec{\alpha}_0,\ \vec{C}_m'(0) = \vec{\alpha}_1. \end{aligned} \tag{16}$$

Relations (16) represent the Cauchy problem for the system of ordinary differential equations, which is solvable on the segment $[0,\, T_m]$. In order to verify the existence of the solution on $[0,\, T]$, we obtain a priori estimates.

Lemma 2. If $u \in \overset{0}{W}{}_2^1(\Omega)$,$1 < \sigma \leq 2$, then the following inequality is performed

$$\int\limits_\Omega |u_m|^\sigma dx \leq \left(\int\limits_\Omega |u|^2 dx\right)^{\frac{\sigma}{2}} |\Omega|^{\frac{2-\sigma}{2}} \leq C_0 \left(\int\limits_\Omega |u|^2 dx + \chi \int\limits_\Omega |\nabla u|^2 dx\right).$$

Lemma 3. If $u \in \overset{0}{W}{}_2^1(\Omega), 2 < \beta < \frac{2N}{N-2},\ N \geq 3,$, then the following inequality is performed

$$\|u\|_{\beta,\Omega}^2 \leq C_0^2 \|\nabla u\|_{2,\Omega}^{2\alpha} \|u\|_{2,\Omega}^{2(1-\alpha)} \leq \chi \|\nabla u\|_{2,\Omega}^2 + \frac{(1-\alpha)\alpha^{\frac{\alpha}{1-\alpha}} C_0^{\frac{2}{1-\alpha}}}{\chi^{\frac{\alpha}{1-\alpha}}} \|u\|_{2,\Omega}^2,$$

where $C_0 = \left(\frac{2(N-1)}{N-2}\right)^\alpha,\ \ \alpha = \frac{(\beta-2)N}{2\beta},\ \ 0 < \alpha < 1.$

We multiply the equality (13) by $C_{mj}'(t)$ and summarize over $j = \overline{1, m}$. As a result, we take

$$\begin{aligned} &\tfrac{1}{2}\tfrac{d}{dt}\int\limits_\Omega |\partial_t u_m(t)|^2 dx + \chi \int\limits_\Omega |\partial_t \nabla u_m|^2 dx + \tfrac{a_0}{2}\tfrac{d}{dt}\|\nabla u_m\|_{2,\Omega}^2 + \\ &+ \tfrac{a_1}{2r+2}\tfrac{d}{dt}\|\nabla u_m\|_{2,\Omega}^{2r+2} + \int\limits_\Omega |\partial_t u_m(t)|^q dx = \\ &= \int\limits_\Omega b(x,t)|u_m|^{p-2} u_m \partial_t u_m dx + \int\limits_\Omega F(t,u_m) h \partial_t u_m dx. \end{aligned} \tag{17}$$

We integrate with respect to τ from 0 to t, then we get the relation

$$\begin{aligned} &\tfrac{1}{2}\int\limits_\Omega |\partial_t u_m(t)|^2 dx + \chi \int\limits_0^t \int\limits_\Omega |\partial_\tau \nabla u_m|^2 dx d\tau + \tfrac{a_0}{2}\|\nabla u_m\|_{2,\Omega}^2 + \\ &+ \tfrac{a_1}{2r+2}\|\nabla u_m\|_{2,\Omega}^{2r+2} + \int\limits_0^t \int\limits_\Omega |\partial_\tau u_m|^q dx d\tau = \\ &= \tfrac{1}{2}\int\limits_\Omega |\partial_t u_m(x,0)|^2 dx + \tfrac{a_0}{2}\|\nabla u_m(x,0)\|_{2,\Omega}^2 + \tfrac{a_1}{2r+2}\|\nabla u_m(x,0)\|_{2,\Omega}^{2r+2} + \\ &+ \int\limits_0^t \int\limits_\Omega b(x,\tau)|u_m|^{p-2} u_m \partial_\tau u_m dx d\tau + \int\limits_0^t \int\limits_\Omega F(t,u_m) h \partial_\tau u_m dx d\tau. \end{aligned} \tag{18}$$

Denote by

$$y(t)=\frac{1}{2}\int\limits_{\Omega}|\partial_t u_m(t)|^2dx+\frac{a_0}{2}\left\|\nabla u_m\right\|_{2,\Omega}^2+\frac{a_1}{2r+2}\left\|\nabla u_m\right\|_{2,\Omega}^{2r+2}.$$

Estimating the right-hand side of (18) using Lemma 2 and 3, as well as the Hölder and Young inequality, we obtain

$$\begin{aligned}&\left|\int\limits_0^t\int\limits_\Omega b(x,\tau)|u_m|^{p-2}u_m\partial_\tau u_m dxd\tau\right|\le b_0\int\limits_0^t\int\limits_\Omega|u_m|^{p-1}\partial_\tau u_m dxd\tau\le\\&\le\left\|\partial_\tau u_m\right\|_{2,Q_t}\left(\int\limits_0^t\int\limits_\Omega|u_m|^{\frac{2n}{n-2}}dxd\tau\right)^{\frac{n-2}{2n}}\left(\int\limits_0^t\int\limits_\Omega|u_m|^{(p-2)n}dxd\tau\right)^{\frac{1}{n}}\le\\&\le\left\|\partial_\tau u_m\right\|_{2,Q_t}^2+C_1\left\|\nabla u_m\right\|_{2,Q_t}^{\frac{2n}{n-2}}.\end{aligned}\tag{19}$$

$$\begin{aligned}&\left|\chi\int\limits_0^t\int\limits_\Omega\tfrac{1}{h_1(\tau)}\int\limits_\Omega\partial_\tau\nabla u_m\nabla\omega dx\,h\partial_\tau u_m dxd\tau\right|\le\\&\le\chi\int\limits_0^t\tfrac{1}{h_1(\tau)}\left\|\partial_\tau\nabla u_m\right\|_{2,\Omega}\left\|\nabla\omega\right\|_{2,\Omega}\left\|h\right\|_{2,\Omega}\left\|\partial_\tau u_m\right\|_{2,\Omega}d\tau\le\\&\le\tfrac{\chi}{2}\int\limits_0^t\left\|\partial_\tau\nabla u_m\right\|_{2,\Omega}^2d\tau+\tfrac{\chi}{2}\left\|\nabla\omega\right\|_{2,\Omega}^2\sup_{0\le t\le T}\tfrac{\|h(x,t)\|_{2,\Omega}^2}{|h_1(t)|^2}\int\limits_0^t\left\|\partial_\tau u_m\right\|_{2,\Omega}^2d\tau.\end{aligned}$$

$$\begin{aligned}&\left|\int\limits_0^t\int\limits_\Omega\tfrac{1}{h_1(\tau)}\left(a_0+a_1\left\|\nabla u_m\right\|_{2,\Omega}^{2r}\right)\int_\Omega\nabla u_m\nabla\omega dx\,h\partial_\tau u_m dxd\tau\right|\le\\&\le\int\limits_0^t\tfrac{1}{h_1(\tau)}\left(a_0+a_1\left\|\nabla u_m\right\|_{2,\Omega}^{2r}\right)\left\|\nabla u_m\right\|_{2,\Omega}\left\|\nabla\omega\right\|_{2,\Omega}\left\|h\right\|_{2,\Omega}\left\|\partial_\tau u_m\right\|_{2,\Omega}d\tau\le\\&\le a_0\int\limits_0^t\left\|\nabla u_m\right\|_{2,\Omega}^2d\tau+a_0\left\|\nabla\omega\right\|_{2,\Omega}^2\int\limits_0^t\tfrac{1}{h_1^2(\tau)}\left\|h\right\|_{2,\Omega}^2\left\|\partial_\tau u_m\right\|_{2,\Omega}^2d\tau+\\&+a_1\int\limits_0^t\left\|\nabla u_m\right\|_{2,\Omega}^{2r+2}d\tau+C_2\left\|\nabla\omega\right\|_{2,\Omega}^{2r+2}\sup_{0\le t\le T}\tfrac{\|h(x,t)\|_{2,\Omega}^{2r+2}}{|h_1(t)|^{2r+2}}\int\limits_0^t\left\|\partial_\tau u_m\right\|_{2,\Omega}^{2r+2}d\tau,\end{aligned}$$

$$C_2=\frac{a_1(2r+1)^{2r+1}}{(2r+2)^{2r+2}}.$$

$$\begin{aligned}&\left|\int\limits_0^t\int\limits_\Omega\tfrac{1}{h_1(\tau)}\int\limits_\Omega|\partial_\tau u_m|^{q-2}\partial_\tau u_m\omega dx\,h\partial_\tau u_m dxd\tau\right|\le\\&\le\int\limits_0^t\tfrac{1}{h_1(\tau)}\left\|\partial_\tau u_m\right\|_{q,\Omega}^{q-1}\left\|\omega\right\|_{q,\Omega}\left\|h\right\|_{2,\Omega}\left\|\partial_\tau u_m\right\|_{2,\Omega}d\tau\le\\&\le\tfrac{1}{2}\int\limits_0^t\left\|\partial_\tau u_m\right\|_{q,\Omega}^q d\tau+C_3\left\|\omega\right\|_{q,\Omega}^q\sup_{0\le t\le T}\tfrac{\|h(x,t)\|_{2,\Omega}^q}{|h_1(t)|^q}\int\limits_0^t\left\|\partial_\tau u_m\right\|_{2,\Omega}^q d\tau,\end{aligned}$$

$$C_3=\frac{q-1}{q\left(\frac{q}{2}\right)^{\frac{1}{q-1}}}.$$

$$\left|\int_0^t\int_\Omega \frac{1}{h_1(\tau)}\int_\Omega b(x,t)|u_m|^{p-2}u_m\omega dx\, h\partial_\tau u_m dx d\tau\right| \leq$$
$$\leq b_0\left\|\omega\right\|_{p,\Omega}\sup_{0\leq t\leq T}\frac{\|h(x,t)\|_{2,\Omega}}{h_1(t)}\int_0^t\left\|\partial_\tau u_m\right\|_{2,\Omega}\left\|u_m\right\|_{p,\Omega}^{p-1}d\tau \leq$$
$$\leq b_0\left\|\omega\right\|_{p,\Omega}\sup_{0\leq t\leq T}\frac{\|h(x,t)\|_{2,\Omega}}{h_1(t)}\left(\int_0^t\left\|u_m\right\|_{p,\Omega}^{p}d\tau+\int_0^t\left\|\partial_\tau u_m\right\|_{2,\Omega}^{p}d\tau\right) \leq$$
$$\leq b_0\left\|\omega\right\|_{p,\Omega}\sup_{0\leq t\leq T}\frac{\|h(x,t)\|_{2,\Omega}}{h_1(t)}\int_0^t\left(\left\|u_m\right\|_{2,\Omega}^{2}+\chi\left\|\nabla u_m\right\|_{2,\Omega}^{2}\right)^{\frac{p}{2}}d\tau+$$
$$+b_0\left\|\omega\right\|_{p,\Omega}\sup_{0\leq t\leq T}\frac{\|h(x,t)\|_{2,\Omega}}{h_1(t)}\int_0^t\left\|\partial_\tau u_m\right\|_{2,\Omega}^{p}d\tau.$$

Denote by

$$y(t)=\frac{1}{2}\int_\Omega|\partial_t u_m(t)|^2dx+\frac{a_0}{4}C(\Omega)\left\|u_m\right\|_{2,\Omega}^2+\frac{a_0}{4}\left\|\nabla u_m\right\|_{2,\Omega}^2+\frac{a_1}{2r+2}\left\|\nabla u_m\right\|_{2,\Omega}^{2r+2}.$$

$$d=\max\left\{\frac{n}{n-2},\ \frac{p}{2},\ \frac{q}{2},\ r+1\right\}.$$

Then from the relation (18), we get

$$y(t)\leq C_4+C_5\int_0^t[y(\tau)]^d d\tau.$$

Applying for this the generalized Bihari lemma, then the next inequality is true

$$y(t)\leq\frac{C_4}{\left[1-(d-1)C_5C_4^{d-1}t\right]^{\frac{1}{d-1}}},$$

i.Pμ. $\frac{1}{2}\int_\Omega|\partial_t u_m(t)|^2dx+\frac{a_0}{4}C(\Omega)\left\|u_m\right\|_{2,\Omega}^2+\frac{a_0}{4}\left\|\nabla u_m\right\|_{2,\Omega}^2+\frac{a_1}{2r+2}\left\|\nabla u_m\right\|_{2,\Omega}^{2r+2}\leq\frac{!_4}{\left[1-(d-1)C_5C_4^{d-1}t\right]^{\frac{1}{d-1}}}$.

From this estimate we can conclude that there exists $T_0>0$ such that

$$\begin{aligned}&\int_\Omega|\partial_t u_m(t)|^2dx+\left\|u_m\right\|_{2,\Omega}^2+\left\|\nabla u_m\right\|_{2,\Omega}^2+\\&+\left\|\nabla u_m\right\|_{2,\Omega}^{2r+2}+\int_0^T\int_\Omega|\partial_\tau\nabla u_m|^2dxd\tau+\int_0^T\int_\Omega|\partial_\tau u_m|^q dxd\tau\leq C_6,\end{aligned}\tag{20}$$

for all $t\in[0,T],\ T<T_0$, where C_6 is constant which does not depend on $m\in N$.

We multiply the relation (13) by $\lambda_j C_{mj}(t)$ and $C''_{mj}(t)$, then summarize over $j=\overline{1,m}$. As a result, we get the next relations

$$-\frac{d}{dt}\int_\Omega\partial_t u_m(t)\Delta u_m(t)dx-\left\|\partial_t\nabla u_m\right\|_{2,\Omega}^2+\frac{\chi}{2}\frac{d}{dt}\left\|\Delta u_m\right\|_{2,\Omega}^2+$$
$$+\left(a_0+a_1\left\|\nabla u_m\right\|_{2,\Omega}^{2r}\right)\left\|\Delta u_m\right\|_{2,\Omega}^2-\int_\Omega|\partial_t u_m(t)|^{q-2}\partial_t u_m(t)\Delta u_m dx=$$
$$=-\int_\Omega b(x,t)|u_m|^{p-2}u_m\Delta u_m dx-\int_\Omega F(t,u_m)h\Delta u_m dx.$$

$$\int\limits_\Omega |\partial_t^2 u_m(t)|^2 dx + \frac{\chi}{2}\frac{d}{dt}\int\limits_\Omega |\partial_t \nabla u_m|^2 dx - \left(a_0 + a_1 \|\nabla u_m\|_{2,\Omega}^{2r}\right)\int\limits_\Omega \Delta u_m \partial_t^2 u_m dx +$$
$$+\frac{1}{q}\frac{d}{dt}\int\limits_\Omega |\partial_t u_m|^q dx = \int\limits_\Omega b(x,t)\,|u_m|^{p-2}u_m \partial_t^2 u_m dx + \int\limits_\Omega F(t,u_m)\partial_t^2 u_m dx.$$

By integrating these relations from 0 to t, we get

$$\frac{\chi}{2}\|\Delta u_m\|_{2,\Omega}^2 + \int\limits_0^t \left(a_0 + a_1\|\nabla u_m\|_{2,\Omega}^{2r}\right)\|\Delta u_m\|_{2,\Omega}^2 d\tau = \frac{\chi}{2}\|\Delta u_m(0)\|_{2,\Omega}^2 + \int\limits_\Omega \partial_t u_m(t)\Delta u_m(t)dx -$$
$$-\int\limits_\Omega \partial_t u_m(0)\Delta u_m(0)dx + \int\limits_0^t \|\partial_\tau \nabla u_m\|_{2,\Omega}^2 d\tau + \int\limits_0^t\int\limits_\Omega |\partial_\tau u_m(\tau)|^{q-2}\partial_\tau u_m(\tau)\Delta u_m dx d\tau -$$
$$-\int\limits_0^t\int\limits_\Omega b(x,\tau)|u_m|^{p-2}u_m\Delta u_m dx d\tau - \int\limits_0^t\int\limits_\Omega F(\tau,u_m)h\Delta u_m dx d\tau. \tag{21}$$

$$\frac{\chi}{2}\int\limits_\Omega |\partial_t \nabla u_m(t)|^2 dx + \frac{1}{q}\int\limits_\Omega |\partial_t u_m(t)|^q dx + \int\limits_0^t\int\limits_\Omega |\partial_\tau^2 u_m(x,\tau)|^2 dx d\tau =$$
$$= \int\limits_0^t \left(a_0 + a_1\|\nabla u_m\|_{2,\Omega}^{2r}\right)\int\limits_\Omega \Delta u_m \partial_\tau^2 u_m dx d\tau + \tag{22}$$
$$+\int_0^t\int\limits_\Omega b(x,\tau)\,|u_m|^{p-2}u_m\partial_\tau^2 u_m dx d\tau + \int\limits_0^t\int\limits_\Omega F(\tau,u_m)\partial_\tau^2 u_m dx d\tau.$$

Analogically, we estimate the right-hand side of (21) and (22), applying lemmas 2 and 3, Hölder and Young inequalities, Bihari's lemma and a priori estimate (20), as a result we obtain

$$\|\Delta u_m\|_{2,\Omega}^2 + \int\limits_0^T \left(a_0 + a_1\|\nabla u_m\|_{2,\Omega}^{2r}\right)\|\Delta u_m\|_{2,\Omega}^2 dt \le C_7, \text{ for all } t\in[0,T],\ T<T_0, \tag{23}$$

$$\int\limits_\Omega |\partial_t \nabla u_m(t)|^2 dx + \int\limits_\Omega |\partial_t u_m(t)|^q dx + \int\limits_0^T\int\limits_\Omega |\partial_\tau^2 u_m(x,\tau)|^2 dx dt \le C_8, \text{ for all } t\in[0,T],\ T<T_0, \tag{24}$$

where C_7 and C_8 are constants which does not depend on $m \in N$.

From the obtained estimates (20), (23) and (3) follows the estimate

$$\int\limits_0^T \|\partial_t \Delta u_m\|_{2,\Omega}^2 dt \le C_9, \text{ for all } t\in[0,T],\ T<T_0,\ m\in N. \tag{25}$$

Then by using (20), (23), (3) and (25), considering the conditions of the theorem, we can show the existence of the derivative $u_{xx} \in L_2(Q_T)$. In this way, $\Delta u,\ \Delta u_t,\ u_{tt} \in L_2(Q_T)$.

4 Uniqueness of the generalized solution

Theorem 2. *Let the conditions* (5), $r>2,\ q>2,\ 2<p\le 2+\frac{1}{N-2},\quad N\ge 3$, *are performed. Then the generalized solution of the problem* (1)-(3) *on the segment* $(0,T)$ *is unique.*

Proof. Assume that the problem (7)-(8) has two generalized solutions: $u_1(x,t)$ and $u_2(x,t)$. Let us put $u(x,t)=u_1(x,t)-u_2(x,t)$. Then there are the following equalities

$$\begin{aligned}&u_{tt}-\chi\Delta u_t-a_0\Delta u-a_1\left(\|\nabla u_1\|_{2,\Omega}^{2r}\Delta u_1-\|\nabla u_2\|_{2,\Omega}^{2r}\Delta u_2\right)+\\&+|u_{1t}|^{q-2}u_{1t}-|u_{2t}|^{q-2}u_{2t}=b(x,t)\left(|u_1|^{p-2}u_1-|u_2|^{p-2}u_2\right)+\\&+h(x,t)\left(F(t,u_1)-F(t,u_2)\right),\quad x\in\Omega,\quad t>0,\end{aligned}\tag{26}$$

$$u(x,0)=0,\ u_t(x,0)=0,\ x\in\Omega,\quad u\big|_S=0.\tag{27}$$

We consider the equality

$$\begin{aligned}&\int_0^t\int_\Omega\left[u_{\tau\tau}-\chi\Delta u_\tau-a_0\Delta u-a_1\left(\|\nabla u_1\|_{2,\Omega}^{2r}\Delta u_1-\|\nabla u_2\|_{2,\Omega}^{2r}\Delta u_2\right)+\right.\\&\left.+|u_{1\tau}|^{q-2}u_{1\tau}-|u_{2\tau}|^{q-2}u_{2\tau}\right]u_\tau dxd\tau=\int_0^t\int_\Omega\left[b(x,\tau)\left(|u_1|^{p-2}u_1-|u_2|^{p-2}u_2\right)+\right.\\&\left.+h(x,\tau)\left(F(\tau,u_1)-F(\tau,u_2)\right)\right]u_\tau dxd\tau.\end{aligned}$$

By applying the next inequalities

$$\left||u_1|^q u_1-|u_2|^q u_2\right|\le(q+1)\left(|u_1|^q+|u_2|^q\right)|u_1-u_2|\ \text{at}\ q>0,$$

$$\left|\left(|u_1|^q u_1-|u_2|^q u_2\right)(u_1-u_2)\right|\ge|u_1-u_2|^{q+2}\ \text{at}\ q>0.$$

As a result, we obtain the inequality

$$\begin{aligned}&\tfrac12\int_\Omega u_t^2(t)dx+\chi\int_0^t\int_\Omega|\nabla u_\tau|^2dxd\tau+\tfrac{a_0}{2}\int_\Omega|\nabla u|^2dx+\int_0^t\int_\Omega|u_\tau|^q dxd\tau\le\\&\le-a_1\int_0^t\left(\|\nabla u_1\|_{2,\Omega}^{2r}\int_\Omega\nabla u\nabla u_\tau dx-\left(\|\nabla u_1\|_{2,\Omega}^{2r}-\|\nabla u_2\|_{2,\Omega}^{2r}\right)\int_\Omega\nabla u_2\nabla u_\tau dx\right)d\tau+\\&+\int_0^t\int_\Omega b(x,\tau)\left(|u_1|^{p-2}+|u_2|^{p-2}\right)u_\tau dxd\tau+\int_0^t\int_\Omega h(x,\tau)\left(F(\tau,u_1)-F(\tau,u_2)\right)u_\tau dxd\tau.\end{aligned}\tag{28}$$

We estimate the right-hand side of the inequality (28), applying the Hölder's inequality

$$\begin{aligned}&\left|\int_0^t\int_\Omega b(x,\tau)\left(|u_1|^{p-2}u_1-|u_2|^{p-2}u_2\right)u_\tau dxd\tau\right|\le b_1(p-1)\int_0^t\int_\Omega\left(|u_1|^{p-2}+|u_2|^{p-2}\right)uu_\tau dxd\tau\le\\&\le b_1(p-1)\left(\left(\int_0^t\int_\Omega|u_1|^{\frac{2r(p-2)}{r-2}}dxd\tau\right)^{\frac{r-2}{2r}}+\left(\int_0^t\int_\Omega|u_2|^{\frac{2r(p-2)}{r-2}}dxd\tau\right)^{\frac{r-2}{2r}}\right)\times\\&\times\left(\int_0^t\int_\Omega u^r dxd\tau\right)^{\frac1r}\left(\int_0^t\int_\Omega u_\tau^2dxd\tau\right)^{\frac12}.\end{aligned}$$

Let us put $r = \frac{2N}{N-2}$, $p \leq 2 + \frac{1}{N-2}$, $N \geq 3$. Then by the Sobolev embedding theorem $H^1(\Omega) \to\to L_r(\Omega)$ and $H^1(\Omega) \to\to L_{2r(p-2)/(r-2)}(\Omega)$. In this case, taking into account the smoothness class of solutions $u_1(x,t)$ and $u_2(x,t)$, we come to the estimate

$$\left|\int_0^t \int_\Omega b(x,\tau)\left(|u_1|^{p-2}u_1 - |u_2|^{p-2}u_2\right)u_\tau dx d\tau\right| \leq C_1 \int_0^t \left(\|u_\tau\|_{2,\Omega}^2 + \|\nabla u\|_{2,\Omega}^2 + \|u\|_{2,\Omega}^2\right) d\tau. \quad (29)$$

Let us estimate the first term

$$\left|a_1 \int_0^t \left(\|\nabla u_1\|_{2,\Omega}^{2r} \int_\Omega \nabla u \nabla u_\tau dx - \left(\|\nabla u_1\|_{2,\Omega}^{2r} - \|\nabla u_2\|_{2,\Omega}^{2r}\right)\int_\Omega \nabla u_2 \nabla u_\tau dx\right) d\tau\right| \leq$$
$$\leq a_1 \int_0^t \|\nabla u_1\|_{2,\Omega}^{2r} \|\nabla u\|_{2,\Omega} \|\nabla u_\tau\|_{2,\Omega} d\tau + a_1 \int_0^t \left(\|\nabla u_1\|_{2,\Omega}^{2r-2} + \|\nabla u_2\|_{2,\Omega}^{2r-2}\right)\|\nabla u_2\|_{2,\Omega}\|\nabla u_\tau\|_{2,\Omega} \times$$
$$\times \int_\Omega \left(|\nabla u_1|^2 - |\nabla u_2|^2\right) dx d\tau \leq a_1 \int_0^t \|\nabla u_1\|_{2,\Omega}^{2r} \|\nabla u\|_{2,\Omega} \|\nabla u_\tau\|_{2,\Omega} d\tau +$$
$$+ a_1 C_2' \int_0^t \||\nabla u_1| + |\nabla u_2|\|_{2,\Omega} \|\nabla u\|_{2,\Omega} \|\nabla u_\tau\|_{2,\Omega} \leq$$
$$\leq \frac{\chi}{4}\int_0^t \|\nabla u_\tau\|_{2,\Omega}^2 d\tau + C_2 \int_0^t \left(\|u_\tau\|_{2,\Omega}^2 + \|\nabla u\|_{2,\Omega}^2 + \|u\|_{2,\Omega}^2\right) d\tau.$$

The third term is estimated in a similar way. From the obtained estimates, we get

$$\int_\Omega u_t^2(t)dx + C_0 \int_\Omega |u|^2 dx + a_0 \int_\Omega |\nabla u|^2 dx + \chi \int_0^t \int_\Omega |\nabla u_\tau|^2 dx d\tau + \int_0^t \int_\Omega |u_\tau|^q dx d\tau \leq$$
$$\leq C_4 \int_0^t \left(\|u_\tau\|_{2,\Omega}^2 + a_0\|\nabla u\|_{2,\Omega}^2 + C_0\|u\|_{2,\Omega}^2\right) d\tau + C_5 \int_0^t \left(\|u_\tau\|_{2,\Omega}^2 + a_0\|\nabla u\|_{2,\Omega}^2 + C_0\|u\|_{2,\Omega}^2\right)^d d\tau,$$

where $d > 1$.

From the last inequality follows that

$$\int_\Omega u_t^2(t)dx + C_0 \int_\Omega |u|^2 dx + a_0 \int_\Omega |\nabla u|^2 dx \leq C_4 \int_0^t \left(\|u_\tau\|_{2,\Omega}^2 + a_0\|\nabla u\|_{2,\Omega}^2 + C_0\|u\|_{2,\Omega}^2\right) d\tau +$$
$$+ C_5 \int_0^t \left(\|u_\tau\|_{2,\Omega}^2 + a_0\|\nabla u\|_{2,\Omega}^2 + C_0\|u\|_{2,\Omega}^2\right)^d d\tau,$$

where by Bihari's lemma, implies $\int_\Omega u_t^2(t)dx + C_0 \int_\Omega |u|^2 dx + a_0 \int_\Omega |\nabla u|^2 dx = 0$ almost everywhere on the time interval $(0,T)$, which means that the generalized solution is unique.

5 Conclusion

In the paper, we investigated the solvability of the inverse problem of determining the solution of the pseudohyperbolic equation, also an unknown coefficient of a special form which identifies the external source. The methods used are based on the transition from the original problem to the equivalent problem for the loaded nonlinear pseudohyperbolic equation. For this problem we use Galerkin's method to prove the existence of a strong generalized solution. The obtained results on the solvability of the inverse problem are new and can be useful to study another problems in the given area.

6 Acknowledgment

This research has been funded by the Science Committee of the Ministry of Education and Science of the Republic of Kazakhstan, Grant No. AP08052425.

References

[1] Greiner W., Relativistic Quantum Mechanics - Wave Equations, 3rd ed. Berlin: Springer; 2000.

[2] Dodd R. K. Solitons and nonlinear wave equations / R. K. Dodd, I. C. Eilbeck, J. D. Gibbon,H. C. Morris. L.: Acad. Press, 1982.

[3] Strauss W, Vazquez L. Numerical solution of a nonlinear Klein-Gordon equation.JComput Phys. 1978;28(2):271-278.

[4] JimГ©nez S, Vazquez L. Analysis of four numerical schemes for a nonlinear Klein-Gordon equation.Appl Math Comput. 1990;35(1):61-94.

[5] Wong YS, Qianshun C, Lianger G. An initial-boundary value problem of a nonlinear Klein-Gordon equation.Appl Math Comput.1997;84(1):77-93.

[6] Wazwaz A. M. Compactons, solitons and periodic solutions for some forms of nonlinear Klein-Gordon equations // Chaos Solitons Fractals. 2006. V. 28. P. 1005вЂ"1013.

[7] Sassaman R., Biswas A. Soliton perturbation theory four phi-four model and nonlinear Klein-Gordon equations // Comm. Nonlinear Sci. Numer. Simulat. 2009. V. 14. P. 3239вЂ"3249.

[8] Sassaman R., Biswas A. Topological and non-topological solitons of the generalized Klein-Gordon equations // Appl. Math. Comput. 2009. V. 215. P. 212вЂ"220.

[9] Bratsos A. G., Petrakis L. A. A modified predictor-corrector scheme for the Klein-Gordonequation // Intern. J. Comput. Math. 2010. V. 87, N 8. P. 1892вЂ"1904.

[10] Böhme C, Michael R. A scale-invariant Klein-Gordon model with time-dependent potential.Ann Univ Ferrara. 2012;58(2):229-250.

[11] BГ¶hme C, Michael R. Energy bounds for Klein-Gordon equations with time-dependent potential.Ann Univ Ferrara. 2013;59(1):31-55.

[12] Hao Cheng, Xiyu Mu, GreenвЂ™s function for the boundary value problem of the static Klein-Gordon equation stated on a rectangular region and its convergence analysis. Boundary Value Problems (2017) 2017:72.

[13] Tekin I, Mehraliyev YT, Ismailov MI. Existence and uniqueness of an inverse problemfor nonlinear Klein-Gordon equation.Math Meth Appl Sci. 2019;1-15.https://doi.org/10.1002/mma.5609

[14] Sobolev, S. L.; On certain new problem of mathematical physics. Izvestia Acad. Nauk.Mathem. 1954, no. 18, p. 3-50.

[15] Kozhanov, A.I., Comparison Theorems and Solvability of Boundary Value Problems for Some Classes of Evolution Equations Like Pseudoparabolic and Pseudohyperbolic Ones, Preprint Inst. Mat. Acad. Sci., Novosibirsk, 1990, no. 17

[16] Demidenko, G., The Cauchy problem for pseudohyperbolic equations, Selcuk J Appl Math, Vol. 1,No.1, pp.47-62, 2000.

[17] Sveshnikov, A. G.; Alshin, A. B.; Korpusov, O. M.; Pletner, Yu. D.; Linear and nonlinearSobolev type equations. Moscow.,Phizmatlit. 2007.

[18] Korpusov, O. M., Blow-up in nonclassical wave equations. Moscow, URSS. 2010.

[19] Pulkina, L.S. Solution to nonlocal problems of pseudohyperbolic equations // Electronic Journal of Differential Equations, Vol. 2012 (2012), No. 116, pp. 1вЂ"9.

[20] Yamamoto, M., Zhang, X.: Global uniqueness and stability for a class of multidimensional inverse hyperbolic problems with two unknowns. Appl. Math. Optim. 48, 211вЂ"228 (2003)

[21] Oussaeif,T.E., Bouziani, A.: Inverse problem of hyperbolic equationwith an integral overdetermination condition. Electron J. Diff. Equ. 2016(138), 1вЂ"7 (2016)

[22] Shahrouzi, M.: On behavior of solutions to a class of nonlinear hyperbolic inverse source problem. Acta. Math. Sin-English Ser. 32(6), 683вЂ“698 (2016)

[23] Shahrouzi, M.: Blow up of solutions to a class of damped viscoelastic inverse source problem. Differ. Equ. Dyn. Syst. 28, 889вЂ“899 (2020)

[24] Shahrouzi, M.: General decay and blow up of solutions for a class of inverse problem with elasticity term and variable-exponent nonlinearities. Math Meth Appl Sci, (2021), Early view, https://doi.org/10.1002/mma.7891

ISSN 1563–0277, eISSN 2617-4871 JMMCS. №3(115) 2022 https://bm.kaznu.kz

IRSTI 27.29.25 DOI: https://doi.org/10.26577/JMMCS.2022.v115.i3.02

E.A. Bakirova[1,2*], **Zh.M. Kadirbayeva**[1,3], **A.N Nessipbayeva**[2]
[1]Institute of Mathematics and Mathematical Modeling, Kazakhstan, Almaty
[2]Kazakh National Women's Teacher Training University, Kazakhstan, Almaty
[3]International Information Technology University, Kazakhstan, Almaty
*e-mail: bakirova1974@mail.ru

MODIFICATION OF THE PARAMETRIZATION METHOD FOR SOLVING A BOUNDARY VALUE PROBLEM FOR LOADED DEPCAG

The functional differential equation plays important role in mathematical modeling of biological problems. In the present research work, we investigate a boundary value problem (BVP) for a functional differential equation. This equation includes loaded terms and a term with generalized piecewise constant argument. We apply a modified version of the Dzhumabaev parameterization method. The method's goal is to lead the original problem into an equivalent multi-point BVP for ordinary differential equations with parameters, which is composed of a problem with initial and additional conditions. The multi-point BVP is leaded to a system of linear algebraic equations in parameters, which are introduced as the values of the desired solution at the dividing points. The found parameters are plugged into auxiliary Cauchy problems on the partition subintervals, whose solutions are the restrictions of the solution to the original problem. The obtained results are verified by a numerical example. Numerical analysis showed high efficiency of the constructed modified version of the Dzhumabaev parameterization method.
Key words: load, piecewise-constant argument, two-point boundary value problem, parametrization method, numerical solution.

Э.А. Бакирова[1,2*], Ж.М. Қадырбаева[1,3], А.Н. Несiпбаева[2]
[1]Математика және математикалық моделдеу институты, Қазақстан, Алматы қ.
[2]Қазақ ұлттық қыздар педагогикалық университетi , Қазақстан, Алматы қ.
[3]Халықаралық ақпараттық технологиялар университетi, Қазақстан, Алматы қ.
*e-mail: bakirova1974@mail.ru

Жалпыланған түрдегi бөлiктi-тұрақты аргументi бар жүктелген дифференциалдық теңдеу үшiн шеттiк есептi шешудiң параметрлеу әдiсiнiң модификациясы

Функционалдық-дифференциалдық теңдеу биологиялық есептердi математикалық модельдеуде маңызды рөл атқарады. Осы жұмыста функционалдық-дифференциалдық теңдеу үшiн шеттiк есеп (ШЕ) қарастырылады. Бұл теңдеу жүктелген мүшелер мен жалпыланған түрдегi бөлiктi-тұрақты аргументi бар қосылғыштан тұрады. Жұмабаевтың параметрлеу әдiсiнiң модификацияланған нұсқасы қолданылады. Әдiстiң мақсаты - берiлген есептi бастапқы және қосымша шарттардан тұратын эквиваленттi параметрлерi бар жәй дифференциалдық теңдеулер жүйесi үшiн көп нүктелi ШЕ келтiрiлуi болып табылады. Көп нүктелi ШЕ бөлу нүктелерiнде iзделiндi шешiмнiң мәнi ретiнде енгiзiлетiн параметрлерi бар сызықтық алгебралық теңдеулер жүйесiне келтiрiледi. Табылған параметрлер бөлiктеудiң iшкi интервалдарындағы қосымша Коши есептерiне қойылады, олардың шешiмдерi бастапқы шеттiк есептiң шешiмдерiнiң сығылуы болып табылады. Алынған нәтижелер сандық мысалмен тексерiледi. Сандық талдау Жұмабаевтың параметрлеу әдiсiнiң құрастырылған модификациясының жоғары тиiмдiлiгiн көрсеттi.
Түйiн сөздер: жүктеу, бөлiктiтұрақты аргумент, екi нүктелi шеттiк есеп, параметрлеу әдiсi, сандық шешiм.

Э.А. Бакирова[1,2*], Ж.М. Кадирбаева[1,3], А.Н. Несипбаева[2]
[1]Институт математики и математического моделирования, Казахстан, г. Алматы
[2]Казахский национальный женский педагогический университет, Казахстан, г. Алматы
[3]Международный университет информационных технологий, Казахстан, г. Алматы
*e-mail: bakirova1974@mail.ru

Модификация метода параметризации решения краевой задачи для нагруженных дифференциальных уравнений с кусочно-постоянным аргументом обобщенного типа

Функционально-дифференциальное уравнение играет важную роль в математическом моделировании биологических задач. В настоящей работе исследуется краевая задача (КЗ) для функционально-дифференциального уравнения. В это уравнение входят нагруженные члены и член с обобщенным кусочно-постоянным аргументом. Применим модифицированный вариант метода параметризации Джумабаева. Цель метода - привести исходную задачу к эквивалентной многоточечной КЗ для обыкновенных дифференциальных уравнений с параметрами, состоящей из задачи с начальными и дополнительными условиями. Многоточечная КЗ приводится к системе линейных алгебраических уравнений с параметрами, которые вводятся как значения искомого решения в точках деления. Найденные параметры подставляются во вспомогательные задачи Коши на подинтервалах разбиения, решения которых являются сужениями решения исходной задачи. Полученные результаты проверяются на численном примере. Численный анализ показал высокую эффективность построенной модифицированной версии метода параметризации Джумабаева.
Ключевые слова: нагрузка, кусочно-постоянный аргумент, двухточечная краевая задача, метод параметризации, численное решение.

1 Introduction and preliminaries

The theory's creators, K. Cook, J. Wiener and S. Busenberg, suggested using differential equations with the piecewise constant argument for investigations in [1], [2]. Within the final four decades, numerous interesting results have been found, and applications have been realized in this theory. Numerous additional theoretical issues, such as existence and uniqueness of solutions, oscillations and stability, integral manifolds and periodic solutions, as well as many more, have been thoroughly discussed. Information about differential equations with piecewise constant argument of generalized type (DEPCAG) can be found in books [3], [4] and papers [5], [6].

This article's basic objective is to broaden the modification of Dzhumabaev parametrization method [7], [8] to the boundary value problem for the system of loaded DEPCAG. For this purpose, we have developed computational method solving a boundary-value problem for the system of loaded DEPCAG.

Loaded differential equations (LDE) were investigated in [9], [10] and the references therewith. Numerous problems for LDE and methods for solving problems for LDE are considered in [11]– [16].

We consider the following system of loaded DEPCAG

$$\frac{dx}{dt} = A_0(t)x + K(t)x(\gamma(t)) + \sum_{i=1}^{m+1} M_i(t)x(\theta_{i-1}) + f(t), \quad x \in R^n, \quad t \in (0,T), \tag{1}$$

subject to the two-point boundary condition

$$B_0x(0) + C_0x(T) = d, \quad d \in R^n, \tag{2}$$

where $A_0(t)$, $K(t)$, $M_i(t)$, $(i=\overline{1,m+1})$, are of dimensions $(n\times n)$ and are continuous on $[0,T]$, and the n-vector-function $f(t)$ are piecewise continuous on $[0,T]$ with possible discontinuities of the first kind at the points $t=\theta_j$, $(j=\overline{1,m})$; B_0 and C_0 are $(n\times n)$ constant matrices, $\|x\|=\max\limits_{i=\overline{1,n}}|x_i|$.

The argument $\gamma(t)$ is a step function defined as $\gamma(t)=\xi_{i-1}$ if $t\in[\theta_{i-1},\theta_i)$, $i=\overline{1,m+1}$; $\theta_{i-1}<\xi_{i-1}<\theta_i$ for all $i=\overline{1,m+1}$; where $0=\theta_0<\theta_1<...<\theta_m<\theta_{m+1}=T$.

A function $x(t)$ is called a solution to problem (1) and (2) if:

(i) the $x(t)$ is continuous on $[0,T]$;

(ii) the $x(t)$ is differentiable on $[0,T]$ with the possible exception of the θ_j, $j=\overline{0,m}$, where the one-sided derivatives exist;

(iii) the $x(t)$ satisfies (1) on each interval (θ_{i-1},θ_i), $i=\overline{1,m+1}$; at the θ_j, Eq. (1) is satisfied by the right-hand derivatives of $x(t)$;

(iv) the $x(t)$ satisfies the boundary condition (2).

2 Materials and methods

We employ the approach proposed in [17] to solve the boundary-value problem for the system of loaded DEPCAG (1) and (2). This approach is based on the algorithms of the modified version of the Dzhumabaev parameterization method and numerical methods for solving Cauchy problems.

By using loading points, the interval $[0,T]$ is split into subintervals: $[0,T)=\bigcup\limits_{s=1}^{m+1}[\theta_{s-1},\theta_s)$.

$C([0,T],\theta,\mathbb{R}^{n(m+1)})$ be the space of functions systems $x[t]=(x_1(t),x_2(t),\ldots,x_{m+1}(t))'$, where $x_s:[\theta_{s-1},\theta_s)\to\mathbb{R}^n$ are continuous and have finite left-hand side limits $\lim\limits_{t\to\theta_s-0}x_s(t)$, $s=\overline{1,m+1}$ with norm $\|x[\cdot]\|_2=\max\limits_{s=\overline{1,m+1}}\sup\limits_{t\in[\theta_{s-1},\theta_s)}|x_s(t)|$.

Denote by $x_r(t)$ a restriction of function $x(t)$ on r-th interval $[\theta_{r-1},\theta_r)$, i.e.

$$x_r(t)=x(t)\quad\text{for}\quad t\in[\theta_{r-1},\theta_r),\quad r=\overline{1,m+1}.$$

Then the function system $x[t]=(x_1(t),x_2(t),\ldots,x_{m+1}(t))\in C([0,T],\theta,R^{n(m+1)})$, and its elements $x_r(t)$, $r=\overline{1,m+1}$, satisfy the following boundary value problem for system of loaded DEPCAG

$$\frac{dx_r}{dt}=A_0(t)x_r+K(t)x_r(\xi_{r-1})+\sum_{i=1}^{m+1}M_i(t)x_i(\theta_{i-1})+f(t),\quad t\in[\theta_{r-1},\theta_r),\quad r=\overline{1,m+1},\tag{3}$$

$$B_0x_1(0)+C_0\lim_{t\to T-0}x_{m+1}(t)=d,\tag{4}$$

$$\lim_{t\to\theta_s-0}x_s(t)=x_{s+1}(\theta_s),\quad s=\overline{1,m}.\tag{5}$$

Introduce parameters $\lambda_r=x_r(\theta_{r-1})$ and $\mu_r=x_r(\xi_{r-1})$ for all $r=\overline{1,m+1}$. The following problem with parameters is obtained by substituting $v_r(t)=x_r(t)-\lambda_r$ on every r-th interval

$[\theta_{r-1},\theta_r)$:

$$\frac{dv_r}{dt} = A_0(t)(v_r+\lambda_r) + K(t)\mu_r + \sum_{i=1}^{m+1} M_i(t)\lambda_i + f(t), \quad t\in[\theta_{r-1},\theta_r), \tag{6}$$

$$v_r(\theta_{r-1}) = 0, \qquad r=\overline{1,m+1}, \tag{7}$$

$$B_0\lambda_1 + C_0\lambda_{m+1} + C_0\lim_{t\to T-0} v_{m+1}(t) = d, \tag{8}$$

$$\lambda_s + \lim_{t\to\theta_s-0} v_s(t) = \lambda_{s+1}, \quad s=\overline{1,m}, \tag{9}$$

$$\mu_r = v_r(\xi_{r-1}) + \lambda_r, \qquad r=\overline{1,m+1}. \tag{10}$$

A solution to problem (6)–(10) is a triple $(\lambda^*,\mu^*,v^*[t])$, with elements $\lambda^* = (\lambda_1^*,\lambda_2^*,\dots,\lambda_{m+1}^*)$, $\mu^* = (\mu_1^*,\mu_2^*,\dots,\mu_{m+1}^*)$, $v^*[t] = \big(v_1^*(t),v_2^*(t),\dots,v_{m+1}^*(t)\big)$, where $v_r^*(t)$ are continuously differentiable on $[\theta_{r-1},\theta_r)$, $r=\overline{1,m+1}$, and satisfying the system (6), conditions (7)- (10) at the $\lambda_r=\lambda_r^*$, $\mu_r=\mu_r^*$, $j=\overline{1,m+1}$.

The original problem (1), (2) and problem with parameters (6)–(10) are equivalent.

Let consider $\Phi_r(t)$ a fundamental matrix of the differential equation $\dfrac{dx_r}{dt} = A_0(t)x_r(t)$ on $[\theta_{r-1},\theta_r]$, $r=\overline{1,m+1}$.

Consequently, the solution to the Cauchy problem (6), (7) may be expressed as follows

$$v_r(t) = \Phi_r(t)\int_{\theta_{r-1}}^{t} \Phi_r^{-1}(\tau)\Big[A_0(\tau)\lambda_r + K(\tau)\mu_r + \sum_{i=1}^{m+1} M_i(\tau)\lambda_i\Big]d\tau +$$
$$+\,\Phi_r(t)\int_{\theta_{r-1}}^{t} \Phi_r^{-1}(\tau)f(\tau)d\tau, \quad t\in[\theta_{r-1},\theta_r), \quad r=\overline{1,m+1}. \tag{11}$$

Consider the Cauchy problems for ordinary differential equations on the subintervals

$$\frac{dy}{dt} = A_0(t)y + D(t), \qquad y(\theta_{r-1}) = 0, \qquad t\in[\theta_{r-1},\theta_r], \qquad r=\overline{1,m+1}, \tag{12}$$

where $P(t)$ is a $(n\times n)$-matrix or a n-vector, piecewise continuous on $[0,T]$ with possible discontinuities of the first kind at the $t=\theta_j$, $(j=\overline{1,m})$. On each r-th interval, denote by $P_r(D,t)$ a unique solution to the Cauchy problem (12). The uniqueness of the solution to the Cauchy problem yields

$$P_r(D,t) = \Phi_r(t)\int_{\theta_{r-1}}^{t} \Phi_r^{-1}(\tau)D(\tau)d\tau, \qquad t\in[\theta_{r-1},\theta_r], \qquad r=\overline{1,m+1}. \tag{13}$$

The following system of linear algebraic equations is get by substituting the right-hand side of (11) using (13) into conditions (8)-(10):

$$B_0\lambda_1 + C_0\lambda_{m+1} + C_0P_{m+1}(A_0,T)\lambda_{m+1} + C_0P_{m+1}(K,T)\mu_{m+1}+ \\ + C_0\sum_{i=1}^{m+1}P_i(M_i,T)\lambda_i = d - C_0P_{m+1}(f,T), \quad (14)$$

$$\lambda_s + P_s(A_0,\theta_s)\lambda_s + P_s(K,\theta_s)\mu_s + \sum_{i=1}^{m+1}P_i(M_i,\theta_s)\lambda_i - \lambda_{s+1} = -P_s(f,\theta_s), \quad s=\overline{1,m}, \quad (15)$$

$$\mu_r - P_r(K,\xi_{r-1})\mu_r - P_r(A_0,\xi_{r-1})\lambda_r - \lambda_r - \\ - \sum_{i=1}^{m+1}P_i(M_i,\xi_{r-1})\lambda_i = P_r(f,\xi_{r-1}), \quad r=\overline{1,m+1}. \quad (16)$$

Symbolized by $Q(\theta)$ - $\Big(2n(m+1)\times 2n(m+1)\Big)$ matrix corresponding to the system's left side (14) - (16) and write the system as

$$Q(\theta)(\lambda,\mu) = F(\theta), \qquad \lambda\in\mathbb{R}^{n(m+1)}, \quad \mu\in\mathbb{R}^{n(m+1)}, \quad (17)$$

where $(\lambda,\mu) = \big(\lambda_1,\lambda_2,...,\lambda_{m+1},\mu_1,\mu_2,...,\mu_{m+1}\big)'$,
$F(\theta) = \big(d - C_0P_{m+1}(f,T), -P_1(f,\theta_1), -P_2(f,\theta_2),\ldots,-P_m(f,\theta_m), P_1(f,\xi_0), P_2(f,\xi_1),\ldots,$
$P_{m+1}(f,\xi_m)\big)\in\mathbb{R}^{2n(m+1)}$.

It is simple to establish that the solvability of the boundary value problem (1) and (2) is equivalent to the solvability of the system (17). The solution of the system (17) is a pair of vectors $(\lambda,\mu) = \Big(\lambda_1,\lambda_2,\ldots,\lambda_{m+1},\mu_1,\mu_2,...,\mu_{m+1}\Big)'\in R^{2n(m+1)}$ consists of the values of the solutions of the problem (1) and (2), i.e. $\lambda_r = x(\theta_{r-1})$, $\mu_r = x(\xi_{r-1})$, $r=\overline{1,m+1}$.

3 The Main results

We offer the following formulation of an algorithm for solving problem (1) and (2) based on the solving of Cauchy problems.

Step 1. Split up each r-th interval $[\theta_{r-1},\theta_r]$, $r=\overline{1,m+1}$, into N_r parts. Determine the approximate values of coefficients and system's right side (17) of via solutions to the following Cauchy matrix and vector problems obtained using the fourth-order Runge-Kutta method with step $h_r = (\theta_r-\theta_{r-1})/N_r$, $r=\overline{1,m+1}$:

$$\frac{dy}{dt} = A_0(t)y + A_0(t), \qquad y(\theta_{r-1}) = 0, \quad t\in[\theta_{r-1},\theta_r], \quad r=\overline{1,m+1},$$

$$\frac{dy}{dt} = A_0(t)y + K(t), \qquad y(\theta_{r-1}) = 0, \qquad t\in[\theta_{r-1},\theta_r], \quad r=\overline{1,m+1},$$

$$\frac{dy}{dt} = A_0(t)y + M_i(t), \qquad y(\theta_{r-1}) = 0, \quad t \in [\theta_{r-1}, \theta_r], \quad i = \overline{1, m+1}, \quad r = \overline{1, m+1},$$

$$\frac{dy}{dt} = A_0(t)y + f(t), \qquad y(\theta_{r-1}) = 0, \qquad t \in [\theta_{r-1}, \theta_r], \quad r = \overline{1, m+1}.$$

Step 2. Then we have the approximate system of algebraic equations with respect to parameters λ and μ:

$$Q_*(\theta)(\lambda^*, \mu^*) = F_*(\theta), \qquad \lambda^* \in \mathbb{R}^{n(m+1)}, \quad \mu^* \in \mathbb{R}^{n(m+1)}. \tag{18}$$

Solve the system (18) and we find $(\lambda^*, \mu^*) = (\lambda_1^*, \lambda_2^*, \ldots, \lambda_{m+1}^*, \mu_1^*, \mu_2^*, \ldots, \mu_{m+1}^*)' \in \mathbb{R}^{2n(m+1)}$. Note that the elements of λ^* and μ^* are the values of the solution to problem (1) and (2): $\lambda_r^* = x^*(\theta_{r-1})$, $\mu_r^* = x^*(\xi_{r-1})$, $r = \overline{1, m+1}$.

Step 3. Solve the following Cauchy problems

$$\frac{dy}{dt} = A_0(t)y + K(t)\mu_r^* + \sum_{i=1}^{m+1} M_i(t)\lambda_i^* + f(t),$$

$$y(\theta_{r-1}) = \lambda_r^*, \qquad t \in [\theta_{r-1}, \theta_r), \quad r = \overline{1, m+1},$$

and determine the values of the solution $x^*(t)$ at the remaining points of the subintervals.

Hence, the offered algorithm provides us with the numerical solution to the problem for the system of loaded DEPCAG (1) and (2).

Consider the following example to demonstrate the proposed approach of the numerical solving of problem (1) and (2) based on the modification of Dzhumabaev parametrization method.

4 Example

We consider the problem for the system of loaded DEPCAG:

$$\begin{aligned} \frac{dx}{dt} = \begin{pmatrix} t & t^2 \\ 4t^3 & 4 \end{pmatrix} x + \begin{pmatrix} t^3 & t+3 \\ 2 & t^2 \end{pmatrix} x(\gamma(t)) + \\ + \begin{pmatrix} 2 & t-4 \\ t^3 & 3t \end{pmatrix} x(\theta_0) + \begin{pmatrix} 6t^2 & 3 \\ -6t & 1 \end{pmatrix} x(\theta_1) + f(t), \quad x \in R^2, \quad t \in (0, T), \end{aligned} \tag{19}$$

$$\begin{pmatrix} 2 & 5 \\ -7 & 1 \end{pmatrix} x(0) + \begin{pmatrix} -7 & 1 \\ 1 & 9 \end{pmatrix} x(T) = \begin{pmatrix} -36 \\ -29 \end{pmatrix}, \tag{20}$$

where $\theta_0 = 0, \quad \theta_1 = \dfrac{1}{2}, \quad \theta_2 = T = 1,$

$$\gamma(t) = \zeta_0 = \frac{1}{4}, \quad f(t) = \begin{pmatrix} \frac{57}{8}t^3 - 9t^4 + 18t^2 - \frac{129}{16}t + \frac{97}{16} \\ 16t^5 - 32t^6 - 5t^3 - \frac{49}{16}t^2 + 24t - 4 \end{pmatrix}, \quad t \in \left[0, \frac{1}{2}\right),$$

$$\gamma(t) = \zeta_1 = \frac{3}{4}, \quad f(t) = \begin{pmatrix} \frac{47}{8}t^3 - 9t^4 + 18t^2 - \frac{105}{16}t + \frac{169}{16} \\ 16t^5 - 32t^6 - 5t^3 - \frac{25}{16}t^2 + 24t - \frac{13}{2} \end{pmatrix}, \quad t \in \left[\frac{1}{2}, 1\right).$$

Here we have two subintervals: $\left[0,\frac{1}{2}\right),\quad \left[\frac{1}{2},1\right)$. Applying the scheme of the modification of Dzhumabaev parametrization method, introduce parameters $\lambda_1=x_1(0),\quad \lambda_2=x_2\left(\frac{1}{2}\right)$, $\mu_1=x_1\left(\frac{1}{4}\right),\quad \mu_2=x_2\left(\frac{3}{4}\right)$. Making the substitution

$$v_1(t)=x_1(t)-\lambda_1,\quad t\in\left[0,\frac{1}{2}\right),\quad v_2(t)=x_2(t)-\lambda_2,\quad t\in\left[\frac{1}{2},1\right),$$

we get the boundary value problem with parameters:

$$\frac{dv_1}{dt}=\begin{pmatrix} t & t^2\\ 4t^3 & 4\end{pmatrix}(v_1+\lambda_1)+\begin{pmatrix} t^3 & t+3\\ 2 & t^2\end{pmatrix}\mu_1+\begin{pmatrix} 2 & t-4\\ t^3 & 3t\end{pmatrix}\lambda_1+\begin{pmatrix} 6t^2 & 3\\ -6t & 1\end{pmatrix}\lambda_2+$$
$$+\begin{pmatrix} \frac{57}{8}t^3-9t^4+18t^2-\frac{129}{16}t+\frac{97}{16}\\ 16t^5-32t^6-5t^3-\frac{49}{16}t^2+24t-4\end{pmatrix},\quad t\in\left[0,\frac{1}{2}\right),\tag{21}$$

$$v_1(0)=0,\tag{22}$$

$$\frac{dv_2}{dt}=\begin{pmatrix} t & t^2\\ 4t^3 & 4\end{pmatrix}(v_2+\lambda_2)+\begin{pmatrix} t^3 & t+3\\ 2 & t^2\end{pmatrix}\mu_2+\begin{pmatrix} 2 & t-4\\ t^3 & 3t\end{pmatrix}\lambda_1+\begin{pmatrix} 6t^2 & 3\\ -6t & 1\end{pmatrix}\lambda_2+$$
$$+\begin{pmatrix} \frac{47}{8}t^3-9t^4+18t^2-\frac{105}{16}t+\frac{169}{16}\\ 16t^5-32t^6-5t^3-\frac{25}{16}t^2+24t-\frac{13}{2}\end{pmatrix},\quad t\in\left[\frac{1}{2},1\right),\tag{23}$$

$$v_2\left(\frac{1}{2}\right)=0,\tag{24}$$

$$\begin{pmatrix} 2 & 5\\ -7 & 1\end{pmatrix}\lambda_1+\begin{pmatrix} -7 & 1\\ 1 & 9\end{pmatrix}\lambda_2+\begin{pmatrix} 7 & 1\\ 1 & 9\end{pmatrix}\lim_{t\to 1-0}v_2(t)=\begin{pmatrix} -36\\ -29\end{pmatrix},\tag{25}$$

$$\lambda_1+\lim_{t\to\theta_1-0}v_1(t)=\lambda_2,\tag{26}$$

$$\mu_1=v_1\left(\frac{1}{4}\right)+\lambda_1,\quad \mu_2=v_2\left(\frac{3}{4}\right)+\lambda_2.\tag{27}$$

By dividing the subintervals $\left[0,\frac{1}{2}\right),\left[\frac{1}{2},1\right)$, with step $h=0.05$ we give the results of the numerical implementation of algorithm

Using equivalent problem (21)-(27) and solving the relevant system of linear algebraic equations (18) we get

$$\lambda_1^*=\begin{pmatrix} -0.999996216\\ 0.000001946\end{pmatrix},\quad \lambda_2^*=\begin{pmatrix} 0.999995267\\ -1.749996275\end{pmatrix},$$

Таблица 1: Comparison of exact and numerical solutions to problem (19), (20)

t	$\lvert x_1^*(t)-\widetilde{x}_1(t)\rvert$	$\lvert x_2^*(t)-\widetilde{x}_2(t)\rvert$	t	$\lvert x_1^*(t)-\widetilde{x}_1(t)\rvert$	$\lvert x_2^*(t)-\widetilde{x}_2(t)\rvert$
0	0.000003784	0.000001946	0.5	0.000004733	0.000003725
0.05	0.000003775	0.000001785	0.55	0.000003786	0.000004755
0.1	0.000003772	0.000001679	0.6	0.000002852	0.00000587
0.15	0.00000378	0.000001638	0.65	0.000001937	0.000007011
0.2	0.000003807	0.000001674	0.7	0.000001054	0.000008079
0.25	0.000003858	0.000001796	0.75	0.00000023	0.000008925
0.3	0.000003941	0.000002011	0.8	0.000000489	0.000009324
0.35	0.000004063	0.000002324	0.85	0.000001025	0.000008942
0.4	0.00000423	0.000002728	0.9	0.000001251	0.000007283
0.45	0.000004451	0.000003207	0.95	0.000000967	0.000003617
0.5	0.000004733	0.000003725	1	0.00000014	0.000003144

$$\mu_1^* = \begin{pmatrix} 0.874996142 \\ -0.937498204 \end{pmatrix}, \quad \mu_2^* = \begin{pmatrix} 2.12499977 \\ -2.437491075 \end{pmatrix}.$$

Then, using the found values λ_1^*, λ_2^*, μ_1^*, μ_2^*, we solve the Cauchy problems by the fourth-order Runge-Kutta method

$$\frac{d\widetilde{x}_1}{dt} = \begin{pmatrix} t & t^2 \\ 4t^3 & 4 \end{pmatrix}\widetilde{x}_1 + \begin{pmatrix} t^3 & t+3 \\ 2 & t^2 \end{pmatrix}\cdot\begin{pmatrix} 0.874996142 \\ -0.937498204 \end{pmatrix} +$$
$$+\begin{pmatrix} 2 & t-4 \\ t^3 & 3t \end{pmatrix}\cdot\begin{pmatrix} -0.999996216 \\ 0.000001946 \end{pmatrix} + \begin{pmatrix} 6t^2 & 3 \\ -6t & 1 \end{pmatrix}\cdot\begin{pmatrix} 0.999995267 \\ -1.749996275 \end{pmatrix} +$$
$$+\begin{pmatrix} \frac{57}{8}t^3 - 9t^4 + 18t^2 - \frac{129}{16}t + \frac{97}{16} \\ 16t^5 - 32t^6 - 5t^3 - \frac{49}{16}t^2 + 24t - 4 \end{pmatrix}, \quad \widetilde{x}_1(0) = \begin{pmatrix} -0.999996216 \\ 0.000001946 \end{pmatrix}, \quad t\in\Big[0,\frac{1}{2}\Big),$$

$$\frac{d\widetilde{x}_2}{dt} = \begin{pmatrix} t & t^2 \\ 4t^3 & 4 \end{pmatrix}\widetilde{x}_2 + \begin{pmatrix} t^3 & t+3 \\ 2 & t^2 \end{pmatrix}\cdot\begin{pmatrix} 2.12499977 \\ -2.437491075 \end{pmatrix} +$$
$$+\begin{pmatrix} 2 & t-4 \\ t^3 & 3t \end{pmatrix}\cdot\begin{pmatrix} -0.999996216 \\ 0.000001946 \end{pmatrix} + \begin{pmatrix} 6t^2 & 3 \\ -6t & 1 \end{pmatrix}\cdot\begin{pmatrix} 0.999995267 \\ -1.749996275 \end{pmatrix} +$$
$$+\begin{pmatrix} \frac{47}{8}t^3 - 9t^4 + 18t^2 - \frac{105}{16}t + \frac{169}{16} \\ 16t^5 - 32t^6 - 5t^3 - \frac{25}{16}t^2 + 24t - \frac{13}{2} \end{pmatrix}, \quad \widetilde{x}_2\Big(\frac{1}{2}\Big) = \begin{pmatrix} 0.999995267 \\ -1.749996275 \end{pmatrix}, \quad t\in\Big[\frac{1}{2},1\Big).$$

and we find numerical solution of the problem (19) and (20).

Exact solution of the (19) and (20) is $x^*(t) = \begin{pmatrix} 8t^3 - 4t^2 + 1 \\ t^2 - 4t \end{pmatrix}$.

In Table 1, difference between the exact solution $x^*(t_k)$ and numerical solution $\widetilde{x}(t_k)$, $k=\overline{0,20}$, are shown.

5 Acknowledgement

This research is funded by the Science Committee of the Ministry of Education and Science of the Republic of Kazakhstan (Grant No. AP08855726).

References

[1] Busenberg S., Cooke K.L., "Models of vertically transmitted disease with sequential continuous dynamics" , *Lakshmikantham V (ed) Nonlinear phenomena in mathematical sciences,* (1982): 179-189.

[2] Cooke K.L., Wiener J., "Retarded differential equations with piecewise constant delays" , *J Math Anal Appl,* 99, (1984): 179-189.

[3] Akhmet M., Yilmaz E., "Neural Networks with Discontinuous/Impact Activations" , (Springer, New York, 2014).

[4] Akhmet M.U., "Nonlinear hybrid continuous/discrete time models" , (Atlantis, Amsterdam-Paris, 2011).

[5] Akhmet M.U., "Almost periodic solution of differential equations with piecewise-constant argument of generalized type" , *Nonlinear Analysis-Hybrid Systems,* 2, (2008): 456-467.

[6] Akhmet M.U., "On the reduction principle for differential equations with piecewise-constant argument of generalized type" , *J. Math. Anal. Appl.,* 1, (2007): 646-663.

[7] Dzhumabaev D.S., "Criteria for the unique solvability of a linear boundary-value problem for an ordinary differential equation" , *USSR Comput. Math. Math. Phys.,* 1, (1989): 34–46.

[8] Temesheva S.M., Dzhumabaev D.S., Kabdrakhova S.S. "On One Algorithm To Find a Solution to a Linear Two-Point Boundary Value Problem" , *Lobachevskii J. of Math.,* 42, (2021): 606-612.

[9] Nakhushev A.M., "Loaded equations and their applications" , (Nauka, Moscow, 2012) (in Russian).

[10] Nakhushev A.M., "An approximation method for solving boundary value problems for differential equations with applications to the dynamics of soil moisture and groundwater" , *Differential Equations,* 18, (1982): 72–81.

[11] Dzhenaliev M.T. "Loaded equations with periodic boundary conditions" , *Differential Equations,* 37, (2001): 51-57.

[12] Abdullaev V.M., Aida-zade K.R. "Numerical method of solution to loaded nonlocal boundary value problems for ordinary differential equations" , *Comput. Math. Math. Phys.,* 54, (2014): 1096-1109.

[13] Assanova A.T., Kadirbayeva Zh.M. "Periodic problem for an impulsive system of the loaded hyperbolic equations" , *Electronic Journal of Differential Equations,* 72, (2018): 1-8.

[14] Assanova A.T., Imanchiyev A.E., Kadirbayeva Zh.M. "Numerical solution of systems of loaded ordinary differential equations with multipoint conditions" , *Comput. Math. Math. Phys.,* 58, (2018): 508–516.

[15] Assanova A.T., Kadirbayeva Zh.M. "On the numerical algorithms of parametrization method for solving a two-point boundary-value problem for impulsive systems of loaded differential equations" , *Comp. and Applied Math.,* 37, (2018): 4966–4976.

[16] Kadirbayeva Zh.M., Kabdrakhova S.S., Mynbayeva S.T. "A Computational Method for Solving the Boundary Value Problem for Impulsive Systems of Essentially Loaded Differential Equations" , *Lobachevskii J. of Math.,* 42, (2021): 3675-3683.

[17] Dzhumabaev D.S. "On one approach to solve the linear boundary value problems for Fredholm integro-differential equations" , *J. of Comp. and Applied Math.,* 294, (2016): 342-357.

IRSTI 27.39.19 DOI: https://doi.org/10.26577/JMMCS.2022.v115.i3.03

A.B. Birgebaev[1*], **M.B. Muratbekov**[2]
[1]Abai Kazakh National Pedagogical University, Kazakhstan, Almaty
[2]M.Kh. Dulati Taraz Regional University, Kazakhstan, Taraz
*e-mail: ahtai@mail.ru

SMOOTHNESS OF SOLUTIONS (SEPARABILITY) OF THE NONLINEAR STATIONARY SCHRÖDINGER EQUATION

The equation of motion of a microparticle in various force fields is the Schrödinger wave equation. Many questions of quantum mechanics, in particular the thermal radiation of electromagnetic waves, lead to the problem of separability of singular differential operators. One such operator is the above Schrödinger operator. In this paper, the named operator is studied by the methods of functional analysis. Found sufficient conditions for the existence of a solution and the separability of an operator in a Hilbert space. All theorems were originally proved for the model Sturm-Liouville equation and extended to a more general case.
In §1-2, for the nonlinear Sturm-Liouville equation, sufficient conditions are found that ensure the existence of an estimate for coercivity, and estimates of weight norms are obtained for the first derivative of the solution. In Sections 3-4 the results of Sections 1-2 are generalized for the Schrödinger equation in the case $m = 3$.
Key words: Nonlinear equations, continuous operator, equivalence, potential function.

Биргебаев А.Б[1*], Муратбеков М.Б.[2]
[1]Абай атындағы Қазақ ұлттық педагогикалық университеті, Қазақстан, Алматы қ.
[3]М.Х. Дулати атындағы Тараз өңірлік университеті, Қазақстан, Тараз қ.
*e-mail: ahtai@mail.ru

Шредингер теңдеуінің сызықты емес стационарлық теңдеуінің шешімдерінің тегістілігі (бөлімділігі)

Микробөлшектердің әртүрлі күш өрістеріндегі қозғалыс теңдеуі Шредингер толқынының теңдеуі болып табылады. Кванттық механиканың көптеген сұрақтары, атап айтқанда электромагниттік толқындардың жылулық сәулеленуі сингулярлы дифференциалдық операторлардың бөліну мәселесіне әкеледі. Осындай операторлардың бірі жоғарыдағы Шредингер операторы болып табылады. Бұл жұмыста аталған оператор функционалдық талдау әдістерімен зерттеледі. Шешімнің болуы және Гильберт кеңістігіндегі оператордың бөлінуі үшін жеткілікті шарттар табылды. Барлық теоремалар бастапқыда Штурм-Лиувилл теңдеуінің үлгісі үшін дәлелденді және жалпы жағдайға дейін кеңейтілді.
§1-2-де сызықты емес Штурм-Лиувилл теңдеуі үшін коэрцивтілік бағасының болуын қамтамасыз ететін жеткілікті шарттар табылды және шешімнің бірінші туындысы үшін салмақ нормаларының бағалаулары алынды. 3-4 бөлімдерде 1-2 бөлімдердің нәтижелері $m = 3$ жағдайындағы Шредингер теңдеуі үшін жалпыланған.
Түйін сөздер: Сызықты емес теңдеулер, үздіксіз оператор, эквиваленттілік, потенциалдық функция.

Біргебаев А.Б[1*], Муратбеков М.Б.[2]
[1]Казахский национальный педагогический университет имени Абая, Казахстан, г. Алматы
[2]Таразский региональный университет имени М.Х. Дулати, Казахстан, г. Тараз
*e-mail: ahtai@mail.ru

Гладкость решений (разделимость) нелинейного стационарного уравнения Шредингера

Уравнением движения микрочастицы в различных силовых полях является волновое уравнение Шредингера. Многие вопросы квантовой механики в частности тепловое излучение электромагнитных волн приводят к задаче разделимости сингулярных дифференциальных операторов. Одним из таких операторов является вышеуказанный оператор Шредингера. Данной работе исследуется названный оператор методами функционального анализа. Найденный достаточные условия существовании решении и разделимости оператора в Гильбербовом пространстве. Все теоремы первоначально доказаны для модельного уравнение Штурма -Лиувилля и распространено на более общий случай.
В §1-2 для нелинейного уравнения Штурма-Лиувилля найдены достаточные условия, обеспечивающие наличие оценки коэрцитивности, а для первой производной решения получены оценки весовых норм. В §3-4 обобщены результаты §1-2 для уравнения Шредингера в случае $m = 3$.
Ключевые слова: Нелинейные уравнения, непрерывный оператор, эквивалентность, потенциальная функция.

1 Introduction

In this paper, the smoothness of solutions to the nonlinear equation is considered

$$Lu = -\Delta u + q(x, u)u = f(x) \in L_2(R^m)$$

In [1,2] for the nonlinear Sturm-Liouville equation, sufficient conditions are found that ensure the existence of an estimate for the coecitivity, and for the first derivative of the solution, estimates for the weight norms were obtained. In [1,2] generalized the results of §1-2 for the Schrödinger equation in the case $m = 3$.

2 Materials and methods

For simplicity, we present one result for the Sturm-Liouville equation.

Theorem 1 *Let the following conditions are satisfied:*

a) $q(x, y) \geq \delta\rangle 0$*;*

b) $q(x, y)$ *is a continuous function on the set of variables in* R^2*;*

c) $\sup\limits_{[x-\eta)\leq 1} \quad \sup\limits_{|C_0-C_1|\leq A} \quad \sup\limits_{|C_0|\leq A} \dfrac{q(x, C_0)}{q(x, C_1)} < \infty$, *where is any finite value. Then for any* $f(x) \in L_2(R^m)$ *there is a solution* (x) *to the equation*

$$Ly = -y''(x) + q(x, y)y = f,$$

which has quadratically summable second derivative, i.e. $y''(x) \in L_2(R^m)$*.*

The proof of this theorem belongs to Muratbekov M.B. [3]. Unfortunately, in the work [8] the author was incorrectly specified. Please apologize for inaccuracy. As we will see later (in Section 2.4), such results hold for a wide class of nonlinear operators. For linear operators of similar work was considered in [1-3, 5-7, 9, 11, 12, 13]

Let us enter the following designations: R^m is Euclidean m-dimensional real space of points $x = (x_1, x_2, \ldots, x_m)$. $\bar{\Omega}$ is a closure of Ω where Ω is an open set in R^m, $\|\cdot\|_{p,\Omega}$. is a norm of the element $L_\rho(\Omega)$. Instead of $\|\cdot\|_{p,\Omega}$ at $\Omega = R^m$ we will write $\|\cdot\|_\rho$, if $p = 2$ in designations $\|\cdot\|_{\rho,\Omega}$ and $\|\cdot\|_p$ we will omit ρ.

$$D_u^\alpha = \frac{\partial^{|\alpha|} u}{\partial x_1^{\alpha_1} \ldots \partial x_m^{\alpha_m}},$$

$\alpha = (\alpha_1, \ldots, \alpha_m)$ – multiindex, $|\alpha| = \alpha_1 + \alpha_2 + \ldots + \alpha_m$. $C_1, C_2, \ldots$ are various constants constants, the exact value of which does not interest us.

2.1 Existence of the solution

In the given section the following equation is considered

$$Ly = -y''(x) + q(x,y)y = f(x) \in L_2(R), \tag{1}$$

where $R = (-\infty, \infty)$.

The function $y \in L_2(R)$ is called the weak solution of equation (1), if there is a sequence $\{y_n\} \subset W_2^1(R) \bigcap W_{2,l>c}^2(R)$ such that

$$\|y_n - y\|_{\alpha_{2,loc(R)}} \to 0, \quad \|Ly_n - f\|_{L_{2,loc(R)}} \to 0, \quad n \to \infty.$$

It is said that the sequence $\{\eta_n\}_{n=1}^\infty$ of basic functions from $C_0^\infty(R^m)$ converges to (1) in R^m, if:

a) for any compact $K \subset R^m$ there will be such a number N, that $\eta_n(x) = 1$ at all $x \in K$ and $n \geq N$

b) functions $\{\eta_n\}$ uniformly limited in R^m, $|\eta_n(x)| \leq 1$, $x \in R^m$, $n = 1, 2, \ldots$ [8].

Lemma 1 *Let $q(x,y) \geq \delta < 0$ and is continuous on both arguments in R^2, then for any $f \in L_2(R)$ there is a weak solution of the equation (1) in the space $W_2^1(R)$.*

Proof. Since, according to the assumption, the function $q(x,y)$ is limited from below, then, without losing the generality of reasoning, we can assume that the condition $q(x,y) \geq 1$ is hold.

First, we will be engaged in proving the existence of a solution to the first boundary value problem

$$L_{n_\varepsilon} y_{n_\varepsilon} = -y''_{n_\varepsilon} + y_{n_\varepsilon} + \frac{(q(x, y_{n_\varepsilon}) - 1) y_{n_\varepsilon}}{(1 + \varepsilon(q)(x, y_{n_\varepsilon}) - 1) + \varepsilon \|b(x, y_{n_\varepsilon})\|_{2,(-a_n, a_n)}} = f\eta_n, \tag{2}$$

$$y_{n_\varepsilon}(+a) = y_{n_\varepsilon}(a) = 0, \tag{3}$$

where $[-a_n, a_n]$ – $\sup p\eta_n$, and $b(x, y_{n_e}) = (q(x, y_{n_\varepsilon}) - 1) y_{n_\varepsilon}$ in the space $W_{2,0}^2[-a_n, a_n]$; $W_{2,0}^2[-a_n, a_n]$ – is space of functions $z \in W_2^2$ и $z(-a_n) = z(a_n) = 0$.

We will reduce problem (2) – (3) to an equivalent integral equation, to which we then apply the Schauder principle [9].

Let us denote by L_0 the operator defined on $W_{2,0}^2[-a_n, a_n]$ with the equality

$$L_0 y = -y''(x) + y(x).$$

Due to the known theorems for the Sturm-Liouville operator there is a completely continuous inverse operator L_0^{-1}, defined all over space $L_2[-a_n, a_n]$. We need Lemma.

Lemma 2 *The problem (2) – (3) is equivalent to the integral equation*

$$z_{n_\varepsilon} = \frac{(q(x, L_0^{-1} z_{n_\varepsilon}) - 1) L_0^{-1} z_{n_\varepsilon}}{1 + \varepsilon (q(x, L_0^{-1} z_{n_\varepsilon}) - 1) + \varepsilon \| b(x, L_0^{-1} z_{n_\varepsilon}) \|_2^2} + f\eta_n, \tag{4}$$

$$z_{n_\varepsilon}, f\eta_n \in L_2[-a_n, a_n].$$

The proof is obvious.

Let us denote by A the operator which acts on the following formula:

$$A(z) = \frac{(q(x, L_0^{-1} z) - 1) L_0^{-1} z}{1 + \varepsilon (q(x, L_0^{-1} z) - 1) + \varepsilon \| b(x, L_0^{-1} z) \|_{2,[-a_n, a_n]}^2} + f\eta_n.$$

Further we denote

$$\overline{S}(0; N) = \left\{ \vartheta \in L_2(-a_n, a_n) : \|\vartheta\|_2 \le N = \frac{1}{\sqrt{\varepsilon}} \right\},$$

where $\vartheta = z - f\eta_n$. Consider the operator on this ball

$$A(\vartheta) = A(z) - f\eta_n = A(\vartheta + f\eta_n) - f\eta_n =$$

$$= \frac{(q(x, L_0^{-1}(\vartheta + f\eta_n)) - 1) L_0^{-1}(\vartheta + f\eta_n)}{1 + \varepsilon (q(x, L_0^{-1}(\vartheta + f\eta_n)) - 1) + \varepsilon \| b(x, L_0^{-1}(\vartheta + f\eta_n)) \|_{2,(-a_n, a_n)}^2}.$$

It is obvious that, if ϑ_0 – is a fixed point of operator $_m$, then $\vartheta_0 + f\eta_n$ – is a fixed point of operator . Therefore, in the future instead of operator A, it is enough to consider A_0.

Let us prove that $_0$ reflects the ball $\overline{S}(0; N) \in L_2[-a_n, a_n]$ in itself. Let $\vartheta \in \overline{S}(0; N)$. We will consider two cases:

1.

$$\| (q(x, L_0^{-1}(\vartheta + f\eta_n)) - 1) L_0^{-1}(\vartheta + f\eta_n) \|_{2,(-a_n, a_n)}^2 \le N = \frac{1}{\sqrt{\varepsilon}}.$$

Then

$$\|A_0(\vartheta)\|_2 = \left\| \frac{(q(x, L_0^{-1} z) - 1) L_0^{-1} z}{1 + \varepsilon (q(x, L_0^{-1} z) - 1) + \varepsilon \| b(x, L_0^{-1} z) \|_2^2} \right\|_{2,(-a_n, a_n)} \le$$

$$\le \| (q(x, L_0^{-1}(\vartheta + f\eta_n)) - 1) L_0^{-1}(\vartheta + f\eta_n) \| \le N = \frac{1}{\sqrt{\varepsilon}}$$

2.

$$(q(x, L_0^{-1}(\vartheta + f\eta_n)) - 1) L_0^{-1})(\vartheta + f\eta_{n)} \| \ge N.$$

Then

$$A_0(\vartheta)_2 \le \frac{\| (q(x, L_0^{-1}(\vartheta + f\eta_n)) - 1) L_0^{-1}(\vartheta + f\eta_n) \|_{2,(-a_n, a_n)}}{\varepsilon \| (q(x, L_0^{-1}(\vartheta + f\eta_n)) - 1) L_0^{-1}(\vartheta + f\eta_n) \|_{2,(-a_n, a_n)}^2} =$$

$$= \frac{1}{\varepsilon \| (q(x, L_0^{-1}(\vartheta + f\eta_n)) - 1) L_0^{-1}(\vartheta + f\eta_n) \|_{2,(-a_n, a_n)}} \le \frac{1}{\varepsilon N} = \frac{1}{\sqrt{\varepsilon}}.$$

Therefore,

$$\|A(\vartheta)\|_{2,(-a_n, a_n)} \le N, \quad \forall \vartheta \in \overline{S}(0; N). \tag{5}$$

Now we will show that $_m$ – is completely continuous operator at $\overline{S}(0;N)$. Continuity is obvious. Further, by virtue of Riesz theorem, it is enough to prove that the set of functions $\{A_0\vartheta : \vartheta \in \overline{S}(0;N)\}$ is uniformly limited and the relation is performed

$$\lim_{h\to 0} \|(A_0(\vartheta))(x+h) + (A_0(\vartheta))(x)\|_{2,(-a_n,a_n)} = 0$$

uniformly on $\vartheta \in \overline{S}$.

Due to estimate (5) the set of functions $\{A_0(\vartheta) : \vartheta \in \overline{S}(0;N)\}$ is uniformly bounded.

Due to the continuity of $q(x,y)$ on combination of variables and properties of the operator L_0^{-1}, the relation $q(x,y)$

$$\|(A_0(\vartheta)(x+h) - A_0(\vartheta))(x)\|^2_{2,(-a_n,a_n)} \to 0$$

uniformly at $h \to 0$ on $\vartheta \in \overline{S}(0;N)$.

Thus, the operator A_m is completely continuous and reflects $\overline{S}(0;N)$ in itself. Therefore, according to the Schauder principle; integral equation (4) has at least one solution in the ball $\overline{S}(0;N)$. Hence, by virtue of Lemma 2, it follows that there exists a solution to problem (2) – (3) belonging to the space W_2^2.

Further $\|y_{n_\varepsilon}\|_{W_2^1[-a_n,a_n]}$ is estimated from above by constant independent of n, ε.

To prove this fact, let us take the linear operator

$$\ell_{n_\varepsilon} y = y''(x) + (1 + \frac{\tilde{q}(x) - 1}{1 + \varepsilon(\tilde{q}(x) - 1) + \varepsilon\|(q(x, y_{n_\varepsilon}) - 1)y_{n_\varepsilon}\|_2^2})y(x),$$

Defined on a set $W_{2,0}^2(-a_n, a_n)$, where $\tilde{q}(x) = q(x, y_{n_\varepsilon})$, and y_{n_ε} – is a solution of the problem (2) – (3) with the right side $f\eta_n$. Let us construct a scalar product $\langle \ell n_\varepsilon, y_{n_\varepsilon}, y_{n_\varepsilon}\rangle$. Integrating in parts and taking into account that non-integral members disappear due to (3), we obtain

$$\|y_{n_\varepsilon}\|_{W_2^1[-a_n,a_n]} \le 2^{1/2} \left(\int_{-\infty}^{\infty} |f|^2 dx \right)^{1/2}.$$

Assume that $C = 2^{1/2} \left(\int_{-\infty}^{\infty} |f|^2 dx \right)^{1/2}$, then

$$\|y_{n_\varepsilon}\|_{W_2^1[-a_n,a_n]} \le C. \tag{6}$$

Let us choose some sequence $\{y_{n_{\varepsilon_k}}\}$ of solutions belonging to a bounded set $\{y_{n_\varepsilon}\}$, so that

$$\|y_{n_{\varepsilon_k}}\|_{W_2^1[-a_n,a_n]} \le C, \tag{7}$$

where $\varepsilon_k \to 0$ at $k \to \infty$.

By virtue of (7) from the sequence $\{y_{n_{\varepsilon_k}}\}$ we can select subsequence, denote it again by $\{y_{n_{\varepsilon_k}}\}$, so that

$$y_{n_{\varepsilon_k}} \to y_n \quad \text{weakly in} \quad W_2^1(-a_n, a_n),$$

$$y_{n_{\varepsilon_k}} \to y_n \ \text{ weakly in } \ L_2(-a_n, a_n).$$

From (7) we have $\|y_n\|_{W_2^1(-a_n,a_n)} \leq C$, and it is not difficult to see that y_n satisfies the equation

$$L_n y_n = -y_n''(x) + q(x, y_n) y_n = f\eta_n \ \text{ and } \ y_n(-a_n) = y_n(a_n) = 0.$$

Next, each y_n we continue with zero outside of $[-a_n, a_n]$, continuation denote by $\tilde{y}_n$. With this continuation, we obtain elements $W_2^1(R)$, norms of which are limited:

$$\|\tilde{y}_{n_\varepsilon}\|_{W_2^1(R)} \leq C.$$

Therefore, from the sequence, we can select a subsequence $\tilde{y}_{n_k}$, such that

$$\tilde{y}_{n_k} \to y \ \text{ weakly in } \ W_2^1(R) \tag{8}$$

$$\tilde{y}_{n_k} \to y \ \text{ weakly in } \ L_{2,\ell\, oc}(R), \tag{9}$$

and besides

$$\|y\|_{W_2^1(R)} \leq C. \tag{10}$$

Let $[\alpha, \beta]$ is any fixed segment in R. Then for any $\varepsilon\rangle 0$ there exists such number N, that at $k = N(\alpha, \beta) \in \sup p\tilde{y}_{n_k}$ and by virtue (8)

$$\|L\tilde{y}_{n_k} - f\|_{2,(\alpha,\beta)} \langle \varepsilon.$$

From here and (9) we get that $y(x)$ is a weal solution of the equation (1). Lemma is proved.

2.2 Smoothness of the solution

In this section we will show that all solutions from $W_2^1(R)$ will be elements from $W_2^2(R)$, as soon as a potential function known in it has some properties.

Theorem 2 *Let the following conditions hold;*

a) $q(x, y) \geq \delta\rangle 0$;

b) $q(x, y)$ *is continuous function on a set of variables in* R^2;

c) $\sup\limits_{|x-\eta|\leq 1} \ \sup\limits_{|C_1 - C_2| \leq A} \ \sup\limits_{|C_1|\leq A} \dfrac{q(x, C_1)}{q(x, C_2)} < \infty$,

where A is any finite value. Then for any $f \in L_2(R)$ there exists the solution $y(x) \in L_2(R)$ of the equation (1), such that $y''(x) \in L_2(R)$.

Theorem 3 *Let the conditions hold:*

a) $q(x, y) \geq \delta\rangle 0$;

b) $q(x, y)$ *are continuous on a set of variables in* R^2;

c) $\sup\limits_{x\in R} \ \sup\limits_{|C_1 - C_2| \leq A} \ \sup\limits_{|C_2|\leq A} \dfrac{q(x, c_1)}{\theta^2(x, c_2)} < \infty$, *where*

$$\theta(x, C_1) = \inf_{d\rangle 0 |x-t|\leq 10} (d^{-1} + \int\limits_{|t-h|\leq d} q(\eta, C_2) d\eta),$$

A is any finite value. Then for any $f \in L_2(R)$ there exists the solution $y(x) \in L_2(R)$ of the equation (1) such that $y''(x) \in L_2(R)$.

Theorem 4 *Let the conditions a)-c) of theorem 2 are held and $r(x)$ is a continuous, such that* $\sup\limits_{|x-y|\le 1} \dfrac{r(y)}{r(x)} < \infty$.

If for any $k > 0$ the value

$$B = \sup_{x\in R} \sup_{|C_1|\le K} \sup_{0<\eta\le m^{-1}(x,C_1)} \left[\eta^{-p} \int\limits_{|t-x|\le\eta} |r(t)|^0 dt \right]^{1/\theta}$$

is finite, then for any $f \in L_2(R)$ function

$$r(x)\frac{d}{dx}y(x) \in L_2(R), \quad (2 \le \theta < \infty,\ p = -\frac{\theta}{2}, m(x,C_1) = (q(x,C_1))^{1/q}),$$

here $y(x)$ is the solution of the equation (1) from $L_2(R)$.

Proof of Theorems 2-4. At any function $f \in L_2(R)$ by virtue of Lemma 1 for the equation there exists a solution $y(x)$ such that $y(x) \in W_2^1(R)$. Therefore, by Sobolev's embedding theorem [10] $y(x) \in C(R)$. Then according to the condition b)

$$q(x, y(x)) \in C_{\ell oc}(R). \tag{11}$$

Let $y_0(x)$ is a weak solution of the equation (1) with the right side $f_0 \in L_2(R)$. Since $y_0(x) \in W_2^1(R)$, then

$$y_0(t) - y_0(\eta) = \int\limits_{\eta}^{t} \frac{dy_0}{dx}dx.$$

By the Bunyakovsky inequality and by (10), we have

$$|y_0(t) - y_0(\eta)| \le (|t-\eta|)^{1/2}\|f\|_{2,R}. \tag{12}$$

Assume that $\tilde{q}(x) = q(x, y_0(x))$ and denote by $\tilde{L}$ closure in norm of L_2 operator, given on $C_0^\infty(R)$ by equality $L_0 y = -y''(x) + \tilde{q}(x)y$.

Lemma 3 *Operator $\tilde{L}$ is self adjoint and positive defined.*

Proof. The positive definiteness of $\tilde{L}$ follows from condition a) of Theorem 2. Self-adjointness follows from (2) and from the results of [2]. The lemma is proved.

Now, assuming that $y_0(t) = C_2$, $y_0(\eta) = C_1$, $A = 2\|f\|_2 \ge \sqrt{A\eta\|f\|_2}$, from (12) we obtain $|C_2 - C_1| \le A$. From here, due to conditions a)-c) of Theorem 2, for operator $\tilde{L}$ all conditions of the Theorem 3, 4 are satisfied. Therefore, the operator L is separable, i.e.

$$\|y''\|_2 + \|\tilde{q}(x)y\|_2 \le C(\|\tilde{L}y\| + \|y\|_2),$$

where does not depend on $y \in D(\tilde{L})$, where $D(\cdot)$ is the definition area, and $\|\cdot\|$ is the norm in $L_2(D)$.

It remains for us to show that $y_0(x) \in D(\tilde{L})$. Suppose the contrary, that $y_0(x) \notin D(\tilde{L})$. By virtue of Lemma 2, there exists $y_1(x) \in W_2^1(R)$ such that $y_1(x) = \tilde{L}^{-1} f_0$. So, it is assumed that $y_0(x) \in W_2^1(R)$ is a solution of equation (1) with the right side of $f_o(x)$, then

$$\tilde{L} y_2 = 0, y_2 = y_1 - y_0 \in L_2(R).$$

To complete the proof of the theorem, we need a lemma.

Lemma 4 *Let the conditions a) and b) of theorem 2 be satisfied. Then the equation* $\tilde{L}y = 0$ *does not have a solution* $y(x) \in L_2(R)$.

Proof. It is well known that if $\tilde{q}(x) \geq \delta > 0$, then the solution of the equation $y''(x) = q(x)y$ exponentially grows both at $x \to -\infty$, and at $x \to +\infty$. Therefore, this solution cannot belong to $L_2(R)$. The Lemma is proved.

From this lemma we obtain that $y_0(x) = y_1(x)$. We get a contradiction. The theorem 2. is completely proved.

Theorems 3, 4 are proved in the same way.

2.3 Nonlinear Schrödinger-type operator in $L_2(R^3)$

Now let us consider the equation

$$-\Delta u + q(x, u)u = f(x) \tag{13}$$

in the space $L_2(R^3)$.

Lemma 5 *Let* $q(x, u) \geq \delta > 0$ *and is continuous on both arguments in* R^2*, then for each* $f \in L_3(R^3)$ *there is a weak solution to equation* (13) *in space* $W_2^1(R^3)$.

This lemma is proved in the same way as the lemma 1.

Lemma 6 *Let* $q(x, u) \geq \delta > 0$ *and is continuous on both arguments in* R^2*, then for each* $f \in L_2(R^3)$ *there is a weak solution to equation* (13) *and the following inequality holds*

$$\|u\|_{L_\infty(R^3)} + \|u\|_{W_2^1(R^3)} \leq C\|f\|_{L_2(R^3)}, \tag{14}$$

Where the constant C *does not depend on* u *and* f.

Proof. Let

$$q_N(x, u) = \begin{cases} q(x, u), & \text{if} \quad q(x, u) \leq N, \\ N, & \text{if} \quad q(x, u) \geq N \end{cases}$$

The existence of a solution to the equation

$$-\Delta u + q_N(x, u)u = f_N \tag{15}$$

follows from lemma 5.

Let $u_x \in W_2^1(R^3)$is a solution to equation (15). Let us consider the equation

$$L_u = f_N, \tag{16}$$

where $L = -\Delta + \tilde{q}_N(x)$,

Since $q_N(x, u_N)$ are limited and $\tilde{q}_N(x)$, then on the theorem (3), [see 11] operator L is self-adjoint and the equation (16) has a unique solution that coincides with u_N.

It is known, if $q_1(x) \le q_2(x)$, then $Q_1(x, y) \ge 0$ and $Q_2(x, y) \ge 0$, and $Q_1(x, y) \ge Q_2(x, y)$, where $Q_1(x, y)$ and $\tilde{N}_2(x, y)$ are Green functions of operators $-\Delta + q_1(x)$, $-\Delta + q_2(x)$.

Let $Q_N(x, y)$is the Green function of the operator L, then it follows from the above fact that

$$Q_N(x, y) \le Q_0(x, y), \tag{17}$$

where $Q_0(x, y)$ is Green function of the operator $-\Delta + 1$. It follows from this and (17) that

$$|u_x(x)| = \left| \int\limits_{R^3} Q_N(x, y) f(y) dy \right| \le \int\limits_{R^3} Q_N(x, y) f(y) dy \le \int\limits_{R^3} Q_0(x, y) |f(y)| dy.$$

It is known that the operator

$$(Qf)(x) = u_0(x) = \int\limits_{R^3} Q_0(x, y) |f(y)| dy \tag{18}$$

acts from $L_2(R^3)$ in $W_2^2(R^3)$. Therefore, by virtue of the Sobolev embedding theorems [10], we have

$$\|u_N(x)\|_{L_\infty(R^3)} \le C_0 \|f\|_{L_2(R^3)}, \tag{19}$$

where C_0 does not depend on N and f.

On the other hand, here is an estimation

$$\|u_N(x)\|_{W_2^1(R^3)} \le C_1 \|f\|_{L_2(R^3)}, \tag{20}$$

where C_1 does not depend on N and f.

Indeed, we will compose a scalar product $\langle Lu_N, u_N \rangle$. Integrating in parts, we obtain (20).

From (19) and (20) we will have

$$\|u_N(x)\|_{L_\infty(R^3)} + \|u_N\|_{W_2^1(R^3)} \le C_2 \|f\|, \tag{21}$$

where $C_2 = \max(C_1, C_2)$.

Moving to limit at $N \to \infty$ we get

$$\|u(x)\|_{L_\infty(R^3)} + \|u(x)\|_{W_2^1(R^3)} \le C_2 \|f\|_{L_2(R^3)}.$$

It is not difficult to check that $u(x)$ is the weak solution to equation (13) (see lemma 2). The lemma is proved.

2.4 Smoothness of the solution

Theorem 5 *Let the following conditions be satisfied: a)* $q(x,y) \geq \delta > 0$*; b)* $q(x,y)$ *is a continuous function on a set of variables in* R^2 *and*

$$\sup_{|x-y|\leq 1} \sup_{|C_1-C_2|\leq A |C_1|\leq A} \frac{q(x,C_1)}{q(y,C_2)} < \infty,$$

where A *is any finite value. Then: a) for any right side of* $f \in L_2(R^3)$ *there exists a solution* $u(x)$ *of the equation* (13) *such that* $\Delta u \in L_2(R^3)$*; b) let* $r(x)$ *is continuous function in* R^3*, if for any* $k\rangle 0$ *the value*

$$B = \sup_{x\in R} \sup_{|C_1|\leq K} \sup_{0<\eta\leq m^{-1}(x,C_1)} \left[\eta^{-p} \int\limits_{|t-x|<\eta} |r(t)|^{\theta} dt\right]^{1/\theta}$$

Is finite, then

$$r(x)D^2u(x) = L_{\theta}(R^3),$$

$$(2 \leq \theta < \infty, \quad p = -\frac{\theta}{2}, \quad m(x,C_1) = (q(x,C_1))^{1/2}.$$

Let us enter the function

$$q_{\varepsilon}^*(t,C_0) = \inf\left\{ d^{-1}; d \geq \inf_{e\in F_d^{(\varepsilon)}(t)} \int\limits_{\theta_d(t)|_e} q(x,C_0)dx \right\},$$

where $F_d^{(\varepsilon)}(t)$ *is a set of all compact subsets of cube* $\theta_d(t)$*, satisfying the following inequality*

$$mese \leq \varepsilon d^n, \quad \varepsilon \in (0,1).$$

Theorem 6 *Let the conditions a), b) of the theorem 5 be satisfied and*

$$\sup_{|x-y|\leq 1} \sup_{|C_0-C_1|\leq A} \frac{q_{\varepsilon}^*(x,C_0)}{q_{\varepsilon}^*(x,C_1)} < \infty,$$

Let us denote $m(x,C_0) = q_{\varepsilon}^*(x,C_0)$*, and by* $A_p(x,C_0)$ *– the function which is defined with the equality*

$$A_p(x,C_0) = m^{-1-\beta}(x,C_0) \sup_{|C_1|\leq K} \sup_{0<\eta<m^{-1}(x,C_1)} \eta^{-\beta} \int\limits_{|x-t|<\eta} q(t,C_1)dt,$$

where κ *is any value,* $\beta = 2(\frac{3}{p} - 1)$*,* p *– is any number from the interval (1,2). Then, if at some* $p \in (1,2)$ *the value*

$$A_p = \sup_{|C_0|\leq K} \sup_{x\in R^3} A_p(x,C_0)$$

Is finite, then for any $f(x) \in L_2(R^3)$ *there exists a solution* $u(x) \in L_2(R^3)$ *of the equation* (13)*, such that* $\Delta u \in L_2(R^3)$.

Theorems 5, 6 are proved in the same way as theorems 2-4, based on results of work [7].

3 Discussion

For differential equations one of the important questions is finding solutions in function spaces. In this paper, using operator methods, a sufficient condition for the existence of solutions to the nonlinear Sturm-Liouville and Schrodinger equations is found. Research methods and results can be used in the study of other nonlinear differential equations.

4 Conclusion

The issues of separability of operators and coercive estimates, and also the existence of a solution to differential equations, are solved in combination. The results of this work are new and generalize previously published works.

References

[1] Everitt W.H., Yiertz M., "Some propereties of certein operators" , *Proc. London Math. Soc.* 23(3) (1971): 301–304.

[2] Everitt W.N., Yiertz M., "Some ineqalies assocated with certein differential equations" , *Math. Z.* 126 (1972): 308–326.

[3] Muratbekov M.B., *Separability theorems and spectral properties of a class of differential operators with irregular coefficients. Abstract of the thesis. ... doctors of physical and mathematical sciences: 01.01.02* (Almaty, 1994): 30 (in russian).

[4] Birgebaev A., "Smoothness of solutions of the nonlinear Sturm-Liouville equation" , *Austrian Journal of Science and Technology* 1-2 (2015).

[5] Muratbekov M.B., Otelbayev M., "Smoothness and Approximative Properties of Solutions to a Class of Nonlinear Schrödinger Type Equations" , *Proceedings of universities. Series Mathematics* 3 (1999): 44–47. (in russian).

[6] Otelbayev M., "On the separability of elliptic operators" , *Reports of the Academy of Sciences of the USSR* 234 (3) (1977): 540–543 (in russian).

[7] Otelbayev M., "Coercive estimates and separability theorems for elliptic equations in R^m" , *Proceedings of MIAN* (1983) (in russian).

[8] Vladimirov B.C., *Equations of mathematical physics* (M.: Nauka, 1973) (in russian).

[9] Birgebaev A., *Smoothness of solutions of non-linear differential equations and separability theorems: dissertation ... candidate of physical and mathematical sciences: 01.01.02* (Almaty, 1984): 100 (in russian).

[10] Sobolev S.L., "Some applications of functional analysis in mathematical physics" , *Leningrad: LSU* (1952) (in russian).

[11] Otelbayev M., "On conditions for the self-adjointness of the Schr?dinger operator with the operator potential" , *Ukrainian Mathematical Journal* 280 (6) (1976) (in russian).

[12] Zayed M.E., Omran S.A., "Separation of the Tricomi Differential Operator in Hilbert Space with Application to the Existence and Uniqueness Theorem" , *International journal of Contemp Mathematical Sciences* 6 (8) (2011): 353–364.

[13] Berdyshev A.S., Birgebaev A.B., Cabada A., "On the smoothness of solutions of the third order nonlinear differential equation" , *Boundary value problems* (2017): 1–11. DOI 10.1186/s13661-017-0799-4 Impact factor 0,642 https://boundary valueproblems.springeropen.com/articles/10.1186/s13661-017-0799-4

IRSTI 27.39.21 DOI: https://doi.org/10.26577/JMMCS.2022.v115.i3.04

M.T. Jenaliyev[1], **A.S. Kassymbekova**[2], **M.G. Yergaliyev**[1,2*]
[1]Institute of Mathematics and Mathematical Modeling, Kazakhstan, Almaty
[2]Al-Farabi Kazakh National University, Kazakhstan, Almaty
*e-mail: ergaliev@math.kz

ON A BOUNDARY VALUE PROBLEM FOR A BOUSSINESQ-TYPE EQUATION IN A TRIANGLE

Earlier, we considered an initial-boundary value problem for a one-dimensional Boussinesq-type equation in a domain that is a trapezoid, in which the theorems on its unique weak solvability in Sobolev classes were established by the methods of the theory of monotone operators. In this article, we continue research in this direction and study the issues of correct formulation of the boundary value problem for a one-dimensional Boussinesq-type equation in a degenerate domain, which is a triangle. A scalar product is proposed with the help of which the monotonicity of the main operators is shown, and uniform a priori estimates are obtained. Further, using the methods of the theory of monotone operators and a priori estimates, theorems on its unique weak solvability in Sobolev classes are established. A theorem on increasing the smoothness of a weak solution is established. In proving the smoothness enhancement theorem, we use a generalization of the classical result on compactness in Banach spaces proved by Yu.I. Dubinsky ("Weak convergence in nonlinear elliptic and parabolic equations Sbornik: Mathematics, 67 (109): 4 (1965)) in the presence of a bounded set from a semi-normed space instead of a normed one. It is also shown that the solution may have a singularity at the point of degeneracy of the domain. The order of this feature is determined, and the corresponding theorem is proved.
Key words: Boussinesq equation, degenerating domain, a priori estimates, Sobolev space.

М.Т. Жиенәлиев[1], А.С. Қасымбекова[2], М.Ғ. Ерғалиев[1,2*]
[1]Математика және математикалық модельдеу институты, Қазақстан, Алматы қ.
[2]Әл-Фараби атындағы Қазақ Ұлттық Университеті, Қазақстан, Алматы қ.
*e-mail: ergaliev@math.kz

Үшбұрыштағы Буссинеск типтес теңдеуіне қойылған шекаралық есеп

Осыған дейін біз трапециялы облыстағы бірөлшемді Буссинеск типтес теңдеуі үшін қойылған бастапқы-шекаралық есепті қарастырдық. Есептің соболев кеңістіктеріндегі бірмәнді әлсіз шешімділігі туралы теоремалар монотонды операторлар теориясы әдісімен дәлелденді. Осы мақалада біз осы бағыттағы зерттеулерді жалғастырып, азғындалатын үшбұрышты облыстағы бірөлшемді Буссинеск типтес теңдеуі үшін қойылған шекаралық есептің қисынды қойылуын қарастырамыз. Негізгі операторлардың монотондылығын көрсету кезінде қолданылған скалярлы көбейтінді ұсынылып, бірқалыпты априорлы бағалаулар алынды. Әрі қарай монотонды операторлар теориясы және априорлы бағалаулар көмегімен есептің соболев кластарындағы бірмәнді әлсіз шешімділігі туралы теоремалар дәлелденді. Әлсіз шешімнің дифференциалдық қасиеттерін жақсартатын теорема дәлелденді. Шешімнің дифференциалдық қасиеттерін жақсартатын теореманы дәлелдеу кезінде біз банах кеңістіктеріндегі компактылық туралы классикалық нәтиженің нормаланған кеңістіктің орнына полунормаланған кеңістіктегі шенелген жиын бар болу жағдайының Ю.И. Дубинский ("Weak convergence in nonlinear elliptic and parabolic equations Sbornik: Mathematics, 67 (109): 4 (1965)) дәлелдеген жалпылауын пайдаландық. Бұған қоса облыстың азғындалу нүктесінде шешімнің ерекшігі бар екендігі көрсетілген. Осы ерекшеліктің реті анықталып, сәйкес теорема дәлелденді.
Түйін сөздер: Буссинеск теңдеуі, азғындалатын облыс, априорлы бағалаулар, Соболев кеңістігі.

М.Т. Дженалиев[1], А.С. Касымбекова[2], М.Г. Ергалиев[1,2*]
[1]Институт математики и математического моделирования, Казахстан, г. Алматы
[2]КазНУ имени аль-Фараби, Казахстан, г. Алматы
*e-mail: ergaliev@math.kz

Граничная задача для уравнения типа Буссинеска в треугольнике

Ранее нами была рассмотрена начально-граничная задача для одномерного уравнения типа Буссинеска в области, представляющей собой трапецию, в которой методами теории монотонных операторов установлены теоремы об её однозначной слабой разрешимости в соболевских классах. В этой статье мы продолжаем исследования в данном направлении и изучаем вопросы корректной постановки граничной задачи для одномерного уравнения типа Буссинеска в вырождающейся области, представляющей собой треугольник. Предложено скалярное произведение с помощью которого показана монотонность основных операторов, и получены равномерные априорные оценки. Далее методами теории монотонных операторов и априорных оценок установлены теоремы об её однозначной слабой разрешимости в соболевских классах. Установлена теорема о повышении гладкости слабого решения. При доказательстве теоремы о повышении гладкости мы используем обобщение классического результата о компактности в банаховых пространствах, доказанного Ю.И. Дубинским ("Weak convergence in nonlinear elliptic and parabolic equations Sbornik: Mathematics, 67 (109): 4 (1965)) при наличии ограниченного множества из полунормированного пространства вместо нормированного. Также показано, что решение может иметь особенность в точке вырождения области. Порядок данной особенности определен, и доказана соответствующая теорема.

Ключевые слова: уравнение Буссинеска, вырождающаяся область, априорные оценки, пространство Соболева.

Introduction

The theory of Boussinesq equations and its modifications always attracts the attention of both mathematicians and applied scientists. The Boussinesq equation, as well as their modifications, occupy an important place in describing the motion of liquid and gas, including in the theory of non-stationary filtration in porous media [1]– [13]. Additionally, here we note only the works [14]– [19]. In recent years, boundary value problems for these equations have been actively studied, since they model processes in porous media. These problems acquire particular importance for deep understanding and comprehension in the tasks of exploration and effective development of oil and gas fields.

In this paper, we study questions of the correct formulation of boundary value problems for a one-dimensional Boussinesq-type equation in a degenerating domain. The domain is represented by a triangle. Using the method of monotone operators, we prove theorems on the unique weak solvability of the considered boundary value problems, and also establish a theorem on improving the smoothness of a weak solution.

1 Statement of the boundary value problem and the main result

Let $\Omega_t = \{0 < x < t\}$ and $\partial\Omega_t$ be the boundary of the Ω_t, $0 < t < T < \infty$. In the domain $Q_{xt} = \{x, t |\, x \in \Omega_t,\ t \in (0, T)\}$, which is a triangle, we consider the following boundary value problem for a Boussinesq-type equation

$$\partial_t u - \partial_x \left(|u|\partial_x u\right) = f, \quad \{x, t\} \in Q_{xt}, \tag{1}$$

with boundary conditions

$$u = 0, \ \ \{x, t\} \in \Sigma_{xt} = \partial\Omega_t \times (0, T), \tag{2}$$

where $f(x, t)$ is a given function.

It can be directly shown that the nonlinear operator $A_0(v) = -\partial_x (|v|\partial_x v)$ of boundary value problem (1)–(2) has the following properties:

$$A_0(v) : L_3(\Omega_t) \to L_{3/2}(\Omega_t) \text{ is a hemicontinuous operator,} \tag{3}$$

$$\|A_0(v)\|_{L_{3/2}(\Omega_t)} \le c\|v\|^2_{L_3(\Omega_t)}, \ \ c > 0, \ \ \forall\, v \in L_3(\Omega_t), \tag{4}$$

$$\langle A_0(v), v\rangle \ge \alpha\|v\|^3_{L_3(\Omega_t)}, \ \ \alpha > 0, \ \ \forall\, v \in L_3(\Omega_t). \tag{5}$$

We have established the following theorems.

Theorem 1 (Main result) *Let*

$$f \in L_{3/2}((0, T); W^{-1}_{3/2}(\Omega_t)). \tag{6}$$

Then boundary value problem (1)–(2) *has a unique solution*

$$u \in L_3((0, T); L_3(\Omega_t)) \cap L_\infty((0, T); H^{-1}(\Omega_t)), \tag{7}$$

moreover, at $x \to 0+, \ x \to t - 0, \ t \to 0+$ *we have*

$$\begin{cases} u(x, t) = \mathcal{O}\left(x^{-\alpha_0}(t - x)^{-\alpha+\alpha_0}t^{-\beta}\right), \\ 0 < \alpha < \frac{1}{3}, \ \ \beta > 0, \ \ \alpha + \beta < \frac{2}{3}, \ \ 0 \le \alpha_0 \le \alpha. \end{cases} \tag{8}$$

Theorem 2 (On smoothness) *Let*

$$f \in L_{3/2}((0, T); L_{3/2}(\Omega_t)). \tag{9}$$

Then the solution of boundary value problem (1)–(2) *admits additional smoothness, i.e.,*

$$u \in L_\infty((0, T); L_2(\Omega_t)), \tag{10}$$

$$|u|^{1/2}u \in L_2((0, T); H^1_0(\Omega_t)), \tag{11}$$

$$\partial_t u \in L_{3/2}((0, T); W^{-1}_{3/2}(\Omega_t)). \tag{12}$$

2 Auxiliary initial boundary value problems in trapezoids

To prove Theorem 1, we first consider auxiliary initial boundary value problems. Let $\Omega_t = \{0 < x < t\}$ and $\partial\Omega_t$ be the domain of the Ω_t, $\varepsilon_m < t < T < \infty$, $\varepsilon_1 > \varepsilon_2 > ... > \varepsilon_m > ...$, $\varepsilon_m \to 0$ at $m \to \infty$. In the domain $Q_{xt}^m = \{x, t | \, x \in \Omega_t, \ t \in (\varepsilon_m, T)\}$, which is a trapezoid, we consider the following boundary value problems for a Boussinesq-type equation

$$\partial_t u_m - \partial_x \left(|u_m| \partial_x u_m\right) = f_m, \ \ \{x, t\} \in Q_{xt}^m, \tag{13}$$

with boundary

$$u_m = 0, \ \ \{x, t\} \in \Sigma_{xt}^m = \partial\Omega_t \times (\varepsilon_m, T), \tag{14}$$

and initial conditions

$$u_m = 0, \ \ x \in \Omega_{\varepsilon_m} = (0, \varepsilon_m), \tag{15}$$

where $f_m(x,t)$ are the narrowing of function $f(x,t)$ (6), which is given in the triangle Q_{xt}, into trapezoids Q_{xt}^m.

Earlier, in [1]– [2], we established the following theorems.

Theorem 3 *Let*

$$f_m \in L_{3/2}((\varepsilon_m, T); W_{3/2}^{-1}(\Omega_t)). \tag{16}$$

Then initial boundary value problem (13)–(15) *has a unique solution*

$$u_m \in L_3((\varepsilon_m, T); L_3(\Omega_t)) \cap L_\infty((\varepsilon_m, T); H^{-1}(\Omega_t)). \tag{17}$$

Theorem 4 *Let*

$$f_m \in L_{3/2}((\varepsilon_m, T); L_{3/2}(\Omega_t)). \tag{18}$$

Then the solution of initial boundary value problem (13)–(15) *admits additional smoothness, i.e.,*

$$u_m \in L_\infty((\varepsilon_m, T); L_2(\Omega_t)), \tag{19}$$

$$|u_m|^{1/2} u_m \in L_2((\varepsilon_m, T); H_0^1(\Omega_t)), \tag{20}$$

$$\partial_t u_m \in L_2((\varepsilon_m, T); W_{3/2}^{-1}(\Omega_t)). \tag{21}$$

Note that results similar to Theorem 4 for cylindrical domains are also available in [21]–[22].

3 Proof of Theorem 1. Existence

First of all, for each m and the corresponding given function $f_m(x,t)$, according to the statement of Theorem 3, we have established the existence of a unique solution $u_m(x,t)$ of initial boundary value problem (13)–(15).

We continue functions $u_m(x,t)$, $f_m(x,t)$ from the trapezoid Q^m_{xt} by zero to the entire triangle Q_{xt} and denote them by $\tilde{u}_m(x,t)$, $\tilde{f}_m(x,t)$. These functions will satisfy equations

$$\partial_t \tilde{u}_m - \partial_x \left(|\tilde{u}_m| \partial_x \tilde{u}_m \right) = \tilde{f}_m, \;\; \{x,t\} \in Q_{xt}, \tag{22}$$

with boundary conditions

$$\tilde{u}_m = 0, \;\; \{x,t\} \in \Sigma_{xt}. \tag{23}$$

From (22) we obtain

$$\langle \partial_t \tilde{u}_m(t), v \rangle + a_0(t, \tilde{u}_m(t), v) = \langle \tilde{f}_m(t), v \rangle, \;\; \forall\, v \in H^{-1}(\Omega_t), \;\; t \in (0,T), \tag{24}$$

where $a_0(t,\tilde{u}_m,v) = \langle A_0(t,\tilde{u}_m), v\rangle$, $A_0(t,\tilde{u}_m) = -\partial_x \left(|\tilde{u}_m| \partial_x \tilde{u}_m \right)$ and $\langle \cdot,\cdot \rangle$ is the scalar product defined by formula

$$\langle \varphi, \psi \rangle = \int_{\Omega_t} \varphi \left[\left(-d^{\,2}_x \right)^{-1} \psi \right] dx, \;\; \forall\, \varphi, \psi \in H^{-1}(\Omega_t), \;\; t \in (\varepsilon_m, T), \tag{25}$$

where $d^{\,2}_x = \frac{d^2}{dx^2}$, $\tilde{\psi} = \left(-d^{\,2}_x \right)^{-1} \psi : \; -d^{\,2}_x \tilde{\psi} = \psi$, $\tilde{\psi}(0) = \tilde{\psi}(t) = 0, \forall\, \psi \in H^{-1}(\Omega_t)$.

Note that concepts close to scalar product (25) have already been used in works [21], [22].

The operator $A_0(t,\tilde{u}_m)$ has the monotonicity property in accordance with scalar product (25). For solutions $\{\tilde{u}_m(t)\}_{m=1}^{\infty}$, we establish a priori estimates that are uniform in the index m. From (22)–(25) we will have:

$$\frac{1}{2} \|\tilde{u}_m(t)\|^2_{H^{-1}(\Omega_t)} + \alpha \int_0^t \|\tilde{u}_m(\tau)\|^3_{L_3(\Omega_t)} d\tau \le \int_0^t \|\tilde{f}_m(\tau)\|_{L_{3/2}(\Omega_t)} \|\tilde{u}_m(\tau)\|_{L_3(\Omega_t)} d\tau \le$$

$$\le \frac{2}{3} \sqrt{\frac{2}{3\,\alpha}} \int_0^t \|\tilde{f}_m(\tau)\|^{3/2}_{L_{3/2}(\Omega_t)} d\tau + \frac{\alpha}{2} \int_0^t \|\tilde{u}_m(\tau)\|^3_{L_3(\Omega_t)} d\tau \le$$

$$\le \frac{2}{3} \sqrt{\frac{2}{3\,\alpha}} \int_0^T \|f(t)\|^{3/2}_{L_{3/2}(\Omega_t)} dt + \frac{\alpha}{2} \int_0^t \|\tilde{u}_m(\tau)\|^3_{L_3(\Omega_t)} d\tau. \tag{26}$$

From here we get

$$\|\tilde{u}_m(t)\|^2_{H^{-1}(\Omega_t)} + \alpha \int_0^t \|\tilde{u}_m(\tau)\|^3_{L_3(\Omega_t)} d\tau \le \frac{4}{3} \sqrt{\frac{2}{3\,\alpha}} \|f(t)\|^{3/2}_{L_{3/2}(Q_{xt})}, \;\; t \in (0,T]. \tag{27}$$

In (26) we used the following relations

$$\frac{1}{2} \frac{d}{dt} \|\tilde{u}_m(t)\|^2_{H^{-1}(\Omega_t)} = \langle \tilde{u}'_m(t), \tilde{u}_m(t) \rangle, \;\; \text{since} \;\; \tilde{u}_m(t) \equiv 0 \text{ on } \Sigma_{xt},$$

$$\tilde{f}_m(t)\|_{L_{3/2}(\Omega_t)} \leq \|f(t)\|_{L_{3/2}(\Omega_t)},$$

as well as Young's inequality $(p^{-1}+q^{-1}=1)$:

$$|DE| = \left|\left(d^{1/p}D\right)\left(d^{1/q}\frac{E}{d}\right)\right| \leq \frac{d}{p}\,|D|^p + \frac{d}{qd^q}\,|E|^q\,,$$

where

$$D = \|w_m(t)\|_{L_{3/2}(\Omega)}\,,\ \ E = \|w_m(t)\|_{L_3(\Omega)}\,,\ \ d = \sqrt{\frac{2}{3\,\alpha}},\ \ p=3/2,\ \ q=3.$$

Finally, the relations

$$\tilde{u}_\mu \to u\ \ \star-\text{weak in}\ \ L_\infty((0,T);H^{-1}(\Omega_t)), \tag{28}$$

$$\tilde{u}_\mu \to u\ \text{weak in}\ \ L_3(Q_{xt}), \tag{29}$$

$$\tilde{u}_\mu(T) \to \eta\ \text{weak in}\ \ H^{-1}(\Omega_T), \tag{30}$$

$$A_0(t,\tilde{u}_\mu) \to h(t)\ \text{weak in}\ \ L_{3/2}((0,T);L_{3/2}(\Omega_t)). \tag{31}$$

follow from (27) and inequality

$$\|A_0(t,\tilde{u}_\mu)\|_{L_{3/2}(\Omega_t)} \leq c\|\tilde{u}_\mu\|^2_{L_3(\Omega_t)}.$$

Now we continue functions $\tilde{u}_m(t),\ A_0(t,\tilde{u}_m(t)),...,$ from domain Q_{xt} by zero to the infinite domain $\bar{Q}_{xt}$, where

$$\bar{Q}_{xt} = \begin{cases} x=0, & t\leq 0,\\ x\in\Omega_t, & t\in(0,T],\\ x\in\Omega_T, & t>T; \end{cases}$$

and denote these continuations by $\bar{\tilde{u}}_m(t),\ \bar{A}_0(t,\bar{\tilde{u}}_m(t)),...,$ i.e.,

$$\bar{\tilde{u}}_m(t) = \begin{cases} 0, & t\leq 0,\\ \tilde{u}_m(t)\in H^{-1}(\Omega_t), & t\in(0,T],\\ 0, & t>T; \end{cases} \qquad \bar{v}(t) = \begin{cases} 0, & t\leq 0,\\ v(t)\in H^{-1}(\Omega_t), & t\in(0,T],\\ 0, & t>T. \end{cases} \tag{32}$$

As a result, for continuations (32) we will have:

$$\langle\bar{\tilde{u}}'_m(t),\bar{v}(t)\rangle + \langle\bar{A}_0(t,\bar{\tilde{u}}_m(t)),\bar{v}(t)\rangle = \langle\bar{\tilde{f}}_m(t),\bar{v}(t)\rangle - \langle\tilde{u}_m(T),\bar{v}(t)\rangle\delta(t-T),\ \ t\in R^1. \tag{33}$$

Further, choosing from $\{\bar{\tilde{u}}_m(t)\}_{m=1}^\infty$ a weakly convergent subsequence $\{\bar{\tilde{u}}_\mu(t)\}_{\mu=1}^\infty$ and passing to the limit at $\mu\to\infty$, we obtain

$$\langle\bar{u}'(t),\bar{v}(t)\rangle + \langle\bar{h}(t),\bar{v}(t)\rangle = \langle\bar{f}(t),\bar{v}(t)\rangle - \langle\eta,\bar{v}(t)\rangle\delta(t-T),\ \ t\in R^1,$$

where $\bar{u}(t)$, $\bar{h}(t)$ and $\bar{f}(t)$ are continuations of functions $u(t)$ (28), $h(t)$ (31) and $f(t)$ to R^1, that is, from here we get

$$\bar{u}'(t) + \bar{h}(t) = \bar{f}(t) - \eta\delta(t-T), \;\; t \in R^1. \tag{34}$$

Now, narrowing equality (34) to the time interval $(0,T)$, we obtain

$$u'(t) + h(t) = f(t), \;\; t \in (0,T), \tag{35}$$

$$u'(t) \in L_{3/2}((0,T); L_{3/2}(\Omega_t)). \tag{36}$$

Further, on the one hand, from the monotonicity condition of the operator $A_0(t,v)$ we will have

$$Y_\mu \equiv \int_0^T \langle A_0(t,\tilde{u}_\mu(t)) - A_0(t,v(t)), \tilde{u}_\mu(t) - v(t)\rangle\, dt \geq 0 \;\; \forall\, v \in L_3((0,T); L_3(\Omega_t)), \tag{37}$$

on the other hand, from (24) we get

$$\int_0^T \langle A_0(t,\tilde{u}_\mu(t)), \tilde{u}_\mu(t)\rangle\, dt = \int_0^T \langle \tilde{f}_\mu(t)), \tilde{u}_\mu(t)\rangle\, dt - \frac{1}{2}\|\tilde{u}_\mu(T)\|^2_{H^{-1}(\Omega_T)}. \tag{38}$$

Thus, it follows from relations (37)–(38) that

$$Y_\mu \equiv \int_0^T \langle \tilde{f}_\mu(t), \tilde{u}_\mu(t)\rangle\, dt - \frac{1}{2}\|\tilde{u}_\mu(T)\|^2_{H^{-1}(\Omega_T)} - \int_0^T \langle A_0(t,\tilde{u}_\mu(t)), v(t)\rangle\, dt -$$

$$- \int_0^T \langle A_0(t,v(t)), \tilde{u}_\mu(t) - v(t)\rangle\, dt \;\; \forall\, v \in L_3((0,T); L_3(\Omega_t)). \tag{39}$$

Now, using the property of weak lower semicontinuity of the norm in a Banach space

$$\liminf \|\tilde{u}_\mu(T)\|^2_{H^{-1}(\Omega_T)} \geq \|\tilde{u}(T)\|^2_{H^{-1}(\Omega_T)},$$

we have

$$0 \leq \limsup Y_\mu \leq \int_0^T \langle f(t), u(t)\rangle\, dt - \frac{1}{2}\|u(T)\|^2_{H^{-1}(\Omega_T)} - \int_0^T \langle h(t), v(t)\rangle\, dt -$$

$$- \int_0^T \langle A_0(t,v(t)), u(t) - v(t)\rangle\, dt \;\; \forall\, v \in L_3((0,T); L_3(\Omega_t)). \tag{40}$$

In turn, from (35) we get

$$\int_0^T \langle f(t), u(t)\rangle\, dt = \int_0^T \langle h(t), u(t)\rangle\, dt + \frac{1}{2}\|u(T)\|^2_{H^{-1}(\Omega_T)}. \tag{41}$$

Substituting the expression for $\int_0^T \langle f(t), u(t)\rangle \, dt$ from (41) into inequality (40), we establish the following inequality

$$\int_0^T \langle h(t) - A_0(t, v(t)), u(t) - v(t)\rangle \, dt \geq 0 \ \ \forall \, v(t) \in L_3((0,T); L_3(\Omega_t)). \tag{42}$$

Now, to complete the proof of Theorem 1, i.e. the existence of a solution to boundary value problem (1)–(2), our goal is: to show the validity of the following equality

$$h(t) = A_0(u(t)). \tag{43}$$

We use the property of hemicontinuity of the operator $A_0(t, v)$ (3). Replacing $v(t) = u(t) - \lambda w(t), \ \ \lambda > 0, \ \ w \in L_3(Q_{xt})$ in (42), we obtain

$$\int_0^T \langle h(t) - A_0(t, u(t) - \lambda w(t)), w(t)\rangle \, dt \geq 0 \ \ \forall \, w(t) \in L_3(Q_{xt}).$$

Hence, at $\lambda \to 0+$, we obtain the required equality (43). The existence part of the solution in Theorem 1 is proved.

4 Proof of the Theorem 1. Uniqueness

Let us show that the operator $A_0(t, u)$ in problem (1)–(2) will have the property of monotonicity if the scalar product is introduced in an appropriate way. For this purpose, we take as the scalar product

$$\langle \varphi, \psi\rangle = \int_{\Omega_t} \varphi \left[\left(-d_x^{\,2}\right)^{-1} \psi \right] dy, \ \ \forall \, \varphi, \psi \in H^{-1}(\Omega_t), \ \ \forall \, t \in (0,T), \tag{44}$$

where $d_x^{\,2} = \frac{d^2}{dx^2}$, $\tilde{\psi} = \left(-d_x^{\,2}\right)^{-1} \psi \ : \ -d_x^{\,2} \tilde{\psi} = \psi$, $\tilde{\psi}(0) = \tilde{\psi}(t) = 0, \forall \, \psi \in H^{-1}(\Omega_t), \ \forall \, t \in (0,T)$.

The following lemma is valid.

Lemma 1 *Operator* $A_0(t, u)$ *is monotone in the sense of scalar product* (44) *in space* $H^{-1}(\Omega_t)$, *i.e. the following inequality is valid:*

$$\langle A_0(t, u_1) - A_0(t, u_2), u_1 - u_2\rangle \geq 0, \ \forall \, u_1, u_2 \in \mathcal{D}(\Omega_t), \ \ \forall \, t \in (0,T). \tag{45}$$

To the proof of Lemma 1 For each $t \in (0,T)$, operator $A_0(t, u) = -\partial_x \left(|u| \partial_x u\right)$ is monotonic and condition (45) is satisfied (according to [20], chap. 2, s. 3.1). Indeed, on the one hand, we have

$$\langle A_0(t, \varphi) - A_0(t, \psi), \varphi - \psi\rangle = \frac{1}{2} \int_{\Omega_t} \left(-d_x^2\right) \left(|\varphi|\varphi - |\psi|\psi\right) \left(-d_x^2\right)^{-1} (\varphi - \psi) \, d\, x =$$

$$= \frac{1}{2} \int_{\Omega_t} (|\varphi|\varphi - |\psi|\psi)(\varphi - \psi) \, d\, x, \ \ \forall \, \varphi, \, \psi \in \mathcal{D}(\Omega_t), \ \ \forall \, t \in (t_0, T).$$

On the other hand, from the convexity condition of the functional $J(t,\varphi) = \frac{1}{3}\int_{\Omega_t}|\varphi(x)|^3\,dx,\ \ \varphi\in\mathcal{D}(\Omega_t),\ \ \forall\, t\in(0,T)$, it follows

$$\langle J'(t,\varphi)-J'(t,\psi),\varphi-\psi\rangle \ge 0,\ \ \forall\,\varphi,\,\psi\in\mathcal{D}(\Omega_t),\ \ \forall\, t\in(0,T).$$

Thus, we get

$$\int\limits_{\Omega_t}(|\varphi|\varphi-|\psi|\psi)(\varphi-\psi)\,dx\ \ge 0,\ \ \forall\,\varphi,\,\psi\in\mathcal{D}(\Omega_t),\ \ \forall\, t\in(0,T),$$

that is, inequality (45) is established. Lemma 1 is proved.

Now we are ready to show the uniqueness of the solution in problem (1)–(2). Let $u_1(t)$ and $u_2(t)$ be two solutions to problem (1)–(2). Then their difference $u(t) = u_1(t)-u_2(t)$ satisfies the homogeneous problem:

$$u'(t)+A_0(t,u_1(t))-A_0(t,u_2(t))=0,$$
$$\langle u'(t),u(t)\rangle+\langle (A_0(t,u_1(t))-A_0(t,u_2(t)),u_1(t)-u_2(t)\rangle=0$$

and, due to the monotonicity property of the operator $A_0(t,u)$, we have:

$$\langle u'(t),u(t)\rangle=\frac{d}{2\,dt}\|u(t)\|^2_{H^{-1}(\Omega_t)}\le 0,\ \ \text{i.e.}\ \ u(t)\equiv 0.$$

The uniqueness of the solution to problem (1)–(2) is proved.

5 Proof of Theorem 1. Singularity of the solution

We show that the solution $u(x,t)$ of boundary value problem (1)–(2) having a singularity of the order specified in (8) will belong to the space $L_3(Q^{t_0}_{xt})$, where $Q^{t_0}_{xt}=\{x,t|\,0<x<t,\ 0<t<t_0\ll T\}$. For this purpose, it suffices to show that the following integral is bounded when $t_0\to 0+$:

$$\int_{Q^{t_0}_{xt}} x^{-3\alpha_0}(t-x)^{-3\alpha+3\alpha_0}t^{-3\beta}\,dx\,dt. \tag{46}$$

We have

$$\int_0^{t_0} t^{-3\beta}\int_0^t x^{-3\alpha_0}(t-x)^{-3\alpha+3\alpha_0}\,dx\,dt = \left\|\begin{array}{l} x=t\sin^2\theta\\ 0<\theta<\pi/2\\ dx=2\sin\theta\cos\theta\,d\theta\end{array}\right\| =$$

$$=2\int_0^{t_0} t^{1-3\alpha-3\beta}\int_0^{\pi/2}\sin^{1-6\alpha_0}\theta\cos^{1-6\alpha+6\alpha_0}\theta\,d\theta\,dt.$$

It is not difficult to verify that under the conditions of Theorem 1 in the last expression, the inner integral takes a finite value. Calculating the outer integral, we have

$$\int_0^{t_0} t^{1-3\alpha-3\beta}\,dt=\frac{1}{2-3(\alpha+\beta)}t_0^{2-3(\alpha+\beta)},$$

which, under the conditions of Theorem 1, is also bounded from above.

Note that if the order of the singularity of solution $u(x,t)$ is higher than in (8), then this function is no longer an element of space $L_3(Q^{t_0}_{xt})$.

This completes the proof of Theorem 1.

6 Proof of Theorem 2

It suffice for us to show the existence of a solution, and the uniqueness follows from Theorem 1.

First of all, for each m and the corresponding given function $f_m(x,t)$, according to the statement of Theorem 4 we have established the existence of a smoother (than in Theorem 3) unique solution $u_m(x,t)$ of initial boundary value problem (13)–(15) for the corresponding trapezoid Q_{xt}^m.

We continue functions $u_m(x,t),\ f_m(x,t)$ from the trapezoid Q_{xt}^m by zero to the entire triangle Q_{xt} and denote them by $\tilde{u}_m(x,t),\ \tilde{f}_m(x,t)$. These functions will satisfy equations

$$\partial_t \tilde{u}_m - \partial_x\left(|\tilde{u}_m|\partial_x \tilde{u}_m\right) = \tilde{f}_m,\ \ \{x,t\}\in Q_{xt}, \tag{47}$$

with boundary conditions

$$\tilde{u}_m = 0,\ \ \{x,t\}\in \Sigma_{xt}. \tag{48}$$

From (47) we obtain

$$\langle \partial_t \tilde{u}_m(t), v\rangle + a_0(t,\tilde{u}_m(t),v) = \langle \tilde{f}_m(t), v\rangle,\ \ \forall\, v\in H^{-1}(\Omega_t),\ \ t\in(0,T), \tag{49}$$

where $a_0(t,\tilde{u}_m,v) = \langle A_0(t,\tilde{u}_m),v\rangle$, $A_0(t,\tilde{u}_m) = -\partial_x\left(|\tilde{u}_m|\partial_x\tilde{u}_m\right)$ and $\langle\cdot,\cdot\rangle$ is a scalar product

$$\langle\varphi,\psi\rangle = \int_{\Omega_t}\varphi\left[\left(-d_x^{\,2}\right)^{-1}\psi\right]dx,\ \ \forall\,\varphi,\psi\in H^{-1}(\Omega_t),\ \ t\in(\varepsilon_m,T),$$

where $d_x^{\,2} = \frac{d^2}{dx^2}$, $\tilde{\psi} = \left(-d_x^{\,2}\right)^{-1}\psi:\ -d_x^{\,2}\tilde{\psi} = \psi,\ \tilde{\psi}(0)=\tilde{\psi}(t)=0, \forall\,\psi\in H^{-1}(\Omega_t)$.

Let us rewrite equation (49) in the form

$$\left(\partial_t\tilde{u}_m(t), \left(-\partial_x^{\,2}\right)^{-1}v\right) + \frac{1}{2}\left(|\tilde{u}_m(t)|\,\tilde{u}_m(t), v\right) = \left(\tilde{f}_m(t), \left(-\partial_x^{\,2}\right)^{-1}v\right),\ \ \forall\, v\in H^1_{0,\Delta}(\Omega_t),\ \ t\in(0,T),$$

where $H^1_{0,\Delta}(\Omega_t) = \{\varphi|\,\varphi,\ \partial_x^2\varphi\in H_0^1(\Omega_t)\}$, or

$$(\partial_t\tilde{u}_m(t),\tilde{v}) + \frac{1}{2}\left(|\tilde{u}_m(t)|\,\tilde{u}_m(t), v\right) = \left(\tilde{f}_m(t),\tilde{v}\right),\ \ \forall\,\tilde{v} = \left(-\partial_x^{\,2}\right)^{-1}v\in H_0^1(\Omega_t),\ \ t\in(0,T). \tag{50}$$

Further, from (50) we obtain the following equality

$$\langle\partial_t\tilde{u}_m(t),\tilde{u}_m(t)\rangle + \frac{1}{2}\left(|\tilde{u}_m(t)|\,\tilde{u}_m(t), -\partial_x^2\tilde{u}_m(t)\right) = \langle\tilde{f}_m(t),\tilde{u}_m(t)\rangle,\ \ t\in(0,T), \tag{51}$$

and from (51), therefore, we will have

$$\frac{1}{2}\frac{d}{dt}\|\tilde{u}_m(t)\|^2_{L_2(\Omega_t)} + \frac{4}{9}\int_{\Omega_t}\left[\partial_x\left(|\tilde{u}_m(t)|^{1/2}\,\tilde{u}_m(t)\right)\right]^2dx = \langle\tilde{f}_m(t),\tilde{u}_m(t)\rangle,\ \ t\in(0,T),$$

or

$$\frac{1}{2}\|\tilde{u}_m(t)\|^2_{L_2(\Omega_t)} + \frac{4}{9}\int_0^t\int_{\Omega_\tau}\left[\partial_x\left(|\tilde{u}_m(\tau)|^{1/2}\,\tilde{u}_m(\tau)\right)\right]^2dx\,d\tau = \int_0^t\langle\tilde{f}_m(\tau),\tilde{u}_m(\tau)\rangle\,d\tau,\ \ t\in(0,T).$$

(52)

Here we use the following equality

$$-\frac{1}{2}\int_{\Omega_t}|\tilde{u}_m(t)|\tilde{u}_m(t)\partial_x^2\tilde{u}_m(t)\,dx = \frac{4}{9}\int_{\Omega_t}\left[\partial_x\left(|\tilde{u}_m(t)|^{1/2}\tilde{u}_m(t)\right)\right]^2 dx,\ \ t\in(0,T). \tag{53}$$

Let us show its justice. First, we transform the left side of equality (53). Let us show that equality

$$-\frac{1}{2}\int_{\Omega_t}|\tilde{u}_m(t)|\tilde{u}_m(t)\partial_x^2\tilde{u}_m(t)\,dx = \int_{\Omega_t}|\tilde{u}_m(t)|\left[\partial_x\tilde{u}_m(t)\right]^2\,dx, \tag{54}$$

holds. Indeed, we have:

$$|\tilde{u}_m|\tilde{u}_m = \begin{cases}[\tilde{u}_m]^2, & \text{at } \tilde{u}_m>0,\\ 0, & \text{at } \tilde{u}_m(t)=0,\\ -[-\tilde{u}_m]^2, & \text{at } \tilde{u}_m<0,\end{cases}\quad \partial_x\left(|\tilde{u}_m|\tilde{u}_m\right) = \begin{cases}2\,\tilde{u}_m\partial_x\tilde{u}_m, & \text{at } \tilde{u}_m>0,\\ 0, & \text{at } \tilde{u}_m(t)=0,\\ 2\,[-\tilde{u}_m]\partial_x\tilde{u}_m, & \text{at } \tilde{u}_m<0.\end{cases}$$

Thus, from here we obtain: $\partial_x\left(|\tilde{u}_m(t)|\tilde{u}_m(t)\right) = 2\,|\tilde{u}_m(t)|\partial_x\tilde{u}_m(t)$, i.e. equality (54).

The same holds for the right side of equality (53). We get

$$|\tilde{u}_m|^{1/2}\tilde{u}_m = \begin{cases}[\tilde{u}_m]^{3/2}, & \text{at } \tilde{u}_m>0,\\ 0, & \text{at } \tilde{u}_m=0,\\ -[-\tilde{u}_m]^{3/2}, & \text{at } \tilde{u}_m<0,\end{cases}\quad \partial_x\left(|\tilde{u}_m|^{1/2}\tilde{u}_m\right) = \begin{cases}\frac{3}{2}[\tilde{u}_m]^{1/2}\partial_x\tilde{u}_m, & \text{at } \tilde{u}_m>0,\\ 0, & \text{at } \tilde{u}_m=0,\\ \frac{3}{2}[-\tilde{u}_m]^{1/2}\partial_x\tilde{u}_m, & \text{at } \tilde{u}_m<0.\end{cases}$$

Thus, from here we get: $\partial_x\left(|\tilde{u}_m(t)|^{1/2}\tilde{u}_m(t)\right) = \frac{3}{2}|\tilde{u}_m(t)|^{1/2}\partial_x\tilde{u}_m(t)$, that is, the following equality is true:

$$\frac{4}{9}\int_{\Omega_t}\left[\partial_x\left(|\tilde{u}_m(t)|^{1/2}\tilde{u}_m(t)\right)\right]^2 dx = \int_{\Omega_t}|\tilde{u}_m(t)|\left[\partial_x\tilde{u}_m(t)\right]^2 dx.$$

Thus, we have shown the validity of equality (53).

Since from Theorem 3 we have that the functions $\tilde{u}_m(t)$ are bounded in $L_3(Q_{xt})$, therefore the right part of (52) is bounded when condition (6) of Theorem 1 is fulfilled. Hence from (52) we deduce that

$$\tilde{u}_m\ \text{ are bounded in }\ L_\infty((0,T);L_2(\Omega_t)), \tag{55}$$

$$\partial_x\left(|\tilde{u}_m|\,\tilde{u}_m\right)\ \text{ are bounded in }\ L_2(Q_{xt}),\ \text{ i.e. }\ |\tilde{u}_m|^{1/2}\,\tilde{u}_m\in L_2((0,T);H_0^1(\Omega_t)). \tag{56}$$

From relations (55)–(56), equation (47) and conditions (4), (18) we establish an estimate for the time derivative t

$$\partial_t\tilde{u}_m\ \text{ are bounded in }\ L_{3/2}((0,T);W_{3/2}^{-1}(\Omega_t)). \tag{57}$$

Hence, we can write

$$\tilde{u}_m \to u \text{ weakly in } L_\infty((0,T);L_2(\Omega_t)), \tag{58}$$

$$|\tilde{u}_m|^{1/2}\tilde{u}_m \to \chi \text{ weakly in } L_2((0,T);H_0^1(\Omega_t)). \tag{59}$$

Thus, on the basis of relations (57)–(59) we establish

$$\tilde{u}_m \to u \text{ strongly in } L_3((0,T);L_3(\Omega_t)) \text{ and almost everywhere,}$$

and, further, using (56) and applying Theorem 12.1 and Proposition 12.1 from ([20], chapter 1, 12.2), as well as Lemma 1.3 from ([20], chapter 1, 1.4), as a result we have

$$|\tilde{u}_m|^{1/2}\tilde{u}_m \to |u|^{1/2}u \text{ weakly in } L_2((0,T);H_0^1(\Omega_t)), \text{ i.e. } \chi = |u|^{1/2}u. \tag{60}$$

Lemma 1.3 ([20], chapter 1, 1.4). *Let $\mathcal{O}$ is a bounded domain in $\mathbb{R}_x^n \times \mathbb{R}_t^1$, g_μ and g are functions from $L_q(\mathcal{O})$, $1 < q < \infty$, such that*

$$\|g_\mu\|_{L_q(\mathcal{O})} \leq C, \ \ g_\mu \to g \ \textit{ a.e. in } \mathcal{O}.$$

Then $g_\mu \to g$ weakly in $L_q(\mathcal{O})$.

From (57), (59) and (60) we obtain the required statement (10)–(12). Theorem 2 is completely proved.

Conclusion

In this paper, we study boundary problems for a one-dimensional Boussinesq-type equation in a domain that is a triangle. Using the methods of the theory of monotone operators and a priori estimates, we prove theorems on their unique weak solvability in Sobolev classes, as well as theorems on improving the smoothness of a weak solution.

Acknowledgments

This research is funded by the Science Committee of the Ministry of Education and Science of the Republic of Kazakhstan (Grants No. AP09258892, 2021-2023).

References

[1] M.T. Jenaliyev, A.S. Kasymbekova, M.G. Yergaliyev, "On initial boundary value problems for the Boussinesq-type equation", *The Tradit. Int. April Math. Conf. in honor of the Day of Science workers of the RK. Abstracts of reports.* (Almaty: Publ. IMMM. (2022), 76–77).

[2] M.T. Jenaliyev, A.S. Kasymbekova, M.G. Yergaliyev, A.A. Assetov, "An Initial Boundary Value Problem for the Boussinesq Equation in a Trapezoid", *Bulletin of the Karaganda University. Mathematics,* 106: 2 (2022), 11p.

[3] H. P. McKean, "Boussinesq's Equation on the Circle", *Communications on Pure and Applied Mathematics,* XXXIV (1981): 599–691.

[4] Z. Y. Yan, F.D. Xie, H.Q. Zhang, "Symmetry Reductions, Integrability and Solitary Wave Solutions to Higher-Order Modified Boussinesq Equations with Damping Term", *Communications in Theoretical Physics,* 36: 1 (2001): 1–6.

[5] V. F. Baklanovskaya, A. N. Gaipova, :On a two-dimensional problem of nonlinear filtration", *Zh. vychisl. math. i math. phiz.,* 6: 4 (1966): 237–241 (in Russian).

[6] J. L. Vazquez, *The Porous Medium Equation. Mathematical Theory,* (Oxford University Press, Oxford (2007). XXII+625p).

[7] P. Ya. Polubarinova-Kochina, "On a nonlinear differential equation encountered in the theory of infiltration", *Dokl. Akad. Nauk SSSR,* 63: 6 (1948): 623–627.

[8] P.Ya. Polubarinova-Kochina, *Theory of Groundwater Movement,* (Princeton Univ. Press, Princeton (1962)).

[9] Ya. B. Zel'dovich and A. S. Kompaneets, *Towards a theory of heat conduction with thermal conductivity depending on the temperature,* (In Collection of Papers Dedicated to 70th Anniversary of A. F. Ioffe. Izd. Akad. Nauk SSSR, Moscow (1950), 61–72).

[10] Ya. B. Zel'dovich and G. I. Barenblatt, "On the dipole-type solution in the problems of a polytropic gas flow in porous medium", *Appl. Math. Mech.,* 21: 5 (1957): 718–720.

[11] Ya. B. Zel'dovich and G. I. Barenblatt, "The asymptotic properties of self-modelling solutions of the nonstationary gas filtration equations:, *Sov. Phys. Doklady,* 3 (1958): 44–47.

[12] R. E. Showalter, *Monotone Operators in Banach Space and Nonlinear Partial Differential Equations,* (Amer. Math. Soc., Providence (1997). XIII+270=283p).

[13] M. M. Vainberg, *Variational Method and Method of Monotone Operators in the Theory of Nonlinear Equations,* (Wiley, New York (1973)).

[14] X. Zhong, "Strong solutions to the nonhomogeneous Boussinesq equations for magnetohydrodynamics convection without thermal diffusion", *Electronic Journal of Qualitative Theory Differential Equations,* 2020: 24: 1–23.

[15] H. Zhang, Q. Hu, G. Liu, "Global existence, asymptotic stability and blow-up of solutions for the generalized Boussinesq equation with nonlinear boundary condition", *Mathematische Nachrichten,* 293: 2 (2020): 386–404.

[16] G. Oruc, G. M. Muslu, "Existence and uniqueness of solutions to initial boundary value problem for the higher order Boussinesq equation", *Nonlinear Analysis – Real World Applications,* 47 (2019): 436–445.

[17] W. Ding, Zh.-A. Wang, "Global existence and asymptotic behavier of the Boussinesq-Burgers system", *Journal of Mathematical Analysis and Applications,* 424: 1 (2015): 584–597.

[18] N. Zhu, Zh. Liu, K. Zhao, "On the Boussinesq-Burgers equations driven by dynamic boundary conditions", *Journal of Differential Equations,* 264: 3 (2018): 2287–2309.

[19] J. Crank, *Free and Moving Boundary Problems,* (Oxford University Press, 1984).

[20] J.-L. Lions, *Quelques methodes de resolution des problemes aux limites non lineaires,* (Dunod Gauthier-Villars, Paris (1969)).

[21] Yu.A. Dubinsky, "Weak convergence in nonlinear elliptic and parabolic equations", *Sbornik:Math.,* 67(109): 4 (1965): 609–642.

[22] P.A. Raviart, "Sur la resolution et l'approximation de certaines equations paraboliques non lineaires degenerees", *Archive Rat. Mech. Anal.,* 25 (1967): 64–80.

IRSTI 27.17.21 DOI: https://doi.org/10.26577/JMMCS.2022.v115.i3.05

S.M. Lutsak*, **O.A. Voronina**, **G.K. Nurakhmetova**
M. Kozybayev North Kazakhstan University, Kazakhstan, Petropavlovsk
*e-mail: sveta_lutsak@mail.ru

ON QUASI-IDENTITIES OF FINITE MODULAR LATTICES

In 1970 R. McKenzie proved that any finite lattice has a finite basis of identities. However the similar result for quasi-identities is not true. That is, there is a finite lattice that has no finite basis of quasi-identities. The problem "Which finite lattices have finite bases of quasi-identities?" was suggested by V.A. Gorbunov and D.M. Smirnov. In 1984 V.I. Tumanov found a sufficient condition consisting of two parts under which a locally finite quasivariety of lattices has no finite (independent) basis of quasi-identities. Also he conjectured that a finite (modular) lattice has a finite basis of quasi-identities if and only if a quasivariety generated by this lattice is a variety. In general, the conjecture is not true. W. Dziobiak found a finite lattice that generates a finitely axiomatizable proper quasivariety. Tumanov's problem is still unsolved for modular lattices. We construct a finite modular lattice that does not satisfy one of Tumanov's conditions but the quasivariety generated by this lattice is not finitely based.
Key words: Lattice, quasivariety, finite basis of quasi-identities.

С.М. Луцак*, О.А. Воронина, Г.К. Нурахметова
М. Қозыбаев атындағы Солтүстік Қазақстан университеті, Қазақстан, Петропавл қ.
*e-mail: sveta_lutsak@mail.ru

Соңғы модулярлық торлардың квази-сәйкестіктері туралы

1970 жылы Р. Маккензи кез-келген соңғы тордың түпкілікті сәйкестендіру негізі бар екенін дәлелдеді. Алайда, квази-сәйкестендіру үшін ұқсас нәтиже дұрыс емес. Яғни, квази-сәйкестендірудің түпкілікті негізі жоқ соңғы тор бар. Мәселе "Квази-сәйкестендірудің соңғы негіздері қандай соңғы торларға ие?" В.А. Горбунов және Д.М. Смирнов ұсынды.
1984 жылы В.И. Туманов екі бөліктен тұратын жеткілікті жағдайды тапты: жергілікті түрде, соңғы квазикөпбейне торларда квази-сәйкестендірудің соңғы (тәуелсіз) негізі жоқ.
Сондай-ақ, ол ақырғы (модулярлық) тордың квази-сәйкестендірудің соңғы негізі бар деп ұсынды содан кейін және тек осы тордан пайда болған квазикөпбейне бұл көпбейне. Жалпы жағдайда гипотеза дұрыс емес. В. Дзебяк ақырлы торды тапты, ол аксиоматизацияланатын өзіндік квазикөпбейнені тудырады. Тумановтың мәселесі әлі де модулярлық торлар үшін шешілген жоқ. Біз Тумановтың бір жағдайын қанағаттандырмайтын соңғы модулярлық торды саламыз, бірақ осы тордан пайда болған квазикөпбейненің түпкі негізі жоқ.
Түйін сөздер: Тор, квазикөпбейне, квази-сәйкестіктердің соңғы базисі.

С.М. Луцак*, О.А. Воронина, Г.К. Нурахметова
Северо-Казахстанский университет имени М. Козыбаева, Казахстан, г. Петропавловск
*e-mail: sveta_lutsak@mail.ru

О квазитождествах конечных модулярных решеток

В 1970 году Р. Маккензи доказал, что любая конечная решетка имеет конечный базис тождеств. Однако аналогичный результат для квазитождеств неверен. То есть существует конечная решетка, которая не имеет конечного базиса квазитождеств. Проблема "Какие конечные решетки имеют конечные базисы квазитождеств?" была предложена В.А. Горбуновым и Д.М. Смирновым.

В 1984 году В.И. Туманов нашел достаточное условие, состоящее из двух частей, при котором локально конечное квазимногообразие решеток не имеет конечного (независимого) базиса квазитождеств. Также он предположил, что конечная (модулярная) решетка имеет конечный базис квазитождеств тогда и только тогда, когда квазимногообразие, порожденное этой решеткой, является многообразием. В общем случае гипотеза неверна. В. Дзебяк нашел конечную решетку, которая порождает конечно аксиоматизируемое собственное квазимногообразие. Проблема Туманова до сих пор не решена для модулярных решеток. Мы строим конечную модулярную решетку, которая не удовлетворяет одному из условий Туманова, но квазимногообразие, порожденное этой решеткой, не является конечно базируемым.
Ключевые слова: Решетка, квазимногообразие, конечный базис квазитождеств.

1 Introduction

Questions concerning finite basability are among the most researched and relevant topics in universal algebra. It is well known that the finite based results begin with R.C. Lyndon, who in 1951 proved that the algebras on a two-element universe are always finitely based. R. McKenzie [1] in 1970 established that every finite lattice is finitely based, and generalizing this result, K.A. Baker in 1976 proved that every finite algebra generating a congruence-distributive variety is finitely based. There are two major directions in which Baker's theorem was generalized. In congruence-modular direction there was a series of results by R. Freese and R. McKenzie, the final result by McKenzie published in 1987 states that every finite algebra generating a congruence-modular residually finite variety is finitely based. In congruence meet-semidistributive direction, R. Willard in 2000 proved that every finite algebra generating a congruence meet-semidistributive residually strictly finite variety is finitely based.

Thus, according to R. McKenzie, any finite lattice has a finite basis of identities. The similar result for quasi-identities is not true, that was established by V.P. Belkin [2]. In 1979 he proved that there is a finite lattice that has no finite basis of quasi-identities. In particular, the smallest lattice that does not have a finite basis of quasi-identities is the ten-element modular lattice M_{3-3}. In this regard, the following question naturally arises. Which finite lattices have finite bases of quasi-identities? This problem was suggested by V.A. Gorbunov and D.M. Smirnov [3] in 1979. V.I. Tumanov [4] in 1984 found sufficient condition consisting of two parts under which the locally finite quasivariety of lattices has no finite (independent) basis for quasi-identities. Also he conjectured that a finite (modular) lattice has a finite basis of quasi-identities if and only if a quasivariety generated by this lattice is a variety. In general, the conjecture is not true. W. Dziobiak [5] found a finite lattice that generates finitely axiomatizable proper quasivariety. Also we would like to point out that Tumanov's problem is still unsolved for modular lattices.

The main goal of the paper is to present a finite modular lattice that does not satisfy one of Tumanov's conditions but the quasivariety generated by this lattice is not finitely based (has no finite basis of quasi-identities).

2 Material and methods

We recall some basic definitions and results for quasivarieties that we will refer to. For more information on the basic notions of general algebra introduced below and used throughout this paper, we refer to [6] and [7].

A *quasivariety* is a class of lattices that is closed with respect to subalgebras, direct products, and ultraproducts. Equivalently, a quasivariety is the same thing as a class of lattices axiomatized by a set of quasi-identities. A *quasi-identity* means a universal Horn sentence with the non-empty positive part, that is of the form

$$(\forall \bar{x})[p_1(\bar{x}) \approx q_1(\bar{x}) \wedge \cdots \wedge p_n(\bar{x}) \approx q_n(\bar{x}) \rightarrow p(\bar{x}) \approx q(\bar{x})]$$

where $p, q,\ p_1, q_1, \ldots, p_n, q_n$ are lattice's terms. A *variety* is a quasivariety which is closed under homomorphisms. According to Birkhoff theorem [8], a variety is a class of similar algebras axiomatized by a set of identities, where by an identity we mean a sentence of the form $(\forall \bar{x})[s(\bar{x}) \approx t(\bar{x})]$ for some terms $s(\bar{x})$ and $t(\bar{x})$.

By $\mathbf{Q}(\mathbf{K})$ ($\mathbf{V}(\mathbf{K})$) we denote the smallest quasivariety (variety) containing a class $\mathbf{K}$. If $\mathbf{K}$ is a finite family of finite algebras then $\mathbf{Q}(\mathbf{K})$ is called finitely generated. In case when $\mathbf{K} = \{\mathcal{A}\}$ we write $\mathbf{Q}(\mathcal{A})$ instead of $\mathbf{Q}(\{\mathcal{A}\})$.

Let $\mathbf{K}$ be a quasivariety. A congruence α on algebra $\mathcal{A}$ is called a $\mathbf{K}$-*congruence* or *relative congruence* provided $\mathcal{A}/\alpha \in \mathbf{K}$. The set $\mathrm{Con}_{\mathbf{K}}\mathcal{A}$ of all $\mathbf{K}$-congruences of $\mathcal{A}$ forms an algebraic lattice with respect to inclusion $\subseteq$ which is called a *relative congruence lattice*.

The least $\mathbf{K}$-congruence $\theta_{\mathbf{K}}(a, b)$ on algebra $\mathcal{A} \in \mathbf{K}$ containing pair $(a, b) \in A \times A$ is called a *principal* $\mathbf{K}$-*congruence* or a *relative principal congruence*. In case when $\mathbf{K}$ is a variety, relative congruence $\theta_{\mathbf{K}}(a, b)$ is usual principal congruence that we denote by $\theta(a, b)$.

An algebra $\mathcal{A}$ belonging to a quasivariety $\mathbf{K}$ is *(finitely) subdirectly irreducible relative to* $\mathbf{K}$, or *(finitely) subdirectly* $\mathbf{K}$-*irreducible*, if intersection of any (finite) number of nontrivial $\mathbf{K}$-congruences is again nontrivial; in other words, the trivial congruence 0_A is a (meet-irreducible) completely meet-irreducible element of $\mathrm{Con}_{\mathbf{K}}\mathcal{A}$.

Let $(a] = \{x \in L \mid x \leq a\}$ ($[a) = \{x \in L \mid x \geq a\}$) be a principal ideal (coideal) of a lattice $\mathcal{L}$. A pair $(a, b) \in L \times L$ is called *dividing* (*semi-dividing*) if $L = (a] \cup [b)$ and $(a] \cap [b) = \varnothing$ ($L = (a] \cup [b)$ and $(a] \cap [b) \neq \varnothing$).

For any semi-dividing pair (a, b) of a lattice $\mathcal{M}$ we define a lattice

$$\mathcal{M}_{a-b} = \langle \{(x, 0), (y, 1) \in M \times 2 \mid x \in (a], y \in [b)\}; \vee, \wedge \rangle \leq_s \mathcal{M} \times \mathbf{2},$$

where $\mathbf{2} = \langle \{0, 1\}; \vee, \wedge \rangle$ is a two element lattice.

Theorem 1 (Tumanov's theorem [4]) *Let* $\mathbf{M}$, $\mathbf{N}$ ($\mathbf{N} \subset \mathbf{M}$) *be locally finite quasivarieties of lattices satisfying the following conditions:*

a) in any finitely subdirectly $\mathbf{M}$-*irreducible lattice* $\mathcal{M} \in \mathbf{M} \backslash \mathbf{N}$ *there is a semi-dividing pair* (a, b) *such that* $\mathcal{M}_{a-b} \in \mathbf{N}$*;*

b) there exists a finite simple lattice $\mathcal{P} \in \mathbf{N}$ *which is not a proper homomorphic image of any subdirectly* $\mathbf{N}$-*irreducible lattice.*

Then the quasivariety $\mathbf{N}$ *has no coverings in the lattice of subquasivarieties of* $\mathbf{M}$*. In particular,* $\mathbf{N}$ *has no finite basis of quasi-identities provided* $\mathbf{M}$ *is finitely axiomatizable.*

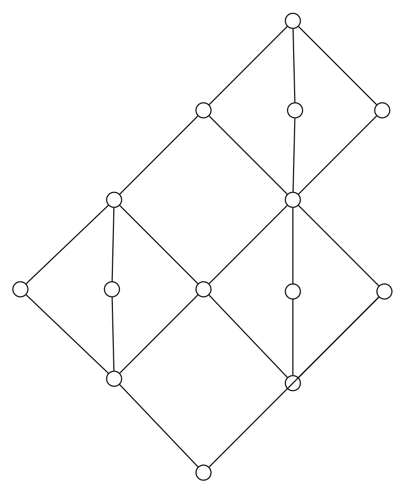
In the next section, the algebra $\mathcal{L}$ and its carrier (its main set) L will be identified and denoted by the same way, namely L.

3 Results and discussion

Let T be a modular lattice displayed in Figure 1. And let $\mathbf{N} = \mathbf{Q}(T)$ and $\mathbf{M} = \mathbf{V}(T)$ be the quasivariety and variety generated by T, respectively. Since every subdirectly $\mathbf{N}$-irreducible lattice is a sublattice of T, we have that a class $\mathbf{N}_{si}$ of all subdirectly $\mathbf{N}$-irreducible lattices consists of the lattices $\mathbf{2}$, M_3, M_{3-3} and T (see Figures 1 and 2). It easy to see that M_3 is a unique simple lattice in $\mathbf{N}_{si}$ and is a homomorphic image of T. Thus, the condition *a*) of Tumanov's theorem is not valid for quasivarieties $\mathbf{N} \subset \mathbf{M}$. We show

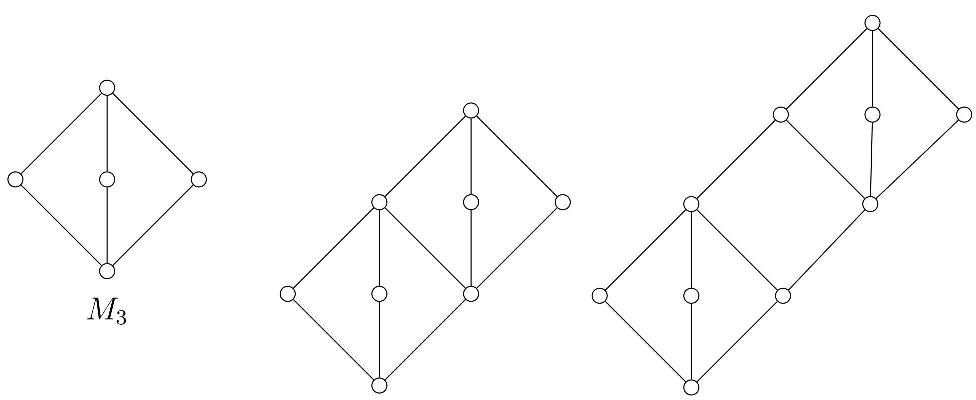
Theorem 2 *Quasivariety $\mathbf{Q}(T)$ generated by the lattice T is not finitely based.*

To prove the theorem we modify the proof of the second part of Theorem 3.4 from [9].



T

Figure 1: Lattice T



M_3 $M_{3,3}$ M_{3-3}

Figure 2: Lattices M_3, $M_{3,3}$ and M_{3-3}

Let S be a non-empty subset of a lattice L. Denote by $\langle S \rangle$ the sublattice of L generated by S.

We define a modular lattice L_n by induction:

$n = 1$. $L_1 \cong M_{3-3}$ and $L_1 = \langle\{a_1, b_1, c_1, e, d\}\rangle$ (see Figure 3);

$n = 2$. L_2 is a modular lattice generated by $L_1 \cup \{a_2, b_2, c_2, d\}$ such that $b_1 = c_2$, $\langle\{a_2, b_2, c_2, e, b_1\}\rangle \cong M_3$, and $a_2 \vee b_2 = e \wedge d_1$, $d \vee b_1 = d_1$, and $b_2 < d$ (see Figure 3).

$n > 2$. L_n is a modular lattice generated by the set $\{a_i, b_i, c_i \mid i \leq n\} \cup \{e, d\}$ such that a_i is not comparable with a_j and b_k for all $j \neq i$ and $k \leq n$, $b_{i-1} = c_i$, $\langle\{a_i, b_i, c_i\}\rangle \cong M_3$ for all $i < n$, $b_i \vee d = d_i$ for all $i < n$, and $b_n < d$ (see Figure 4).

One can see that L_n is a subdirect product of the lattices L_{n-1} and M_3 for any $n > 2$.

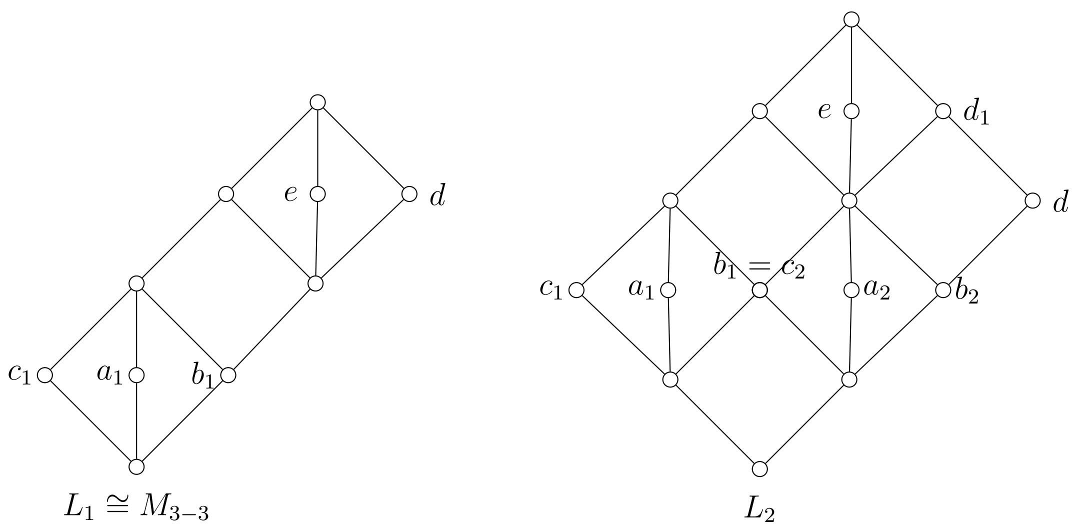


Figure 3: Lattices L_1, L_2

Let L_n^- be a sublattice of L_n generated by the set $\{a_i, b_i, c_i \mid i \leq n\}$.

Lemma 1 *For any $n > 1$ and a non-trivial congruence $\theta \in \mathrm{Con}L_n$ there is $1 < m < n$ such that $L_n/\theta \cong L_m$ or $L_n/\theta \cong M_{3,3}$ provided $(a_1, b_1) \notin \theta$, otherwise $L_n/\theta \cong L_m^-$.*

Proof of Lemma 1.

We prove by induction on $n > 2$. One can check that it is true for $n = 3$ because of $L_3/\theta \cong L_2$ or $L_3/\theta \cong M_{3,3}$ if $(a_1, b_1) \notin \theta$ and $L_3/\theta \cong L_2^-$ or $L_3/\theta \cong M_3$ for any non-trivial congruence $\theta \in \mathrm{Con}L_3$.

Let $n > 3$. And let u cover v in L_n and $\theta(u, v) \subseteq \theta$. By construction of L_n, we have $L_n/\theta(u, v) \cong L_{n-1}$ or $L_n/\theta(u, v) \cong L_{n-1}^-$.

Assume $(a_1, b_1) \notin \theta$. Since for every non-trivial congruence $\theta \in \mathrm{Con}L_n$ there are $u, v \in L_n$ such that u covers v and $\theta(u, v) \subseteq \theta$, we get

$$L_n/\theta \cong (L_n/\theta(u, v))/(\theta/\theta(u, v)).$$

Since $L_n/\theta(u, v) \cong L_{n-1}$ we obtain

$$L_n/\theta \cong (L_n/\theta(u, v))/(\theta/\theta(u, v)) \cong L_{n-1}/\theta',$$

for some $\theta' \in \mathrm{Con}(L_{n-1})$. And, by induction, $L_{n-1}/\theta' \cong L_m$ or $L_{n-1}/\theta' \cong M_{3,3}$ for some $m > 0$. Thus $L_n/\theta \cong L_m$ or $L_n/\theta \cong M_{3,3}$.

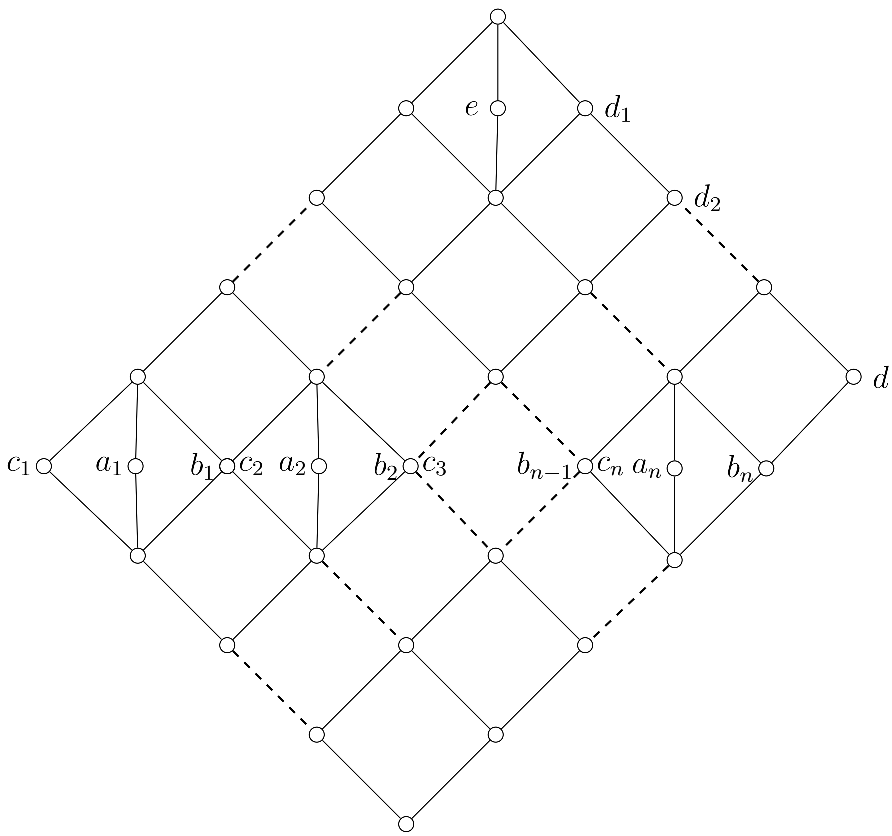


Figure 4: Lattice L_n, $n \geq 2$

Now assume $(a_1, b_1) \in \theta$. Then $\theta(a_1, b_1) = \theta(u, v)$ and $L_n/\theta(u, v) \cong L_n^-$. Hence

$$L_n/\theta \cong (L_n/\theta(u, v))/(\theta/\theta(u, v)) \cong L_n^-/\theta',$$

for some $\theta' \in \mathrm{Con}(L_n^-)$. It is not difficult to check that $L_n^-/\theta' \cong L_m^-$ for some $m > 0$ (see Lemma 3.1 [9]). Thus $L_n/\theta \cong L_m$ or $L_n/\theta \cong L_m^-$.

Corollary 1 *For all $n > 1$, there is no proper homomorphism from L_n to M_{3-3} and T.*

Proof of Corollary 1.

We provide the proof for a proper homomorphism from L_n into M_{3-3}. It is not difficult to check that the same arguments hold for a proper homomorphism from L_n into T.

Assume $h : L_n \to M_{3-3}$, $n > 1$, is a proper homomorphism. Hence $\ker h$ is not a trivial congruence on L_n. By Lemma 1, $L_n/\ker h \cong L_m$ or $L_n/\theta \cong M_{3,3}$ or $L_n/\ker h \cong L_m^-$ for some $m > 1$. Thus $L_m = h(L_n) \leq M_{3-3}$. It is impossible because, by definition of L_m, $|L_m| > |M_{3-3}|$ for all $m > 1$, hence L_n is not a sublattice of M_{3-3}. Obviously, $M_{3,3}$ and L_M^- are not sublattices of M_{3-3}. Thus there is no such homomorphism h.

Lemma 2 *For every $n > 2$, a lattice L_n has the following properties:*

i) $L_n \leq_s L_{n-1} \times L_{n-1}$;

ii) $L_n \in \mathbf{V}(M_{3,3}) = \mathbf{V}(T)$;

iii) $L_n \notin \mathbf{Q}(T)$;

iv) Every proper subalgebra of L_n *belongs to* $\mathbf{Q}(T)$.

Proof of Lemma 2.

i). One can check that $L_n/\theta(a_i, b_i) \cong L_{n-1}$ for all $1 < i \leq n$. Since $n > 2$ then $\theta(a_2, b_2), \theta(a_3, b_3) \in \mathrm{Con}L_n$ and $\theta(a_2, b_2) \cap \theta(a_3, b_3) = \Delta$. This means that $L_n \leq_s L_{n-1} \times L_{n-1}$.

ii). One can see that T is a subdirect product of M_3 and $M_{3,3}$. Hence $T \in \mathbf{V}(M_{3,3})$. On the other hand, by Jonsson lemma [10], every subdirectly irreducible lattice in $\mathbf{V}(T)$ is a homomorphic image of some sublattice of T. Hence $M_{3,3} \in \mathbf{V}(T)$. Thus $\mathbf{V}(M_{3,3}) = \mathbf{V}(T)$, and, by *i)* and induction on n, we get $L_n \in \mathbf{V}(T)$.

iii). Suppose $L_n \in \mathbf{Q}(T)$ for some $n > 1$. Then L_n is a subdirect product of subdirectly $\mathbf{Q}(T)$-irreducible algebras. Since every subdirectly $\mathbf{Q}(T)$-irreducible algebra is a subalgebra of T, we get that L_n is a subdirect product of subalgebras of T. By Lemma 1, there is no proper homomorphism from L_n onto T or M_{3-3}. Hence $L_n \in \mathbf{Q}(M_3)$ for all $n > 1$. It is impossible because $M_{3-3} \leq L_n$ and $M_{3-3} \notin \mathbf{Q}(M_3)$.

iv). We prove by induction on n. It is true for $n \leq 2$ by manual checking. Let $n > 2$ and let S be a maximal sublattice of L_n. Since the lattice L_n is generated by the set of double irreducible elements $\{a_1, \ldots, a_n, c_1, e, d\}$, there is $0 < i \leq n$ such that $a_i \notin S$ or $c_1 \notin S$ or $e \notin S$ or $d \notin S$.

Suppose $c_1 \notin S$. One can see that $\langle S \rangle \leq_s \mathbf{2} \times M_3 \times L^-_{n-1}$. Since $L_{n-1} \leq_s M_3^{n-1}$ we get $\langle S \rangle \in \mathbf{Q}(M_3) \subset \mathbf{Q}(T)$.

Suppose $e \notin S$. Then $\langle S \rangle \leq_s \mathbf{2} \times L^-_n \leq_s \mathbf{2} \times M_3^n \in \mathbf{Q}(M_3) \subset \mathbf{Q}(T)$.

Suppose $d \notin S$. Put $S_m = \{\{a_1, \ldots, a_m, c_1, e\}$, $m < n$, and $T_m = \langle S_m \rangle$. One can see that $T_m/\theta(a_i, b_i) \cong T_{m-1}$ for all $1 < i < m$. And $T_m/\theta(a_1, b_1) \cong L^-_{m-1}$. Since $\theta(a_1, b_1) \cap \theta(a_i, b_i) = \Delta$, by distributivity of $\mathrm{Con}T_m$, we have $\theta(a_1, b_1) \cap (\bigvee\{\theta(a_i, b_i) \mid 1 < i < m\}) = \Delta$. Since $T_m/(\bigvee\{\theta(a_i, b_i) \mid 1 < i < m\}) \cong T$ we obtain $\langle S_m \rangle \leq_s T \times L^-_{n-1} \leq_s T \times M_3^{n-1} \in \mathbf{Q}(T)$.

Suppose $a_i \notin S$. Since $n > 1$ and S is a maximal sublattice, then there are $i \neq k \neq l \neq i$ such that $\theta(b_k, c_k), \theta(b_l, c_l) \in \mathrm{Con}L_n$,

$$\theta(b_k, c_k) \cap \theta(b_l, c_l) = \Delta.$$

and

$$L_n/\theta(b_k, c_k) \cong L_n/\theta(b_l, c_l) \cong L_{n-1} \quad \text{or} \quad \{L_n/\theta(b_k, c_k), L_n/\theta(b_l, c_l)\} = \{L_{n-1}, L^-_{n-1}\}.$$

We provide the proof for the first case, $L_n/\theta(b_k, c_k) \cong L_n/\theta(b_l, c_l) \cong L_{n-1}$. These isomorphisms mean that $L_n \leq_s L_{n-1} \times L_{n-1}$ and $S \leq L_{n-1} \times L_{n-1}$. Let $h_k : L_n \to L_{n-1}$ and $h_l : L_n \to L_{n-1}$ are homomorphisms such that $\ker h_k = \theta(b_k, c_k)$ and $\ker h_l = \theta(b_l, c_l)$. Since $(a_i, b_i) \notin \theta(b_k, c_k) \cup \theta(b_l, c_l)$ then $h_k(S)$, $h_l(S)$ are proper sublattices of L_{n-1}. And, by induction, $h_k(S), h_l(S) \in \mathbf{Q}(T)$. As $b_k, c_k, b_l, c_l \in S$, the restrictions of congruences $\theta(b_k, c_k)|_S$ and $\theta(b_l, c_l)|_S$ on the algebra S are not trivial congruences on S. Moreover $\theta(b_k, c_k)|_S \cap \theta(b_l, c_l)|_S = \Delta$. It means $S \leq_s h_k(S) \times h_l(S)$. Hence $S \in \mathbf{Q}(T)$. Since every maximal proper subalgebra of L_n belongs to $\mathbf{Q}(T)$ then every proper subalgebra of L_n belongs to $\mathbf{Q}(T)$.

It is not difficult to check that for $\{L_n/\theta(b_k, c_k), L_n/\theta(b_l, c_l)\} = \{L_{n-1}, L^-_{n-1}\}$ the same arguments hold.

Now we prove the main result, Theorem 2.

We use the following folklore fact which provides non-finite axiomatizability: A locally finite quasivariety $\mathbf{K}$ is not finitely axiomatizable if for any positive integer $n \in N$ there is a finite algebra L_n such that $L_n \notin \mathbf{K}$ and every n-generated subalgebra of L_n belongs to $\mathbf{K}$.

We show that for quasivariety $\mathbf{Q}(T)$, the lattice L_n satisfies the conditions of this fact. Indeed, by Lemma 2(iii), $L_n \notin \mathbf{Q}(T)$ for all $n > 1$. Since L_n is generated by at least $n+1$ double irreducible elements then every n-generated subalgebra of L_n is a proper subalgebra. By Lemma 2(iv), every n-generated subalgebra of L_n belongs to $\mathbf{Q}(T)$. Hence $\mathbf{Q}(T)$ has no finite basis of quasi-identities.

We note that there is an infinite number of lattices similar to the lattice T.

The proof of Theorem 2 give us more general result:

Theorem 3 *Suppose L is a finite lattice such that $M_{3,3} \not\leq L$, $T \leq L$ and $L_n \not\leq L$ for all $n > 1$. Then the quasivariety $\mathbf{Q}(L)$ is not finitely based.*

4 Conclusion

There are three measures of the highest complexity of the structure of quasivariety lattices: Q-universality, property (N) or non-computability of the set of finite sublattices, and an existence of continuum of quasivarieties without covers in a given quasivariety lattice. The presence in the quasivariety lattices of a continuum of elements that do not have coverings indicates the complexity of the structure of these lattices; in this case, there is a continuum of subquasivarieties of a given quasivariety $\mathbf{K}$ that do not have an independent basis of quasi-identities with respect to $\mathbf{K}$. In [11] a sufficient condition for a quasivariety $\mathbf{K}$ to be Q-universal, to have continuum many subclasses with the property (N), continuum many Q-universal subquasivarieties and continuum many subquasivarieties with no upper covers in the lattice $Lq(\mathbf{K})$ was provided. In [12] a sufficient condition for a class $\mathbf{K}$ to have continuum many subclasses with the property (N) but which are not Q-universal was established. In [13] it was proved that almost all known Q-universal quasivarieties contain classes having property (N).

In this paper we construct a finite modular lattice that does not satisfy one of Tumanov's conditions but the quasivariety generated by this lattice is not finitely based. It has no finite basis of quasi-identities.

5 Acknowledgments

The authors thank A.M. Nurakunov for his attention and useful remarks.

This research is funded by the Science Committee of the Ministry of Education and Science of the Republic of Kazakhstan (Grant No. AP09058390).

References

[1] McKenzie R., "Equational bases for lattice theories", *Math. Scand.,* 27 (1970): 24-38.

[2] Belkin V.P., "Quasi-identities of finite rings and lattices", *Algebra and Logic,* 17 (1979): 171-179.

[3] Gorbunov V.A., Smirnov D.M., "Finite algebras and the general theory of quasivarieties" , *Colloq. Mathem. Soc. Janos Bolyai. Finite Algebra and Multipli-valued Logic,* 28 (1979): 325-332.

[4] Tumanov V.I., "On finite lattices having no independent bases of quasi-identities" , *Math. Notes,* 36 (1984): 625-634.

[5] Dziobiak W., "Finitely generated congruence distributive quasivarieties of algebras" , *Fund. Math.,* 133 (1989): 47-57.

[6] Burris S., Sankappanavar H.P., *A Course in Universal Algebra* (New York: Springer, 1980): 315.

[7] Gorbunov V.A., *Algebraic Theory of Quasivarieties* (New York: Consultants Bureau, 1998): 368.

[8] Birkhoff G., "Subdirect union in universal algebra" , *Bull. Amer. Math. Soc.,* 50 (1944): 764-768.

[9] Basheyeva A.O., Mustafa M., Nurakunov A.M., "Properties not retained by pointed enrichments of finite lattices" , *Algebra Univers.,* 81:56 (2020): 1-11.

[10] Jonsson B., "Algebras whose congruence lattices are distributive" , *Math. Scand.,* 21 (1967): 110-121.

[11] Kravchenko A.V., Nurakunov A.M., Schwidefsky M.V., "Structure of quasivariety lattices. I. Independent axiomatizability" , *Algebra and Logic,* 57:6 (2019): 445-462.

[12] Lutsak S.M., "On complexity of quasivariety lattices" , *Sib. El. Math. Rep.,* 14 (2017): 92-97.

[13] Schwidefsky M.V., "Complexity of quasivariety lattices" , *Algebra and Logic,* 54:3 (2015): 245-257.

IRSTI 27.39.21 DOI: https://doi.org/10.26577/JMMCS.2022.v115.i3.06

D. Serikbaev
Institute of Mathematics and Mathematical Modeling, Kazakhstan, Almaty
e-mail: serykbaev.daurenbek@gmail.com

MULTI-TERM TIME-FRACTIONAL DERIVATIVE HEAT EQUATION FOR ONE-DIMENSIONAL DUNKL OPERATOR

In this paper, we investigate the well-posedness for Cauchy problem for multi-term time-fractional heat equation associated with Dunkl operator. The equation under consideration includes a linear combination of Caputo derivatives in time with decreasing orders in $(0, 1)$ and positive constant coefficients and one-dimensional Dunkl operator. To show solvability of this problem we use several important properties of multinomial Mittag-Leffler functions and Dunkl transforms, since various estimates follow from the explicit solutions in form of these special functions and transforms. Then we prove the uniqueness and existence results. To achieve our goals, we use methods corresponding to the different areas of mathematics such as the theory of partial differential equations, mathematical physics, hypoelliptic operators theory and functional analysis. In particular, we use the direct and inverse Dunkl transform to establish the existence and uniqueness of solutions to this problem on the abstract Hilbert space. The generalized solutions of this problem are studied.
Key words: Dunkl operator, heat equation, Cauchy problem, Caputo fractional derivative.

Д. Серикбаев
Математика және математикалық модельдеу институты, Қазақстан, Алматы қ.
e-mail: serykbaev.daurenbek@gmail.com

Бір өлшемді Данкл операторы үшін уақыт бойынша көпмүшелік бөлшек туындысы бар жылу өткізгіштік теңдеуі

Бұл мақалада біз Данкл операторымен байланысты уақыт бойынша көпмүшелік бөлшек туындысы бар жылу өткізгіштік теңдеуі үшін Коши есебінің қисынды екенін зерттейміз. Қарастырылып отырған теңдеу уақыт бойынша Капуто туындыларының сызықтық комбинациясы $(0, 1)$ оң коэффициенттерімен және де бір өлшемді Данкл операторынан туындаған. Бұл есептің шешімділігін көрсету үшін біз Миттаг-Леффлер көпмүшелік арнайы функциялары және Данкл түрлендіруінің маңызды қасиеттерін қолданамыз, өйткені әртүрлі бағалаулар осы арнайы функциялармен түрлендірулер түріндегі нақты шешімдерден туындайды. Содан кейін біз осы есептің шешімі бар және жалғыз екенін дәлелдейміз. Осы айтылғанды дәлелдеу үшін біз математиканың әртүрлі салаларына сәйкес келетін әдістерді қолданамыз, атап айтқанда дербес туындылы дифференциалдық теңдеулер теориясы, математикалық физика теңдеулері, гипоэллиптикалық операторлар теориясы және функционалдық талдау. Қарастырып отырған есептің шешімі бар және жалғыз болатынын абстрактты Гильберт кеңістігінде дәлелдейміз, ол үшін біз негізгі әдіс ретінде тура және кері Данкл түрлендіруін қолданамыз. Бұл есепте жалпылама шешім қарастырылады.
Түйін сөздер: Данкл операторы, жылу өткізгіштік теңдеуі, Коши есебі, Капуто бөлшек туындысы.

Д. Серикбаев
Институт математики и математического моделирования, Казахстан, г. Алматы
e-mail: serykbaev.daurenbek@gmail.com

Уравнение теплопроводности с многочленной дробной производной по времени для одномерного оператора Данкля

В этой статье мы исследуем корректность задачи Коши для уравнения теплопроводности с многочленным дробным производным по временем связанного с оператором Данкла. Рассматриваемое уравнение включает линейную комбинацию производных Капуто по времени с убывающими порядками в $(0,1)$ и положительными коэффициентами и одномерным оператором Данкла. Чтобы показать разрешимость этой задачи, мы используем несколько важных свойств многочленных функций Миттага-Леффлера и преобразований Данкла, поскольку из явных решений в виде этих специальных функций и преобразований вытекают различные оценки. Затем мы докажем единственность и существования решения этой задачи. Для достижения наших цель мы используем методы, соответствующие различным областям математики, таким как теория дифференциальных уравнений в частных производных, математическая физика, теория гипоэллиптических операторов и функциональный анализ. В частности, мы используем прямое и обратное преобразование Данкла, чтобы установить существование и единственность решений этой задачи в абстрактном Гильбертовом пространстве. Изучаются обобщенные решения этой задачи.

Ключевые слова: Оператор Данкла, уравнение теплопроводности, задача Коши, дробная производная Капуто.

1 Introduction

Let γ be $0 < \gamma < 1$. For a fixed positive integer m, $a_j \in \mathbb{R}$ and γ_j $(j = 1, ..., m)$ be constants such that $1 > \gamma > \gamma_1 >, ..., > \gamma_m > 0$. We consider the following equation

$$\partial_t^{\gamma} u(t,x) - \sum_{j=1}^{m} a_j \partial_t^{\gamma_j} u(t,x) - \Lambda_{\alpha,x}^2 u(t,x) = f(t,x) \tag{1}$$

in the domain $(t,x) \in Q_T$, under the initial condition

$$u(0,x) = g(x), \quad x \in \mathbb{R}, \tag{2}$$

where f and g are sufficiently smooth functions.

Here $\partial_t^{\alpha_j}$ denotes the Caputo derivative defined by

$$\partial_t^{\alpha_j} u(t) := \frac{1}{\Gamma(1-\alpha_j)} \int_0^t \frac{u'(s)}{(t-s)^{\alpha_j}} ds,$$

where $\Gamma(\cdot)$ is a usual Gamma function. For various properties of the Caputo derivative, we refer to Kilbas et al. [9], Podlubny [10].

The operator Λ_α is called the Dunkl operator which was introduced in 1989 by C. Dunkl [2], where $\alpha \geq 1/2$. The Dunkl operator is associated with the reflection group $\mathbb{Z}_2$ on $\mathbb{R}$. The Dunkl operators are very important in pure mathematics and physics. Solution of the spectral problem generated by the Dunkl operator is called the Dunkl kernel $E_\alpha(ix\lambda)$ which is used to define the Dunkl transform $\mathcal{F}_\alpha$ [4]. Main properties of the Dunkl transform is given by M.F.E. de Jeu in 1993 [5]. For more information about harmonic analysis associated with the operator Λ_α, we refer the readers to the papers [1, 3, 5, 6].

A general solution of problem (1)–(2) is the function $u \in C^\alpha([0,T], L^2(\mathbb{R}, \mu_\alpha)) \cap C([0,T], W_\alpha^{2,2}(\mathbb{R}, \mu_\alpha))$ satisfying the equation (1). Let us denote that by $\mathcal{D}_t^\gamma := \partial_t^\gamma - \sum_{j=1}^m a_j \partial_t^{\gamma_j}$.

2 Auxiliary materials

In this section we introduce the Dunkl operator and it's necessary properties to our research.

2.1 The Dunkl operator and the Dunkl transform

The first-order singular differential-difference operator $\Lambda_\alpha, \alpha \geq -1/2$, given by

$$\Lambda_\alpha y(x) = \frac{d}{dx} y(x) + \frac{2\alpha+1}{x} \left(\frac{y(x) - y(-x)}{2} \right), \quad y \in C^1(\mathbb{R})$$

called the Dunkl operator, associated with the reflexion group $\mathbb{Z}_2$ on $\mathbb{R}$. If $\alpha = -1/2$, the Dunkl operator turns into the ordinary differential operator $\Lambda_{-1/2} = \frac{d}{dx}$.

For $\alpha \geq -1/2$ and $\lambda \in \mathbb{R}$ the spectral problem associated with Dunkl operator

$$\begin{cases} \Lambda_\alpha y(x) - (i\lambda) y(x) = 0, \\ y(0) = 1. \end{cases}$$

has a unique solution $y(x) = D_\alpha(ix\lambda)$ called Dunkl kernel given by

$$D_\alpha(ix\lambda) = j_\alpha(ix\lambda) + \frac{ix\lambda}{2(\alpha+1)} j_{\alpha+1}(ix\lambda), \quad x \in \mathbb{R},$$

where

$$j_\alpha(ix\lambda) = \Gamma(\alpha+1) \sum_{k=0}^{+\infty} \frac{1}{k!} \frac{(ix\lambda/2)^{2k}}{\Gamma(k+\alpha+1)}$$

is the normalized Bessel function of order α.

Замечание 1 *For $\alpha = -\frac{1}{2}$, we have*

$$\begin{cases} \frac{d}{dx} y(x) - (i\lambda) y(x) = 0, \\ y(0) = 1. \end{cases}$$

The solution of this problem is

$$D_{-1/2}(ix\lambda) = e^{ix\lambda}.$$

Definition 1 *We denote by $L^p(\mathbb{R}, \mu_\alpha), 1 \leq p \leq +\infty$, the space of measurable functions h on $\mathbb{R}$ such that*

$$\|h\|_{p,\alpha} = \left(\int_{\mathbb{R}} |h(x)|^p d\mu_\alpha(x) \right)^{\frac{1}{p}} < +\infty, \quad 1 \leq p < +\infty,$$

$$\|h\|_\infty = \sup_{x \in \mathbb{R}} |h(x)| < +\infty.$$

Here μ_α is the measure defined on $\mathbb{R}$ by

$$d\mu_\alpha(x) = \frac{|x|^{2\alpha+1}}{2^{\alpha+1}\Gamma(\alpha+1)} dx, \quad \alpha \geq -1/2.$$

For $h \in L^1(\mathbb{R}, \mu_\alpha)$ the Dunkl transform is defined by

$$\mathcal{F}_\alpha(h)(\lambda) = \widehat{h}(\lambda) := \int_{\mathbb{R}} h(x) D_\alpha(-ix\lambda) d\mu_\alpha(x), \quad \lambda \in \mathbb{R}. \tag{3}$$

This transform has the following properties ([5]):
i) For all $h \in \mathcal{S}(\mathbb{R})$, we have

$$\mathcal{F}_\alpha(\Lambda_\alpha h)(\lambda) = i\lambda \mathcal{F}_\alpha(h)(\lambda), \quad \lambda \in \mathbb{R}. \tag{4}$$

ii) For all $h \in L^1(\mathbb{R}, \mu_\alpha)$, the Dunkl transform $\mathcal{F}_\alpha$ is a continuous function on $\mathbb{R}$ satisfying

$$\|\mathcal{F}_\alpha(h)\|_\infty \leq \|h\|_{1,\alpha}.$$

iii) (L^1-inversion) For all $h \in L^1(\mathbb{R}, \mu_\alpha)$ with $\mathcal{F}_\alpha(h) \in L^1(\mathbb{R}, \mu_\alpha)$, we have

$$h(x) = \int_{\mathbb{R}} \mathcal{F}_\alpha(h)(\lambda) D_\alpha(ix\lambda) d\mu_\alpha(\lambda). \tag{5}$$

iv) $\mathcal{F}_\alpha$ is a topological isomorphism on $\mathcal{S}(\mathbb{R})$ which extends to a topological isomorphism on $\mathcal{S}'(\mathbb{R})$.
v) (Plancherel theorem) The Dunkl transform $\mathcal{F}_\alpha$ is an isometric isomorphism of $L^2(\mathbb{R}, \mu_\alpha)$. In particular,

$$\|\mathcal{F}_\alpha(h)\|_{2,\alpha} = \|h\|_{2,\alpha}. \tag{6}$$

Notation. ([6, p. 22]) For $s \in \mathbb{R}$ we denote by

$$W_\alpha^{s,2}(\mathbb{R}, \mu_\alpha) := \{h \in \mathcal{S}'(\mathbb{R}) : \quad \|h\|^2_{W_\alpha^{s,2}(\mathbb{R},\mu_\alpha)} = \int_{\mathbb{R}} (1+\lambda^2)^s |\mathcal{F}_\alpha(h)(\lambda)|^2 d\mu_\alpha(\lambda) < \infty\}$$

the usual Sobolev space on $\mathbb{R}$.

3 Main results and methods

Theorem 1 *Let $g \in W_\alpha^{2,2}(\mathbb{R}, \mu_\alpha)$, $f \in C([0,T], W_\alpha^{2,2}(\mathbb{R}, \mu_\alpha))$ and $0 < \gamma < 1$. Then there exists a unique solution of problem (1)–(2). Moreover, it is given by the expression*

$$u(t,x) = \int_{\mathbb{R}} \int_{\mathbb{R}} g(y) \left(1 - \lambda^2 t^\gamma E_{\gamma+1}(t)\right) D_\alpha(ix\lambda) D_\alpha(-iy\lambda) d\mu_\alpha(y) d\mu_\alpha(\lambda)$$

$$+ \int_{\mathbb{R}} \int_{\mathbb{R}} \int_0^t \frac{d}{d\tau} \{\tau^\gamma E_{1+\gamma}(\tau)\} f(t-\tau, y) D_\alpha(ix\lambda) D_\alpha(-iy\lambda) d\tau d\mu_\alpha(y) d\mu_\alpha(\lambda),$$

where

$$E_{\gamma+1}(t) = E_{(\gamma-\gamma_1,\ldots,\gamma-\gamma_m,\gamma),1+\gamma}(a_1 t^{\gamma-\gamma_1}, \ldots, a_m t^{\gamma-\gamma_m}, -\lambda^2 t^\gamma).$$

Here

$$E_{(\gamma_1,\ldots,\gamma_{m+1}),\beta(z_1,\ldots,z_{m+1})} = \sum_{k=0}^{\infty} \sum_{\substack{l_1+l_2+\cdots+l_{m+1}=k, \\ l_1 \geq 0,\ldots,l_{m+1} \geq 0}} \frac{k!}{l_1! \ldots l_{m+1}!} \frac{\prod_{j=1}^{m+1} z_j^{l_j}}{\Gamma\left(\beta + \sum_{j=1}^{m+1} \gamma_j l_j\right)} \tag{7}$$

is the multivariate Mittag-Leffler function [7].

Existence of the solution. Now to show that there is a generalised solution to problem (1)–(2), we apply the Dunkl transform $\mathcal{F}_\alpha$ to the equation (1) and the initial condition (2). It gives us

$$\partial_t^\gamma \widehat{u}(t,\lambda) - \sum_{j=1}^{m} a_j \partial_t^{\gamma_j} \widehat{u}(t,\lambda) + \lambda^2 \widehat{u}(t,\lambda) = \widehat{f}(t,\lambda), \tag{8}$$

and

$$\widehat{u}(0,\lambda) = \widehat{g}(\lambda), \tag{9}$$

for all $\lambda \in \mathbb{R}$, where $\widehat{u}(\cdot,\lambda)$ is an unknown function. Then by solving the equation (14) under the initial condition (15) (see [7]), we get

$$\begin{aligned}\widehat{u}(t,\lambda) &= \widehat{g}(\lambda)\left(1 - \lambda^2 t^\gamma E_{\gamma+1}(t)\right) \\ &\quad + \int_0^t \tau^{\gamma-1} E_\gamma(\tau)\, \widehat{f}(t-\tau,\lambda) d\tau.\end{aligned} \tag{10}$$

Lemma 1 *[8] Let $0 < \gamma < 1$. Then*

$$\frac{d}{dt}\{t^\gamma E_{1+\gamma}(t)\} = t^{\gamma-1} E_\gamma(t),\ t > 0.$$

Using Lemma 1 we can rewrite the formula (10) in the form:

$$\begin{aligned}\widehat{u}(t,\lambda) &= \widehat{g}(\lambda)\left(1 - \lambda^2 t^\gamma E_{\gamma+1}(t)\right) \\ &\quad + \int_0^t \frac{d}{d\tau}\{\tau^\gamma E_{1+\gamma}(\tau)\} \widehat{f}(t-\tau,\lambda) d\tau.\end{aligned} \tag{11}$$

Consequently, by using the inverse Dunkl transform (5) to (11), one obtains the solution of problem (1)–(2)

$$\begin{aligned}u(t,x) &= \int_{\mathbb{R}} \int_{\mathbb{R}} g(y)\left(1 - \lambda^2 t^\gamma E_{\gamma+1}(t)\right) D_\alpha(ix\lambda) D_\alpha(-iy\lambda) d\mu_\alpha(y) d\mu_\alpha(\lambda) \\ &+ \int_{\mathbb{R}} \int_{\mathbb{R}} \int_0^t \frac{d}{d\tau}\{\tau^\gamma E_{1+\gamma}(\tau)\} f(t-\tau,y) D_\alpha(ix\lambda) D_\alpha(-iy\lambda) d\tau d\mu_\alpha(y) d\mu_\alpha(\lambda).\end{aligned} \tag{12}$$

Here, we prove convergence of the obtained solution (12) and it's derivatives $\mathcal{D}_t^\gamma u(t,x)$, $\Lambda_\alpha^2 u(t,x)$. To prove the convergence of these, we use the estimate for the multivariate Mittag–Leffler function (7), obtained in [8], of the form

$$|E_{(\gamma-\gamma_1,\ldots,\gamma-\gamma_m,\gamma),1+\gamma}(a_1 t^{\gamma-\gamma_1}, \ldots, a_m t^{\gamma-\gamma_m}, -\lambda^2 t^\gamma)| \le \frac{C}{1+\lambda^2 t^\gamma}.$$

Let us to show absolute convergence the first term of (11):

$$\widehat{g}(\lambda)\left(1 - \lambda^2 t^\gamma E_{\gamma+1}(t)\right) \le C|\widehat{g}(\lambda)|.$$

Let us to show the convergence of the integral term of (11):

$$\begin{aligned} &\left| \int_0^t \frac{d}{d\tau}\{\tau^\gamma E_{1+\gamma}(\tau)\} \widehat{f}(t-\tau,\lambda) d\tau \right| \\ &\leq \max_{0\leq\tau\leq t} |\widehat{f}(t-\tau,\lambda)| t^{\gamma-1} E_\gamma(t) \\ &\leq \max_{0\leq t\leq T} |\widehat{f}(t,\lambda)| T^{\gamma-1} E_\gamma(T) \leq C \|\widehat{f}(\cdot,\lambda)\|_{C([0,T])}. \end{aligned} \tag{13}$$

Now using obtained above inequalities and in view of Plancherel theorem we have

$$\begin{aligned} \|u(t,\cdot)\|^2_{2,\alpha} = \|\widehat{u}(t,\cdot)\|^2_{2,\alpha} &= \int_{\mathbb{R}} |\widehat{u}(t,\lambda)|^2 d\mu_\alpha(\lambda) \\ &\leq C\|g\|^2_{2,\alpha} + C\|f\|^2_{C([0,T];L^2(\mathbb{R},\mu_\alpha))}. \end{aligned}$$

This implies $u(t,x) \in C([0,T]; L^2(\mathbb{R},\mu_\alpha))$. Similarly we can obtain

$$\begin{aligned} \|\Lambda^2_\alpha u(t,\cdot)\|^2_{2,\alpha} = \|\lambda^2 \widehat{u}(t,\cdot)\|^2_{2,\alpha} &= \int_{\mathbb{R}} |\lambda^2 \widehat{u}(t,\lambda)|^2 d\mu_\alpha(\lambda) \\ &\leq C\|g\|^2_{W^{2,2}_\alpha(\mathbb{R},\mu_\alpha)} + C\|f\|^2_{C([0,T];W^{2,2}_\alpha(\mathbb{R},\mu_\alpha))}. \end{aligned}$$

This implies immediately $\Lambda^2_\alpha u(t,x) \in C([0,T]; L^2(\mathbb{R},\mu_\alpha))$. Finally, we have

$$\begin{aligned} \|\mathcal{D}^\gamma_t u(t,\cdot)\|^2_{2,\alpha} = \|\mathcal{D}^\gamma_t \widehat{u}(t,\cdot)\|^2_{2,\alpha} &= \|\widehat{f}(t,\cdot) - \lambda^2 \widehat{u}(t,\cdot)\|^2_{2,\alpha} \\ &\leq C\|\widehat{f}(t,\cdot)\|^2_{2,\alpha} + C\|\lambda^2 \widehat{u}(t,\cdot)\|^2_{2,\alpha} \\ &\leq C\|g\|^2_{W^{2,2}_\alpha(\mathbb{R},\mu_\alpha)} + C\|f\|^2_{C([0,T];W^{2,2}_\alpha(\mathbb{R},\mu_\alpha))}. \end{aligned}$$

This is $\mathcal{D}^\gamma_t u(t,x) \in C([0,T]; L^2(\mathbb{R},\mu_\alpha))$. So we finally proved existence of the generalized solution of the problem (1)–(2) and it belongs to the class $u(t,x) \in C^\alpha([0,T], L^2(\mathbb{R},\mu_\alpha)) \cap C([0,T], W^{2,2}_\alpha(\mathbb{R},\mu_\alpha))$.

Now, we are in a position to show the uniqueness of the solutions. Suppose that there are two solutions u_1 and u_2 of the problem (1)–(2). Denote

$$u(t,x) = u_1(t,x) - u_2(t,x).$$

Then the function u satisfies the equation

$$\partial^\gamma_t u(t,x) - \sum_{j=1}^m a_j \partial^{\gamma_j}_t u(t,x) - \Lambda^2_{\alpha,x} u(t,x) = 0, \tag{14}$$

with homogeneous condition

$$u(0,x) = 0. \tag{15}$$

Then by applying the Dunkl transform $\mathcal{F}_\alpha$ to the problem (14)–(15) one obtains

$$\mathcal{D}^\gamma_{0+,t} \widehat{u}(t,\lambda) + \lambda^2 \widehat{u}(t,\lambda) = \widehat{f}(t,\lambda), \quad \widehat{u}(0,\lambda) = 0.$$

Then we have the trivial solution, i.e. $\widehat{u}(t,\lambda) \equiv 0$. Then by acting inverse Dunkl transform to this trivial solution we see that the solution u of the problem (14)–(15) is equal to zero. This means $u_1 \equiv u_2$. It contradicts to the our assumption, so the solution of the problem (1)–(2) is unique.

4 Conclusion

In this paper we showed existence and uniqueness of the solution to the (1)–(2) by using direct and inverse Dunkl transforms and using property of the multivariate Mittag-Leffler function. Further investigation problems will be application this technique for other type of equations such as Multi-term time-fractional derivative wave equation for Dunkl operator.

Acknowledgements

This research was funded by the Science Committee of the Ministry of Education and Science of the Republic of Kazakhstan (Grant No. AP08052425).

References

[1] Dachraoui A. "Psedidifferential-difference operators associated with Dunkl operators". *Integral Transforms and Special Functions*, (2001); 12(2): 161-178.

[2] Dunkl C.F. "Differential-difference operators associated to reflection group". *Trans. Amer. Math. Soc.*, (1989); 311: 167-183.

[3] Dunkl C.F. "Integral kernels with reflection group invariant". *Canad. J. Math.*, (1991); 43: 1213-1227.

[4] Dunkl C.F. "Hankel transforms associated to finite reflection groups". *Contemp. Math.*, (1992); 138: 123-138.

[5] De Jeu M.F.E. "The Dunkl transform". *Invent. Math.*, (1993); 113: 147-162.

[6] Soltani F. "L^p-Fourier multipliers for the Dunkl operator on the real line". *Journal of Function Analysis*, (2004); 209: 16-35.

[7] Luchko Y., Gorenflo R. "An operational method for solving fractional differential equations with the Caputo derivatives". *Acta Math. Vietnam.*, (1999); 24: 207-233.

[8] Li Z., Liu Y., Yamamoto M. "Initial-boundary value problems for multi-term time-fractional diffusion equations with positive constant coefficients". *Appl. Math. Comput.*, (2015); 257: 381–397.

[9] Kilbas A. A., Srivastava H. M., Trujillo J. J. "Theory and Applications of Fractional Differential Equations". *Mathematics studies North-Holland: Elsevier*, (2006).

[10] Podlubny I. "Fractional differential equations. An introduction to fractional derivatives, fractional differential equations, to methods of their solution and some of their applications". *Mathematics in Science and Engineering, 198. Academic Press, Inc., San Diego, CA*, (1999).

2-бөлім **Раздел 2** **Section 2**

Механика **Механика** **Mechanics**

IRSTI 30.17.35, 44.31.41, 27.35.25 DOI: https://doi.org/10.26577/JMMCS.2022.v115.i3.07

T. Amanzholov[1], Ye. Belyayev[1], M. Mohanraj[2], A. Toleukhanov[1*]
[1]Satbayev University, Kazakhstan, Almaty
[2]Hindusthan College of Engineering and Technology, India, Coimbatore
[*]e-mail: aman.toleukhanov@gmail.com

STUDY OF BOREHOLE HEAT EXCHANGER HEAT TRANSFER ENHANCEMENT PARAMETERS

This paper discusses the study of parameters for improving the heat transfer of a borehole heat exchanger for a ground source heat pump application. The study of efficiency parameters was carried out based on an experimental prototype of a ground source heat pump developed by the authors. A mathematical model has been developed for calculating the efficiency of a ground heat exchanger based on three-dimensional equations of heat and mass transfer in a porous medium. The numerical solution was carried out using the COMSOL Multiphysics software. The numerical calculation algorithm was verified by comparison with experimental data from the created prototype. Calculations were made of the efficiency of a borehole heat exchanger with various geometric configurations of the pipes in the well. With an increase in the tube diameter, the heat transfer increases. With a tube diameter of 40 mm, the thermal efficiency of the heat exchanger was 42.4 W/m in the heat charging mode, which is 24% more with a diameter of 20 mm. With increasing well depth, the heat transfer efficiency increases. The influence of the thermal conductivity coefficients of the pipe material, grout material and various types of ground on the heat transfer efficiency was also studied. It was shown that with an increase in the thermal conductivity coefficients of grout and ground, the heat flux increases, but above 6.0 W/m K, the heat flux practically does not change. When the coefficient of thermal conductivity of the pipe material is higher than 1.0 W/m K, the heat fluxes almost do not change. In general, materials containing plastics are used for piping of ground heat exchangers, the thermal conductivity coefficients of which vary between 0.24-0.42 W/m K.

Key words: borehole heat exchanger, ground source heat pump, thermal efficiency, heat and mass transfer in porous media, thermal conductivity, heat exchanger geometry, mathematical model, numerical solver.

Т. Е. Аманжолов[1], Е.К. Беляев[1], М. Моханрадж[2], А.Е. Толеуханов[1*]
[1]Сәтбаев университеті, Қазақстан, Алматы қ.
[2]Үндістан инженерия және технологиялар колледжі, Үндістан, Коимбатор қ.
[*]e-mail: aman.toleukhanov@gmail.com

Жерасты ұңғымалы жылу алмастырғыштың жылу беруін жақсарту параметрлерін зерттеу

Бұл жұмыста жер жылу сорғысында пайдалану үшін жерасты ұңғымалы жылу алмастырғышының жылу беруін жақсарту параметрлерін зерттеу қарастырылады. Тиімділік параметрлерін зерттеу авторлар әзірлеген жер жылу сорғысының тәжірибелік прототипі негізінде жүргізілді. Кеуекті ортадағы жылу мен масса алмасудың үш өлшемді теңдеулері негізінде жерасты жылу алмастырғыш өнімділігін есептеудің математикалық моделі жасалды. Сандық шешім COMSOL Multiphysics бағдарламалық жасақтамасы арқылы жүзеге асырылды. Сандық есептеу алгоритмі жасалған прототиптің тәжірибелік деректерімен салыстыру арқылы тексерілді. ұңғымадағы құбыршалардың әртүрлі геометриялық конфигурациялары бар жерасты ұңғымалы жылу алмастырғышының өнімділгі есептеулері жүргізілді. Құбыршаның диаметрі ұлғайан сайын жылу беру артады.

Құбырша диаметрі 40 мм, жылу алмастырғыштың жылу өнімділігі жылу айдау режимінде 42,4 Вт/м құрады, бұл диаметрі 20 мм болғандағы жағдайға қарағанда 24 % артық. Ұңғыманың тереңдігі артқан сайын жылу беру тиімділігі артады. Құбырша материалының, грут материалының және әр түрлі жерасты материалдар (топырақ) түрлерінің жылу өткізгіштік коэффициенттерінің жылу беру өнімділігіне әсері зерттелді. Жерасты бетон мен топырақ жылу өткізгіштік коэффициенттерінің жоғарылауымен жылу ағыны артады, алайда 6,0 Вт/м К-ден жоғары жылу ағыны айтарлықтай өзгермейді. Құбырша материалының жылу өткізгіштік коэффициенті 1,0 Вт/м К жоғары болғанда, жылу ағындары айтарлықтай өзгермейді. Жалпы, жылу өткізгіштік коэффициенттері 0,24-0,42 Вт/м К аралығында өзгеретін жерасты жылу алмастырғыштардың түтіктері үшін қңрамында пластмасса бар материалдар қолданылады.

Түйін сөздер: жерасты ұңғымалы жылуалмастрығыш, жер жылу сорғысы, жылу өнімділігі, кеуекті ортадағы жылу және масса тасымалы, жылуөткізгіштік, жылуалмастырғыш геометриясы, математикалық модель, сандық шешім.

Т.Е. Аманжолов[1], Е.К. Беляев[1], М. Моханрадж[2], А.Е. Толеуханов[1]*
[1]Satbayev University, Казахстан, г. Алматы
[2]Хиндустанский колледж инженерии и технологий, Индия, г. Коимбатор
*e-mail: aman.toleukhanov@gmail.com

Исследование параметров улучшения теплопередачи скважинного грунтового теплообменника

В данной работе рассматривается исследование параметров улучшения теплопередачи скважинного теплообменника для применения в грунтовом тепловом насосе. Исследование параметров эффективности проведено на основе разработанного авторами экспериментального прототипа грунтового теплового насоса. Разработана математическая модель расчета эффективности грунтового теплообменника на основе трехмерных уравнений тепломассопереноса в пористой среде. Численное решение было осуществлено на ПО COMSOL Multiphysics. Численный алгоритм расчета был верифицирован путем сравнения с экспериментальными данными из созданного прототипа. Проведены расчеты эффективности грунтового скважинного теплообменника с различными геометрическими конфигурациями расположения трубок в скважине. С увеличением диаметра трубки теплообмен увеличивается. При диаметре трубки 40 мм тепловая эффективность теплообменника составила 42,4 W/m в режиме закачки тепла, что на 24 % больше при диаметре 20 мм. С увеличением глубины скважины увеличивается эффективность теплопередачи. Исследовано влияние коэффициентов теплопроводности материала трубки, материала грута и различных типов грунта на эффективность теплообмена. С увеличением коэффициентов теплопроводности грута и грунта, увеличивается тепловой поток, однако выше 6,0 Вт/м К тепловой поток практически не меняется. При коэффициенте теплопроводности материала трубки выше 1,0 Вт/м К тепловые потоки практически не меняются. В основном для трубок грунтовых теплообменников используются материалы, содержащие пластик, коэффициенты теплопроводности которых варьируются между 0,24-0,42 Вт/м К.

Ключевые слова: грунтовый скважинный теплообменник, грунтовый тепловой насос, тепловая производительность, тепломассоперенос в пористой среде, теплопроводность, геометрия теплообменников, математическая модель, численный решатель.

1 Introduction

A heat pump system is more efficient when connected to a ground heat exchanger (GHE) than a conventional air source heat exchanger based heat pump. This is because the ground has a relatively more stable temperature and is generally warmer in winter and cooler in summer than the fluctuating ambient air temperatures. As a result, GHE as part of a ground source heat pump (GSHP) system is a critical element that determines its overall performance. GHEs are mainly classified as either horizontal or vertical according to their configurations. Vertical downhole GHEs, which also the so-called borehole heat exchangers (BHE), are more widely used comparing to other GHEs. Since BHEs can provide high heat transfer capacity on a limited surface area [1] and less influenced by ambient air temperature. On the other hand, when there is enough land and digging trenches is not difficult [2], horizontal GHE could be economically attractive since vertically well drilling is avoided. For climate conditions with a predominance of the heating season, horizontal GHEs are less suitable, because the influence of atmospheric air on such heat exchangers is significant. From this point of view, BHEs are more versatile. The most used BHEs are single U-shaped (one loop per well) and double U-shaped (two loops per well) heat pipes, which are used as part of a heat pump for heating and cooling [3].

Three main models for predicting the BHE heat transfer efficiency are widely available in the literature: analytical, numerical, and semi-analytical models. Compared to numerical models, analytical models are easier to implement. However, for simulations with small time intervals, discrete numerical models are most suitable. This is applicable for example for hourly energy analysis and optimal control of the GSHP depending on the local meteorological and hydrogeological conditions. Also, with the help of a computational tool, it is possible, for example, to simulate complex physical processes of heat and mass transfer in a porous medium. On the other hand, such calculations require large computational resources, especially for time variable year-round modeling and BHE life cycle modeling [4].

Analytical approaches include the line-source (LS) model [5] and the cylindrical-source (CS) model [6]. In general, the LS and CS models give a rough estimate of the actual heat transfer in the BHE; they are easy to implement and provide quick solutions. However, these models are limited to only radial conductive heat transfer and neglect heat transfer in the other coordinate direction. In addition, the BHE internal thermal resistance and heat capacity are neglected, which restricted wellbore thermal resistance prediction in short-term time interval. As a result, these models later improved by various researchers, for example in [7] non-uniform heat flow in a well was considered. As another approach, in [8] to estimate borehole wall temperature g-function dimensionless temperature response coefficient was proposed. The g-function provides the response of a single BHE to a single thermal step to predict the long-term performance of the GSHP. In [9] a more accurate two-dimensional soil heat transfer model, which is called a finite line-source (FLS) model was developed. As an improvement of basic analytical models, two-dimensional analytical (semi-analytical) models have been developed. Although they are still not suitable as a numerical simulation tool. For example, in [10] a robust two-dimensional analytical model considering U-shaped BHEs thermal interactions have been developed. In [10] by combining analytical and numerical approaches borehole thermal energy storage simulation model have been developed. Combination of analytical and numerical models for the double U-shaped BHE inner and outer regions have been studied in [11]. However, the BHE internal heat capacity was not considered, so the model cannot be applied to predict non-stationary heat transfer inside the BHE. In the recent decade, there have been research works on the study of heat transfer in single and double U-shaped GHEs. Few studies have been devoted to the study of heat transfer for more complex BHE geometries that consider the influence of thermal properties of the ground, heat exchanger material, grout, and other BHE parameters.

In [12] a detailed overview of the design aspects of various GHEs with an emphasis on improving performance and overall manufacturing costs. According to this research the most important factors influencing the GHE design are pipes geometrical configuration, GHE location, the wellbore and pipes length and diameter, pipe connections (serial or parallel), ground and grout properties, experimental methodology and mathematical modeling tool. In [13] the CFD tool was used to study the effect of linear spacing on the BHEs thermal performance with a shallow wellbore. It was reported that the temperature drop is the smallest at a small pitch and the maximum at the largest pitch of the shank. The authors also shoed that with increasing liner spacing the improvement in thermal performance decreases significantly. However, in their study, there is no effect of wellbore spacing on BHE performance with a large well depth, as well as a combined increase in the thermal conductivity of the cement slurry with other parameters. A simple analytical model for calculation of the average fluid temperature and hence improving the BHE thermal resistance estimation accuracy for a single U-shaped heat exchanger-based wellbore was proposed in [14]. The authors investigated the effect of well depth and volumetric flow rate on the estimation of RMS distribution between well thermal resistivity and efficient well thermal resistivity. Additionally, the relative deviation between the two resistances for the specific flow rate was estimated. However, a comprehensive sensitivity analysis regarding the combined effects of wellbore spacing, cement slurry thermal conductivity, well depth, wellbore and pipe diameters on the wellbore thermal resistance estimate has not been performed. As discussed above, most of the factors affecting BHE performance have been investigated. However, most of these research papers are not detailed as they deal with the influence of only one or a few parameters. A comprehensive analysis of all major influencing parameters along with a comparative performance analysis of single and double U-shaped BHEs using the same simulation model is lacking in the literature. A detailed analysis of the influencing factors, combined with a comparison of the thermal performance of single- and double-pipe BHEs in terms of thermal efficiency, heat transfer per unit wellbore depth, and wellbore thermal resistance under various conditions, is missing from the previous reports.

Therefore, in this research, a numerical analysis of the thermal performance of BHEs was carried out based on a verified mathematical and numerical model of heat transfer. The combined effect of borehole diameter, borehole depth, pipe diameter, grout/ground/pipe material thermal conductivity on the thermal performance of BHE was studied, including with various geometric configurations of the piping arrangement. Geometric configurations include single U-shaped, double U-shaped and spiral types of heat exchangers. The analysis was carried out and conclusions were drawn on the influence of these parameters on the efficiency of heat transfer between BHE and the surrounding ground both heat charging and discharging modes.

2 Physical formulation

Heat flow in a geothermal system includes heat conduction and convection occurring in well heat exchangers and the surrounding soil mass. Thermal conductivity in the soil mass occurs as a result of the transfer of thermal energy due to temperature gradients between the bottom layers of the earth, air and borehole heat exchangers. Thermal convection occurs as a result of diffusion and advection of heat due to the flow of groundwater. Temperatures and temperature gradients in geothermal systems are relatively low, on the order of 5-30°C. In the presence of groundwater, the soil mass is considered as a saturated two-phase porous material consisting of solid particles and water. Dry soil is considered as a single-phase material.

The borehole heat exchanger is one of the most important components of a ground source heat pump system. Due to the complex nature of heat transfer in BHE, an efficient thermal design that meets the required requirements is a challenge. When designing a BHE, thermal performance is an important parameter that determines the effective transfer of heat between the ground and the system. Moreover, the thermal performance of the BHE also determines the operational efficiency and operating costs of the system integrated in the BHE. In this regard, a comprehensive study of BHE is needed, especially from the point of view of the complex impact of factors affecting its thermal performance. In addition to the thermal performance of the BHE, the wellbore thermal resistance, which is an important parameter in the design and analysis of BHE heat transfer, will also be analyzed for the single U-tube BHE; and a comparative analysis will be carried out between different types of BHE in terms of heat transfer, efficiency and thermal resistance of the wellbore. The thermal performance of a BHE is influenced by various factors: geometric, thermal, geological and operational parameters. This article will discuss the combined effect of the main parameters that affect the performance of a BHE single U-tube heat exchanger.

In this work, a new model was described for simulating downhole heat exchangers consisting of a single U-tube. In the first part, the theory of building a finite element of a downhole heat exchanger was presented. The work begins with the definition of the general equation for the balance of flow and heat transfer within each element of the heat exchanger. The generated numerical model is for single U exchangers. It can also be adapted for two- or multi-pipe BHEs as shown in the results of this article. The downhole heat exchanger is modeled as a one-dimensional finite element with many degrees of freedom. It is necessary to take into account the heat exchange between the individual sections of the heat exchanger. To obtain a more accurate model, the division of the region into three subregions was introduced. The first area is the pipe, which we model as a line, the second area is the cement, which we model as a solid, and the third, the soil, which we model as a solid with porosity.

3 General equations

The nonisothermal pipe flow is used to compute the temperature, velocity, and pressure fields in pipes and channels of different shapes. It approximates the pipe flow profile by 1D assumptions in curve segments, or lines. These lines drawn in 3D and represent simplifications of hollow tubes.

The heat equation to model nonisothermal pipe flow:

$$\frac{\partial \rho_f}{\partial t} + \nabla(\rho_f u) = 0,, \tag{1}$$

$$\rho_f \frac{\partial u}{\partial t} = -\nabla p - \frac{1}{2} f_D \frac{\rho_f}{d_h} |u| u + F, \tag{2}$$

where ρ_f - density of fluid (kg/m^3),u$-velocity(m/s)$,p$-pressure(Pa)$,f$_D$ - Darcy friction factor, F - volume force (H), d_h - parametric value (m).

The heat equation to model nonisothermal pipe heat transfer:

$$\rho_f c_{p,f} \frac{\partial T_f}{\partial t} + \rho_f c_{p,f} u \nabla T_f = \nabla(k_f \nabla T_f) + \frac{1}{2} f_D \frac{\rho}{d_h} |u| u^2 + Q_w, \tag{3}$$

where $c_{p,f}$ - specific heat capacity of fluid (J/(kg K)), k_f - thermal conductivity of fluid (W/(m K)), T_f - temperature of fluid (K), Q_w - wall heat source (J).

In this work the wall heat transfer node to set up heat exchange across the pipe wall was used for define the external temperature and the nature of the heat transfer.

$$Q_w = (hZ)_{eff}(T_{ext} - T_f), \tag{4}$$

where $(hZ)_{eff}$ is an effective value of the heat transfer coefficient h (W/(m^2 K)) times the wall perimeter Z (m) of the pipe. T_{ext} (K) the external temperature outside of the pipe. Q_{wall} appears as a source term in the pipe heat transfer equation.

$$(hZ)_{eff} = \frac{2\pi}{\frac{1}{r_0 h_{int}} + \frac{1}{r_N h_{ext}} + \Sigma(\frac{ln(\frac{r_n}{r_n - 1})}{k_{wall,n}})}, \tag{5}$$

where r_n is the outer radius of wall (m), r_0 - inner radius (m), Z - is the outer perimeter of wall (m), h_{int} and h_{ext} are the film heat transfer coefficients on the inside and outside of the tube, respectively $(W/(m^2K))$.

The heat transfer in solids is used to model heat transfer in solids by conduction, convection, and radiation. The temperature equation defined in solid domains corresponds to the differential form of the Fourier's law that may contain additional contributions like heat sources.

The heat equation to model heat transfer in solids:

$$\rho_s c_{p,s} \frac{\partial T_s}{\partial t} + \rho_s c_{p,s} u \nabla T_s = \nabla(k_s \nabla T_s), \tag{6}$$

where ρ_s - density of solid (kg/m^3),c$_{p,s}$ - specific heat capacity of solid (J/(kg K)), k_s - thermal conductivity of solid (W/(m K)), T_s - temperature of solid (K), u - velocity (m/s).

The heat transfer in porous media is used to model heat transfer by conduction, convection, and radiation in porous media. The temperature equation defined in porous media domains corresponds to the convection-diffusion equation with thermodynamic properties averaging models to account for both solid matrix and fluid properties. This equation is valid when the temperatures into the porous matrix and the fluid are in equilibrium. The heat equation to model heat transfer in porous media:

$$(\rho c_p)_{eff} \frac{\partial T_p}{\partial t} + (\rho c_p)_{eff} u \nabla T_p = \nabla(k_{eff} \nabla T_p), \tag{7}$$

$$(\rho c_p)_{eff} = \theta \rho_s c_{p,s} + (1 - \theta) \rho_f c_{p,f}, \tag{8}$$

$$k_{eff} = \theta k_s + (1 - \theta) k_f, \tag{9}$$

where T_p - temperature of porous media (K), θ - porosity, ρ_s - density of solid (kg/m^3),c$_{p,s}$ - specific heat capacity of solid (J/(kg K)), k_s - thermal conductivity of solid (W/(m K)), $c_{p,f}$ - specific heat capacity of fluid (J/(kg K)), k_f - thermal conductivity of fluid (W/(m K)), ρ_f - density of fluid (kg/m^3).

3.1 Initial and boundary condition

Two modes of operation are considered: ground charging and discharging. Ground charging is understood as heat transfer to the ground, where the working fluid inlet temperature is set as 45 °C. Ground discharging is understood as heat transfer from the ground, i.e., heat extraction. In discharging mode working fluid inlet temperature is set as 5 °C.

As an initial condition for the ground temperature constant undisturbed soil temperature is assumed. The undisturbed soil temperature is equal to 15 °C for both charging and discharging cases.

At the boundaries of the computational domain constant temperature is assumed because there is no influence of temperature boundary condition on the BHE temperature distribution.

4 Results and discussion

The numerical implementation of the indicated mathematical model with the corresponding initial and boundary conditions was carried out on the COMSOL Multiphysics software. To verify this numerical tool, a comparison was made with the experimental data of the thermal response test [15]. Figure 1 shows this comparison. According to Figure 1, the comparison was carried out according to the working fluid temperature T_{out}. The relative error does not exceed 2-3%, which indicates a very good agreement between the results.

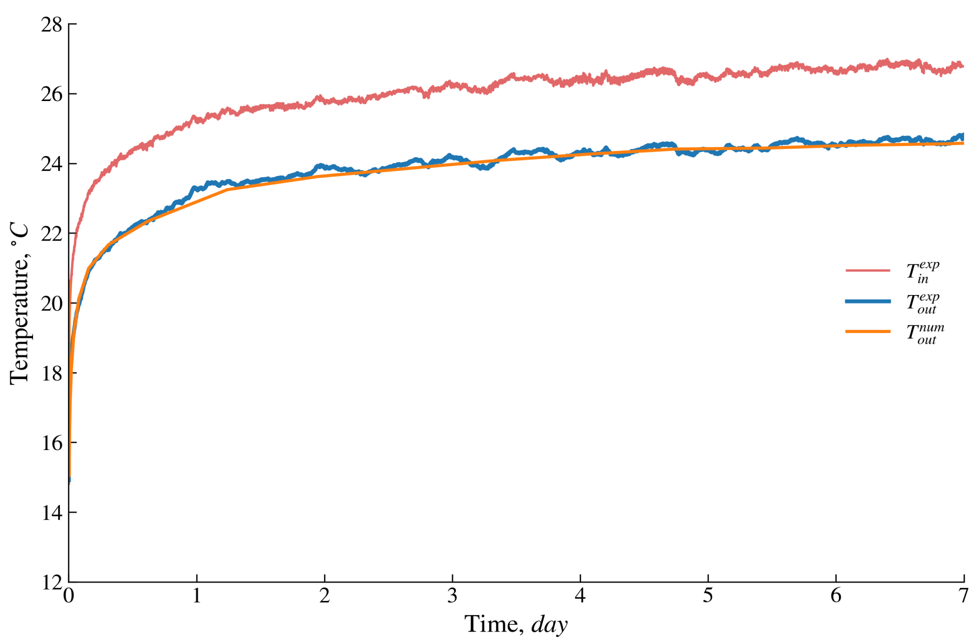


Figure 1: Verification of the numerical calculation algorithm

To study the efficiency of various heat exchanger configurations, 4 types were selected: (1) single U - shaped (U), (2) - double U-shaped cross (dU-x), (3) - double U-shaped parallel (dU-u), and (4) spiral (Spiral). Figure 2 shows these geometrical configurations. These configurations have been proposed to increase BHE thermal efficiency. Of course, the most common is the single U-configuration, whereas more complex versions are costly and laborious to install. However, with the use of more complex configurations, it is possible to save on the depth of drilling a well.

Wellbore depth is one of the important BHE geometrical parameters that affects the total amount of heat supplied (in cooling mode) and removed (in heating mode) to/from the well. Hence, it is very important to investigate its impact on BHE performance. In order to obtain the result of thermal performance in response to a change in well depth, the input parameters of the numerical model, which were fixed, are specified as: distance between inlet and outlet pipes within one BHE, X_c = 0.05 m, well radius r_b = 0.017 m, soil thermal conductivity k_s = 1.2 W/m.K, working fluid inlet temperature $T_{f,in}$ = 45 °C , 5 °C for charging and discharging modes, flow rate $\dot{m}$= 0.6 m^3/h, pipe thermal conductivity k_p = 0.4 W/m.K. Figure 3 illustrates the effect of well depth on the overall heat transfer rate and thermal efficiency of the BHE. With the increasing well depth, the heat transfer per unit length of the well tends to decrease, while thermal efficiency improves significantly. The deeper the well depth, the smaller the temperature difference between the working fluid and the surrounding soil, and this leads to a decrease in the heat transfer rate per unit depth of the well. The increase in thermal performance may be since as the well depth increases, more heat enters the well (in cooling mode); consequently, the outlet liquid temperature decreases. This, in turn, increases the difference between the inlet and outlet temperature of the working fluid, which leads to an increase in the thermal efficiency of the BHE. However, with a deep depth of the well, it is not economically feasible. This is due to an increase in drilling cost (which depends on geological conditions) and installation cost, as well as the cost of materials. As a result, when designing with a large well depth, a trade-off must be found between thermal performance and total cost. In addition, a BHE with a large well depth requires more pump power to circulate the working fluid and therefore requires more electricity consumption, which again leads to increased costs.

Pipe diameter is another factor to consider when investigating the impact of pipe parameters on BHE performance. The effect of pipe diameter on BHE thermal performance is briefly discussed here. Conventional pipe outer diameters (from 15 mm to 40 mm) were taken to evaluate the effect of pipe diameter on heat transfer rate, thermal efficiency and thermal resistance of the wellbore. The effect of pipe diameter on the overall heat transfer coefficient per unit depth of the well and the thermal efficiency of the BHE is shown in Figure 4. Heat transfer rates and efficiency increase with larger BHE pipe diameters, especially in high thermal conductivity grounds. BHE with pipe diameter 40 mm has the highest heat transfer rate and thermal efficiency than BHE with pipe diameter 25 mm and 32 mm. The average heat transfer coefficient per unit of well depth and thermal efficiency of BHE with 40 mm pipe is 42.4 W/m (charging mode), which is higher than that of BHE with 25 mm pipe diameter. Thus, according to Figure 4, BHEs with a larger diameter pipe are more efficient and improve the transfer of more heat. This can be explained by a change in the heat transfer area of the BHE with a well configuration and a change in the pipe diameter. Therefore, convective heat transfer improves as the heat exchange surface area increases.

BHE consists of a U-shaped pipe, grout material, and, accordingly, the BHE surrounding ground. Since pipe and grout materials are considered solid, the influence of their thermal conductivity coefficient on the BHE thermal efficiency should be considered. Since surrounding ground is a porous medium with predominantly conductive heat transfer

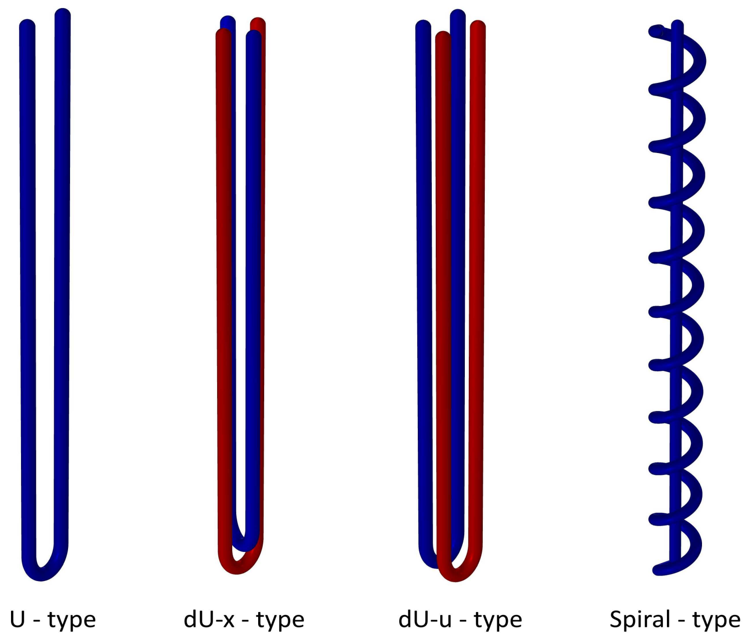


Figure 2: BHE geometrical configurations

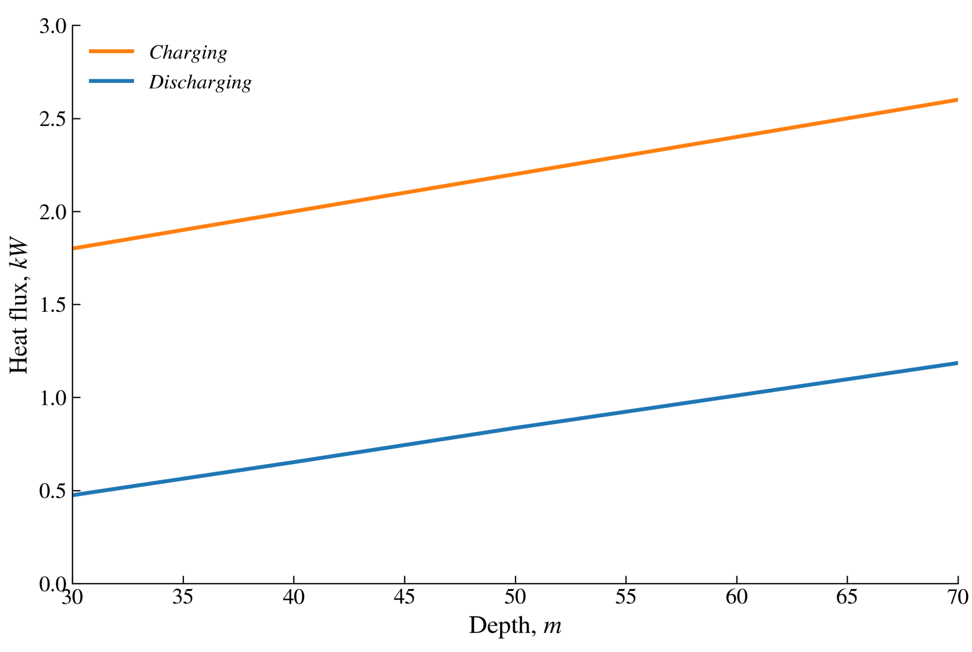


Figure 3: Influence of well depth on heat fluxes in BHE

mechanism, then influence of the ground thermal conductivity is also interesting to test. Since the surrounding ground is a porous medium with a predominantly conductive heat transfer mechanism, the influence of the ground's thermal conductivity is also interesting to test. Figure 5 shows the influence of mentioned thermal conductivity coefficients on the BHE

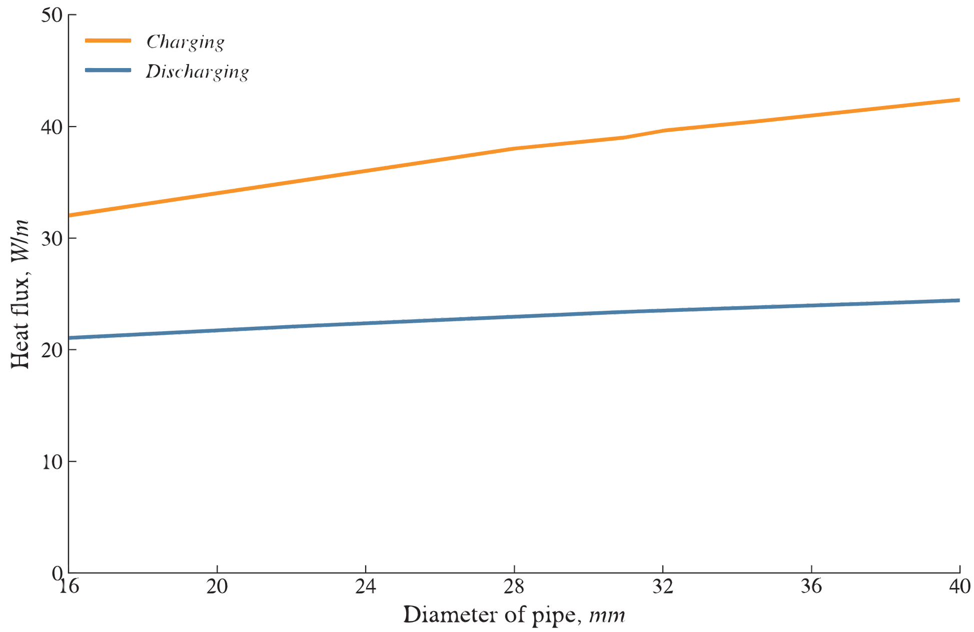


Figure 4: Influence of pipe diameter on heat fluxes in BHE

heat fluxes. A high-density polyethylene (HDPE), polyvinyl chloride (PVC), polyethylene, polyamide, steel, and copper are some of the common piping materials. Grout material is the cement slurry, which is in the ratio of 70% - water, 24% - cement, and 6% - bentonite. Underground materials could be unconsolidated ground type (clay/silt, sand, gravel/stones, till/loam), sedimentary rocks (clay/silt stones, limestones, dolomitic rocks, etc.), magmatic and metamorphic rocks (basalt, granite, quartzite, etc.). According to Figure 5, with the piping material's thermal conductivity above 1.2 W/m K, there is no change in heat flux. It is known that the HDPE, PVC, and polyethylene thermal conductivity is less than 1.0 W/m K, and because of the flexibility, durability, service life, and the piping material cost they are the most used in BHE. According to Figure 5, the influence of the thermal conductivity coefficients of the grout and ground material is almost the same. This means that with an increase in the thermal conductivity, heat fluxes increase, but above 6.0 W/m K this change is insignificant.

Additionally, calculations were carried out on the effect of the well diameter on the thermal efficiency of BHE. With an increase in the borehole diameter from 100 mm to 200 mm, the BHEs heat transfer increased by 5.5 W/m; on the other hand, the corresponding thermal efficiency is somewhat reduced by 3.7 % for BHE. The result shows that as the well diameter increases, more heat can be injected into the well as the heat transfer area increases. However, the improvement in thermal performance with borehole diameter is not as significant as the change in thermal performance with parameters such as borehole depth, inlet fluid temperature, and soil thermal conductivity. However, from an economic standpoint, a BHE with a larger borehole diameter may have a higher capital cost and therefore may not be feasible compared to a BHE with a smaller borehole diameter.

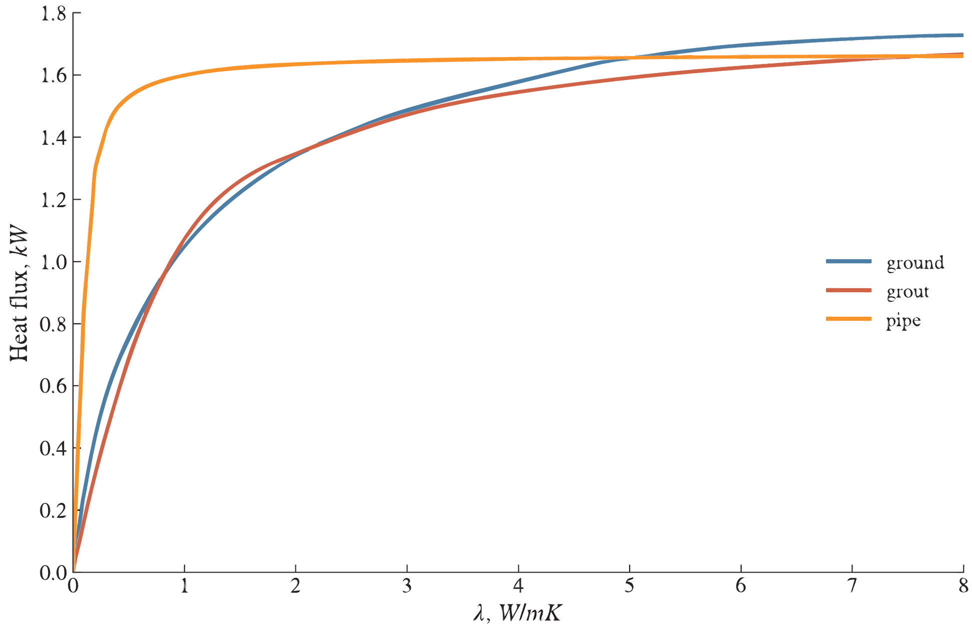


Figure 5: Influence of different thermal conductivity coefficients on heat fluxes in BHE

5 Conclusion

This paper discusses the study of parameters for improving the heat transfer of a borehole heat exchanger for a ground source heat pump application. The study of efficiency parameters was carried out based on an experimental prototype of a ground source heat pump developed by the authors. A mathematical model has been developed for calculating the efficiency of a ground heat exchanger based on three-dimensional equations of heat and mass transfer in a porous medium. The numerical solution was carried out using the COMSOL Multiphysics software. The numerical calculation algorithm was verified by comparison with experimental data from the created prototype. Calculations were made of the efficiency of a downhole heat exchanger with various geometric configurations of the pipes in the well. The study of the influence of the tube diameter on the heat transfer efficiency showed that with an increase in the tube diameter, the heat transfer increases. With a tube diameter of 40 mm, the thermal efficiency of the heat exchanger was 42.4 W/m in the heat charging mode, which is 24% more with a diameter of 20 mm. It has also been shown that with increasing well depth, the heat transfer efficiency increases. However, it is not possible to excessively increase the depth of the well and the diameter of the pipe for economic reasons. The influence of the thermal conductivity coefficients of the pipe material, grout material and various types of ground on the heat transfer efficiency was also studied. It was shown that with an increase in the thermal conductivity coefficients of grout and ground, the heat flux increases, but above 6.0 W/m K, the heat flux practically does not change. When the coefficient of thermal conductivity of the pipe material is higher than 1.0 W/m K, the heat fluxes almost do not change. In general, materials containing plastics are used for piping of ground heat exchangers, the thermal conductivity coefficients of which vary between 0.24-0.42 W/m K.

6 Acknowledgement

This research is funded by the Science Committee of the Ministry of Education and Science of the Republic of Kazakhstan Grant No. AP08857319 "Study of heat transfer enhancement mechanisms of vertical type borehole heat exchanger to ensure high heat pump performance". Postdoctoral Research Programme, Al-Farabi Kazakh National University, Almaty, Kazakhstan.

References

[1] Florides G., Kalogirou S., "Ground heat exchangers: a review of systems, models and applications" , *Renew. Energy* 32 (2007): 2461-2478.

[2] Javed S., Spitler J., "ccuracy of borehole thermal resistance calculation methods for grouted single U-tube ground heat exchangers" , *Applied Energy* 187 (2017): 790-806.

[3] Fang L., Fang Zh., Zhu K., "Study on the efficiency of single and double U-tube heat exchangers" , *Procedia Engineering* 205, (2017): 4045-4051.

[4] Ma Zh., Xia L., Gong X., Kokogiannakis G., Wang Sh., Zhou X., "Recent advances and development in optimal design and control of ground source heat pump systems" , *Renewable and Sustainable Energy - Reviews* 131 (2020): 110001

[5] Ingersoll L.R., Zoel O.J., Ingersoll A.C., "Heat Conduction: with Engineering, Geological and Other Applications" , *Madison: Madison University of Wisconsin Press* (1954): 325

[6] Carslaw, H.S., Jaeger, J.C., "Conduction of Heat in Solids" , *Claremore Press - Oxford* (1959).

[7] Soni S.K., Pandey M., Bartaria V.N., "Around coupled heat exchangers: a review and applications" , *Renew. Sustain. Energy Rev.* 47 (2015): 83-92.

[8] Hecht-Mendez J., de Paly M., Beck M., Bayer P., "Optimization of energy extraction for vertical closed-loop geothermal systems considering groundwater flow" , *Energy Convers Manag.* 66 (2013): 1-1.

[9] Zeng H.Y., Diao N.R., Fang Z.H., "A finite line-source model for boreholes in geothermal heat exchangers" , *Heat Transf. Asian Res.* 31 (2002).

[10] Pu L., Qi D., Li K., Tan H., Li Y., "Simulation study on the thermal performance of vertical U-tube heat exchangers for ground source heat pump system" , *Appl. Therm. Eng.* 79 (2015): 202-213.

[11] Salilih E., Abu-Hamdeh N., Oztop H., "Analysis of double U-tube ground heat exchanger for renewable energy applications with two-region simulation model by combining analytical and numerical techniques" , *Int. Commun. Heat Mass Transfer* 123 (2021).

[12] Molina-Giraldo N., Blum P., Zhu K., Bayer P., Fang Z., "A moving finite line source model to simulate borehole heat exchangers with groundwater advection" , *Int J. Therm. Sci.* 50 (2011): 2506-2513.

[13] Madani H., Claesson J., Lundqvist P., "Capacity control in ground source heat pump systems part II: comparative analysis between on/off controlled and variable capacity systems" , *Int. J. Refrig.* 34 (2011): 1934-1942.

[14] Badescu V., "Economic aspects of using ground thermal energy for passive house heating" , *Renew. Energy* 32 (2007): 895–903.

[15] Amanzholov T., Akhmetov B., Georgiev A., Kaltayev A., Popov R., Dzhonova-Atanasova D., Tungatarova M., "Numerical modelling as a supplementary tool for Thermal Response Test" , *Bulg. Chem. Commun.* 48 (2016): 109-114.

Список литературы

[1] Florides G., Kalogirou S., "Ground heat exchangers: a review of systems, models and applications" , *Renew. Energy* 32 (2007): 2461-2478.

[2] Javed S., Spitler J., "ccuracy of borehole thermal resistance calculation methods for grouted single U-tube ground heat exchangers" , *Applied Energy* 187 (2017): 790-806.

[3] Fang L., Fang Zh., Zhu K., "Study on the efficiency of single and double U-tube heat exchangers" , *Procedia Engineering* 205, (2017): 4045-4051.

[4] Ma Zh., Xia L., Gong X., Kokogiannakis G., Wang Sh., Zhou X., "Recent advances and development in optimal design and control of ground source heat pump systems" , *Renewable and Sustainable Energy - Reviews* 131 (2020): 110001

[5] Ingersoll L.R., Zoel O.J., Ingersoll A.C., "Heat Conduction: with Engineering, Geological and Other Applications" , *Madison: Madison University of Wisconsin Press* (1954): 325

[6] Carslaw, H.S., Jaeger, J.C., "Conduction of Heat in Solids" , *Claremore Press - Oxford* (1959).

[7] Soni S.K., Pandey M., Bartaria V.N., "Around coupled heat exchangers: a review and applications" , *Renew. Sustain. Energy Rev.* 47 (2015): 83-92.

[8] Hecht-Mendez J., de Paly M., Beck M., Bayer P., "Optimization of energy extraction for vertical closed-loop geothermal systems considering groundwater flow" , *Energy Convers Manag.* 66 (2013): 1-1.

[9] Zeng H.Y., Diao N.R., Fang Z.H., "A finite line-source model for boreholes in geothermal heat exchangers" , *Heat Transf. Asian Res.* 31 (2002).

[10] Pu L., Qi D., Li K., Tan H., Li Y., "Simulation study on the thermal performance of vertical U-tube heat exchangers for ground source heat pump system" , *Appl. Therm. Eng.* 79 (2015): 202-213.

[11] Salilih E., Abu-Hamdeh N., Oztop H., "Analysis of double U-tube ground heat exchanger for renewable energy applications with two-region simulation model by combining analytical and numerical techniques" , *Int. Commun. Heat Mass Transfer* 123 (2021).

[12] Molina-Giraldo N., Blum P., Zhu K., Bayer P., Fang Z., "A moving finite line source model to simulate borehole heat exchangers with groundwater advection" , *Int J. Therm. Sci.* 50 (2011): 2506-2513.

[13] Madani H., Claesson J., Lundqvist P., "Capacity control in ground source heat pump systems part II: comparative analysis between on/off controlled and variable capacity systems" , *Int. J. Refrig.* 34 (2011): 1934-1942.

[14] Badescu V., "Economic aspects of using ground thermal energy for passive house heating" , *Renew. Energy* 32 (2007): 895–903.

[15] Amanzholov T., Akhmetov B., Georgiev A., Kaltayev A., Popov R., Dzhonova-Atanasova D., Tungatarova M., "Numerical modelling as a supplementary tool for Thermal Response Test" , *Bulg. Chem. Commun.* 48 (2016): 109-114.

IRSTI 30.15.15 DOI: https://doi.org/10.26577/JMMCS.2022.v115.i3.08

A.B. Kydyrbekuly[1], **G.E. Ibrayev**[2*], **S.A. Rakhmatullayeva**[3]
[1]Research Institute of Mathematics and Mechanics, Kazakhstan, Almaty
[2]Al-Farabi Kazakh National University, Kazakhstan, Almaty
[3]L.N. Gumilev Eurasian National University, Kazakhstan, Astana
*e-mail: ybraev.alysher@mail.ru

NONLINEAR VIBRATIONS OF THE "ROTOR – JOURNAL BEARINGS" SYSTEM

The equations of motion of a rotor system mounted on journal bearings with a non-linear characteristic are solved by high-precision analytical methods. A new technique has been developed for solving nonlinear differential equations of motion of rotor systems mounted on journal bearings, taking into account nonlinearity of reaction forces of the lubricating layer.Algebraic systems of equations were obtained that allow us to determine amplitudes of nonlinear oscillations of the rotor and supports, and construct the amplitude-frequency characteristics of the system for varying parameters of the rotor, supports and fluid depending on the angular velocity of the rotor. The conditions and frequency intervals for the presence of self-oscillations of the rotor and supports were determined. The amplitude-frequency characteristics of the nonlinear oscillations of the rotor system are obtained, taking into account nonlinearity of characteristics of journal bearings.The optimal parameters depending on the size of the gap and the oil film, the mass of the supports, the fluids used as a lubricating layer in the journal bearing, with rigidity and damping coefficients, at which the magnitudes of the amplitudes of self-excited oscillations have optimal values, are obtained.

Key words: Nonlinear Vibrations, Harmonic Balance Method, Journal Bearing, Sommerfeld's Hypothesis, Rotor System, Self-Excited Vibrations.

Қыдырбекұлы А.Б.[1], Ибраев Г.Е.[2*], Рахматуллаева С.А.[3]
[1]Математика және механика ғылыми-зерттеу институты, Қазақстан, Алматы қ.
[2]Әл-Фараби атындағы Қазақ ұлттық университеті, Қазақстан, Алматы қ.
[3]Л.Н. Гумилев атындағы Еуразиялық ұлттық университеті, Қазақстан, Астана қ.
*e-mail: ybraev.alysher@mail.ru

"Ротор – сырғу мойынтіректері" жүйесінің бейсызық тербелістері

Жоғары дәлдікті аналитикалық әдістермен сипаттамасы бейсызық болып табылатын сырғу мойынтіректерінде орнатылған роторлық жүйенің қозғалыс теңдеулері шешілді.Майлау қабаты реакция күштерінің бейсызықтығын ескере отырып, сырғу мойынтіректерінде орнатылған роторлық жүйелер қозғалысының бейсызық дифференциалдық теңдеулерін шешудің жаңа әдістемесі жасалды. Ротор мен тіректердің бейсызық тербелістерінің амплитудасын анықтауға және ротордың бұрыштық жылдамдығына қатысты кезіндегі ротордың, тіректердің және сұйықтықтың параметрлерін вариациялау кезінде жүйенің амплитудалық-жиілік сипаттамаларын құруға мүмкіндік беретін алгебралық теңдеулер жүйесі алынды. Сырғу мойынтіректерінің сызықты емес сипаттамаларын ескере отырып роторлық жүйенің бейсызықты тербелістерінің амплитудалық-жиіліктік сипаттамалары тұрғызылды. Жүйенің өздігінен қозатын тербелістер амплитудасының мәні оптимальді мәнге ие болатындай саңылаудың қалыңдығы мен майлауқабаты, тіректердің массасы, сырғумойынтірегінде майлау қабаты ретінде қолданылатын сұйықтықпен, қатаңдықжәне демпферлік коэффициенттермен байланысты оптимальді параметрлер анықталды.

Түйін сөздер: Бейсызық тербелістер, гармоникалық баланс әдісі, сырғу мойынтірегі, Зоммерфельд гипотезасы, роторлық жүйе, өздігінен қозатын тербелістер.

Кыдырбекулы А.Б.[1], Ибраев Г.Е.[2*], Рахматуллаева С.А.[3]
[1]Научно-исследовательский институт математики и механики, Казахстан, г. Алматы
[2]Казахский национальный университет имени аль-Фараби, Казахстан, г. Алматы
[3]Евазийский национальный университет имени Л.Н. Гмилева, Казахстан, г. Астана
*e-mail: ybraev.alysher@mail.ru

Нелинейные колебания системы "Ротор - подшипники скольжения"

Решены уравнения движения роторной системы, установленных на подшипниках скольжения с нелинейной характеристикой высокоточными аналитическими методами. Разработана новая методика решений нелинейных дифференциальных уравнений движения роторных систем, установленных на подшипниках скольжения, с учетом нелинейности сил реакций смазочного слоя. Были получены алгебраические системы уравнений, позволяющие определить амплитуды нелинейных колебаний ротора и опор, и построить амплитудно-частотные характеристики системы при варьировании параметров ротора, опор и жидкости в зависимости от угловой скорости ротора. Были определены условия и интервалы частот наличия автоколебаний ротора и опор. Построены амплитудно-частотные характеристики нелинейных колебаний роторной системы, с учетом нелинейности характеристик подшипников скольжения. Определены оптимальные параметры связанные с толщиной зазора и масленой пленки, массой опор, жидкости использующиеся в качестве смазочного слоя в подшипнике скольжения, с коэффициентами жесткости и демпфирования, при которых величины амплитуд самовозбуждающихся колебаний имеют оптимальные значения.

Ключевые слова: Нелинейные колебания, метод гармонического баланса, подшипник скольжения, гипотеза Зоммерфельда, роторная система, самовозбуждающиеся колебания.

1 Introduction

Journal bearings have a number of significant advantages over rolling bearings. They are resistant to a wide range of loads and dynamic disturbances, capable of operating at higher rotational speeds, have a long service life and low cost, and are easy to operate.

Due to specific properties of hydrodynamic forces caused by the presence of a lubricating layer during rotation of the rotor in journal bearings, self-excited oscillations (self-oscillations) with large amplitudes can arise in a wide range of rotation speeds. Therefore, it is often necessary to develop suppression measures in industry and production and study the behavior of this type of oscillation depending on various physical and geometric parameters of the system.

2 Literature review

At present, journal bearings, used in many rotary machines as key elements and serving to transfer rotational energy, are complex elements for dynamic analysis since under certain geometric and operating parameters they can cause, as mentioned above, self-excited [1-3], parametric [3, 4] and chaotic oscillations [4, 5]. As at operating frequencies of the system similar to the model considered in this paper, self-excited oscillations often occur, the paper studies the conditions for occurrence and further behavior of these oscillations.

One of the first researchers who studied the phenomenon of self-excitation and the reasons for its occurrence was Newkirk in 1924 [6]. Together with Taylor, he conducted the first experimental study of this phenomenon and explained the causes of self-excited oscillations [7]. When studying self-oscillations, in many cases the problem is reduced to studying precessional motion of the system. Approximate solutions, assuming that the load on the stud is sufficiently small, were first obtained by Hagg [8] and Yukio Hori [9]. Works on the analysis of the precessional movement of the stud in the oil-filled bearing were also carried out by Kesten [10].

Conditions for stability of the equilibrium position of the rotor system mounted on journal bearings, as well as the nature of unsteady motion in an unstable position, were studied by Someya [11]. Experimental studies of these phenomena were also carried out by such authors as Hagg, Boecker, Schnittger and Hori [8, 9, 12, 13]. Different results were obtained concerning the influence of oil viscosity and the size of backlash in the bearing. Some authors such as Schnittger have noted the benefits of low viscosity as it contributes to stud stability. Other authors such as Boecker, Schnittger and Pinkus [14] noted that high viscosity is more conducive to stability. According to the third group of authors, such as Hummel [15] and Hagg, both of the above cases are equivalent. Different points of view are also observed when studying the effect of bearing width on system dynamics. However, researchers agree that the unbalance of the rotor has no effect on the occurrence and intensity of self-excited oscillations. Some authors obtained different frequency of self-excited oscillations [16-19]. For most authors, the frequency of self-excited oscillations coincided with the natural frequency of the rotor, in some cases, for example, Pinkus, it increased with increasing speed, while Schnittger experimentally obtained results in which the frequency curve first decreased and then began to increase [13, 14].

Experimental studies of self-excited oscillations as a whole showed not only the complexity of this problem, but also revealed a number of specific features of this phenomenon. The most important of the identified effects is "inertia" (dragging), i.e. self-excited oscillations, after arising at a certain frequency, continue to exist even when the rotor speed decreases below the frequencies of occurrence of self-excited oscillations [20, 21-23]. Another feature is the possibility of occurrence of self-excited oscillations under the action of a short-term pulse, for example, a blow to the rotor, at speeds that are lower than the characteristic speeds at which self-excited oscillations arise [24, 25].

3 Statement of the problem and equations of motion

Consider a vertical solid rotor of mass m symmetrically mounted on a flexible shaft with respect to supports. The shaft is mounted on elastic supports. The rotor system rotates on journal bearings of mass m_0 with an angular velocity ω (Figure 1). Equivalent rigidity of the elastic field of supports is c; δ is the size of the clearance in the bearing; t is the oil temperature in the bearing; μ is the oil viscosity in the bearing; d is the diameter of the bearing spike; L is the length of the bearing; D is the bearing diameter; l is the length of the shaft; k_1, k_2 are damping coefficients; e is the rotor unbalance.

To derive the equations of motion, we introduce the fixed coordinate system Oxy. Let in this system x_1, y_1 be coordinates of O_1 (the center of the elastic support), x_2, y_2 be coordinates of O_2 (the center of the bearing spike), x_3, y_3 be coordinates of O_3 (the center of gravity of the rotor), φ be the polar angle of the line of centers.

Taking into account that

$$x_3 = x_2 + e\cos\omega t, \quad y_3 = y_2 + e\sin\omega t, \tag{1}$$

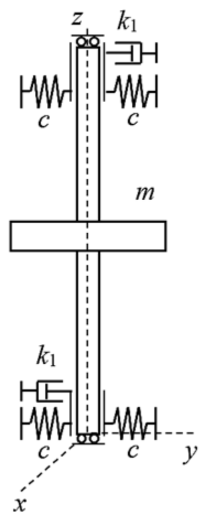


Figure 1: Rotor system rotating on journal bearings

we obtain the differential equations of motion of the system

$$\begin{aligned} m_0\ddot{x}_1 + k_1\dot{x}_1 + cx_1 - 2\left(P_e\cos\varphi + P_\varphi\sin\varphi\right) = 0, \\ m_0\ddot{y}_1 + k_1\dot{y}_1 + cy_1 - 2\left(P_e\sin\varphi - P_\varphi\cos\varphi\right) = 0, \\ m\ddot{x}_2 + k_2\dot{x}_2 + 2\left(P_e\cos\varphi + P_\varphi\sin\varphi\right) = me\omega^2\cos\omega t, \\ m\ddot{y}_2 + k_2\dot{y}_2 + 2\left(P_e\sin\varphi - P_\varphi\cos\varphi\right) = me\omega^2\sin\omega t. \end{aligned} \tag{2}$$

where P_e and P_φ are determined from the Sommerfeld hypothesis, according to which no restrictions are imposed on the length of the lubricating layer between the bearing and the stud and are determined as [26]

$$P_e = \frac{12\pi\mu LR^3\dot{\chi}}{\delta^2\left(1-\chi^2\right)^{3/2}}, \quad P_\varphi = \frac{12\pi\mu LR^3\chi\left(\omega - 2\dot{\varphi}\right)}{\delta^2\left(2+\chi^2\right)\sqrt{1-\chi^2}}.$$

The first two equations of system (2) are equations of motion of the support under the action of elastic forces cx_1, cy_1, damping forces $k_1\dot{x}_1$, $k_1\dot{y}_1$, and reaction forces of the lubricating layer P_e and P_φ, directed in the opposite direction to the forces of the same name shown in Figure 2.

The second two equations of system (2) determine the equations of motion of the rotor under the action of the reaction forces of the lubricating layer P_e and P_φ, and the external damping forces $k_2\dot{x}_2$, $k_2\dot{y}_2$. In order for the equations of system (2) in combination with the equations of hydrodynamic forces to form a closed system, it is necessary to express the eccentricity of the stud center e and the polar angle φ through the coordinates of the center of the elastic support x_1, y_1 and the coordinates of the center of the stud x_2, y_2. Figure 2 shows that

$$x_2 - x_1 = e\cos\varphi, \quad y_2 - y_1 = e\sin\varphi. \tag{3}$$

Then

$$e = \sqrt{\left(x_2 - x_1\right)^2 + \left(y_2 - y_1\right)^2}, \tag{4}$$

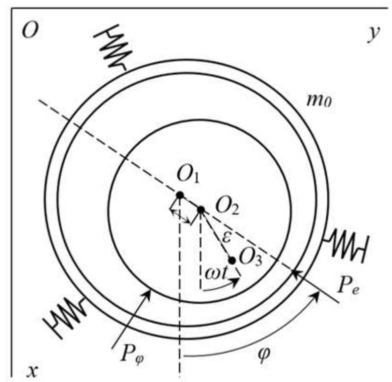


Figure 2: Reaction forces of journal bearings

$$\dot{e} = \frac{(x_2 - x_1)(\dot{x}_2 - \dot{x}_1) + (y_2 - y_1)(\dot{y}_2 - \dot{y}_1)}{e}, \tag{5}$$

$$\sin\varphi = \frac{(y_2 - y_1)}{e}, \quad \cos\varphi = \frac{(x_2 - x_1)}{e}, \tag{6}$$

$$\dot{\varphi} = \frac{(x_2 - x_1)(\dot{y}_2 - \dot{y}_1) - (y_2 - y_1)(\dot{x}_2 - \dot{x}_1)}{e^2}. \tag{7}$$

The system of equations (2) and equations (4)-(7) together with expressions for the reaction forces of the lubricating layer, the form of which depends on the accepted hypothesis, forms a closed system of nonlinear equations, the integration of which in general is not possible. To obtain an approximate solution of the equations of motion (2), we introduce complex variables of the form

$$z_1 = x_1 + iy_1, \quad z_2 = x_2 + iy_2, \quad z_3 = e(\cos\varphi + i\sin\varphi). \tag{8}$$

Then equations (2) and reaction forces can be rewritten as

$$\begin{gathered} m\ddot{z}_2 + c(z_2 - z_3) + k_1(\dot{z}_2 - \dot{z}_3) = 0, \\ c(z_2 - z_3) + k_1(\dot{z}_2 - \dot{z}_3) = 2(P_e - iP_\varphi)e^{i\varphi}, \\ P_e = \frac{6\mu LR^3}{\delta^2}\frac{2\chi^2(\omega - 2\Omega)}{(2+\chi^2)(1-\chi^2)}, \quad P_\varphi = \frac{6\mu LR^3}{\delta^2}\frac{\pi\chi(\omega - 2\Omega)}{(2+\chi^2)\sqrt{(1-\chi^2)}}. \end{gathered} \tag{9}$$

Let the system be "weakly" nonlinear, then its solution can be sought as

$$z_2 = \delta a e^{i(\Omega t - \gamma)}, \quad z_3 = \delta\chi e^{i\Omega t}. \tag{10}$$

Thus, substituting solutions in the form (10) into the equations of motion of system (9) and equating the terms in front of the same harmonics, we obtain a system of algebraic equations for the rotor amplitudes in the form

$$\begin{aligned} &-a\alpha^2\cos\gamma + a\cos\gamma - \chi + Da\alpha\sin\gamma = 0, \\ &a\alpha^2\sin\gamma - a\sin\gamma - D\alpha\chi + Da\alpha\cos\gamma = 0, \end{aligned} \tag{11}$$

where

$$\varphi = \Omega t, \quad A = a\delta, \quad D = k_1/m\Omega, \quad \alpha = 1/\sqrt{1-D^2}.$$

From system (11) we find that

$$a = \frac{\chi^2 \left(1 + D^2\alpha^2\right)}{\left(1-\alpha^2\right)^2 + D^2\alpha^2}, \quad \gamma = \frac{a}{\chi}\frac{D\alpha^3}{1+D^2\alpha^2}. \tag{12}$$

Thus, by varying parameters of the dimensionless damping D, dimensionless frequency of self-excited oscillations α, etc., we obtain amplitude-frequency characteristic for a rotor system mounted on journal bearings, taking into account nonlinearity of the reaction forces of the lubricating layer of journal bearings (Figures 3-13).

4 Results and discussion

The calculations were carried out for a rotor system rotating at a speed of 0 to 20000 rpm. It should be noted that five main parameters vary during the calculation, namely, the viscosity of the fluid in the lubricating layer, the mass of the supports, the damping coefficient, the rigidity coefficient of the equivalent field of elasticity and the size of the gap in the bearing, since these parameters are fundamental in the study of the behavior of self-excited vibrations. The analysis of vibrations was carried out on the basis of the analytical solution of the system of equations (11), with the following initial data: rotor mass $m = 5$ kg, support mass $m_0 = 0.15$ kg, clearance in the bearing $\delta = 0.06$ mm, oil temperature in the bearing $t = 50°$ C, bearing oil viscosity $\mu = 22.39$ mPa.s (turbine oil), bearing stud diameter $d = 20$ mm, bearing length $L = 20$ mm, bearing diameter $D = 20{+}2\delta$ mm, shaft length $l = 650$ mm; the equivalent rigidity of the elastic field of the support $c = 29$ kg/s^2, damping coefficients $k_1 = 42$ kg/s, $k_2 = 6.59$ kg/s.

Figure 3 shows the amplitude-frequency characteristics of the system with a gap of $\delta = 0.06$ mm. It can be seen from the figure that with a rigid fastening (red curve), the system performance is limited by the rotation speed, which is approximately equal to twice the critical speed of the rotor. Starting from 6000 rpm, intense self-oscillations arise in the system in a wide frequency range. With the elastic mounting (blue curve), the vibration level is many times lower. The rotor, mounted on elastic supports, does not have a self-oscillation zone, and the system acquires the ability for stable operation at speeds of 20,000 rpm and higher, i.e. at speeds twenty times the first critical speed. When the rotor starts up after an easy and calm transition through two critical rotation speeds, the first self-centering zone is detected, in which operation with small vibration amplitudes is possible.

The second, even wider self-centering zone is located in the range from 6,000 to 20,000 rpm. Finally, it can be seen from the figure that the range of possible speeds of stable rotation of the rotor due to rotor mounting on elastic supports has increased three times compared to the rigid mounting of bearings, and this is especially important, the upper limit of the speed of rotation of the rotor has no fundamental boundaries. At the same time, it is observed that rotor mounting on elastic supports leads to a decrease in the level of vibrations not only in the areas of self-centering, but also during transition through resonant modes. In this case, the lower the rigidity of the supports, the less the vibration overloads.

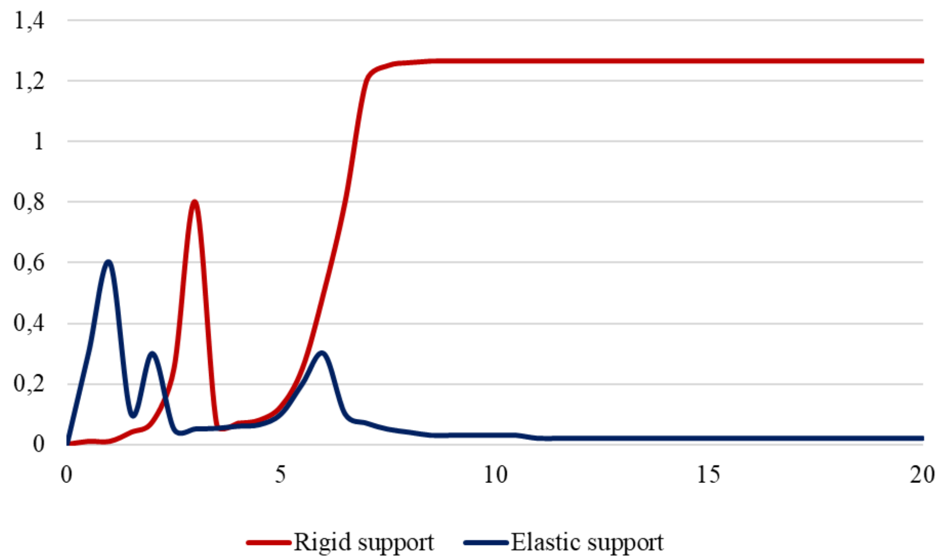


Figure 3: Rotor amplitudes with elastic and rigid mounting in the case when $d = 20$ mm, $l = 650$ mm, $c = 29$ kg/s^2, $\delta = 0.06$ mm, $t = 50°$ C, $k_1 = 42$ kg/s, $k_2 = 6.59$ kg/s, $\mu = 22.39$ mPa.s (turbine oil)

Figures 4 and 5 show the amplitude-frequency characteristics of the rotor and support, depending on the type of oil in the sleeve bearing, when $t = 50°$ C, $\delta = 0.06$ mm, pressure 1 atm. In the first case (red curve), when $\mu = 14.99$ mPa.s (anhydrous glycerol), the amplitudes of both the rotor and the support are maximum. Further, as the viscosity of the liquid increases, the amplitudes decrease and have minimum values at maximum values of viscosity (black curve), i.e. $\mu = 40$ mPa.s (fuel oil). In this case, the optimal values correspond to the case when turbine oil is used, i.e. when $\mu = 22.39$ mPa.s, as further increase in viscosity may lead to violation of the thermal regime in the journal bearing.

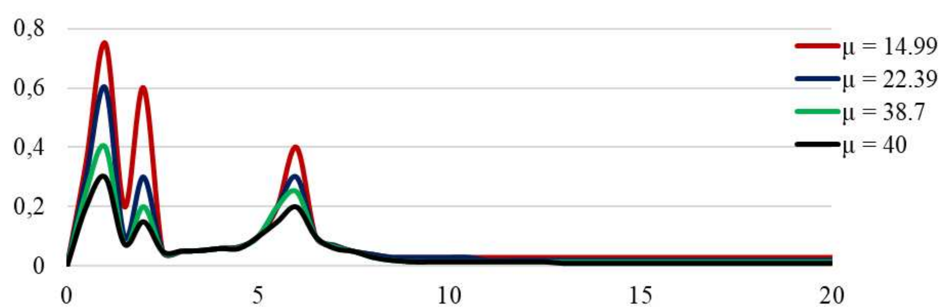


Figure 4: Rotor amplitudes at different values of fluid viscosity in the bearing when $m = 5$ kg, $m_0 = 0.15$ kg, $d = 20$ mm, $l = 650$ mm, $c = 29$ kg/s^2, $\delta = 0.06$ mm, $t = 50°$ C, $k_1 = 42$ kg/s, $k_2 = 6.59$ kg/s, $\mu = 22.39$ mPa.s (turbine oil)

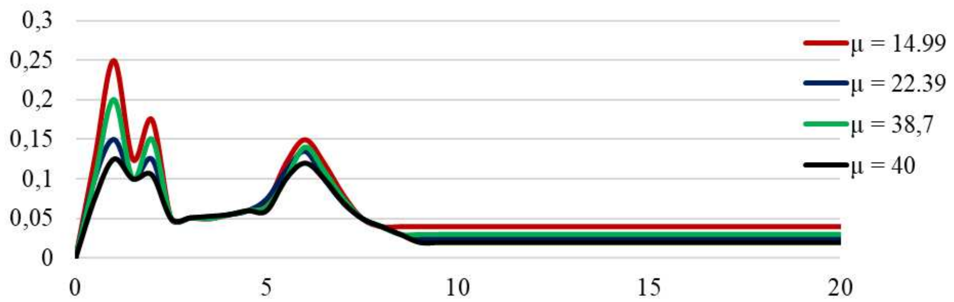


Figure 5: Support amplitudes at different values of fluid viscosity in the bearing when $m = 5$ kg, $m_0 = 0.15$ kg, $d = 20$ mm, $l = 650$ mm, $c = 29$ kg/s^2, $\delta = 0.06$ mm, $t = 50°$ C, $k_1 = 42$ kg/s, $k_2 = 6.59$ kg/s, $\mu = 22.39$ mPa.s (turbine oil)

Figures 6 and 7 show the amplitude-frequency characteristics of the rotor and support depending on the weight of the support. In both cases, the amplitudes of the rotor and support are damped with an increase in the mass of the support, since the support, with a sufficiently large mass, serves as an anti-weight and acts as a vibration damper, i.e. there is an anti-resonance phenomenon, for example, when $m_0 = 1$ kg (black curve). It should be noted that with an increase in the mass of the support, critical frequencies are shifted towards smaller angular velocities, whereas strong displacements of self-centering areas are not observed. With a decrease in the mass of the support, resonance frequencies are shifted towards large angular velocities, and amplitudes also increase, the first section of self-centering is also narrowed, for example, the case when $m_0 = 0.15$ kg (red curve).

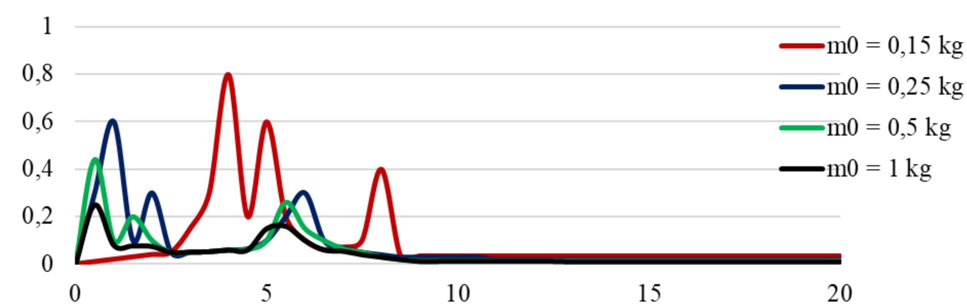


Figure 6: Rotor amplitudes at different values of the support mass in the case when $m = 5$ kg, $d = 20$ mm, $l = 650$ mm, $c = 29$ kg/s^2, $\delta = 0.06$ mm, $t = 50°$ C, $k_1 = 42$ kg/s, $k_2 = 6.59$ kg/s, $\mu = 22.39$ mPa.s (turbine oil)

Figures 8 and 9 show the amplitude-frequency characteristics of the rotor and support depending on the damping coefficient, for gaps $\delta = 0.06$ mm. Here, the amplitudes sharply decrease when passing through resonances. Moreover, the damping effect of the elastic supports is most effective when passing through the first and second critical speeds of the rotor. The influence of damping of supports on the third critical speed is less significant.

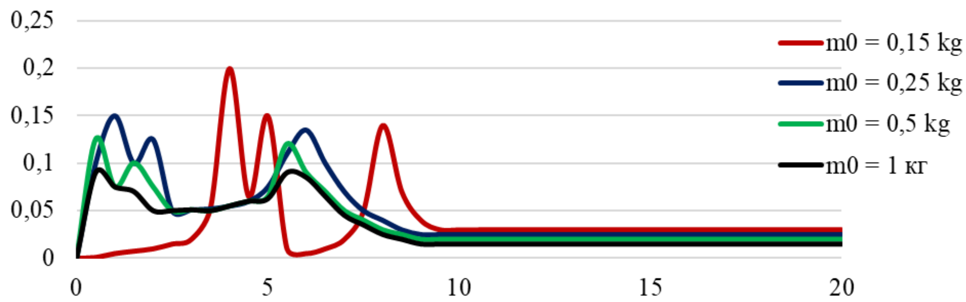


Figure 7: Support amplitudes for different values of the support mass in the case when $m = 5$ kg, $d = 20$ mm, $l = 650$ mm, $c = 29$ kg/s^2, $\delta = 0.06$ mm, $t = 50°$ C, $k_1 = 42$ kg/s, $k_2 = 6.59$ kg/s, $\mu = 22.39$ mPa.s (turbine oil)

An increase in the vibration amplitudes in the self-centering zones is not observed. Smooth operation of the system with low vibration amplitudes is observed in these zones.

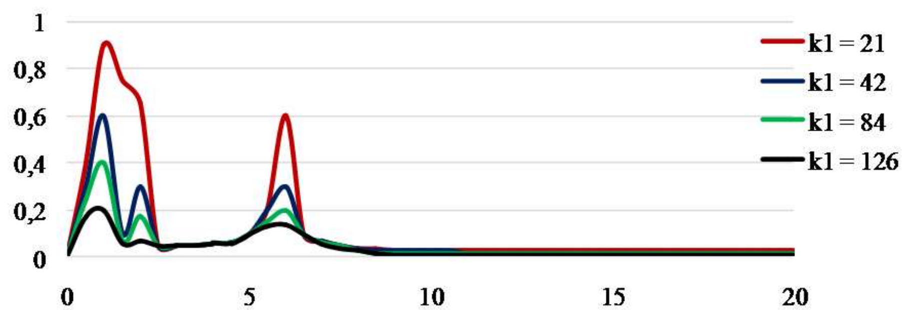


Figure 8: Rotor amplitudes at different values of the damping coefficient k_1 in the case when $m = 5$ kg, $m_0 = 0.15$ kg, $d = 20$ mm, $l = 650$ mm, $c = 29$ kg/s^2, $\delta = 0.06$ mm, $t = 50°$ C, $k_1 = 42$ kg/s, $k_2 = 6.59$ kg/s, $\mu = 22.39$ mPa.s (turbine oil)

At different values of rigidity of the equivalent field of the supports, there is also a shift in the vibration amplitudes along the frequency axis and change in their magnitudes (Figures 10 and 11). For example, with an increase in rigidity, the amplitudes of both the rotor and the supports increase. Also, with an increase in the coefficient, the peaks of the amplitudes are shifted towards higher angular velocities. In general, an increase in rigidity, as was shown initially (Figure 3), does not have a positive effect on the behavior of the system, while with an increase in compliance, the opposite picture is observed.

Figures 12 and 13 show the amplitude-frequency characteristics of the rotor and support, depending on the width of the gap in the journal bearing. As can be seen from the figures, an increase in the width of the gap adversely affects the operation of the system. An increase in the gap width leads to an increase in the amplitude of both the rotor and the support. With a decrease in the gap width, the opposite effect is observed, i.e. the minimum values of δ correspond to the minimum values of the amplitudes. But since, in practice, a small gap

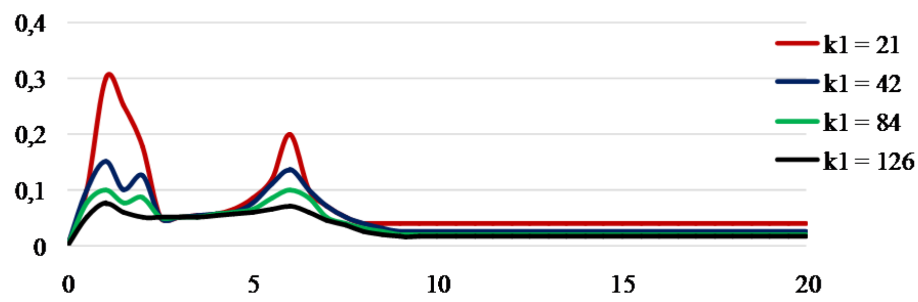


Figure 9: Support amplitudes at different values of the damping coefficient k_1 in the case when $m = 5$ kg, $m_0 = 0.15$ kg, $d = 20$ mm, $l = 650$ mm, $c = 29$ kg/s^2, $\delta = 0.06$ mm, $t = 50°$ C, $k_1 = 42$ kg/s, $k_2 = 6.59$ kg/s, $\mu = 22.39$ mPa.s (turbine oil)

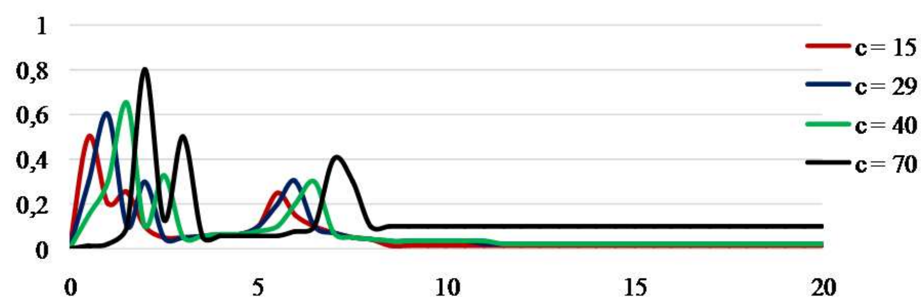


Figure 10: Rotor amplitudes at different values of the rigidity coefficient c in the case when $m = 5$ kg, $m_0 = 0.15$ kg, $d = 20$ mm, $l = 650$ mm, $c = 29$ kg/s^2, $\delta = 0.06$ mm, $t = 50°$ C, $k_1 = 42$ kg/s, $k_2 = 6.59$ kg/s, $\mu = 22.39$ mPa.s (turbine oil)

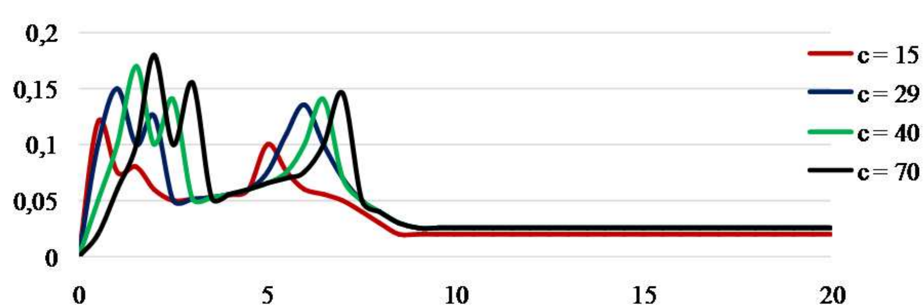


Figure 11: Support amplitudes at different values of the rigidity coefficient c in the case when $m = 5$ kg, $m_0 = 0.15$ kg, $d = 20$ mm, $l = 650$ mm, $c = 29$ kg/s^2, $\delta = 0.06$ mm, $t = 50°$ C, $k_1 = 42$ kg/s, $k_2 = 6.59$ kg/s, $\mu = 22.39$ mPa.s (turbine oil)

width entails violation of the thermal regime due to heating [27], the best option in this case is the gap value $\delta = 0.06$ mm.

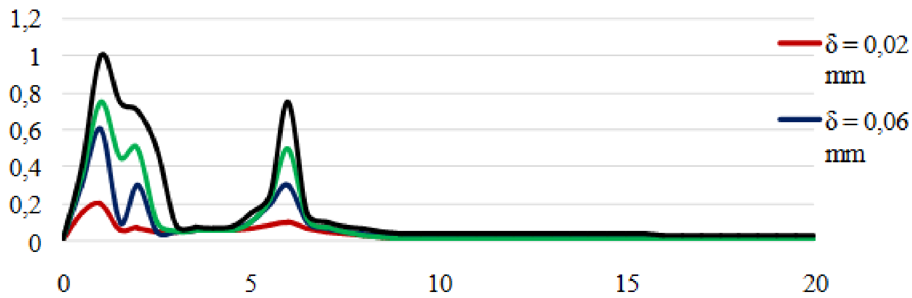


Figure 12: Rotor amplitudes for different values of the gap thickness δ in the case when $m = 5$ kg, $m_0 = 0.15$ kg, $d = 20$ mm, $l = 650$ mm, $c = 29$ kg/s^2, $\delta = 0.06$ mm, $t = 50°$ C, $k_1 = 42$ kg/s, $k_2 = 6.59$ kg/s, $\mu = 22.39$ mPa.s (turbine oil)

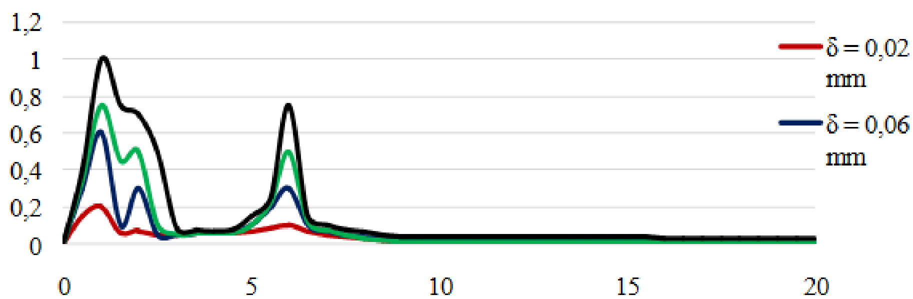


Figure 13: Rotor amplitudes for different values of the gap thickness δ in the case when $m = 5$ kg, $m_0 = 0.15$ kg, $d = 20$ mm, $l = 650$ mm, $c = 29$ kg/s^2, $\delta = 0.06$ mm, $t = 50°$ C, $k_1 = 42$ kg/s, $k_2 = 6.59$ kg/s, $\mu = 22.39$ mPa.s (turbine oil)

In the first resonant zone, the vibrations of the disk and supports occur in phase, i.e. the type of the waveform is cylindrical precession. In the second zone, vibrations of the supports occur in antiphase with respect to each other; in this case, in the region of the disk, vibrations have a node. Thus, in the second zone, the mode of vibrations is a skew-symmetric precession. In the third resonant zone, the vibrations of the supports with respect to each other occur in phase, and near the disk – in antiphase. Thus, the third form of vibrations is a two-node symmetrical form, the type of which resembles the first form of vibrations of an unsupported shaft. It should be noted that the location and types of the first and second modes of vibrations are determined mainly by the compliance of the supports,whereas the third form is caused by bending vibrations of the rotor shaft. Thus, these studies show that the zones of increased vibrations are narrow resonant zones due to dynamic and static imbalances of the rotor.

5 Conclusion

Installation of rotors in elastic supports leads to complete suppression of self-oscillations that occurred during rigid mounting of journal bearings, and oscillations of the system over the entire speed range become purely forced. The damping efficiency of elastic supports is very high and increases with decreasing rigidity. Self-centering of the system in non-resonant zones leads to significant reduction in the magnitude of vibrations and vibration overloads of the system. Installation of the rotor in elastic supports "linearizes"the dynamic system "rotor – supports" . It should also be noted that the main parameter that determines the type of oscillations is the size of the gap of the journal bearing, since with its increase the amplitudes will increase, and at its limiting values, self-excited oscillations will turn into a chaotic type of oscillations, which will negatively affect the stability of the system even at high speeds. According to the theory of self-centering [28], where it is shown that overloads in self-centering areas are determined only by the magnitude of the unbalance and the rigidity of the supports, it can be concluded that vibration overloads of the system will practically not increase even with a significant value of the rotor unbalance. Therefore, with sufficient compliance of the supports, even with large imbalances, one can expect stable operation of the machine with a moderate level of vibration overloads in a wide range of speeds.

References

[1] Muszynska A., *Rotordynamics* (Boca Raton: Taylor & Francis, 2005): 1054.

[2] Greenhill L.M., "Critical Speeds Resulting from Unbalance Excitation of Backward Whirl Modes / L.M. Greenhill, G.A. Cornejo" , *Design Engineering Technical Conferences* (DETC'95), September 17-20, 1995, Boston Massachusetts, USA: Proceedings. Boston: ASME. 3B (DE-Vol. 84-2). (1995): 991–1000.

[3] Yamamoto T., Ishida Y., *Linear and nonlinear rotor dynamics* (New York, John Willey and Sons, 2001): 326.

[4] Leung A. Y. T. and Kuang J. L., "Chaotic Rotations of a Liquid – Filled Solid" , *Journal of Sound and Vibration* 302(3). (2007): 540–563.

[5] Adams M.L., *Rotating machinery vibration* (NY: MarcelDekker, 2001): 354.

[6] Newkirk B.L., *Shaft whipping* (General Electric Review. March, 1924).

[7] Newkirk B.L., Taylor H.D., *Shaft whipping due to oil action in journal bearings* (General Electric Review, August, 1925).

[8] Hagg A.C., "The influence of oil-film journal bearings on the stability of rotating machines" , *Journal of Applied Mechanics* (1946).

[9] Yukio Hori, "A theory of oil whip" , *J. of Applied Mechanics* (1959).

[10] Kesten J., "Stabilité de la position de l'arbre dans un palier ä graissage hydrodynamique" , *Wear,* 5 (1960).

[11] Someya T., "Stabilität einer in zylindrischen Gleitlagern laufenélsen, unwunchtfreien Welle" , *Ingenieur-Archiv,* 33 (1963).

[12] Boecker G.F., "Sternlicht B. Investigation of Translatory fluid whirl in vertical machines" , *Trans, of the ASME,* January (1956).

[13] Schnittger J.R., "Development of a smooth running double-spool, gas-turbine rotor system" , *ASME Paper* 58 (A-l97) (1958).

[14] Pinkus O., "Experimental investigation of resonant whip" , *Trans.of the ASME,* July (1956).

[15] Hummel Ch., *Kritische Drehzahlen als Folge der Nachgiebigkeit des Schmiermittels im Lager* (VDI-Forschungsheft, 1926): 287.

[16] Olimpiev V.I., "About the natural frequencies of the rotor on sliding bearings", *Proceedings of the USSR Academy of Sciences* 93 (1960).

[17] Capriz G., "On the vibrations of shafts rotating on lubricated bearings", *Ann. Mat. Pura Appl.* IV (50) (1960)"223.

[18] Tondl A., *Experimental invesrigation of self-excited vibrations of rotors due to the action of lubricating oil film in journal bearings* Monographs and Memoranda of the National Research Institute of Heat Engineering. No 1 (Prague, 1961).

[19] Tondl A., "Einige Ergebnisse experimenteller Untersuchungen der Zapfenbewegung in Lagern", *Revue de mécanique appliquée,* VI(1) (1961).

[20] Dick J., *Alternating loads on sleeve bearings"*, *Philosophical Magazine* 35 (1944).

[21] Shawki G.S.A., *Whirling of a journal bearing experiments under no-load conditions* (Engineering, 1955).

[22] Shawki G.S.A., "Analytical study of journal-bearing performance under variable loads", *Trans. of the ASME* 3 (1956).

[23] Shawki G.S.A., "Journal bearing performance for combinations of steady, fundamental and harmonic components of load", *Proc. of the Inst. of Mech. Eng.* 171 (28) (1957).

[24] Cameron A., *Oil whirl in bearings* (Engineering, 1955).

[25] Krämer E., 'Der Einfluss des Olfilms von Gleitlagern auf di Schwingungen von Maschinenwellen", *VDI-Berichte* 35 (1959).

[26] Kelzon A.S., *Dynamics of rotors in elastic supports* (Moscow, Science, 1982).

[27] Tondl A., *Dynamics of rotors of turbogenerators* (Publishing house "Energy 1971).

[28] Kelzon A.S., "Self-centering and balancing of a rigid rotor rotating in two elastic bearings", *DAN USSR* 110 (1) (1956): 31–33.

IRSTI 30.03, 55.30.31, 55.30.13 DOI: https://doi.org/10.26577/JMMCS.2022.v115.i3.012

A.B. Baratova[1*], **K.A. Ozhikenov**[1], **A. Adilkhan**[1], **A.N. Mussulmanbekova**[1], **M.K. Zhumadillayev**[2], **N.S. Tolebaev**[2]

[1]U. Joldasbekov Institute of Mechanics and Engineering, Kazakhstan, Almaty

[2]Satbayev University, Kazakhstan, Almaty

*e-mail: ainur.baratova.ab@gmail.com

THE FEATURE OF THE AUTONOMOUS ROBOT FOR CLEANING THE FLOOR IN THE BATHROOM

This article is a new robotic arm for cleaning the floor in the toilet with an increased radius of action of the robotic arm type SCARA. The most common current trends in production include short production cycles, low volumes and a wide variety of orders that can be solved with the help of the SCARA robot. With the advent of the COVID-19 virus in the world, the term "cleaning and disinfection" has become one of the most important tools for preventing the population from becoming infected with the virus. The research focuses on the research and implementation of SCARA-type robots and describes the possibilities of using a SCARA-type robot. This article describes the selection and deployment of a SCARA robot in industrial automation. This research project describes the simulation of a new SCARA-type robotic arm with a long reach and sliding mechanism, we have developed a new multi-joint robotic arm for working in confined spaces with an autonomous toilet floor cleaning system.

Key words: Automation, SCARA robot, forward kinematics, inverse kinematics.

А.Б. Баратова[1*], К.А. Ожикенов[1], А. Әділхан[1], А.Н. Мусулманбекова[1], М.К. Жумадиллаев[2], Н.С. Төлебаев[2]

[1]Ө.А. Жолдасбеков атындағы Механика және машинатану институты, Қазақстан, Алматы қ.

[2]Satbayev University, Қазақстан, Алматы қ.

*e-mail: ainur.baratova.ab@gmail.com

Жуынатын бөлме еденін тазалауға арналған автономды роботтың ерекшелігі

Ұзартылған SCARA типті робот қолымен дәретхана еденін тазартатын жаңа роботтың көрсету қазіргі жағдайда өте маңызды мәселелердың шешімі болып табылады. өндіріске әсер ететін заманауи тенденцияларға қысқа тауарлық циклдар, шағын көлемдер және SCARA роботының көмегімен шешуге болатын тапсырыстардың үлкен алуандығы жатады. Әлемде COVID-19 вирусының пайда болуымен "тазалау және дезинфекция" термині халықтың вирусты жұқтыруының алдын алудың маңызды құралдарының біріне айналды. Соған орай, осы ғылыми жұмыста біз дәретхана еденін тазартатын өзін-өзі басқару жүйесі барә ішкі құрылымы жағынан жаңа көп буынды роботты ұсынылды. SCARA роботтары өздерінің қаттылығы мен жоғары дәлдігіне байланысты салада ең көп қолданылатын роботтардың бірі болып табылады. Жобалау процесі қосылым конструкциясын, сілтеме конструкциясын, контроллер конструкциясын және механикалық таңдауды қамтиды. Ғылыми жұмыс SCARA типті роботын зерттеуге және орындалу процессіне бағытталған және SCARA типті роботын пайдалану мүмкіндіктерін сипаттайды. Бұл мақалада өнеркәсіптік автоматтандыруда SCARA типті роботын таңдау және орналастыру жұмысы туралы айтылады. Жұмыс құрылымында роботтың конструкциясын модельдеу, кинематика, кинематикалық валидация қарастырылған. Кинематикалық валидация арқылы буындардың бұрылу бұрышың, жылдамдығының және үдеудің мәні алынған.

Түйін сөздер: Автоматтандыру, SCARA робот, алға кинематика, кері кинематика.

А.Б. Баратова[1*], К.А. Ожикенов[1], А. Әділхан[1], А.Н. Мусулманбекова[1], М.К. Жумадиллаев[2], Н.С. Төлебаев[2]
[1]Институт механики и машиноведения им. У.А. Джолдасбекова, Казахстан, г. Алматы
[2]Satbayev University, Казахстан, г. Алматы
*e-mail: ainur.baratova.ab@gmail.com

Особенность автономного робота для уборки пола в ванной комнате

В данной статье представлена новая роботизированная рука типа SCARA для уборки пола в туалете с увеличенным радиусом действия. К наиболее распространенным современным тенденциям в производстве относятся короткие циклы продукции, малые объемы и большое разнообразие заказов, которые можно решить с помощью робота SCARA. С появлением в мире вируса COVID-19 термин "очистка и дезинфекция" стал одним из важнейших инструментов предотвращения заражения населения вирусом. В этом исследовательском проекте описывается моделирование нового робота-манипулятора типа SCARA с большим вылетом и раздвижным механизмом. Роботы SCARA являются одними из наиболее широко используемых роботов в промышленности благодаря присущей им жесткости и высокой точности. Процесс проектирования включал проектирование соединения, проектирование звеньев, проектирование контроллера, а также выбор механических и электрических компонентов. Исследование посвящено изучению и внедрению роботов типа SCARA и описывает возможности использования робота типа SCARA. В данной статье описана работа по выбору и внедрению робота типа SCARA в промышленную автоматизацию. Мы разработали новый многошарнирный робот-манипулятор для работы в ограниченном пространстве с автономной системой уборки пола в туалете.

Ключевые слова: Автоматизация, робот SCARA, прямая кинематика, обратная кинематика.

1 Introduction

It has been at least two decades since conventional robotic manipulators became a common production tool in industries ranging from automotive to pharmaceuticals [1]. In many ways, the proven benefits of using robotic manipulators for manufacturing in various industries have motivated scientists and researchers to try to expand the use of it in many different areas. To apply robotics in all areas, scientists had to invent several other types of robots, different from conventional manipulators. New types of robots can be divided into two groups: redundant manipulators and mobile robots. These two groups of robots have greater mobility, allowing them to perform tasks that conventional manipulators cannot. Many engineers have expanded the work with the added mobility of new robots to make them work in tight spaces [1]. In the course of work, the limitations for robotic arms are usually dependent on the working environment, they are changeable. Engineers had to invent different methods to allow robots to automatically cope with various constraints. And an autonomous robot is one that is equipped with those methods that allow it to automatically cope with various environmental constraints while performing the desired task [1].

Autonomous robots must be able to efficiently use and synchronize their limited physical and computational exchequer to operate in a dynamic environment. In each field of activity of progressive complexity, it becomes necessary to impose explicit restrictions on the control of planning, perception and action in order to exclude unexpected interactions between behaviors [2]. Autonomous robots must plan when to act, how to find errors and recover from them, how to deal with conflicting goals when performing complex tasks in any dynamic environment. Following this, robots must precisely coordinate all of their limited dynamic and computational resources [2]. In order to improve the comprehensibility of the system and ensure that the robots perform their tasks, explicit constraints are needed that impose structure on the control of planning, perception, and action as tasks and environments become more complex. Any methodology should be to develop robotic systems consisting of sets of behaviors, which can be independent objects that control actions. Running systems consist of sets of local behaviors that can run without additional awareness of the environment. The main problem is that as the number of additional tasks increases, so does the ratio of complexity between behaviors, which can reach such an extent that it becomes difficult to predict the overall behavior of the control system [2].

The robot arm's arm can move within the three main x, y, and z axes associated with base motion, vertical direction, and horizontal direction. Manipulators are available in various configurations: rectangular, cylindrical, spherical, rotating and horizontally articulated. A robot with a horizontal rotating configuration, the Selective Compliance Articulated Robot Arm (SCARA) has four degrees of freedom, in which the two or three horizontal servo-controlled joints are the wrist, elbow, and shoulder [3]. Most importantly, the last vertical axle is pneumatically controlled. Each working task can be set as pickup, non-contact task (ceiling mounting) and contact task (stuff sorting). SCARA, developed in Japan, is suitable for inserting small parts on assembly lines, such as inserting electronic components [3].

SCARA robots have become popular on packaging and assembly lines with three rotating and one prismatic degrees of freedom [4]. Hiroshi Makino first introduced this type of robot in 1979. Commercial SCARA robots are develop in a variety of sizes, line speeds and payload capacities, thus, the control systems of such robots are intended for general industrial applications [4].

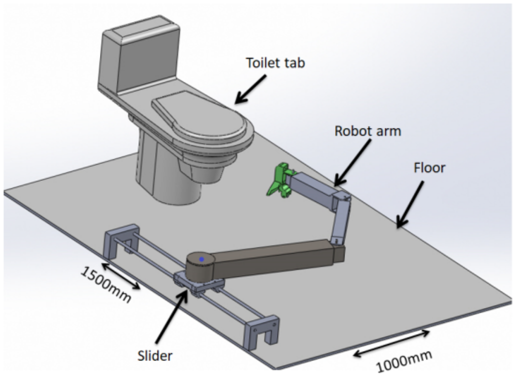
2 Robot design

The workplace and the task set determine the design of the robot. The robot you are designing has several significant parts to learn; the resulting robot can work only in the analyzed and predetermined workplace. The dimensions of the area obtained in this study, i.e. the bath, should be 1000 mm wide and 1500 mm deep. The toilets considered in the study were obtained in accordance with the standards of Western European countries, i.e. the dimensions of the public toilet were 850 mm wide and 1500 mm deep [5].

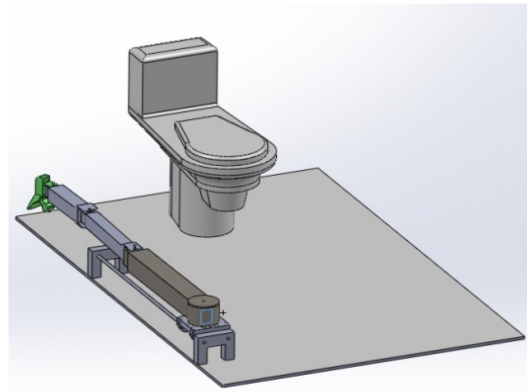
Firstly, the manipulator performing the task must be accessible at any point in the given workspace without dead zones and must be sufficiently compact. Therefore, with this in this study, we proposed a multi-joint arm, which is similar to the structure of the SCARA robot [5]. As shown in Figure 1, the robot arm is aligned along the slide rail after the cleaning process. In general, in such a limited working space, there are individual advantages due to the flexible structure of the continuum arms. The main thing in this task is to have a strong connection with the robot in the hands of a heavy control device [5].

3 Robot arm design

The studied manipulator has the following designs: the manipulator consists of four joints and three links. The robot arm has a particular advantage because the robot is designed to rotate only on the Z axis. The design of the manipulator has been simplified as much as possible



a)



b)

Figure 1: a) The mode of operation of the robot; b) Robot standby mode.

in order to reduce the cost and production time. For the accuracy of work, stepper motors were used to drive the robotic arm [5]. This robotic arm has a total of four stepper motors. A feature of the robot arm is that the robot arm rotates only along the Z axis. Therefore, the design creates a significant dynamic load on the basic drive of the robot. Following these, the dimensions of the engine, such as size and weight, are gradually reduced; on ours, we used a 50 mm frame stepper motor for the base part, and 42 mm and 35 mm frame motors for the middle joints respectively. The end effector that holds and drives the cleaning tool is designed with a 28 mm stepper motor (see Figure 2).

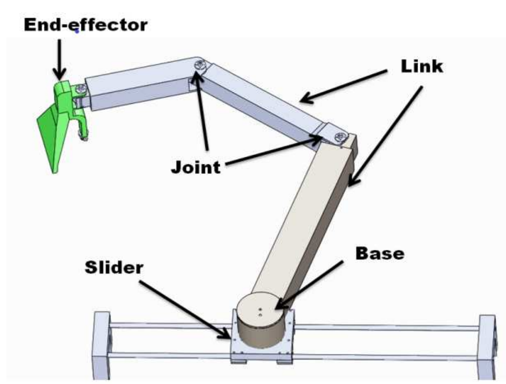


Figure 2: CAD design of the robot.

For the design of the robot, a two-shaft stepper motor was chosen, the effect of skew of the links is important to us. The two-shaft motor has its own characteristics, for example, the use of a two-shaft motor allows the link to be fixed on both sides; side links up and down [5]. In addition, this design simplifies the connection mechanism of the robot.

4 Robot kinematics

4.1 Forward and Inverse Kinematics

The kinematics and dynamics of SCARA robots have also been obtained and modeled using various programs. The experimental results of the SCARA robot were obtained and compared with the simulation results [4].

The researched SCARA robot is widely used as an assembly robot and is a kind of selective picking robot arm. The main features of the robot are the accuracy of the repeating position index and the ease of dynamic execution. The first generation of robots, the serial arm has developed rapidly, and mature designs have already been formed, the connection of which is mainly composed of a servo motor and gearboxes with high speed ratio and good accuracy, such as harmonic reducer [6]. The kinematics of the robot has one translational joint, forming a sequential mechanism, and three rotational joints between the links. The gear mechanism in the rotary joints is a harmonic gear, without shading on the third axis, which makes it possible to obtain a high reduction ratio in sufficient space.

A special advantage of the proposed robot design is the kinematic structure of the robot, which facilitates the kinematic solution of the robot. Because we use serial manipulators, it is much easier to get forward kinematic solutions. As mentioned earlier, the SCARA robot [6] rotates only along the Z axis, and the design of the SCARA robot has the simplest kinematic structure, which means it provides great advantages.

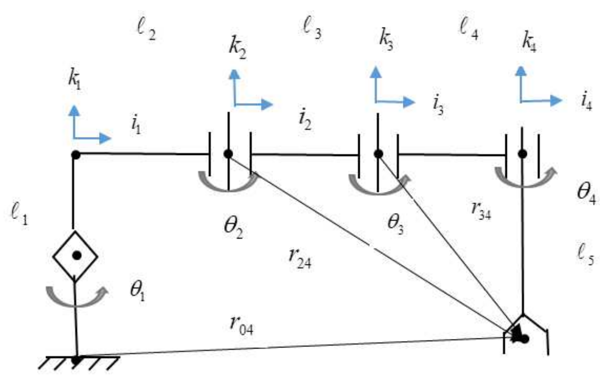


Figure 3: Kinematic structure of the robot.

The robot kinematics starts with Determination of Denavit-Hartenberg parameters. The coordinate systems are directly attached to the robot in accordance with the DH convention [4] and is shown in Table 1.

Table 1: Denavit-Hartenberg parameters

	R_z	R_x	T_x	T_z
	θ	α	h	d
$\sum_{0\to1}$	θ_1	0	0	ℓ_1
$\sum_{1\to2}$	θ_2	0	ℓ_2	0
$\sum_{2\to3}$	θ_3	0	ℓ_3	0
$\sum_{3\to4}$	θ_4	0	0	ℓ_4
$\sum_{4\to5}$	0	0	0	ℓ_5

4.2 Forward kinematics of robot

Table 1 shows the homogeneous transformation formula.

$$H_{01} = \begin{pmatrix} \cos\theta_1 & -\sin\theta_1 & 0 & 0 \\ \sin\theta_1 & \cos\theta_1 & 0 & 0 \\ 0 & 0 & 1 & \ell_1 \\ 0 & 0 & 0 & 1 \end{pmatrix}, \quad H_{12} = \begin{pmatrix} \cos\theta_2 & -\sin\theta_2 & 0 & \ell_2 \\ \sin\theta_2 & \cos\theta_2 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix},$$

$$H_{23} = \begin{pmatrix} \cos\theta_3 & -\sin\theta_3 & 0 & \ell_3 \\ \sin\theta_3 & \cos\theta_3 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}, \quad H_{34} = \begin{pmatrix} \cos\theta_4 & -\sin\theta_4 & 0 & 0 \\ \sin\theta_4 & \cos\theta_4 & 0 & 0 \\ 0 & 0 & 1 & \ell_4 \\ 0 & 0 & 0 & 1 \end{pmatrix}, \quad H_{45} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & \ell_5 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

Here we have obtained a homogeneous transformation, then it is necessary to multiply the matrices from H_{01} to H_{45}.

$$H_{05} = H_{01} * H_{12} * H_{23} * H_{34} * H_{45} = \begin{pmatrix} i_{05} & j_{05} & k_{05} & r_{05} \\ 0 & 0 & 0 & 1 \end{pmatrix} \tag{1}$$

4.3 Inverse kinematics of robot

The robot has inverse kinematics and is quite simple compared to other existing robotic arms [6]. As mentioned earlier, the rotation function of the robot rotates only along the Z axis and this allows us to simplify the calculation and formulation of inverse kinematics (Figure 3).

For inverse kinematics, the robot has four variables as: linear prismatic movement is the main difference from the SCARA robot.

$$q_4 = (\theta_1, \theta_2, \theta_3, \theta_4) = (q_1, q_2, q_3, q_4) \tag{2}$$

$$\nu_4 = \begin{pmatrix} \omega_4 \\ r_{04} \end{pmatrix} \tag{3}$$

Derivation of variables by the Jacobian method

$$\nu_4 = Jq_4 \tag{4}$$

$$j = \begin{pmatrix} k_1 & k_2 & k_3 & k_4 \\ k_1 \times r_{04} & k_2 \times r_{24} & k_3 \times r_{34} & 0 \end{pmatrix} \tag{5}$$

Here r_{04}, r_{24}, r_{34} are the access vectors of the individual rotation. Instead $\vec{r}_{04}$ we can use $\vec{r}_{14}$. Here $r_{44}\vec{k}_4$ parallel connection is equal to 0.

$$k_{0\ldots4} = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \tag{6}$$

$$r_{04} = \begin{pmatrix} \ell_1 \cos(\theta_1) + \ell_2 \cos(\theta_1 + \theta_2) + \ell_3 \cos(\theta_1 + \theta_2 + \theta_3) + \ell_4 \cos(\theta_1 + \theta_2 + \theta_3 + \theta_4) \\ \ell_1 \sin(\theta_1) + \ell_2 \sin(\theta_1 + \theta_2) + \ell_3 \sin(\theta_1 + \theta_2 + \theta_3) + \ell_4 \sin(\theta_1 + \theta_2 + \theta_3 + \theta_4) \\ \ell_1 - \ell_5 \end{pmatrix} \tag{7}$$

Here: $\ell_1 \cos(\theta_1) = \ell_1 c_1$; $\ell_2 \cos(\theta_1 + \theta_2) = \ell_2 c_{12}$; $\ell_1 \sin(\theta_1) = \ell_1 s_1$; $\ell_2 \sin(\theta_1 + \theta_2) = \ell_2 s_{12}$;

$$r_{24} = \begin{pmatrix} \ell_3 \cos(\theta_1 + \theta_2 + \theta_3) + \ell_4 \cos(\theta_1 + \theta_2 + \theta_3 + \theta_4) \\ \ell_3 \sin(\theta_1 + \theta_2 + \theta_3) + \ell_4 \sin(\theta_1 + \theta_2 + \theta_3 + \theta_4) \\ -\ell_5 \end{pmatrix} \tag{8}$$

Here: $\ell_3 \cos(\theta_1 + \theta_2 + \theta_3) = \ell_3 c_{123}$; $\ell_3 \sin(\theta_1 + \theta_2 + \theta_3) = \ell_3 s_{123}$;

$$r_{34} = \begin{pmatrix} \ell_4 \cos(\theta_1 + \theta_2 + \theta_3 + \theta_4) \\ \ell_4 \sin(\theta_1 + \theta_2 + \theta_3 + \theta_4) \\ -\ell_5 \end{pmatrix} \tag{9}$$

Here: $\ell_4 \cos(\theta_1 + \theta_2 + \theta_3 + \theta_4) = \ell_4 c_{1234}$; $\ell_4 \sin(\theta_1 + \theta_2 + \theta_3 + \theta_4) = \ell_4 s_{1234}$;

$$j = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 1 & 1 & 1 & 1 \\ r_{04} & r_{24} & r_{34} & 0 \\ & & & 0 \\ & & & 0 \end{pmatrix} \tag{10}$$

5 Kinematic validation of the robot

In this study, we used RoboAnalyzer to test the robot's kinematics. The proposed robot arm modeling tested in RoboAnalyzer software [7]. Most industrial robots are described geometrically by Denavit-Hartenberg (DH) parameters, which are also difficult for students to perceive. Students will find it easier to study a subject if they can visualize in three dimensions.

The RoboAnalyzer software [7] was developed using Object Oriented Modeling concepts in the Visual C# programming language. 3D graphics are rendered using OpenGL via the Tao Framework. The ZedGraph open source library is used for graphing. The software has been developed in modules, so adding or changing modules does not affect the entire software. The Forward Kinematics module of serial robots with rotating joints has been reported in a paper. It uses wireframe models. The results of the analysis were viewed in the form of animation and a built-in plotting module. The addition of prismatic connections, inverse dynamics and forward dynamics analysis have been reported. Additional modules have been developed here, such as "Visualization of DH parameters and transformations", "Import of 3D CAD models" and "Inverse kinematics".

The software can simultaneously provide the robot's working space and analyze the movement trajectory (see Figure 4).

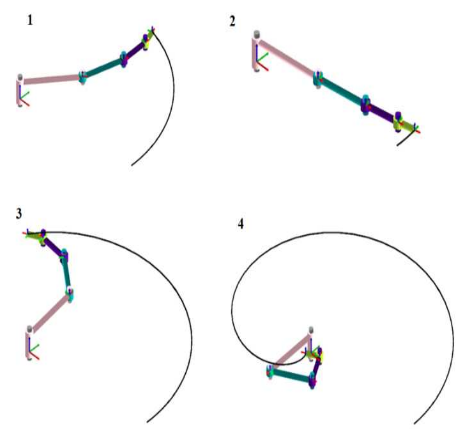
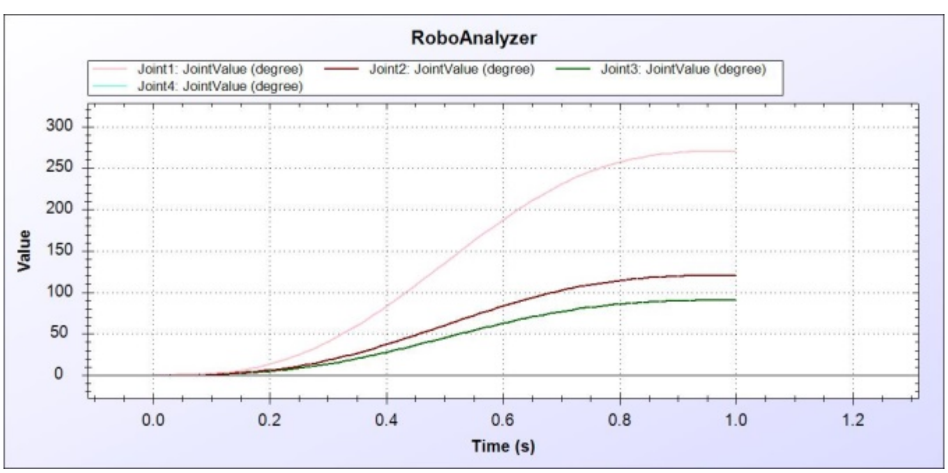
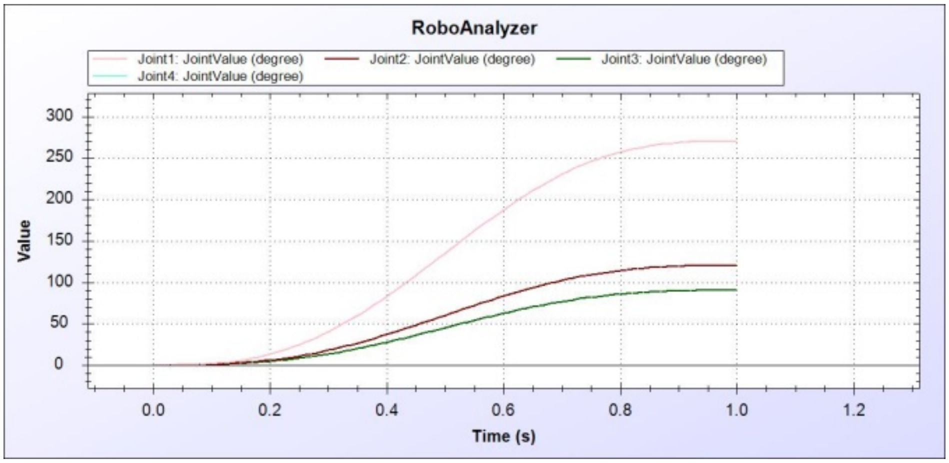


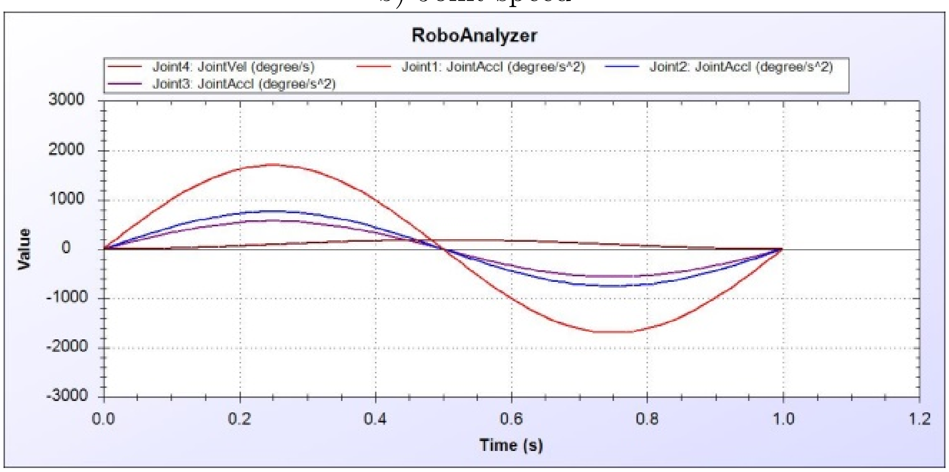
Figure 4: Model of robot movement in the RoboAnalyzer software environment.



a) Joint angle



b) Joint speed



c) Acceleration of the joint

Figure 5: Graph of parameters of robot joints. x-axis describes time; The y-axis describes the angle.

6 Conclusion

This study demonstrates the design and control algorithm of a new robotic system that cleans the bathroom floor. The importance of the robot in quarantine in hospitals is very high. According to the survey, there are bathrooms and toilets in infected areas in public places and hospitals. The world was not ready for COVID-19, and simple places related to hygiene were one of the main drivers of the spread of such infections. In addition, in this study, we took into account human rights.

In the future, we plan to conduct experiments in public places with a laboratory prototype to test the suitability of the proposed robot's signature and system.

References

[1] Fahimi F., "Autonomous Robots Modeling, Path Planning, and Control" , *Springer Science+Business Media* (2009): 2-13.

[2] Reid G. Simmons, "Structured Control for Autonomous Robots" , *IEEE Transactions on robotics and automation* 10(1) (1994).

[3] George A. Bekey, "On autonomous robots" , *The Knowledge Engineering Review* 13(2) (1998): 143–146.

[4] Shariatee M., Akbarzadeh A., Mousavi A., Alimardani S., "Design of an Economical SCARA Robot for Industrial Applications" , *International Conference on Robotics and Mechatronics. Tehran. Iran* (2014).

[5] Azamat Yeshmukhametov, Ainur Baratova, Ainur Salemkhan, Zholdas Buribayev, Kassymbek Ozhikenov, Yedilkhan Amirgaliyev, "Design and modeling of self-sustainable bathroom floor cleaning robot system" , *The 21st International Conference on Control, Automation and Systems (ICCAS 2021) Jeju, Korea* (2021).

[6] Nianfeng Wang, Jinghui Liu, Shuai Wei, Zhijie Xu, and Xianmin Zhang, "The Control System Design of A SCARA Robot" , *Springer International Publishing Switzerland* (2014): 136–145.

[7] Jyoti Bahuguna, Rajeevlochana G. Chittawadigi, Subir K. Saha, "Teaching and Learning of Robot Kinematics Using RoboAnalyzer Software" , *AIR'13. Pune. India* (2013).

ISSN 1563–0277, eISSN 2617-4871 JMMCS. №3(115) 2022 https://bm.kaznu.kz

IRSTI 28.23.25 DOI: https://doi.org/10.26577/JMMCS.2022.v115.i3.09

A.K. Berdaly, Z.M. Abdiahmetova*
Al-Farabi Kazakh National University, Kazakhstan, Almaty
*e-mail: zukhra.abdiakhmetova@gmail.com

PREDICTING HEART DISEASE USING MACHINE LEARNING ALGORITHMS

Increasing the accuracy of detecting heart disease is widely studied in the field of machine learning. Such study is intended to prevent large costs in the field of healthcare and is the reason for the misdiagnosis. As a result, various methods of analyzing disease factors were proposed, aimed at reducing differences in the practice of doctors and reducing medical costs and errors. In this study, 6 classification learning algorithms were used, including machine learning methods such as classification Tree, Close neighborhood method, Naive Bayes, Random forest tree, and Busting methods. These methods were collected by the University of Cleveland. Using heart.csv dataset, they were trained to make an effective and accurate prediction of heart disease. In order to increase the predictive capabilities of algorithms, all methods were trained primarily on non-standardized data. A study was conducted on how much data standardization affects the result using the Standard Scaler method. In the paper, this method helped algorithms such as KNN and SVC improve the result about 25%.

Key words: Classification, Standardization, Training Selection, Metrics, Busting, Confusion Matrix.

А.К. Бердалы, З.М. Абдиахметова*
Әл-Фараби атындағы Қазақ ұлттық университеті, Қазақстан, Алматы қ.
*e-mail: zukhra.abdiakhmetova@gmail.com

МАШИНАЛЫҚ ОҚЫТУ АЛГОРИТМДЕРІ АРҚЫЛЫ ЖҮРЕК АУРУЛАРЫН БОЛЖАУ

Жүрек ауруларын анықтаудың дәлдігін арттыру машиналық оқыту саласында кеңінен зерттелуде. Мұндай зерттеу денсаулық сақтау саласында үлкен шығындардың алдын алу үшін және қате диагноздың қойылу себептерінен туындайды. Нәтижесінде дәрігерлердің тәжірибесиндегі айырмашылықтарды азайтуға және медициналық шығындар мен қателіктерді төмендетуге бағытталған ауру факторларын талдаудың әртүрлі әдістері ұсынылады. Бұл зерттеуде классификациялық оқытудың 6 алгоритмң, соның ішінде атап айтқанда жіктеу ағашы, жақын көршілер әдісі, аңғал Байес, кезейсоқ орман ағашы, бустинг әдістері қолданылды. Осы әдістерді Клевеленд университетінің жинақтаған heart.csv датасетіне қолдану арқылы жүрек аурулары бойынша машинаға тиімді және дәлдігі жоғары болатын болжам жасау үйретілді. Алгоритмдердің болжау қабілетін арттыру мақсатында барлық әдістер бірінші кезекте стандартталмаған деректерге оқытылды. Standart Scaler әдісін қолдану арқылы деректерді стандартизациялау нәтижеге қаншалықты әсер ететініне зерттеу жүргізілді. Зерттеу барысында бұл әдіс KNN мен SVC секілді алгоритмдерге нәтижені шамамен 25%-ға жақсартуға көмек беретіні анықталды.

Түйін сөздер: Классификация, стандартизация, оқыту таңдамалары, метрика, бустинг, шатасу матрицасы.

А.К. Бердалы, З.М. Абдиахметова*
Казахский национальный университет имени аль-Фараби, Казахстан, г. Алматы
*e-mail: zukhra.abdiakhmetova@gmail.com

ПРОГНОЗИРОВАНИЕ ЗАБОЛЕВАНИЙ СЕРДЦА С ПОМОЩЬЮ АЛГОРИТМОВ МАШИННОГО ОБУЧЕНИЯ

Повышение точности выявления заболеваний сердца широко изучается в области машинного обучения. Такое исследование призвано предотвратить большие потери в здравоохранении и привести к неправильному диагнозу. В результате были предложены различные методы анализа факторов заболевания, направленные на снижение различий в опыте врачей и снижение медицинских расходов и ошибок. В данном исследовании были использованы 6 алгоритмов классификационного обучения, в том числе методы машинного обучения, такие как дерево классификации, метод ближайших соседей, наивный Байес, случайное лесное дерево, методы бустинга. Эти методы были обобщены университетом Клевеленда применяя к датасету ССЗ. Они были обучены делать эффективные и высокоточные прогнозы сердечных заболеваний. С целью повышения предсказательной способности алгоритмов все методы были обучены в первую очередь нестандартизированным данным. Проведено исследование того, насколько стандартизация данных с использованием метода Standard Scaler влияет на результат. В ходе исследования данный подход улучшил результаты алгоритмов как KNN и SVC почти на 25%.
Ключевые слова: Классификация, стандартизация, обучающая выборка, метрика, бустинг, матрица путаницы.

1. Introduction

Cardiovascular disease is a disease that poses a risk of death in the modern world and is the biggest problem, as predicted by medicine in terms of growth. According to World statistics, this disease is such a problem that it worries the whole world, which leads to a large mortality factor. According to the World Health Organization, about 20 million people die from heart disease. In England, cardiovascular diseases account for 34% of deaths, while in European countries these statistics reach 40%. According to the latest statistics, the number of deaths from cardiovascular diseases around the world is increasing, the main reason for this forecast is that the statistics of countries with the lowest risk of cardiovascular disease are increasing every year. But according to who forecasts, more than 75% of cardiovascular diseases can be prevented, thereby reducing the burden of developing diseases.
Purpose of the work: selection and description of machine learning methods in Big Data Processing, increasing accuracy in the process of big data learning and reducing machine learning time. Research objectives:

- analysis of the literature on the use of machine learning (ML) methods for data on heart failure;
- analysis of python language libraries and part of machine learning methods;
- initial analysis and pretreatment of data related to cardiac arrhythmias;
- use methods for classifying signs, selecting and filling in missing values;
- analyze obtained results;
- justification of the research results in the subject area.

Object of research: the object of research is the prediction of cardiovascular diseases using machine learning algorithms. Use methods that allow to study and analyze the data used to optimize the process of solving research problems. Using this method, to create a system based on predicting the disease, minimizing human participation in analysis and creating an optimal solution with the participation of machine learning algorithms.

2. Literature Review

Machine learning is an analysis method that allows us to conduct data training and analysis methods that we use to optimize the process of solving research problems. This method is a system based on minimizing human participation in analysis and creating an optimal solution with the participation of artificial intelligence intelligence systems. This article will explore machine learning methods to make predictions in the process of Big Data Processing and analyze some specific methods. Currently, there is an active implementation of machine learning methods in medical information systems (MIS). This is primarily due to the need to analyze a large amount of information about patients in real time, as well as predict whether to seek outpatient care or hospitalization within a given time frame [1]. For the database, there are many open sources for accessing acient records, and research can be conducted to use various computer technologies to identify this disease in order to make the correct diagnosis of the patient and prevent his death [2]. Patients are often diagnosed asymptomatic until death, and even if they are under supervision, trained personnel are required to detect cardiac abnormalities [3]. Heart disease was the cause of 6.2 million deaths between the ages of 30 and 70 in 2019 [4]. These diseases usually occur as a result of stroke, hypertensive heart disease, rheumatic heart disease, artery disease and other defects in the heart vessels and the heart itself [5]. In many countries, there is little experience in cardiovascular research and a significantly higher percentage of misdiagnosed cases, which can be solved by developing accurate and effective methods for predicting heart disease at an early stage through analytical support for clinical decision-making through digital medical records [6].

Amin Ul Hak, Jiang Ping Li, Muhammad Hammad Memon, Shah Nazir and Ruinan Sun were tested on their systems in a Cleveland heart disease dataset. Seven well-known classifiers, such as logistic regression, KN, AN, SM, NB, DT and random forest were used with three algorithms for selecting functions Relief, mRMR, and LASSO, which are used to select important functions. In terms of features SVM (linear) with the selection of functions, the performance of the mrmr algorithm was better than that of other classifiers [7]. Fajr Ibrahim Alarsan and Mamun Yunets received a data set of 205.146 lines, which were randomly divided into two parts: training and testing. They compared the Random Forest and Decision Tree Classifier algorithms in machine learning of this data set. In a random forest, the learning process is faster than in a decision tree and in a decision tree, the testing process is faster than in a random forest. The parameters of both algorithms were changed manually. The optimal values for the configured parameters could be obtained by running cross-checking methods, but the algorithms took a lot of time [8]. Jiang Yi, Zhang X, Ma R, Wang X, Liu J. , Kerman M, Yang Yi, Ma J, Son Yi, Zhang J. , He J, Go C, Go X chose dataset as the data that monitored 1,508 Kazakh subjects in China at the initial level without cardiovascular diseases. All subjects were randomly divided into a study sample (80%) and a test sample (20%). LR, SVM, DT, RF, KNN, NB and XGB were used to predict outcomes in cardiovascular diseases. LR and SVM had better predictive characteristics than other machine learning models in the context of discrimination and calibration. LR was similar to the predicted effectiveness of SVM in predicting the risk of cardiovascular diseases and surpassed other ML models. The sensitivity of LR was higher than that of SVM and the specificity gave the opposite result [9].

3. Problem Settings

This section discusses sorting data from the collected databases, conducting pre-machine learning processing measures, fully studying target variables, dividing them into machine learning and testing stages and learning these information using machine learning method classifiers. Through the selected classifiers, the level of training is evaluated and measures are taken to improve the results. The first step is to access the database used in data training. The dataset taken from the database consists of 14 columns of 303 consecutive factors affecting the symptoms of cardiovascular disease. This database was collected by the Cleveland Clinic, which was connected with the university clinics of Zurich and Basel. The database originally consisted of 72 columns, and as a result of removing columns that did not attach much importance to special processing and research activities, 14 columns were left.

We can show statistical characteristics for numeric attributes in the database. Statistical values are represented as the total number of attributes, the average, standard statistical deviations, the smallest and largest values, as well as indicators of 25%, 30% and 75% on 3 quartils. You can see it in the table below.

Since the indicator of people with cardiovascular diseases was taken as a target variable, the indicators for this variable were visually displayed. Age indicator of the number of people suffering from heart disease according to Figure 3. As we have seen, the most sick people can be called the age range of 40 to 55 years.

4. Materials and Methods

To check the accumulated commands, we first look at whether there are zero elements in the dataset, and if such data is found, fill in the spaces by calculating the median or average value of this column. Disable them because dictionary columns are not involved in training. Algorithms of the machine learning method are used by setting target variables.

The general picture of the work carried out on the methodology is shown in Figure 4:

Until measures to improve the accuracy of the algorithms used give good results, it is necessary to implement such measures as training, avoiding mistakes in the course of excessive training. Classification techniques used to detect cardiovascular diseases are as follows: Decision Tree Classifier, Kneighbors Classifier, Logistic Regression, XGBClassifier, Random Forest Classifier, Support Vector Classfier. Although training is carried out using such techniques, cross – checking with the target variable column of the dataset is carried out in order to increase the result. Cross-validation (verification) is the process of improving the result of an algorithm by training each time with different random values with a random transfer of a target variable to a test set in order to improve the learning efficiency of

Column name	Meaning	Type	Range
Age	Patient age	number	[29, 77]
Sex	Gender	nominal	0 = female, 1 = male
Cp	Type of pain	nominal	1 = normal angina pectoris 2 = abnormal angina pectoris 3 = no angina pectoris 4 = symptoms were not observed
Trestbps	Blood pressure in a calm state	number	[94, 100]
Chol	Cholesterol levels	number	[126, 564]
Fbs	Glucose levels in the blood of a hungry person	nominal	0 = false, 1 = true
Restecg	ECG(Electrocardiography)	nominal	0 = normal, 1 = The history of ST-T is determined 2 = Hypertrophy of the left ventricle according to the Estes criterion
Thalach	Maximum pulse rate	number	[71, 202]
Exang	Angina pectoris during physical exertion	nominal	0 = no, 1 = yes
Oldpeak	St depression level	number	[0, 6.2]
Slope	ECG at maximum load	nominal	1 = high rise 2 = normal 3 = go down
Ca	Fluorescent color important blood vessel number	nominal	[0, 3]
Thal	A type of blood disease called thalassemia	nominal	3 = normal 6 = detected defect 7 = cured defect
target	Heart disease	nominal	0 = no, 1 = yes

Table 1: Database analysis

machine learning operations.

The classifier tree method is a tree-type structure consisting of certain rules that represent the result in the learning process in a hierarchical type. It consists of 2 types of elements, a node and a leaf. Elements to be written in the node if the elements that affect the value of the target variables are written in the Leaf, the functions of the target variables are written in the Leaf. The decision tree is often used because it gives good results in statistical reports, including in medical reports for more probabilistic reports, by classifying the same data, making better predictions, clarity, and simple processing of the data without converting or causing severe distortion [10].

The k nearest neighbors method is an algorithm for classifying objects by class by dividing them into groups previously distributed by region after calculating the distances by weight by vote. This method is considered the simplest of the classification algorithms. It is a classifier algorithm that can be used in cases where there is little information about the

	age	sex	cp	trestbps	chol	fbs	restecg	thalach	exang	oldpeak	slope	ca	thal
count	303.000000	303.000000	303.000000	303.000000	303.000000	303.000000	303.000000	303.000000	303.000000	303.000000	303.000000	303.000000	303.000000
mean	54.366337	0.683168	0.966997	131.623762	246.264026	0.148515	0.528053	149.646865	0.326733	1.039604	1.399340	0.729373	2.313531
std	9.082101	0.466011	1.032052	17.538143	51.830751	0.356198	0.525860	22.905161	0.469794	1.161075	0.616226	1.022606	0.612277
min	29.000000	0.000000	0.000000	94.000000	126.000000	0.000000	0.000000	71.000000	0.000000	0.000000	0.000000	0.000000	0.000000
25%	47.500000	0.000000	0.000000	120.000000	211.000000	0.000000	0.000000	133.500000	0.000000	0.000000	1.000000	0.000000	2.000000
50%	55.000000	1.000000	1.000000	130.000000	240.000000	0.000000	1.000000	153.000000	0.000000	0.800000	1.000000	0.000000	2.000000
75%	61.000000	1.000000	2.000000	140.000000	274.500000	0.000000	1.000000	166.000000	1.000000	1.600000	2.000000	1.000000	3.000000
max	77.000000	1.000000	3.000000	200.000000	564.000000	1.000000	2.000000	202.000000	1.000000	6.200000	2.000000	4.000000	3.000000

Figure 2: Statistical indicators for databases

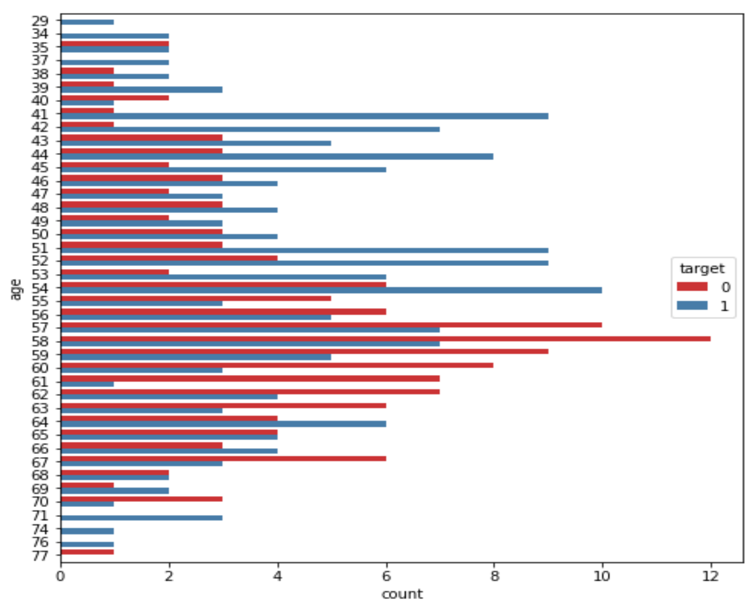


Figure 3: Quantitative indicator for the target variable

data in the preliminary data separation, and is completely unknown. The optimal method for understanding and implementing KNN. Therefore, this situation should be taken into account for any calcification calculations. Its main advantages are that the process is very clear, time-consuming, efficiency and accuracy are often high, and there are methods that eliminate noise in the process that work only for KNN [11].

The logistic regression method is a statistical method that classifies a classification using a linear classification line. The main idea is to determine the optimal line that best divides data through a set of data. The range of logistic regression covers the range from 0 to 1. In addition, this method does not require a connection between input and output data. Logistic

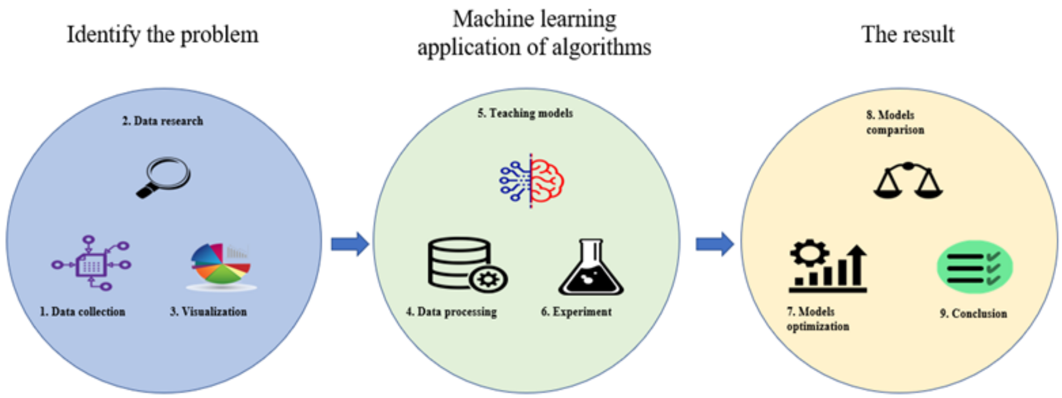


Figure 4: General research plan

regression is a method in medical research that allows you to perform several tests at the same time, minimizing external factors. If the model structure created by the researcher avoids raw data,then the probability of logistic regression is also high [12].
The XGBoost method is a method that belongs to the ensemble method, designed to improve gradient descent, with optimal and high accuracy. This is a method that aims to get the best results by training multiple decision trees in parallel to improve the gradient. Through XGBoost, trees grow rapidly and parallel trees are erected, the final decision is made by an ensemble voice. In this method, random forest trees and decision trees are solved by using models and making comparisons with their parameters [13].
The random forest tree method is another type of algorithm that uses the ensemble method. An algorithm that randomly creates a forest of decision trees, takes forest trees of different selections, matches them to the classifier, and finally takes the average value in order to increase accuracy. The main advantage is the ability to achieve good results when working with large groups and classes, independence from the scale of learning, and the ability to perform high parallelization. Therefore, the random forest tree is an effective predictor [14].
The method of reference vectors is a set consisting of intensive learning algorithms and bringing changes through hyperactivity to a single norm. The idea of the method is that we place data elements consisting of points on the n-dimensional plane, creating hyperspace by creating a classification that best defines classes. SVM also has a core, which is used to convert data entered into the plane by the cores to a large one, taking it as small. It is mainly used for Tex cataloging, recognizing handwritten numbers, finding tones, classifying images, and gene expression using a microchip [15].

In the future, we will use assessment metrics to evaluate the training of these 6 used methods. To do this, in the process of comparing the algorithm-trained results of y with the true value of y in the target variable, a reflection matrix is created, as in Table 5. The result of algorithms based on the generated matrix will be evaluated.

True Positive (TP) – the classifier assumes that the positive result is positive.
True Negative (TN) - the classifier assumes that the negative result is negative.

	Positive	Negative
Positive	TP	FP
Negative	FN	TN

Table 5: Confusion Matrix

False Positive (FP) - the classifier incorrectly predicts a negative result as positive.
False Negative (FN) - the classifier incorrectly predicts a positive result as negative.
The classifier is evaluated using the formulas of the metrics listed below: Accuracy – the total accuracy of the model, the amount of accuracy of classifiers when compared with the main values.

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \tag{1}$$

Precision is an indicator that the classifier finds positive and is actually positive.

$$Precision = \frac{TP}{TP + FP} \tag{2}$$

Recall is an indicator of true positive classes among all positive classes found by the algorithm.

$$Recall = \frac{TP}{TP + FN} \tag{3}$$

F1-score is the hormonal average of accuracy and completeness.

$$F1 - score = \frac{2TP}{2TP + FP + FN} = \frac{2 * Precision * Recall}{Precision + Recall} \tag{4}$$

5. Results

Data source of a 303-row, 14-column collected by Cleveland Medical Center for cardiovascular disease has gone through processing measures that consist of many steps. As a result of the processing measures, no particularly strong outs and zero elements were found in the database. The absence of columns that strongly influence each other on the data was observed through the correlation matrix. After processing, 30 percent of the data was sent for training. Subsequently, 10 classification algorithms were trained. It includes algorithms Decision Tree Classifier, Kneighbors Classifier, Logistic Regression, XGBClassifier, Random Forest Classifier, Support Vector Classfier. During the training of each algorithm, the result was increased by standardization using the Standard Scaler function. The Standard Scaler function tries to show good results by normalizing our data so that the average value does not exceed 0 and the standard deviation does not exceed 1, which gives the opposite effect before applying the algorithm. Algorithms such as Decision Tree Classifier showed a decrease in accuracy from 0.7142% to 0.7023% from the scattered neural structure algorithm, while Kneighbors Classifier helped to increase the accuracy from 0.5934% to 0.7692%. The same result was obtained from the support Vector Classfier algorithm, which increased the accuracy from 0.5604% to 0.8021% by standardization. In addition to standardization, we tried to find the most optimal parameters and increase the result using the GreedSearchCV algorithm. GreedSearchCV refers to a cross – validation operation. It is one of the most powerful tools in machine learning, the main reason for which the correct choice of parameters is the main guarantee of good results. If the parameters are chosen correctly, then, of course, the training will also go well. As for work, it calculates the result for each parameter over the entire connection, providing us with the best indicator. The result was not satisfied, there were significant delays in terms of time, and the result of the algorithm did not show much difference from standardization.

Thus, 6 algorithms were evaluated on 4 metrics. The results were compared among themselves. This can be seen in Table 6. The best indicator for the Accuracy metric was the result of the Random Forest Classifier algorithm. In Figure 7, dynamic comparisons were

	Accuracy	Precision	Recall	F1-score
DecisionTreeClassifier	0.71	0.72	0.70	0.68
KNeighborsClassifier	0.76	0.83	0.77	0.72
LogisticRegression	0.81	0.90	0.82	0.75
XGBClassifier	0.81	0.86	0.81	0.77
RandomForestClassifier	0.82	0.90	0.82	0.76
SupportVector Classfier	0.80	0.88	0.80	0.74

Table 6: Comparison of results of classifiers

made. You can clearly see the real difference through the diagram.

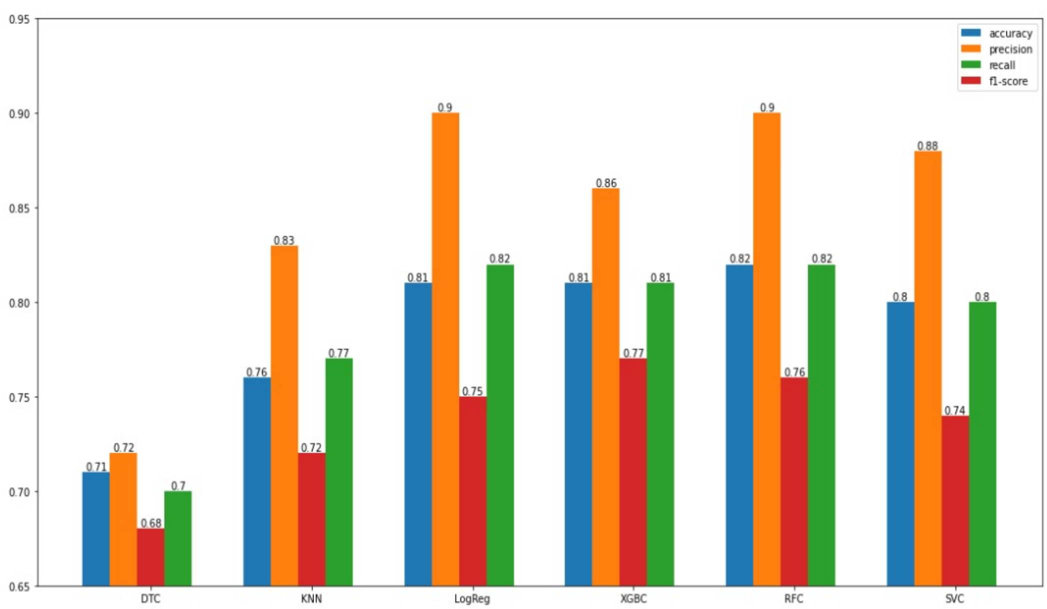


Figure 7: Comparison of results of classifiers

Conclusion

Heart disease is one of the main problems of society, as the number of people with heart diseases is increasing day by day. The growth of Statistics is influenced by many factors, such as the time spent by medicine to predict diseases or the lack of an accurate diagnosis. It is difficult to manually determine the probability of heart disease based on many such factors. But with deep data analysis and machine learning models, it is possible to identify diseases and treat these diseases in a timely manner. For this purpose, relevant data on heart disease collected by the University of Cleveland were studied. Work achieved and done during the study:

- analysis of the literature on the use of machine learning (ML) methods for data on heartbeats was carried out;
- analysis of python language libraries and part of machine learning methods;
- primary analysis of data on heart beauties and pre-processing;
- the marks were stitched, selected and methods of filling in the missing values were used;
- the results obtained were analyzed;
- based on the results, a comparison was made between the models.

According to the conducted research, the classification method showed the highest results. Its metrics showed accuracy = 0.82%, precision = 0.91%, recall = 0.83%, and f1-score = 0.76%. In the future, training of the algorithm on various data will continue, increasing these results given by Random Forest. Further experiments are developed on algorithms and optimal solutions are developed using various methods. Algorithms that have been trained to read various data on heart disease are also good at making predictions.

References

[1] Gusev A .V., Novitskiy R.E., Ivshin.A.A."Machine learning based on laboratory data for disease prediction", *FARMAKOEKONOMIKA. Modern Pharmacoeconomics and Pharmacoepidemiology*(2021): 8-9 p.

[2] R. Bharti [et al.],"Prediction of Heart Disease Using a Combination of Machine Learning and Deep Learning", *Computational Intelligence and Neuroscience*(2021): 3-6 p.

[3] Dimitris Bertsimas, Luca Mingardi, Bartolomeo Stellato "Machine Learning for Real-Time Heart Disease Prediction", *IEEE Journal of Biomedical and Health Informatics* (2021): 12-14 p.

[4] Mensah G.A., Roth G.A., Valentin Fuster, "The Global Burden of Cardiovascular Diseases and Risk Factors: 2020 and Beyond ", *Journal of the American College of Cardiology,* (2019): 2-9 p.

[5] Amini M., Zayeri F., Salehi M., "Trend analysis of cardiovascular disease mortality, incidence, and mortality-to-incidence ratio: results from global burden of disease study 2017", *BMC Public Health,* (2021): 12 p.

[6] Md Mamun Ali [et al.], "Heart disease prediction using supervised machine learning algorithms: Performance analysis and comparison", *Computers in Biology and Medicine.* (2021): 6-7 p.

[7] Amin Ul Haq [et al.], "A hybrid intelligent system framework for the prediction of heart disease using machine learning algorithms", *Mobile Information System* (2018): 11-13 p.

[8] Fajr Ibrahem Alarsan, Mamoon Younes, "Analysis and classification of heart diseases using heartbeat features and machine learning algorithms" , *Journal of Big Data* (2019): 8 p.

[9] Yunxing Jiang [et al.], "Cardiovascular disease prediction by machine learning algorithms based on cytokines in kazakhs of china" , *Clinical Epidemiology* (2021): 5 p.

[10] Yan Yan Song, Ying Lu., "Decision tree methods: applications for classification and prediction" , *Shanghai Archives of Psychiatry* (2015): 1-7 p.

[11] Pбdraig Cunningham, "K-Nearest Neighbour Classifiers-A Tutorial" , *ACM Computing Surveys* (2021): 6-9 p.

[12] Sandro Sperandei, "Understanding logistic regression analysis" , *Biochemia Medica* (2014): 5-9 p.

[13] Candice Bentejac, Anna Csurgo, Gonzalo Martinez-Munoz, "Data classification using support vector machine" , *Artificial Intelligence Review* (2021): 2-7 p.

[14] Leo Breima, "Random forests" , *Machine Learning 45,5-32* (2001)

[15] Durgesh K Srivastava, Lekha Bhambhu, "Data classification using support vector machine" , *Journal of Theoretical and Applied Information Technology* (2010): 3-9 p.

IRSTI 28.23.27 DOI: https://doi.org/10.26577/JMMCS.2022.v115.i3.010

E.B. Zhantileuov[1*], **A. Aibatbek**[1], **A.M. Smaiyl**[1], **A. Albore**[2], **E.A. Aitmukhanbetova**[1], **Sh.M. Saimassayeva**[1]
[1]Astana IT University, Kazakhstan, Astana
[2]IRT Saint Exupery, France, Toulouse
*e-mail: zhantileuov.eldiyar@Astanait.edu.kz

Multi-Agent Learning for the Inverse Kinematics of a Robotic Arm

This paper presents a solution to the inverse kinematics problem for robotic manipulator based on the Adaptive Multi-Agent System (AMAS) approach. In this research, multi-agent system is in charge of controlling a robot arm with four degrees of freedom (DOF) and two motorized wheels, giving appropriate commands, such as rotation angles and velocities, to reach the desired position and orientation of the end effector. The calculation of commands is directly related to the solving of forward and inverse kinematics. Before the learning process of AMOEBA, the rotational angles, θ values, are encoded into a single number N, this parameter is the desired value that we are going to predict in the predicting stage. During the learning phase, the Agnostic MOdEl Builder by self-Adaptation (AMOEBA) builds context agents, which has local models and is able to self-adapt. After the getting the predicted value, N_{pred}, it will be decoded back to get the set of rotational angles that is given to robot end effector. In addition, the robot with all its physical parameters is modeled and simulated in the Robot Operating System (ROS) environment

Key words: Forward kinematics, inverse kinematics, adaptive multi-agent system, agnostic model builder by self-adaptation.

Е.Б. Жантилеуов[1*], А. Айбатбек[1], А.М. Смайыл[1], А. Альборе[2], Э.А. Айтмұханбетова[1], Ш.М. Саймасаева[1]
[1] Astana IT University, Қазақстан, Астана қ.
[2]IRT Saint Exupery, Франция, Тулуза қ.
*e-mail: zhantileuov.eldiyar@Astanait.edu.kz

Робот қолының кері кинематикасы үшін мультиагенттік оқыту

Бұл жұмыста бейімделетін көп агенттік жүйе (Adaptive Multi-Agent System) тәсіліне негізделген роботтық манипуляторға арналған кері кинематика мәселесінің шешімі ұсынылады. Бұл зерттеуде мульти-агенттік жүйе төрт еркіндік дәрежесі (DOF) бар робот қолы мен екі дөңгелегін басқаруға жауапты. Роботтық қажетті позиция және бағдарына жетуі үшін, оның қолы мен дөңгелектеріне айналу бұрышы мен жылдамдық тәрізді тиісті командалар беріледі. Командаларды есептеу тура және кері кинематика есебін шешумен тікелей байланысты. AMOEBA-ның үйрену кезеңіне дейін θ айналу бұрыштары бір N санына шифрланады. Бұл параметр болжау кезеңіндегі біздің болжам жасайтын негізгі мән болып табылады. Үйрену кезеңінде Agnostic MODEl Builder by self-adaptation (AMOEBA) жергілікті үлгілері бар және өзін-өзі бейімдей алатын контекстік агенттерді құрады. Болжамды мән, N_{pred}, есептелініп алынғаннан кейін, айналу бұрыштарының жиынтығын алу үшін кері бағытта шифр ашылады. Бұл жиынтық роботтық атқарушы механизмі, яғни робот қолының саусақ ұшы, қажетті позиция және бағдарға жетуі үшін төрт еркіндік дәрежелі қолы мен екі дөңгелегіне команда ретінде беріледі. Сонымен қатар, робот өзінің барлық физикалық параметрлерімен Robot Operating System (ROS) ортасында модельденеді және имитацияланады.

Түйін сөздер: Кинематика, кері кинематика, адаптивті көп агенттік жүйе, өзін-өзі бейімдеу, агностикалық модель.

Е.Б. Жантилеуов[1*], А. Айбатбек[1], А.М. Смайыл[1], А. Альборе[2], Э.А. Айтмуханбетова[1], Ш.М. Саймасаева[1]
[1] Astana IT University, Казакстан, г. Астана
[2]IRT Saint Exupery, Франция, г. Тулуза
*e-mail: zhantileuov.eldiyar@Astanait.edu.kz

Мультиагентное обучение для обратной кинематики роботизированной руки

В данной статье представлено решение обратной задачи кинематики для робота-манипулятора на основе подхода Adaptive Multi-Agent System (AMAS). В этом исследовании мультиагентная система отвечает за управление манипулятором робота с четырьмя степенями свободы (DOF) и двумя моторизованными колесами, давая соответствующие команды, такие как углы поворота и скорости, для достижения желаемого положения и ориентации исполнительного механизма, то есть концевого эффектора. Расчет команд напрямую связан с решением прямой и обратной кинематики. На этапе обучения Agnostic MOdEL Builder путем самостоятельной адаптации (AMOEBA) создает агенты контекста, которые имеют локальные модели и способны к самостоятельной адаптации. Перед процессом обучения AMOEBA, углы поворота, θ значения, кодируются в одно число N, этот параметр является желаемым значением, которое мы собираемся предсказать на этапе прогнозирования. После получения предсказанного значения N_{pred}, оно будет декодировано обратно, чтобы получить набор углов поворота, заданный концевому исполнительному механизму робота. Кроме того, робот со всеми его физическими параметрами моделируется и симулируется в среде Robot Operating System (ROS).

Ключевые слова: Прямая кинематика, обратная кинематика, адаптивная мультиагентная система, независимый построитель моделей путем самостоятельной адаптации.

1 Introduction

Nowadays the study and development of intelligent robots are becoming an essential part of robotics. Many methods and approaches are aimed at making the robots fully automated and independent of external impacts, such as neural networks and multi agent systems. Major attention is paid to the motion of the robot, which, in turn, involves the study of kinematics. The general objective of this research is to reach the desired point or target with end-effector of robot with precise accuracy. In order to reach the goal, both forward and inverse kinematic problems must be solved. The forward kinematics (FK) involves determining the position and orientation of the robotic end-effector by giving values for each individual joint of robotic manipulator. Vice versa, by knowing the position and orientation of the end effector, the inverse kinematics (IK) is in charge with determination of values that must be set to the joints, in other words, inverse kinematics is the inverse problem of forward kinematics. In comparison with forward and the inverse kinematics, the solution of inverse kinematics is much more complicated. The FK can be easily solved by performing linear algebraic operations on homogeneous transformation matrices and has a unique solution. However, due to the complex IK equations, which is strongly nonlinear, there is no single solution for IK. As we mentioned, the IK is the main issue of robotics, and several methods are proposed for its solution [1]. Many approaches to this problem lie on the analytical, algebraic, or iterative methods, which give approximate results. Recently, much attention has been paid to artificial networks and self-adaptive multi-agent systems. The controlling of the robotic arm is considered as real-world complex problem and it cannot be solved by predefined model and needs learning and self-adaptation. 'Multi-agent systems are particularly suitable to design and implement self-organizing systems' [2]. In this paper, Self-Adaptive Context Learning

(SACL) recurrent pattern is applied to our problem. It consists of two mechanisms: Adaptive mechanism, which perceives information from the environment and dynamically builds a model describing the current context and its transformation Exploitation mechanism, which decides what actions to perform over the environment [2].

For building a dynamic model in adaptive mechanism, Agnostic MOdEl Builder by self-Adaptation (AMOEBA) is used. AMOEBA is based on AMAS approach. In order to be able to build a model, AMOEBA must learn on data provided by simulation or FK problem, which makes it supervised learning. In the final application, the multi-agent system will be integrated with machine learning, the function of which is to process an image, taken from the robot's camera, identify the target point and compute its distance and position relative to the camera. The integration of a machine-learning application with multi-agent system is another key feature of the project. The position of the button and the robot with all its physical parameters are simulated in ROS environment. Motivating Example. Figure 1 shows the real problem of the work. Consider a robot inside an elevator, the starting position and orientation of which are known. The robot's camera, which is attached on the end-effector, takes a picture of buttons in the elevator and the robot needs to press the desired button. Once the picture of elevator buttons is taken, the machine learning software identifies the desired button and calculates its position $(x,\ y,\ z)$ with respect to the camera. The coordinates of the button are then sent to the multi-agent system. Multi-agent system is responsible to control the 6 servo motors: 4 for robot arm and 2 for wheels. Taking the positions received from ML as input data, the multi-agent system solves IK problem to get rotation angles for each joint, $\theta_0,\ \theta_1,\ \theta_2,\ \theta_3$. The servo motors are given an angle setpoints and they rotate and maintain to reach this setpoint:

$$\text{CAMERA} \xrightarrow{image} \text{coord} \xrightarrow{x,y,z} \text{AMAS} \xrightarrow{\theta_0,\theta_1,\theta_2,\theta_3} \text{Robot Arm.}$$

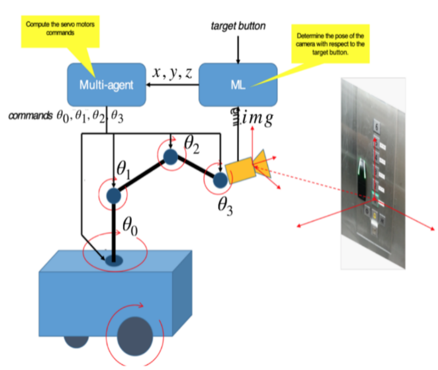


Fig. 1: The robot in an elevator, identifying the desired target and tries to reach it

Figure 2 contains a snapshot of the real robot with four degrees of freedom (DOF) arm which is placed on the platform. The platform has two motorized wheels and one castor wheel.

Fig. 2: The picture of the real robot with 4 DOF named TwIRTee

2 Simulating Physical Model of Robot under Robot Operating System (ROS)

This section describes the simulation model of the TwIRTee robot under ROS/Gazebo. This model includes the robot chassis with its two motorized wheels, the robot arm, and the Light Identification Detection and Ranging (LIDAR). It gives the procedure to setup the environment and to interact with the simulation via the com-mand line and programmatically. The two wheels are identical, so they are modeled using a macro with a parameter, "tY"that

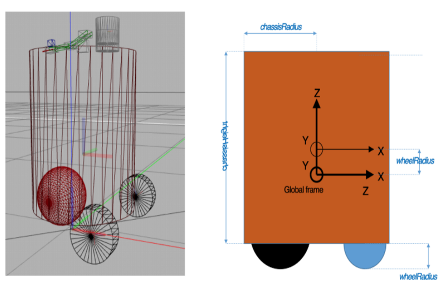


Fig. 3: The local frame for the chassis definition

gives the translation of the wheel with respect to the Y axis. Each wheel is drawn in a local frame that is obtained by a rotation of $\frac{\pi}{2}$ radians along the Y and Z axis with respect to the joint reference frame (see Figure 3). The robot is a set of links (such as the chassis described previously) and joints. Let's take the example of the robot arm that is fitted on top of the robot, as show on Figure 3, with a close-up view on Figure 4.

The arm is composed of: 4 servo motors (the green boxes): link0, link1, link3 and link5; two sets of "bars"(brown colored): link2 and link4; camera (in blue); "finger"(in red, at the tip of the arm):
"link1_joint"joints "link1"and "link2"with a "revolute"joint;
"link3_joint"joints "link2"and "link3"with a "revolute"joint;
"link4_joint"joints "link3"and "link4"with a "fixed"joint.
In the model, the camera is represented by a simple blue box (see Figure 5).

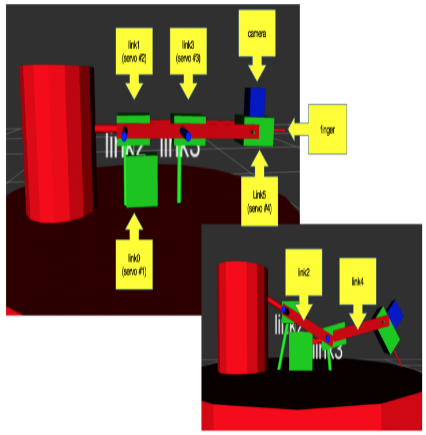


Fig. 4: The robotic arm closed view

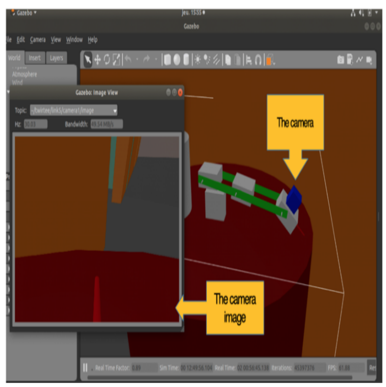


Fig. 5: The illustration of the camera attached to the end-effector (blue box)

This environment allows the complete dynamics of the system to be simulated, in-cluding the effect of inertia: the simulator receives the angles for each joint and com-putes the position of the arm and camera. Figure 6 shows the general idea of integrating the machine learning part with the multi-agent system in ROS environment. There are many related works with image processing and object detection and ML for image processing is quit out of this paper. The main task is to tackle with multi-agent system, to make the multi-agent system learn and self-adapt with precise accuracy.

3 Forward Kinematics

In robotics, forward kinematics is responsible for determining the final coordinates and the direction of the end-effector relative to the global coordinate space. Let's consider that the initial position and orientation of each servo motor is known. Ho-mogeneous transformation matrices with a dimension of 4x4 will be constructed from the base frame to the end effector frame [3]. These matrices consist of a 3x3 rotation matrix, that describes the orientation of

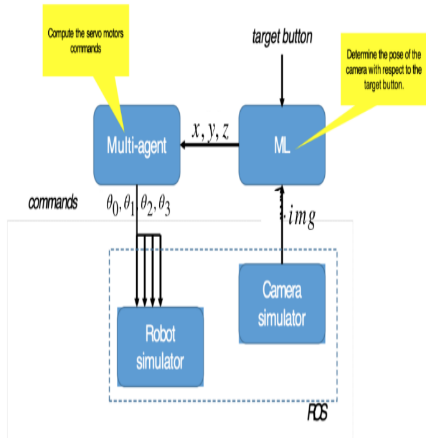


Fig. 6: The illustration of integration of ML and AMAS in ROS environment

joints and their behaviors, and trans-lation vector. Further, linear algebra operations will be performed on matrices to ob-tain FK results. In this section, more detailed solutions are provided for the robot arm.

3.1 Kinematics for 4 DOF Robotic Arm

In our case, the robotic arm has 4 degrees of freedom (DOF). The robot is articulated vertically with 4 joints. It has a stationary base, shoulder, elbow and wrist, where the base joint rotates around the z-axis and the other three rotate around the y-axis. The position of joints is represented in the three-dimensional Cartesian coordinate system and a local reference frame is assigned to each joint. The coordinate frame assign-ment is shown in Figure 7. In addition, it is necessary to assign a global coordinate frame to the base of the robot [4] (see Figure 8). The servo motors in three-dimensional space can have movements of

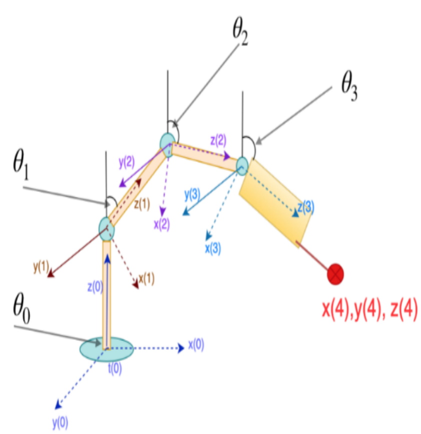


Fig. 7: The coordinate frame assignment of robotic arm

rotation and translation. The homogeneous transformation matrix (H.T.M) with dimension of 4x4 is constructed separately for each joint to describe its position and orientation relative to the world coordinate system. The H.T.M is composed of 3x3 rotation matrix and 3x1 translation vector:

$$\begin{vmatrix} . & . & . & . \\ . & R_{3\times 3} & . & t_{3\times 1} \\ . & . & . & . \\ 0 & 0 & 0 & 1 \end{vmatrix} \tag{1}$$

– Rotational matrix describes the rotation of joints in Euclidean space. The rotation is done about $z,\ y\ and\ x$ axes through a counterclockwise angle θ. The axis rotation matrices for a rotation about $z,\ y\ and\ x$ axes given, respectively [5]:

$$R_z(\theta) = \begin{vmatrix} 1 & 0 & 0 \\ 0 & \cos(\theta) & \sin(\theta) \\ 0 & -\sin(\theta) & \cos(\theta) \end{vmatrix} \tag{2}$$

$$R_y(\theta) = \begin{vmatrix} \cos(\theta) & 0 & \sin(\theta) \\ 0 & 1 & 0 \\ -\sin(\theta) & 0 & \cos(\theta) \end{vmatrix} \tag{3}$$

$$R_x(\theta) = \begin{vmatrix} \cos(\theta) & -\sin(\theta) & 0 \\ \sin(\theta) & \cos(\theta) & 0 \\ 0 & 0 & 1 \end{vmatrix} \tag{4}$$

– Translation or displacement vector shows the location of each joint in the base frame. The final translation vector is the answer of the FK problem. In order to obtain the final translational vector, the transformation matrices of each joint are multiplied. The sequence of multiplication is important, as it results in a trajectory generation step [5].

The transformation matrices of each joint are represented as ${}^{i}_{j}T$ (see Fig.8.):

1. Transformation matrix of base joint, rotates about z-axis:

$${}^{0}_{1}T = \begin{vmatrix} 1 & 0 & 0 & x_0 \\ 0 & \cos(\theta_0) & \sin(\theta_0) & y_0 \\ 0 & -\sin(\theta_0) & \cos(\theta_0) & z_0 \\ 0 & 0 & 0 & 1 \end{vmatrix} \tag{5}$$

2. Transformation matrix of shoulder joint, rotates about y-axis:

$${}^{1}_{2}T = \begin{vmatrix} \cos(\theta_1) & 0 & \sin(\theta_1) & x_1 \\ 0 & 1 & 0 & y_1 \\ -\sin(\theta_1) & 0 & \cos(\theta_1) & z_1 \\ 0 & 0 & 0 & 1 \end{vmatrix} \tag{6}$$

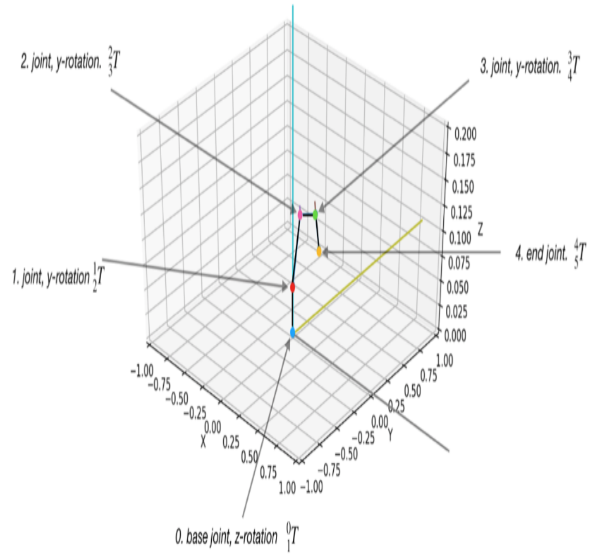


Fig. 8: The coordinate frame assignment of robotic arm in world space

3. Transformation matrix of elbow joint, rotates about y-axis:

$$^2_3T= \begin{vmatrix} \cos(\theta_2) & 0 & \sin(\theta_2) & x_2 \\ 0 & 1 & 0 & y_2 \\ -\sin(\theta_2) & 0 & \cos(\theta_2) & z_2 \\ 0 & 0 & 0 & 1 \end{vmatrix} \tag{7}$$

4. Transformation matrix of wrist joint, rotates about y-axis:

$$^3_4T= \begin{vmatrix} \cos(\theta_3) & 0 & \sin(\theta_3) & x_3 \\ 0 & 1 & 0 & y_3 \\ -\sin(\theta_3) & 0 & \cos(\theta_3) & z_3 \\ 0 & 0 & 0 & 1 \end{vmatrix} \tag{8}$$

5. Transformation matrix of end joint:

$$^4_5T= \begin{vmatrix} 1 & 0 & 0 & x_4 \\ 0 & 1 & 0 & y_4 \\ 0 & 0 & 1 & z_4 \\ 0 & 0 & 0 & 1 \end{vmatrix} \tag{9}$$

Finally, the desired transformation matrix is obtained by multiplying all ^i_jT matrices:

$$^0_5T = {^0_1T} \cdot {^1_2T} \cdot {^2_3T} \cdot {^3_4T} \cdot {^4_5T} \tag{10}$$

where ${}^0_5T$ has the form of:

$$\begin{vmatrix} . & . & . & x \\ . & R_{3\times 3} & . & y \\ . & . & . & z \\ 0 & 0 & 0 & 1 \end{vmatrix}. \tag{11}$$

The $(x,\ y,\ z)$ is the answer to the FK problem. The $(x_4,\ y_4,\ z_4)$ is the position of the end-effector in the local coordinate space (see Figure 7) and $(x,\ y,\ z)$ is the position of the end-effector on the world system, in another words, the coordinates of the end-effector are translated to the global coordinate system.

The Implementation of FK on 4 DOF Robotic Arm Implementing the FK to determine the final position and orientation of the end-effector is done in Python.

Suppose the arm of the robot is raised up initially. The rotation angles are given to each servo motor, i.e. the rotation angle setpoints, $(\theta_0,\ \theta_1,\ \theta_2,\ \theta_3)$, are sent to the base, shoulder, elbow and wrist. Depending on the given angles, the motors begin to rotate. The final location of the end-effector is determined by the translation of the coordinate from the local coordinate system to the global one. Input data is angular setpoint, the output is the coordinates of the end-effector on the global system:

$$\theta_0,\ \theta_1,\ \theta_2,\ \theta_3 \xrightarrow{FK\_4Dof} x,\ y,\ z.$$

The result of the problem is illustrated in Python on Figure 9. Initially, the robot arm is

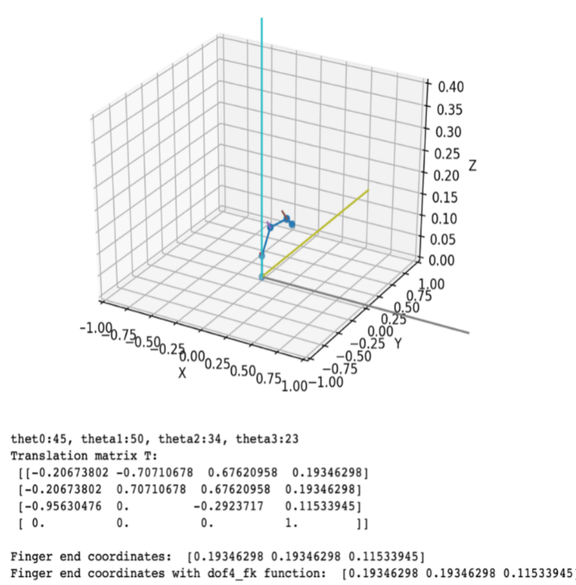


Fig. 9: The result of the example to check the correctness of implementation FK.

raised up and $\theta_0 = 45,\ \theta_1 = 50,\ \theta_2 = 34,\ \theta_3 = 23$ is given to the motors. The dark blue curve is the final position and orientation of the arm manipulators. The final position of the end-effector computed by FK is (0.19346298, 0.19346298, 0.11533945).

4 Inverse Kinematics

Summing up the previous section, we can say that solving forward kinematics for robotic manipulators is a fairly simple task, only linear algebra operations are performed on matrices to determine the final position and orientation of the end-effector. The problem has one and only one solution. However, when the final position and orientation of the manipulators is initially given, and the task is to find the rotation angles for each joint, (θ_0, θ_1, θ_2, θ_3), the problem becomes non-linear and complex. This kind of task in robotics is called Inverse Kinematics Problem. There are many approaches to solving inverse kinematics problem, e.g. analytical solution, numerical methods, artificial neural networks and self-adaptive multi-agent systems. In this paper, we propose Adaptive Multi-Agent System based solution for solving IK problem.

5 Adaptive Multi-Agent System

To solve IK problem, we need to prepare a model which is responsible to predict the revise the degrees of liberty and “rotation time”. However, due to the complexity of the problem, it is difficult and expensive to solve using a predefined model; instead, we will use several agents, an autonomous entities, responsible for predicting the result. A system where the agents are plugged-in should be able to adapt to the environment and learn independently. The Adaptive Multi-Agent Systems (AMAS) approach has been applied to designed and developed self-adaptive multi-agent system. This approach aims at solving problems in dynamic non-linear environments by a bottom-up design of cooperative agents, where cooperation is the engine of the self-organization process [7].

6 The Self-Adaptive Context Learning Pattern

Our self-adaptive system is connected with a dynamic environment by a cycle of observations. The main task of system is to receive the observations coming from the environment and find a proper actions for the current state of inputs, which, in turn, is called the context [8]. This is a context mapping problem. The Self-Adaptive Context Learning (SACL) is recurrent pattern, based on the AMAS approach, the key feature of which is to solve the context-mapping sub-problem. It is composed of two mechanisms, that interacts with the environment:

- Adaptation mechanism, is dynamically building a model, that describes the current context and possible actions in it. [2, 8] It is related to the learning phase of the system and its changes.
- Exploitation mechanism, is in charge with the selecting the most appropriate action in the current context.

7 AMOEBA: Agnostic MOdEl Builder by Self-Adaptation

The building of the model in adaptation mechanism is performed by using Agnostic MOdEl Builder by Self-Adaptation (AMOEBA), based on the AMAS approach. The model explains the interaction that occurs between the mechanism of exploitation and the environment [2]. The model receives a set of input data, we call it percepts, and produces one output. We call the obtained result as prediction and the actual, correct result is called oracle.

There are two types of agents in AMOEBA [9]:

- Percept agents are responsible for the perceiving information from the environment.
- Context agents are in charge of determination the context, where a specific output would be a good one.

AMOEBA learning phase is done by building the context agents. Each context agent has its own validity range and local model. The validity range of the context agent is the interval, where a specific output will be relevant [9]. If the received value of the percept agent is included in the validity range interval, we say the validity range is valid for this percept. The context agents have rectangular shape in two-dimension space (see Figure 10). The local model is built separately for each context agent. When the validity range of the

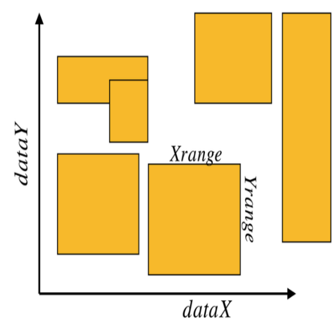


Fig. 10: The context agents in AMOEBA

context agent is valid for the current perceived value, the output is calculated by using the local model of that context agent. In this paper, the linear regression is used as a model. The linear regression function computed using a set of points [9]:

$$\sum_{n=1}^{p} x_n v_n + a \tag{12}$$

where p is the number of percepts, x_n and v_n are the coefficients, a is the real number. The creation of the context agent, the self-organization, the changing of the validity ranges, the changing local model and the destroying itself is deeply described in reference [9].

Working Principle of AMEOBA. At first, AMOEBA must learn from examples with the correct outputs. This approach of learning is called supervised learning. Once, the AMOEBA is learned, it starts to predict the result for a new inputs.

Let's look at the illustration taken from reference [9]:

1. During the learning phase, AMOEBA uses incoming data to adapt and improve itself. The specific data set with the correct result is given to AMOEBA. However, at the

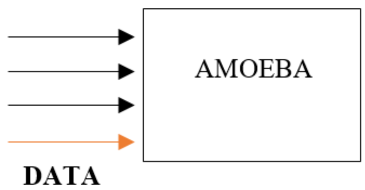


Рис. 11: Learning phase of AMOEBA, with the given oracle (red arrow)

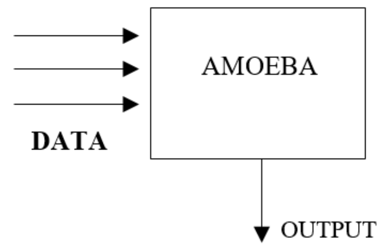


Рис. 12: The exploitation step of AMOEBA, without labeled data

beginning, the oracle, actual result, is “hidden” from AMOEBA. The valid context agent tries to predict the output, and checks the predicted value with the oracle. If it was wrong, it adapts and improves itself by reducing its validity range or changing the local model (see Figure 11).

2. During the exploitation step, AMOEBA receives a set of data without an oracle. Based on the previous knowledge it provides an output (see Figure 12).

AMOEBA for the Inverse Kinematics Problem. In IK problem for the robot arm, the input data are the final coordinates of the end-effector, $(x,\ y,\ z)$, remember that in practice these coordinates are taken from machine learning software. The output is a set of rotation angles for each servo motor, $(\theta_0,\ \theta_1,\ \theta_2,\ \theta_3)$. Then the servo motors execute the given commands to achieve the $(x,\ y,\ z)$ target position. This means that we must predict 4 parameters for the robot arm. However, AMOEBA learns to predict only one parameter at the time. So, using four independent AMOEBAs to perform the learning of each parameter can give physically unreachable commands to the robot arm because the correlation between each parameter would be lost; e.g. to reach point (x, y, z) the arm has many ways to reach desired point by varying its angles, a single solution is given by an arm configuration expressing four angles which depend from each other in each configuration. If the learning process for each angle is independent, the prediction for the angle will be non-correlated to the one of the other angles, resulting in an “impossible” arm configuration. Therefore, in order to preserve the correlation between the joint’s positions of the robotic manipulators, we decided to encode the four angles $(\theta_0, \theta_1, \theta_2, \theta_3)$ to one single number N. This number N is used as an oracle in the learning process. Then, when the (x, y, z) coordinate is provided to the trained system, the number encoding the joints angles is given as output, and decoded for the final application.

In order to AMOEBA to predict values, the system need to be trained. Therefore, a training data set should be provided.

Training Data for AMOEBA. The learning data for AMOEBA is built in Python programming language. Several training sets of 100, 1000 and 5000 examples respectively, are randomly generated in different files. The angle values uniformly cover the fallowing ranges:

$$\theta_0 \in (0;\ \pi)\,,\ \theta_1 \in \left(0;\ \frac{\pi}{2}\right),\ \theta_2 \in \left(0;\ \frac{\pi}{2}\right),\ \theta_3 \in \left(0;\ \frac{\pi}{2}\right).$$

For each example, the final position and orientation of the end-effector is calculated by solving FK problem (see Figure 13). The result of FK problem is exact and stored in vector $(x,\ y,\ z)^{\mathrm{T}}$ form.

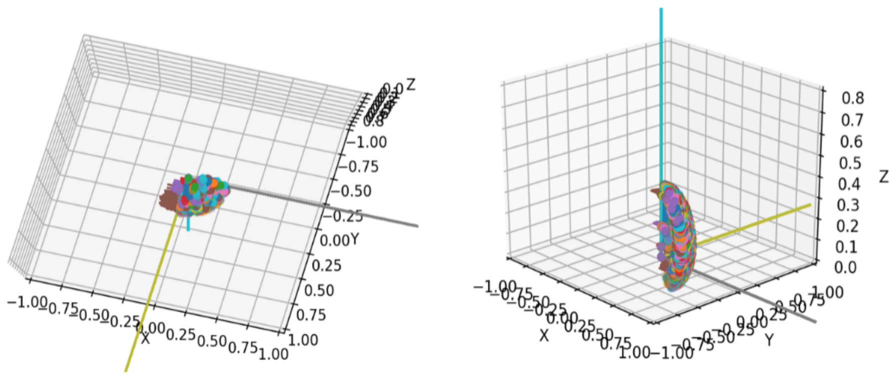


Fig. 13: The resulting positions of the end-effector for 5000 randomly generated set of joint angle values

Finally, each example used in the training file is a row in a table consisting in $\theta_0,\ \theta_1,\ \theta_2,\ \theta_3,\ x,\ y,\ z$ parameters, and the respective encoding of the joint positions, given that the learning ability of AMOEBA is limited by only one parameter. This means that, we feed AMOEBA with data, that has the correct answers or oracles.

Encoding and Decoding of θ Values. The process of encoding θ values into a single number, N, occurs before the learning process of AMOEBA. The number N is used as an oracle at the learning stage (red colored) and at the predicting stage, this is the value that we aim to predict. The value of $N_{predict}$ is then decoded to retrieve $\theta_0 pred$, $\theta_1 pred$, $\theta_2 pred$ and $\theta_3 pred$ (see Figure14).

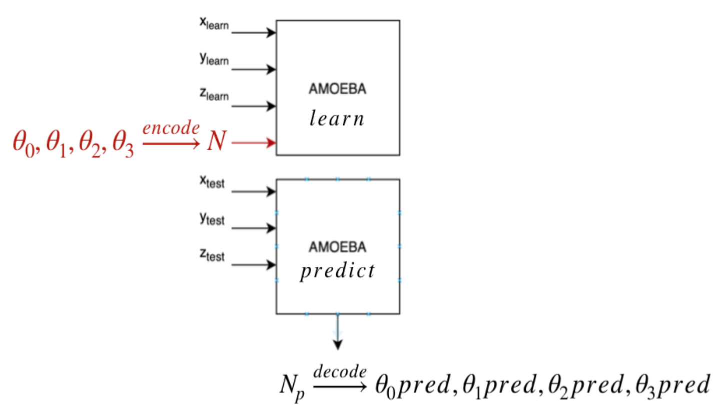


Fig. 14: The role of encoding/decoding in the learning and predicting stages of AMOEBA.

Let's see how θ values are encoded. The movement of joints are limited within the following ranges:

$$\theta_0 \in (0; \pi), \theta_1 \in \left(0; \frac{\pi}{2}\right), \theta_2 \in \left(0; \frac{\pi}{2}\right), \theta_3 \in \left(0; \frac{\pi}{2}\right),$$

and the maximum value that angle can assume is 180^0. This value, incremented by 1, is called the base ($B = 181$). Finally, the value N is calculated:

$$N = \theta_0 \times B^3) + (\theta_1 \times B^2) + (\theta_2 \times B^1) + (\theta_3 \times B^0) \tag{13}$$

To decode N, we divide it by base B. The value in remainder is θ_3. In order to get θ_2, θ_1 and θ_0, the division process is repeated, but instead of N, quotient of previous division is used.
Example 1. Let's encode and decode the set of angles:

$$\theta_0{=}45^0; \theta_1{=}23^0;\ \theta_2{=}\ 54^0; \theta_3{=}89^0$$

$$N{=}\left(45{\times}181^3\right) + \left(23{\times}181^2\right) + \left(54{\times}181^1\right) + \left(89{\times}181^0\right) = 267601711$$

The four values of θ are encoded in one N.
The decoding of N :

$$267601711 \div 181 = 1478462\ \ (remainder\ 89)$$

$$1478462 \div 181 = 8168\,(remainder\ 54)$$

$$8168 \div 181 = 45\,(\text{remainder}23)$$

$$45 \div 181 = 0\,(\text{remainder}45)$$

The values in remainders are our angles, which we encode earlier.

Returning to our training data, let's encode all the joint angles. Table 1 represents several lines from the real dataset.

$\boldsymbol{\theta_0}$	$\boldsymbol{\theta_1}$	$\boldsymbol{\theta_2}$	$\boldsymbol{\theta_3}$	$\boldsymbol{x}$	$\boldsymbol{y}$	$\boldsymbol{z}$	$\boldsymbol{N}$
7	12	78	19	0.216441	0.026575	0.13153	5207239
72	18	36	74	0.059224	0.182274	0.20249	52637114
62	27	11	60	0.087914	0.165342	0.25034	45417750
53	48	83	72	0.101095	0.134158	-0.027521	39033342
...	...	...	...	...	...	...	...
19	83	53	48	0.189071	0.065102	-0.095592	14528118

Table 1. A training data for AMOEBA

Now instead of fourfold training for each θ, AMOEBA will be trained once on the values of N.
Learning Phase of AMOEBA. In the problem of inverse kinematics for the robot arm, AMOEBA starts learning by mapping $(x,\ y,\ z)$ into cartesian plane. Note that: the oracle is N. For each point, AMOEBA randomly produces a value N_{pred}. If this value is closer to the oracle, the validity range of context agent expands, and vice versa, if the difference between the exact value of N and the predicted value of N is large, the range becomes smaller. If

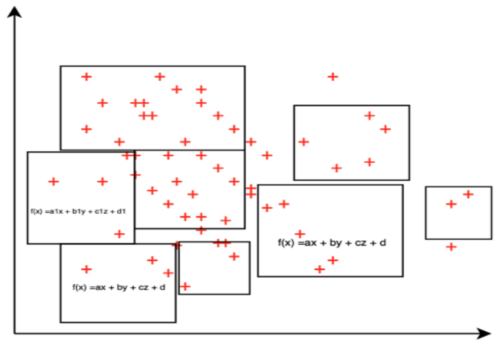


Fig. 15: The illustration of context agents (red crosses are percept agents; the rectangles are context agents). Each.

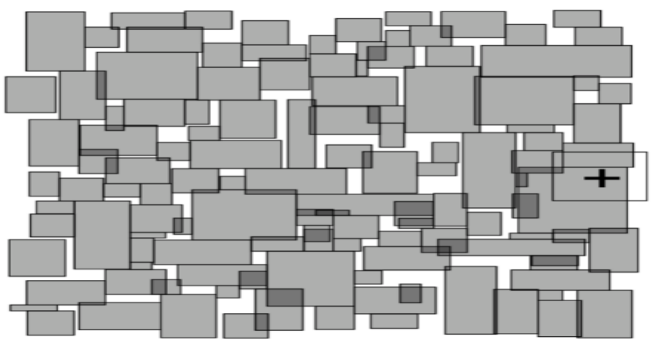


Fig. 16: 2D visualization of learning AMOEBA in JAVA.

the validity range of context agent is too small, AMOEBA decides that it is useless and the context agent will be self-destroyed. Context agents with their local regression models are illustrated on Figure 15.

Validation Phase. To estimate how well AMOEBA was trained, we need to provide a testing dataset. Just like in training dataset, this file consists of 100 lines of $\theta_0,\ \theta_1,\ \theta_2,\ \theta_3,\ x,\ y,\ z,\ N$ *values. However*, at this stage we will use the oracle only to calculate model error. This means that the input for AMOEBA is only $x,\ y,\ z$, remember that at the learning stage, the input was $x,\ y,\ z$ and oracle N. Based on previous knowledge, AMOEBA predicts the value of N, the output is N_{pred}. This output is then decoded to get $\theta_0 pred,\ \theta_1 pred,\ \theta_2 pred$ and $\theta_3 pred$. Next, we simply solve FK problem for predicted joint angles and obtain the predicted coordinates of the end-effector, $(x_{pred},\ y_{pred},\ z_{pred})$. These steps are described in the following scheme:

$$x,\, y,\, z \xrightarrow{input} AMOEBA \xrightarrow{predic} N_{pred} \xrightarrow{decode} \theta_0 pred,\, \theta_1 pred,\, \theta_2 pred,\, \theta_3 pred$$

$$\xrightarrow{solve\ FK} x_{pred},\, y_{pred},\, z_{pred}.$$

The performance of AMOEBA was determined based on the Euclidean distance of two points and the mean squared error (MSE) between the predicted output and the expected output.

8 Experimental results

First of all, I generated 3 training datasets for AMOEBA with 100, 1000 and 5000 rows of θ values. To find the corresponding localizations, the problem of forward kinematics has been solved and for each row, the values of θ were encoded into a single N (see Tab.2). These data were then transmitted to AMOEBA, so that it could learn. After each training with the data of different sizes, another testing dataset is given, to check the correctness the model.

θ_0	θ_1	θ_2	θ_3	x	y	z	N
7	16	19	19	0.15	0.02	0.3	5234329
81	65	63	54	0.03	0.2	-0.05	59581224
78	9	68	17	0.04	0.21	0.18	56941037
30	21	13	12	0.13	0.08	0.3	22041282
...	...	...	...	...	...	...	...
55	2	82	87	0.09	0.13	0.116	40118667

Table 2. An example of learning data for AMOEBA

Once AMOEBA is trained, the validation phase is conducted. Tables 3, 4 and 5 show the results of validation stage after training with 100, 1000 and 5000 data rows, respectively.

x	y	z	N_{pred}	θ_0p	θ_1p	θ_2p	θ_3p	x_p	y_p	z_p	ED	MSE
0.15	0.02	0.3	9720459	13	30	5	9	0.17	0.04	0.3	0.02	0.001
0.03	0.2	-0.05	55375708	75	86	45	58	0.05	0.2	-0.1	0.05	0.002
0.04	0.21	0.18	59999713	82	27	33	43	0.03	0.22	0.2	0.03	0.001
0.13	0.08	0.3	27920717	38	27	0	17	0.12	0.09	0.31	0.02	0.00
...	...	...	...	...	...	...	...	...	...	...	...	...
0.09	0.13	0.116	37979806	52	8	77	76	0.11	0.14	0.11	0.02	0.01

Table 3. After training AMOEBA with 100 rows of data, the mean Euclidean distance is 0.33 and the $\sum MSE = 0.03$.

x	y	z	N_{pred}	θ_0p	θ_1p	θ_2p	θ_3p	x_p	y_p	z_p	ED	MSE
0.15	0.02	0.3	5250410	7	18	17	80	0.16	0.02	0.25	0.05	0.003
0.03	0.2	-0.05	59594295	81	67	28	75	0.04	0.25	0.03	0.09	0.009
0.04	0.21	0.18	57170504	78	38	7	74	0.04	0.21	0.21	0.03	0.001
0.13	0.08	0.3	21849462	29	87	41	72	0.18	0.10	-0.1	0.39	0.147
...	...	...	...	...	...	...	...	...	...	...	...	...
0.09	0.13	0.116	40903374	56	9	71	84	0.10	0.15	0.13	0.02	0.001

Table 4. After training AMOEBA with 1000 rows of data, the mean Euclidean distance is 0.12 and the $\sum MSE = 2.66$.

$\boldsymbol{x}$	$\boldsymbol{y}$	$\boldsymbol{z}$	$\boldsymbol{N_{pred}}$	$\boldsymbol{\theta_0 p}$	$\boldsymbol{\theta_1 p}$	$\boldsymbol{\theta_2 p}$	$\boldsymbol{\theta_3 p}$	$\boldsymbol{x_p}$	$\boldsymbol{y_p}$	$\boldsymbol{z_p}$	ED	MSE
0.15	0.02	0.3	5229175	7	15	51	85	0.19	0.02	0.16	0.14	0.021
0.03	0.2	-0.05	59609975	81	69	23	5	0.05	0.29	0.07	0.15	0.023
0.04	0.21	0.18	56898316	78	4	43	46	0.03	0.16	0.25	0.08	0.008
0.13	0.08	0.3	22157307	30	35	42	27	0.22	0.13	0.15	0.18	0.033
...	...	...	...	...	...	...	...	...	...	...	...	...
0.09	0.13	0.116	40903374	56	9	71	84	0.10	0.15	0.13	0.02	0.001

Table 5. After training AMOEBA with 5000 rows of data, the mean Euclidean distance is 0.1 and the $\sum MSE = 3.12$.

With the help of testing data, we can compute the Euclidean distance between two points, $(x,\ y,\ z)$ and $(x_{pred},\ y_{pred},\ z_{pred})$. With an increase in the training data, the mean Euclidean distance decreased, and the sum of the mean squared error increased (see Figure 17). The explanation for this is closely related to the number of context agents. When we try to train AMOEBA with more data, it also tries to build a perfect model. Thus, it breaks down the initial context agents into several small ones. When we have more context agents than necessary, our model becomes overfitted. On the other hand, we can notice that the values of

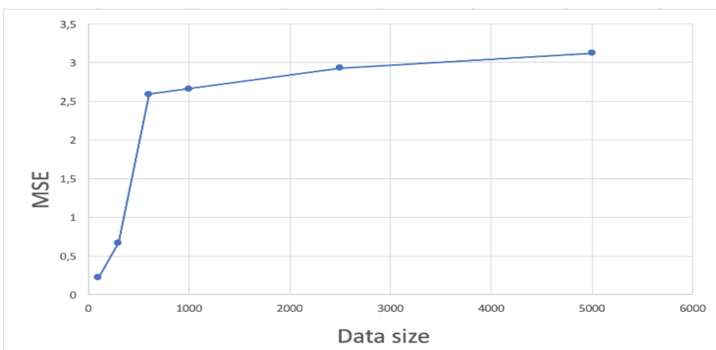


Fig. 17: The graph of mean squared error of different data size.

N_{pred} are approximated to the real values of N and that AMOEBA always perfectly coincides with the first angle. So, I found that the order of θ values at the encoding stage is highly important, since when encoding, the first value is multiplied by the highest base accordance with equation (12).

Returning to the problem, at the encoding stage, we need to encode so that each θ is occurred first in order (see Figure 28).

For each N, we create 4 independent AMOEBAs. Note that in this case all links and relations will be preserved between θ parameters. Further, all other steps will be the same for this learning. Only, at predicting stage, we select the better values of θ from each independent learning. Table 6 shows the testing phase results after training AMOEBA with 5000 data rows.

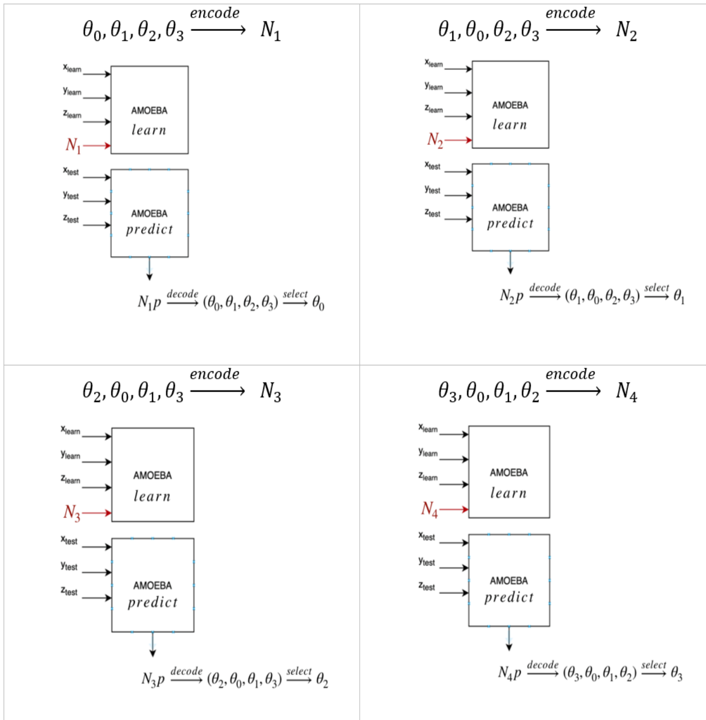


Fig. 18: At the encoding step, we encode so that each θ will be first in order.

x	y	z	$\theta_0 p$	$\theta_1 p$	$\theta_2 p$	$\theta_3 p$	x_p	y_p	z_p	ED	MSE
0.08	0.18	-0.05	67	50	69	31	0.09	0.21	0.00	0.06	0.004
0.04	0.17	-0.11	78	85	56	65	0.04	0.17	-0.1	0.01	0.000
0.20	0.16	-0.04	39	69	26	37	0.22	0.18	0.04	0.09	0.007
0.18	0.11	0.19	31	34	18	56	0.19	0.11	0.21	0.03	0.000
...	...	...	...	...	...	...	...	...	...	...	...
0.27	0.10	0.10	56	9	71	84	0.15	0.06	0.29	0.23	0.052

Table 6. The result of validation phase, after training AMOEBA with 5000 lines of data. To obtain this results, learning is done independently of each other with the oracle N.

The result of latter method is pretty impressive: after 5000 learning, $\sum MSE = 0.008$ and average Euclidean distance is 0.06.

9 Conclusion

This study presented a detailed solution for inverse kinematics problem using an Adaptive Multi-Agent System approach.

To avoid parameter non-correlation, we encoded the output / input of the IK problem as one base-dependent number. Using this approach, we were able to predict the position and orientation of the robot arm joints, given a final position.

The results show that the size of the training set is relevant to the performances, as the bigger it is, the model becomes more complex and the MSE increases. To make the error less, four AMOEBAs were trained with the differently encoded labels.

This applications were aimed as a part of a more complex system, involving machine learning techniques to identified a goal for a robot, and multi-agent system to elaborate the robotic arm position to reach this goal.

References

[1] Samuel R. Buss "Introduction to Inverse Kinematics with Jacobian Transpose, Pseudoinverse and Damped Least Squares methods"IEEE Journal of Robotics and Automation, 16(2004): 1-19.

[2] N. Verstaevel, J. Boes, J. Nigon, D. d'Amico, M. Gleizes "Lifelong Machine Learning with Adaptive Multi-Agent Systems In Proceedings of the 9th International Conference on Agents and Artificial Intelligence, 1(2017): 275-286.

[3] R.R. Serrezuela, A.F.C. Chavarro, M.A.T. Cardozo, A.L. Toquica, L.F.O. Martinez, "Kinematic modelling of a robotic arm manipulator using Matlab ARPN Journal of Engineering and Applied Sciences, 12:7(2017): 2037-2045.

[4] A. Mohammed "Forward and Inverse Kinematic Analysis and Validation of the ABB IRB 140 Industrial Robot"International journal of electronics, mechanical and mechatronics engineering, 7:2(2017): 1383-1401.

[5] Kwon3d.com. (1998). Rotation Matrix. [online] Available at: http://www.kwon3d.com/theory/transform/rot.html [Accessed 28 Aug. 2019].

[6] G. Dudek and M. Jenkin, "Computational Principles of Mobile Robotics"Cambridge University Press, USA, 2nd edition, 2010.

[7] J.-P. Georgé, M.-P. Gleizes, and V. Camps, "Cooperation In Di Marzo G. Serugendo, M.-P. Gleizes, and A. Karageogos, editors, Self-organising Software, Natural Computing Series, pages 7-32. Springer Berlin Heidelberg, 2011.

[8] J. Boes, J. Nigon, N. Verstaevel, M.-P. Gleizes, F. Migeon, "The Self-Adaptive Context Learning Pattern: Overview and Proposal International and Interdisciplinary Conference on Modeling and Using Context (CONTEXT), 9405(2015) in LNAI, Springer, Larnaca, Cyprus, 91-104. .

[9] J. Nigon, E. Glize, D. Dupas, F. Crasnier, J. Boes, "Use Cases of Pervasive Artificial Intelligence for Smart Cities Challenges IEEE Workshop on Smart and Sustainable City (WSSC 2016) associated to the International Conference IEEE UIC (2016): 1021-1027.

Список литературы

[1] Samuel R. Buss "Introduction to Inverse Kinematics with Jacobian Transpose, Pseudoinverse and Damped Least Squares methods"IEEE Journal of Robotics and Automation, 16(2004): 1-19.

[2] N. Verstaevel, J. Boes, J. Nigon, D. d'Amico, M. Gleizes "Lifelong Machine Learning with Adaptive Multi-Agent Systems In Proceedings of the 9th International Conference on Agents and Artificial Intelligence, 1(2017): 275-286.

[3] R.R. Serrezuela, A.F.C. Chavarro, M.A.T. Cardozo, A.L. Toquica, L.F.O. Martinez, "Kinematic modelling of a robotic arm manipulator using Matlab ARPN Journal of Engineering and Applied Sciences, 12:7(2017): 2037-2045.

[4] A. Mohammed "Forward and Inverse Kinematic Analysis and Validation of the ABB IRB 140 Industrial Robot"International journal of electronics, mechanical and mechatronics engineering, 7:2(2017): 1383-1401.

[5] Kwon3d.com. (1998). Rotation Matrix. [online] Available at: http://www.kwon3d.com/theory/transform/rot.html [Accessed 28 Aug. 2019].

[6] G. Dudek and M. Jenkin, "Computational Principles of Mobile Robotics"Cambridge University Press, USA, 2nd edition, 2010.

[7] J.-P. Georgé, M.-P. Gleizes, and V. Camps, "Cooperation In Di Marzo G. Serugendo, M.-P. Gleizes, and A. Karageogos, editors, Self-organising Software, Natural Computing Series, pages 7-32. Springer Berlin Heidelberg, 2011.

[8] J. Boes, J. Nigon, N. Verstaevel, M.-P. Gleizes, F. Migeon, "The Self-Adaptive Context Learning Pattern: Overview and Proposal International and Interdisciplinary Conference on Modeling and Using Context (CONTEXT), 9405(2015) in LNAI, Springer, Larnaca, Cyprus, 91-104. .

[9] J. Nigon, E. Glize, D. Dupas, F. Crasnier, J. Boes, "Use Cases of Pervasive Artificial Intelligence for Smart Cities Challenges IEEE Workshop on Smart and Sustainable City (WSSC 2016) associated to the International Conference IEEE UIC (2016): 1021-1027.

4-бөлім	Раздел 4	Section 4
Қолданылмалы математика	**Прикладная математика**	**Applied mathematics**

IRSTI 27.35.31 DOI: https://doi.org/10.26577/JMMCS.2022.v115.i3.011

L.A. Khajiyeva*, **Yu.F. Sabirova**, **R.F. Sabirova**
Al-Farabi Kazakh National University, Kazakhstan, Almaty
*e-mail: khadle@mail.ru

MODELLING OF HORIZONTAL DRILL STRING MOTION BY THE LUMPED-PARAMETER METHOD

The motion of drill strings is modeled in the drilling of geotechnological wells in the mining industry by the Lumped-Parameter Method (LPM). This method is widely used in structural mechanics and is most justified in modeling dynamic systems with a variable structure. On the example of horizontal drilling of geotechnological wells, longitudinal vibrations of a drill string with a static compressive load at the left end are considered [1]. The contact interaction of the drill string with the borehole walls and the inertia force of the bit on the destructible rock at the right end of the string are taken into account. The analysis of the column splits number, which specifies the dimension of the system of discrete equations, is carried out by verifying the obtained results with the previously known data [1]. For verification, the developed C# software was used, allowed to determine the error of the column splits in comparison with the test data. The optimal number of the drill string splits in terms of "implementation time – calculation error" by the LPM was identified. The numerical implementation of the model is conducted by the fourth-order Runge-Kutta method. In connection with the increase in the implementation time of the program code due to the increase in the dimension of the system, the numerical algorithm is optimized using the parallel programming tools. The expediency of this optimization is analyzed.
Key words: drill string, nonlinear, vibrations, lumped-parameter method, parallel programming.

Л.А. Хаджиева*, Ю.Ф. Сабирова, Р.Ф. Сабирова
Әл-Фараби атындағы Қазақ ұлттық университеті, Қазақстан, Алматы қ.
*e-mail: khadle@mail.ru

Жиынтық параметрлер әдісі бойынша бұрғылау бағананың көлденден қозғалуын модельдеу

Бұрғылау бағананың қозғалысы тау-кен өнеркәсібінде геотехнологиялық ұңғымаларды игеруде жиынтық параметрлер әдісімен (ЖПӘ) модельденеді. Бұл әдіс құрылымдық механикада кеңінен қолданылады және айнымалы құрылымы бар динамикалық жүйелерді модельдеуде барынша негізделген. Геотехнологиялық ұңғымаларды көлденең бұрғылау мысалында оның сол жақ шетінде статикалық қысу жүктемесі бар бұрғылау бағананың бойлық тербелістері қарастырылған [1]. Бұрғылау бағананың ұңғыма қабырғаларымен жанасу әрекеті және тізбенің оң жақ шетінде жойылатын жыныcқа қашау инерция күші ескеріледі. Алынған нәтижелерді бұрын белгілі [1] деректермен тексеру арқылы, дискретті теңдеулер жүйесінің өлшемін белгілейтін бағананың бөлімдер санының талдауы жүргізіледі. Тексеру үшін C# тілінде әзірленген бағдарламалық қамтамасыз ету пайдаланылады, бұл сынақ деректерімен салыстырғанда шығарылған бағана бөлімдерінің қателігін анықтауға мүмкіндік береді. Бағананың бөлімдерінің оңтайлы саны «есептеу уақыты-есептеу қатесі» тұрғысынан анықталады. Модельдің сандық орындалуы 4-ші ретті Рунге-Кутта әдісімен жүзеге асырылды. Жүйе өлшемінің өсуі мен программалық кодты орындау уақытының ұлғаюына байланысты параллельді бағдарламалау құралдарының көмегімен сандық алгоритм оңтайландырылды. Осы оңтайландырудың орындылығына талдау жүргізілді.
Түйін сөздер: бұрғылау бағана, бейсызықтылық, тербелістер, жиынтық параметр әдісі, параллельді бағдарламалау.

Л.А. Хаджиева*, Ю.Ф. Сабирова, Р.Ф. Сабирова
Казахский национальный университет имени аль-Фараби, Казахстан, г. Алматы
*e-mail: khadle@mail.ru

Моделирование горизонтального движения буровой колонны методом сосредоточенных параметров

Моделируется движение буровых колонн при разработке геотехнологических скважин в добывающей промышленности методом сосредоточенных параметров (МСП). Данный метод широко применяется в строительной механике и наиболее оправдан при моделировании динамических систем с переменной структурой. На примере горизонтального бурения геотехнологических скважин рассмотрены продольные колебания буровой колонны со статической сжимающей нагрузкой на ее левом конце [1]. Учтены контактное взаимодействие буровой колонны со стенками скважины и сила инерции долота на разрушаемую породу на правом конце колонны. Посредством верификации полученных результатов с ранее известными данными [1] проведен анализ числа разбиений колонны, задающий размерность системы дискретных уравнений. Для верификации использовалось разработанное программного обеспечение на C#, позволяющее определить погрешность производимых разбиений колонны в сравнении тестовыми данными. Определено оптимальное число разбиений колонны с точки зрения «вычислительное время-погрешность расчета». Численная реализация модели осуществлена методом Рунге-Кутта 4-го порядка. В связи с увеличением времени реализации программного кода за счет роста размерности системы произведена оптимизация численного алгоритма с применением средств параллельного программирования. Проведен анализ целесообразности данной оптимизации.

Ключевые слова: буровая колонна, нелинейность, колебания, метод сосредоточенных параметров, параллельное программирование.

1 Introduction

In the complex process of drilling geotechnological wells in the mining industry, horizontal drilling has become widespread [2–4]. Research in the field of modelling the motion of horizontal drill strings from the point of view of the influence of stochastic processes on the dynamics of drilling equipment was carried out by Ritto T.G. with a group of scientists [1, 5, 6] and the authors of [7]. The authors of [8, 9] created an experimental setup based on the principle of mechanical similarity, and analyzed the accuracy of the theoretical models in accordance with the obtained experimental data. The authors of [10] studied the importance of drilling fluid formulations when drilling horizontal wells and proposed the use of biopolymer-based drilling fluids. In [11], the longitudinal vibrations of the column were modeled by the method of summation of modes, the analysis of the influence of the modes number on the dynamics of the system and their convergence was carried out. The authors of [12] developed a model that takes into account the geometric nonlinearity and the contact of the drill string with the well, based on the geometrically exact beam theory and the method of quadrature elements.

The search and application of alternative solutions in modelling are of scientific and practical interest, since they allow verifying the correctness of the already available results and expanding the class of problems under study. In particular, today, little-studied problems of modelling the dynamics of industrial equipment and machines in complicated conditions are relevant, namely due to the inhomogeneity of physical and mechanical properties, the variable structure of the research object, local and point loads.

Sadler J.P. in his work [13] considered the lumped-parameter method (LPM) for kinetic-elastodynamic analysis of mechanisms, later successfully used for the analysis of nonlinear vibrations of elastic multi-link mechanisms [14–16]. LPM is a special case of the finite element method, when the equation of a one-dimensional continuous medium is replaced by its discrete analogue. The essence of the method lies in the convertion from the model of a continuous medium to its discrete representation at the nodes by a system of ordinary differential equations. This method is widely known in structural mechanics, as well as in the study of the dynamics of flat beam structures [17–19]. Its application is most justified when modelling nonlinear systems with elements of heterogeneous material, variability of cross sections, loading, etc.

The purpose of this work is to identify the optimal number of drill string splits from the point of view of "implementation time-calculation error" by the LPM using parallel programming tools.

2 Mathematical model and its discretization

The horizontal motion of a drill string [1] under the action of a static compressive load at its left end, friction forces of the drill string against the rock, a variable harmonic force, gravitational forces, as well as an interaction force between the bit and the rock at the right end is considered

Figure 1. The lumped-parameter method (LPM) for solving the problem of the dynamics of drilling equipment was applied.

The equation of motion of the drill string with a length L is given in a general form [1]:

$$\begin{gathered}\rho A\frac{\partial^2 u(x,t)}{\partial t^2} - EA\frac{\partial^2 u(x,t)}{\partial x^2} = f_{sta}(x,t) + f_{har}(x,t) + \\ + f_{bit}(\dot{u}(x,t)) + f_{fric}(\dot{u}(x,t)) + f_{mass}(\ddot{u}(x,t))\end{gathered} \tag{1}$$

where $u(x,t)$ is the longitudinal displacement of the drill string, ρ is density of the column material, A is the cross-sectional area, E is Young's modulus. The right-hand side of Eq. (1) contains the forces acting on the drill string.

The constant force f_{sta} acts on the left end of the drill string ($x = 0$) and it is given by

$$f_{sta}(x,t) = F_{sta}\delta(x), \tag{2}$$

where F_{sta} is an amplitude, $\delta(x)$ is the Dirac delta function.

The harmonic force f_{har} is given as:

$$f_{har}(x,t) = F_0 sin(\omega_f t)\delta(x-L), \tag{3}$$

where F_0 is an amplitude, ω_f is the harmonic force frequency.

The bit inertia force and the drill string friction force on the rock are defined, respectively, as:

$$\begin{aligned} f_{mass}(\ddot{u}(x,t)) &= -m_{bit}\ddot{u}(x,t)\delta(x-L), \\ f_{fric}(\dot{u}(x,t)) &= -\mu(x)(\rho A)g\ sgn(\dot{u}(x,t)), \end{aligned} \tag{4}$$

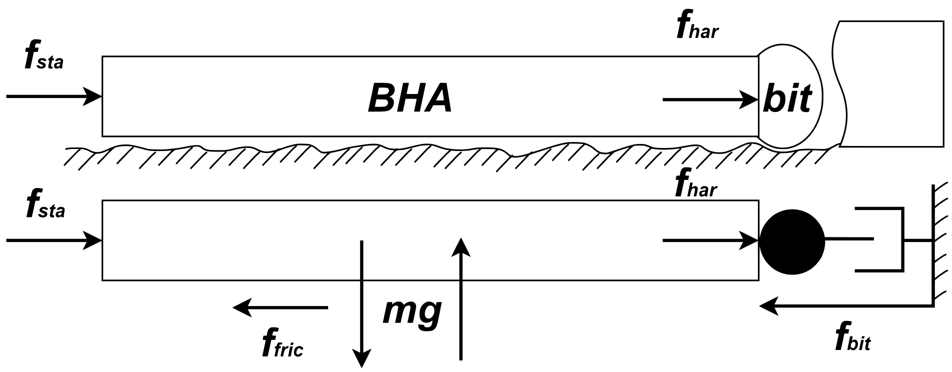


Figure 1: The sketch of forces acting on a drill string

where m_{bit} is the mass of the bit, concentrated at the point $x = L$, $\mu(x)$ is the coefficient of friction against the rock, g is gravitational acceleration.

The static compressive force at the right end of the column is determined in an exponential form as

$$f_{bit}(\dot{u}(x,t)) = \begin{cases} (c_1 exp(-c_2\dot{u}(x,t)) - c_1)\delta(x-L) & for\ \dot{u}(L,t) > 0, \\ 0 & for\ \dot{u}(L,t) \leq 0 \end{cases} \tag{5}$$

where c_1, c_2 are the coefficients of the bit-rock interaction.

The mathematical model Eq. (1-5) was solved by T.G. Ritto et al. in the work [1] by the finite element method. Here, the authors of the work, as in [20], use LPM, which is an effective method for the numerical analysis of such dynamical systems. Due to the inhomogeneity of the drill string loading, the mathematical model is written in accordance with the drilling equipment loading scheme (Figure 1) as follows:

$$\rho A \frac{\partial^2 u(x,t)}{\partial t^2} - EA \frac{\partial^2 u(x,t)}{\partial x^2} = f_{fric}(\dot{u}(x,t)) \tag{6}$$

with the boundary conditions

$$\begin{aligned} x = 0: &\quad EA\frac{\partial u}{\partial x} = -F_{sta}, \\ x = L: &\quad EA\frac{\partial u}{\partial x} = f_{har}(x,t)) + f_{mass}(\ddot{u}(x,t)) + f_{bit}(\dot{u}(x,t)). \end{aligned} \tag{7}$$

The metric is introduced in spatial and time coordinates:

$$u = L\text{U},\ x = L\text{X},\ t = \frac{\tau}{c},\ c = \sqrt{\frac{E}{\rho L^2}} \tag{8}$$

Approximate the derivatives according to the LPM used here:

$$\left(\frac{\partial^2 \text{U}}{\partial \text{X}^2}\right)_j = 2\frac{\Delta\text{X}_{j+1}\text{U}_{j-1} - (\Delta\text{X}_j + \Delta\text{X}_{j+1})\text{U}_j + \Delta\text{X}_j\text{U}_{j+1}}{\Delta\text{X}_{j+1}\Delta\text{X}_j(\Delta\text{X}_j + \Delta\text{X}_{j+1})} \tag{9}$$

$$\left(\frac{\partial \text{U}}{\partial \text{X}}\right)_j = \frac{\text{U}_j - \text{U}_{j-1}}{\Delta \text{X}_j}, \ \left(\frac{\partial \text{U}}{\partial \text{X}}\right)_j = \frac{\text{U}_{j+1} - \text{U}_j}{\Delta \text{X}_{j+1}} \tag{10}$$

$$\left(\frac{\partial \text{U}}{\partial \text{X}}\right)_j = \frac{\text{U}_{j+1} - \text{U}_{j-1}}{\Delta \text{X}_{j+1} + \Delta \text{X}_j} \tag{11}$$

where $\Delta \text{X}_j = \text{X}_j - \text{X}_{j-1}$, $\text{X}_j = \begin{cases} (2j-1)l & for\ 1 \le j \le N \\ 1 & for\ j = N+1 \end{cases}$, $2l = \dfrac{1}{N-1}$, N is the number of the drill string splits.

The model Eq. (6) and its boundary conditions Eq. (7) are represented in the discrete form:

$$\begin{aligned} \frac{\partial^2 \text{U}_1}{\partial \tau^2} - \frac{1}{3l^2}(2\text{U}_0 - 3\text{U}_1 + \text{U}_2) &= -\frac{\mu g}{Lc^2} sgn(\dot{\text{U}}_1)\ for\ j = 1, \\ \frac{\partial^2 \text{U}_j}{\partial \tau^2} - \frac{1}{4l^2}(\text{U}_{j-1} - 2\text{U}_j + \text{U}_{j+1}) &= -\frac{\mu g}{Lc^2} sgn(\dot{\text{U}}_j)\ for\ j = \overline{2, N-2} \\ \frac{\partial^2 \text{U}_{N-1}}{\partial \tau^2} - \frac{1}{3l^2}(\text{U}_{N-2} - 3\text{U}_{N-1} + 2\text{U}_N) &= -\frac{\mu g}{Lc^2} sgn(\dot{\text{U}}_N)\ for\ j = N-1 \end{aligned} \tag{12}$$

$$\begin{aligned} &\text{X} = 0 : \text{U}_1 - \text{U}_0 = -\frac{lF_{sta}}{EA} \\ &\text{X} = 1 : \frac{\partial^2 \text{U}_N}{\partial \tau^2} + \frac{(\rho A)L}{m_{bit}} \frac{(\text{U}_N - \text{U}_{N-1})}{l} = \frac{F_0}{m_{bit} Lc^2} sin\left(\frac{\omega_f}{c}\tau\right) + \frac{1}{m_{bit} Lc^2} f_{bit}(Lc\dot{\text{U}}_N) \end{aligned} \tag{13}$$

As a result, the system of N nonlinear second-order ordinary differential equations with respect to time with one algebraic expression is obtained.

3 Numerical analysis of the model

The numerical analysis of the model was carried out by the fourth-order Runge-Kutta method. The algorithm and the program code for numerical modelling have been developed in the C++ programming language.

The values of the physical and geometric parameters of the drill string, the indicators of the acting loads were taken in accordance with the author's values of [1]: $E = 2.1 \cdot 10^{11}$Pa, $\rho = 7850$kg$\cdot$m^{-3}, $g = 9.81$m$\cdot$s^{-2}, $D_i = 0.10$m (inner diameter), $D_o = 0.15$m (outer diameter), $\dfrac{L}{D_o} = 400$, $m_{bit} = 20$kg, $c_1 = 1.4 \cdot 10^3$N, $c_2 = 400$, $\mu = 0.1$, $\omega_f = 100 \cdot \dfrac{2\pi}{60}$rad $\cdot$ s^{-1}, $t \in [0, 10]$s, $\Delta t = 0.0001$s, $f_{sta} = 5500$N, $F_0 = 550$N.

To evaluate the efficiency of the drilling rig, the ratio of the input power of the drill to the power of the drill at the output was used:

$$\begin{aligned} p_{in}(t) &= f_{sta}\dot{u}(0,t) + f_{har}\dot{u}(L,t) \\ p_{out}(t) &= f_{bit}\dot{u}(L,t) \end{aligned} \tag{14}$$

where $p_{in}(t)$ is the input power, $p_{out}(t)$ is the output power.

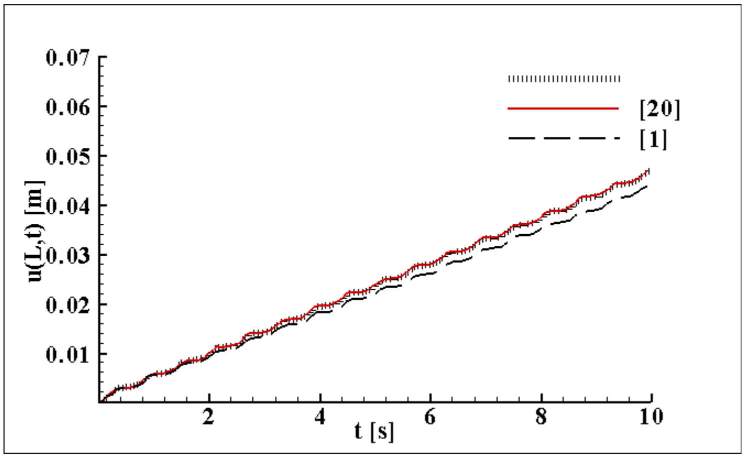


Figure 2: Verification of the obtained results of the longitudinal displacement.

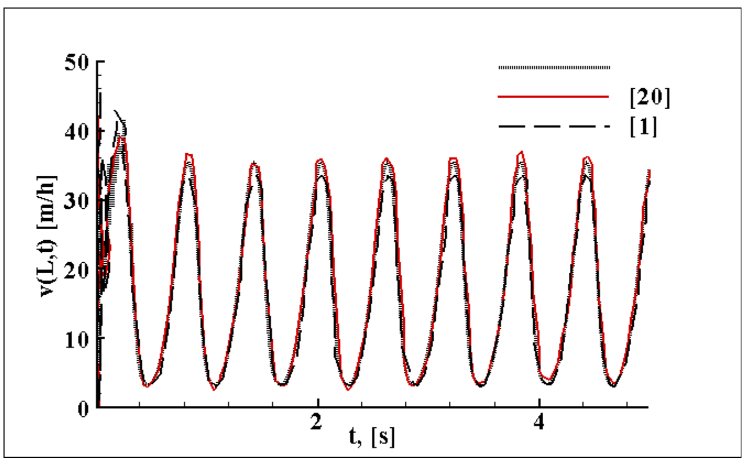


Figure 3: Verification of the obtained results of the bit speed.

The number of the drill string splitting nodes was taken, as in [20], equal to $N = 101$.

The research results, which are longitudinal displacement of the drill string at an interval of 10 s, are shown in Figure 2. The data of the bit speed are demonstrated in Figure 3.

The verification of the obtained results with the results of works [1] and [20] was carried out. In [1], the numerical modelling of the drill string motion was realized by the finite element method. The authors of [20] used LPM, and the numerical solution of the mathematical model was found in the symbolic mathematics package Wolfram Mathematica (WM). Here, numerical modelling was conducted in C++.

It was found that the longitudinal displacement of the drill string at the point $x = L$ increases with time, and the speed of the drill string at the right end is oscillating. It is caused by the presence of loads on the drilling equipment in the model.

The dashed black line shows the results of T.G. Ritto [1], solid red line is the results of L. Khajiyeva, A. Sergaliyev [20], dotted black one is the results of this work.

It is visually clear that the graphs in both figures are qualitatively convergent.

Longitudinal displacements grown by red and dotted black lines coincide completely, while in comparison with dashed black one, the error increases with time. The amplitudes of the speed of motion depicted by red and dotted black lines are slightly higher than the amplitude of dashed black line, which may be caused by the digitization error, the use of various numerical methods, or an insufficient number of nodes in the discrete model.

In Figure 4 the change in the ratio between the output and input power is shown. The higher this ratio, the more efficient the drilling rig is. It can be seen from the graph that this indicator of the drill string does not exceed 25%, which is explained by the fact that the model takes into account the loads affecting the equipment, which are friction forces, the reaction force of the rock on the drill, static compressive force, gravitational forces, etc. The dashed black line shows the results of T.G. Ritto, solid red line is the results of L. Khajiyeva, A. Sergaliyev, dotted black one is the results of this work. Good consistency of the results is observed.

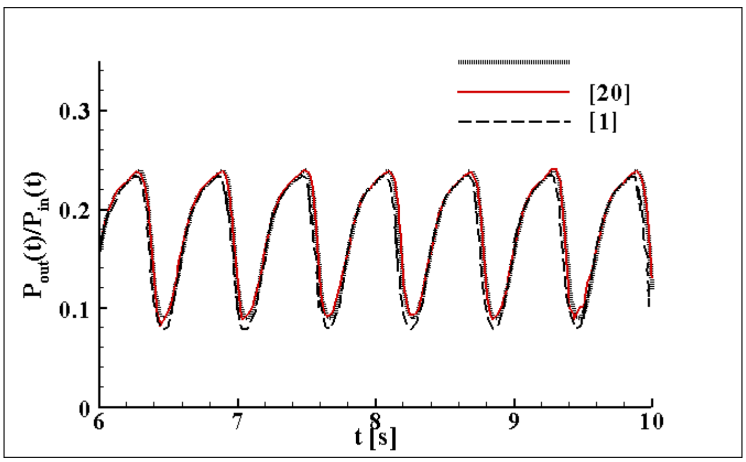


Figure 4: Verification of the obtained results of the ratio between the input and output power

4 Dimension analysis of the discrete ODE system

Obviously, the calculation accuracy depends on the choice of the number of points for dividing the drill string along the length: the spatial steps l decrease with an increase in the nodes in space, the discrete system tends to the continuity equation. However, with an increase in the number of partitions, the program implementation time also increases. This requires additional analysis of the dependence of the computational accuracy on the number of nodes N and the time spent on executing the program code.

The results of T.G. Ritto, who first considered this problem, were taken as a sample to estimate the calculation error. For algebraic verification, a WPF Application was written in the C# language. It compares the digitized data of the work [1] with the results of this work and finds the difference in the data of the loaded files at the closest possible time points.

The results of the longitudinal displacement of the drill string were taken as comparative data. Tables 1 presents the results showing the effect of the number of split points on the calculation error. It is relevant to notice, the accuracy of the results is influenced by the quality of the digitized data from the test graph; the error of the time points at which the difference in the results is located (this indicator does not exceed the time step $dt = 1e - 5s$); the error of the used numerical methods.

Tables 1 shows that the best convergence values were obtained by splitting the column into 1000 segments: the maximum error does not exceed 0.39 mm, while the computation time is no more than 8.5 minutes.

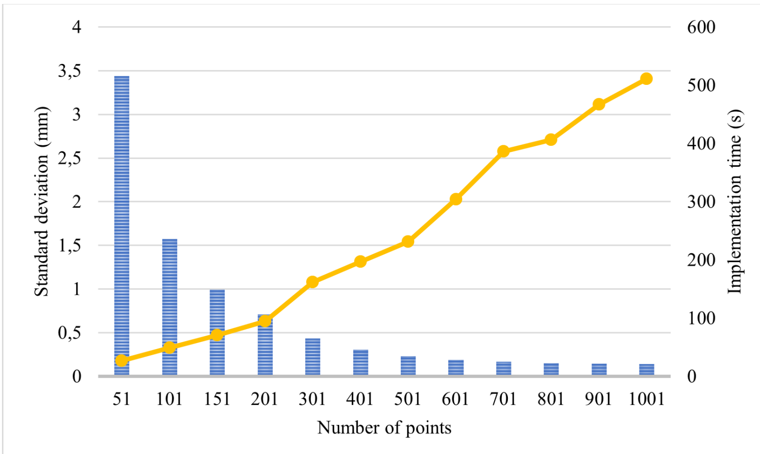


Figure 5: Influence of the number of column splits on the error and implementation time

Table 1: Analysis of the influence of the number of partitioning nodes on time and computation error.

The number of nodes	11	101	201	401	601	801	1001
Maximum error (mm)	47.3475	2.837	1.3428	0.68111	0.46391	0.3847	0.3922
Standard deviation (mm)	26.5089	1.573	0.7076	0.304	0.19071	0.1513	0.1392
Time implementation (s)	6.667	48.9	94.837	196.707	304.165	405.95	510.85

Figure 5 clearly demonstrates the need to use more points, where the bar graph corresponds to the standard deviation for a particular number of splits, and the graph depicts the implementation time. Note that the error for 101 points is more than 1.5 mm, therefore, more than 300 nodes are required to obtain quantitatively accurate values.

If the priority of the research is the accuracy of the calculation with a sufficient amount of time resources, splitting into 1000 or more parts is the most appropriate.

5 Optimization of the numerical algorithm using parallel programming tools

A small time step, the need to use a large number of partitions and, as a consequence, a large number of iterations served as factors for the next stage of the study which is optimization of the program code using parallel programming tools. Parallelization of the C++ code was implemented using the Open Multi-Processing (OpenMP) API. The OpenMP technology, designed for shared memory systems, implements parallelism of calculations due

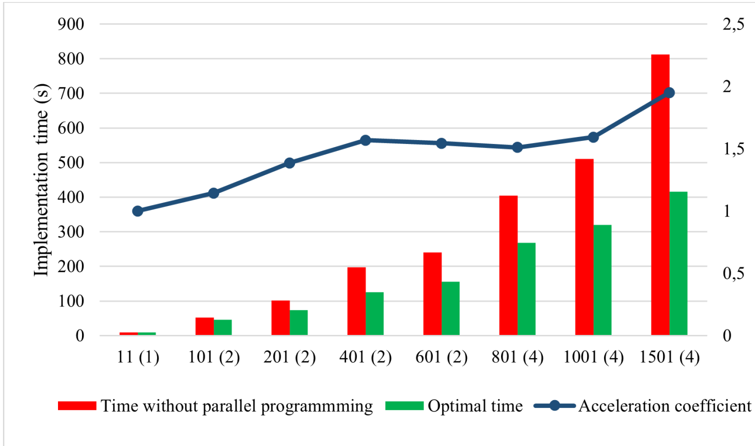


Figure 6: Acceleration coefficient for a different number of points

to multithreading. The master thread creates a number of threads, the task is distributed among them. Due to this technology, the logic of the code does not change compared with MPI, oriented to distributed memory systems, where it is necessary to determine connections between processes. OpenMP allows to find "vulnerable"places in the program and significantly speed up the execution of these blocks, alternating them with a sequential part. In particular, this approach is applied to linear algorithms, which include the fourth-order Runge-Kutta method.

To analyze the advantages of using parallel programming, spent time resources, and the optimal number of threads, the program code was tested for various values of the parameters of the partition nodes and number of threads on the interval of $t = 10$s.

The test results are clearly shown in Figure 6, where the values of the acceleration factor of the program using the OpenMP library are presented for a different number of points. The bar chart shows the implementation time of the code, where the red columns correspond to the execution time of the code by one thread, that is, without using parallel computations, the green columns correspond to the time of the optimal number of threads (in parentheses next to the number of nodes). The line graph shows the acceleration factor as the ratio of the time taken by one thread at the optimal time. Thus, for a smaller number of points, one stream is optimal, but with an increase in the number of points, the use of parallel computation is justified.

It is worth noting that in the further, considering a more complex model and complicating the computational algorithm, using a larger number of nodes, the efficiency indicators will increase accordingly.

6 Conclusion

During the research of the dynamics of longitudinal vibrations of a horizontal drill string the optimal number of the drill string splits by LPM using the developed software in the C# language and parallel programming tools was found. The optimal number of the drill string splits in terms of “implementation time-calculation error” varies within the range of 400-600 nodes. It improves the accuracy of the solution in comparison with the case of splitting the string into 100 elements [20]. A good agreement between the obtained results and the results of T.G. Ritto's work [1] based on the FEM has been established.

In addition, the numerical algorithm implemented in the C++ language allows further refinement of solutions by increasing the number of the drill string splits. At the same time, the increase in the dimension of the discrete lumped model is successfully implemented through the use of parallel programming. Comparative analysis showed the justification of its application for optimization of the numerical algorithm.

In the future, this work of the authors is seen in the use of LPM in modelling nonlinear vibrations of vertical drill strings with spatial type of deformation, inhomogeneous structure, inhomogeneity of loading due to local and point loads, etc.

7 Acknowledgement

This research is funded by the Science Committee of the Ministry of Education and Science of the Republic of Kazakhstan (Grant No. AP08857255).

References

[1] Ritto T.G., "Drill-string horizontal dynamics with uncertainty on the frictional force", *J. Sound. Vib.* 332(1) (2013): 145-153.

[2] , International Association of Drilling Contractors *IADC Drilling Manual*, (USA: Technical Toolboxes Inc, 2000).

[3] Basarygin Yu.M., Bulatov A.I., Proselkov Yu.M. (2002) *Burenie neftianykh i gazovykh skvazhin* [Drilling of oil and gas wells] (Рњ.:Nedra-Biznescentr)

[4] Willoughby D.A., *Horisontal directional drilling*, (USA: The McGraw-Hill Companies Inc, 2005).

[5] Ritto T.G., Sampaio R., "Stochastic drill-string dynamics with uncertainty on the imposed speed and on the bit-rock parameters", *Int. J. Uncertain. Quantification.* 2(2) (2012): 111-124.

[6] Lobo D.M., Ritto T.G., Castello D.A., "Stochastic analysis of torsional drill-string vibrations considering the passage from a soft to a harder rock layer", *J. Braz. Soc. Mech. Sci. Eng.* 39(6) (2017): 2341-2349.

[7] Jialin T., Yinglin Ya., Lin Ya., "Vibration characteristics analysis and experimental study of horizontal drill string with wellbore random friction force", *Arch. Appl. Mech.* 87 (2017): 1439-1451.

[8] Zhanghua L., Qiang Zh., Tiejun L., Fuhui W., "Experimental and numerical study of drill string dynamics in gas drilling of horizontal wells", *J. Nat. Gas Sci. Eng.* 27 (2015): 1412-1420.

[9] Tiejun L., Qiang Zh., Zhanghua L., Zhou X., Tao W. , Gao Li., Jiandong D., "Experimental study on vibrational behaviors of horizontal drillstring", *J. Pet. Sci. Eng.* 164 (2018): 311-319.

[10] Kh.A. Sarsenbayev, B.S. Khamzina, G.A. Koldassova, G.B. Issayeva (2017) *Primenenie biopolimernykh rastvorov dlia effektivnogo vskrytiia produktivnykh gorizontov gorizontalnykh skvazhin* [Application of biopolymer drilling fluid for effective opening productive horizons horisontal wells]. News of the National Academy of Sciences of the Republic of Kazakhstan, Series Physico-mathematical, vol. 312, no. 2, pp. 161-165.

[11] Salehi M.M., Moradi H., Nonlinear multivariable modeling and stability analysis of vibrations in horizontal drill string", *Appl. Math. Model.* 71 (2019): 525-542.

[12] Yiwei H., Zhijiu A., YuchunK., Jianxun L., Wenwu Ya., "Nonlinear dynamic modeling of drill string in horizontal well - A geometrically exact approach", *J. Pet. Sci. Eng.* 172 (2019): 1133-1152.

[13] Sadler J.P., *A lumped parameter approach to the kineto-elastodynamic analysis of mechanisms* (USA: Rensselaer Polytechn. Institute, 1971)

[14] Sadler J.P., Sandor G.N., "A Lumped parameter approach to vibration and stress analysis of elastic linkages" , *J. Eng. Ind.* 95(2) (1973): 549-557.

[15] Sadler J.P., Sandor G.N., "Nonlinear vibration analysis of elastic four-bar linkages", *J. Eng. Ind.* 96(2) (1974): 411-419.

[16] Sadler J.P., "On the Analytical Lumped-Mass Model of an Elastic Four-Bar Mechanism", *J. Eng. Ind.* 97(2) (1975): 561-565.

[17] Wang Y., Huston R.L., "A lumped parameter method in the nonlinear analysis of flexible multibody systems", *Comput. Struct.* 20(3) (1994): 421-432.

[18] Salinic S., Nikolic A., "On the free vibration of a multiplestepped cantilever beam", *Fourth Serbian (29th Yu) Congress on Theoretical and Applied Mechanics, Serbia, 4-7 June 2013.*

[19] Giorgio I., Dionisio D.V., "Non-linear lumped-parameter modeling of planar multi-link manipulators with highly flexible arms", *Robotics* 7(4) (2018): 60.

[20] Khajiyeva L., Sergaliyev A., "About discretization of the horizontal drill-string model with uncertainty on the frictional force by the lumped parameters method", *Appl. Mech. Mater.* 665 (2014): 589-592.

МАЗМҰНЫ – СОДЕРЖАНИЕ – CONTENTS

4-бөлім **Раздел 4** **Section 4**

Қолданбалы математика **Прикладная математика** **Applied Mathematics**