

**КИЇВСЬКИЙ НАЦІОНАЛЬНИЙ УНІВЕРСИТЕТ
ІМЕНІ ТАРАСА ШЕВЧЕНКА**

ВІСНИК

**КИЇВСЬКОГО НАЦІОНАЛЬНОГО
УНІВЕРСИТЕТУ
ІМЕНІ ТАРАСА ШЕВЧЕНКА**

СЕРІЯ ФІЗИКО-МАТЕМАТИЧНІ НАУКИ

ВИПУСК №2 2015

**Вісник Київського національного університету імені Тараса Шевченка,
выпуск №2, 2015,
Серія фізико-математичні науки**

З 1991 року серії вісників Київського університету “Математика і механіка”, “Фізика”, “Моделирование и оптимизация сложных систем” реорганізовано у “Вісник Київського університету. Серія: фізико-математичні науки”. У віснику містяться результати нових досліджень у різних галузях математики, інформатики, механіки, фізики та радіофізики для наукових працівників, викладачів, аспірантів, інженерів і студентів. Друкується за рекомендаціями Вчених Рад фізичного, радіофізичного, механіко-математичного факультетів та факультету кібернетики.

Журнал “Вісник Київського національного університету імені Тараса Шевченка. Серія фізико-математичні науки” включено до переліку фахових видань ВАК України та реферується в Реферативному журналі та базах даних ВИНІТИ, Росія, Москва.

Редакційна колегія:

Анісімов Анатолій Васильович, чл.-кор. НАН України, д.ф.-м.н., проф., **головний редактор**;
Хусаїнов Денис Яхєвич, д.ф.-м.н., проф., **заступник головного редактора, відповідальний за видання**;
Arturs Medvids, Dr. habil. Phys., prof., Riga Technical University, Riga, Latvia;
Miklós Rontó, prof., Doctor of Science, Full Professor at the University of Miskolc, Faculty of Mechanical Engineering and Information Science, Department of Analysis;
Milada Bartlova, Ph.D., Brno University of Technology, Department of Physics, Brno, Czech Republic;
Nickolai Kukhtarev, Research professor, Department of Physics, Alabama A&M University;
Sergei Gorlatch, Dr. habil., prof., professor of Computing Science, University of Muenster, Germany;
Sergey Trofimchuk, prof., Universidad de Talca, Instituto de Matematica y Fisica, Talca, Chile;
Stefan Hudak, DSc, prof., professor of Technical University of Kosice, Slovak Republic;
Toru Aoki, Ph.D., prof., Research Institute of Electronics, Shizuoka University, Japan;
Акіменко Віталій Володимирович, д.т.н., проф.;
Анісімов Ігор Олексійович, д.ф.-м.н., проф.;
Буй Дмитро Борисович, д.ф.-м.н., проф.;
Булавін Леонід Анатолійович, акад. НАН України, д.ф.-м.н., проф.;
Волошин Олексій Федорович, д.т.н., проф.;
Гарашенко Федір Георгійович, д.т.н., проф.;
Єжов Станіслав Миколайович, д.ф.-м.н., проф.;
Жук Ярослав Олександрович, д.ф.-м.н., проф.;
Заславський Володимир Анатолійович, д.т.н., доц.;
Кириченко Володимир Васильович, д.ф.-м.н., проф.;
Козаченко Юрій Васильович, д.ф.-м.н., проф.;
Кудін Володимир Іванович, д.т.н., с.н.с.;
Львов Віктор Анатолійович, д.ф.-м.н., проф.;
Макара Володимир Арсенійович, чл.-кор. НАН України, д.ф.-м.н., проф.;
Макарець Микола Володимирович, д.ф.-м.н., проф.;
Перестюк Микола Олексійович, акад. НАН України, д.ф.-м.н., проф.;
Погорілий Сергій Дем'янович, д.т.н., проф.;
Савенков Сергій Миколайович, д.ф.-м.н., доц.;
Скришевський Валерій Антонович, д.ф.-м.н., проф.

Редакційний відділ:

Анісімова Тетяна Харитонівна, **відповідальний секретар**;
Безущак Оксана Омелянівна, bezusch@univ.kiev.ua;
Стукаленко Вікторія Віталіївна, stu@univ.kiev.ua;
Родіонова Тетяна Василівна, rodtv@univ.kiev.ua;
Хмелюк Надія Кузьмівна, khmeluk@univ.kiev.ua;
Сільвейструк Людмила Миколаївна, **технічний редактор**, slm-klm@ukr.net.

Адреса редакційної колегії:

Факультет кібернетики, Київський національний університет імені Тараса Шевченка,
пр. Глушкова, 4 д, 03680 Тел. (044) 259-01-49

ISBN 978-966-2142

ISSN 1812-5409

ЗМІСТ

АЛГЕБРА, ГЕОМЕТРІЯ ТА ТЕОРІЯ ІМОВІРНОСТЕЙ

Алексеева К.С. Напівгрупа перетворень IO_z	9
Власенко І.Ю. Критерій топологічної спряженості однорідних відображень R^n	14
Іванюк О.М., Білун С.В. Деформації атомів степені 2 на поверхнях з краєм	18
Кочубінська Є.А, Семенюк І.В. Про узагальнений частковий вінцевий добуток напівгруп	24
Петравчук А.П., Степух В.В. Базиси в алгебрах Лі диференціювань	28
Плахотник М.В. Типи траєкторій періодичних точок трикутних відображень, напівгрупа ітерацій яких скінченна	33
Турка Т.В. Центр напівгрупи відповідностей	41
Яневич Т.О. Моделювання даних з великою кількістю нулів при обстеженні капітальних інвестицій України	45

ДИФЕРЕНЦІАЛЬНІ РІВНЯННЯ, МАТЕМАТИЧНА ФІЗИКА ТА МЕХАНІКА

Асроров Ф.А. Інтегральні множини розривних динамічних систем	51
Голиченко О.Л., Малюга В.С. Особливості перемішування рідини в скінченному циліндрі	55
Горошко О.О., Кикоть С.В. Коливання гнучкого трубопроводу на пружній основі при закритичних швидкостях транспортування рідини	61
Денисенко Н.Л. Про періодичні розв'язки систем нелінійних диференціально-функціональних рівнянь з малим параметром	65
Єрьоміна Т.О. Про побудову неперервних розв'язків систем нелінійних різницево-функціональних рівнянь	71
Клімчук Т.В., Острик В.І. Ковзний контакт напівнескінченного штампа з пружною смугою	75
Лимарченко О.С., Маджид М. Динаміка трубопроводу, що обертається, з рідиною	79
Остос О.Х., Острик В.І. Згин пружної смуги зосередженими силами	83

КОМП'ЮТЕРНІ НАУКИ ТА ІНФОРМАТИКА

Анікушин А.В., Номіровський Д.А. Якісний аналіз лінійних диференціальних моделей псевдогіперболічного типу в деяких класах узагальнених функцій	89
Буй Д.Б., Канарська І.С., Сенченко О.С. Критерії збереження активного домену сигнатурними операціями табличних алгебр	94
Буй Д.Б., Компан С.В., Поляков С.А., Карам Д. Алгоритми CLOS та LOOPS лінеаризації класів в мовах програмування: формальна побудова	99
Буй Д.Б., Пузікова А. В. Математична теорія нормалізації: нормальні форми 2-4 порядків	103
Вознюк Т.Г. Побудова класифікатора для вирішення займенникової анафори на основі тензорної моделі	113
Галкін О.А. Дослідження процедур розпізнавання на основі розподільного розташування багатовимірних даних та функцій глибини	117
Гаркуша Н.І. Отримання оцінки області стійкості положення рівноваги моделі популяції з післядією	123
Гордійчук О.В. Однорангові мережі з урахуванням місцезнаходження	128
Донченко В.С., Тарасова О.В. Матрична множинна регресія	133
Завадський І.О. Префіксні (Δ, k) -коди	139
Завадський І.О., Петльована М.В. Квантовий клапан Тоффолі, його реалізація та застосування	143
Катеринич Л.О., Калашнікова А.А., Марголін О.Г. Система прийняття рішень для ідентифікації користувачів	147
Коваль Ю.В. Протокол багатоточкового потоку: означення та застосування в створенні механізмів розпаралелювача та аналізатора подій в системі жестового інтерфейсу	152
Ковальов Д.І. Сервіс дистанційного навчання у форматі соціальної мережі	156
Марченко О. О. Алгоритм виділення явних семантичних зв'язків з векторів семантико-синтаксичної валентності концептів онтології	160

Мащенко С.О., МоренецьВ.І. С-ядро кооперативної гри з нечіткою множиною дозволених коаліцій	165
Насіров Е.М. Блочно-діагональний підхід до невід'ємної факторизації розріджених матриць і тензорів надвеликої розмірності	170
Панченко Т. В. Метод доведення коректності паралельних програм з використанням спрощеного стану та його застосування	174
Скотаренко Ф.М. Алгоритм лінійної дискримінації для матричних просторів та його основи	178
Терещенко В.М., Осадчий Б.О., Петльована М.В., Селезньов А.І. Тара А.М. Архітектура моделі єдиного алгоритмічного середовища (МЄАС) для розв'язання задач обчислювальної геометрії	183
Хомицький В.В., Терещенко Л.М., Фомін В.В., Абрамова Л.П. Застосування чисельної спектральної моделі SWAN для розрахунків параметрів вітрового хвилювання в районі Кривої коси	189
Царук В.І. Про оптимізацію транспортних потоків в моделі Бекмана	193
Шарапов М.М. Статистичне корегування результатів тестів	199
Шликов М.П. Про алгоритм побудови довірчого еліпсоїда Петуніна	203
Яковенко О.А. Про прикладне програмне забезпечення для задач нелінійного програмування	209

РАДІОФІЗИКА

Веклич А.М., Клешич М.М., Борецький В.Ф., Фесенко С.О., Телега В.М. Спектроскопія плазми електричної дуги між Cu-Ni електродами в потоці аргону	215
Козинець О.В. Розрахунок напруги холостого ходу кремнієвого сонячного елемента з модифікованим шаром в області просторового заряду	221
Лясковський В.В., Баужа О.С. Розробка макету мобільного уніфікованого комплексу реєстрації радіаційного випромінювання	225
Манілов А.І. Вдосконалення математичної обробки фотоелектричного відгуку на модульоване освітлення	231
Первак Ю.О. Особливості структури дисперсійних дзеркал з негативною дисперсією групової затримки	235
Слободянюк Д.В., Попов М.О.Малишев В. Ю. Мойсєсенко В. А. Сптворення профілю вихідного сигналу параметричним підсилювачем в багатомодовому режимі	239

СУЧАСНА ФІЗИКА

Гаврюшенко Д. А., Дворецька В. І. Осмотичний тиск бінарного розчину поблизу критичної точки розшарування	247
Горкавенко Т.В., Плющай І.В., Макара В.А. Першопринципний розрахунок рівноважного положення домішки кисню в кремнії	251
Гоцульський В.Я., Свечнікова О.С., Стиранець М.В. Тонка структура низькоконцентраційного піку розсіяння світла водними розчинами етанолу	255
Лисов В.І., Цареградська Т.Л., Саєнко Г.В., Турков О.В. Підвищення термічної стабільності аморфних сплавів за рахунок процесу ініційованої висхідної дифузії	259
Ніколаєнко Т. Ю., Пилипчук Є. В. Визначення фізико-хімічних властивостей рідинних наносистем на основі вимірювання магнітної проникності	263

CONTENTS

ALGEBRA, GEOMETRY AND PROBABILITY THEORY

Alekseeva K.S. Semigroup of transformations IO_z	9
Vlasenko I.Yu. Criterion of topological conjugacy for homogeneous mappings of R^n	14
Ivanyuk O.M., Bilun S.V. Deformations of atoms of complexity 2 on surfaces with boundary	18
Kochubinska E.A., Semeniuk I.V. On a generalized partial wreath product of semigroups	24
Petravchuk A.P., Stepukh V.V. On bases of Lie algebras of derivations	28
Plakhotnyk M.V. Types of periodical points trajectories of triangular maps, whose semigroup of iterations is finite	33
Turka T.V. The Center of Semigroup of Correspondences	41
Ianevych T.O. Modeling data with excess of zeros for capital expenditure surveys in Ukraine	45

DIFFERENTIAL EQUATIONS, MATHEMATICAL PHYSICS AND MECHANICS

Asrorov F.A. Integral sets of discontinuous dynamical systems	51
Golichenko O.L., Malyuga V.S. Mixing characteristics in a finite cylindrical container	55
Goroshko O.A., Kykot' S.V. Oscillations of a flexible conduit on an elastic foundation at overcritical values of transported fluid	61
Denysenko N.L. About periodic solutions of systems of nonlinear differential-functional equations with a small parameter	65
Yeromina T.O. About construction of continuous solutions of nonlinear difference-functional equations	71
Klimchuk T.V., Ostryk V.I. Sliding contact of the semi-infinite stamp with an elastic strip	75
Limarchenko O.S., Majid M. Dynamics of rotating pipeline with liquid	79
Ostos O.H., Ostryk V.I. Bending of elastic strip by concentrated forces	83

COMPUTER SCIENCES AND INFORMATICS

Anikushyn A.V., Nomirovskii D. A. Qualitative analysis of linear differential pseudohyperbolic equations in some classes of distributions	89
Buy D.B., Kanarska I.S., Senchenko A.S. The criteria for preservation of active domain by signature operations in the table algebra	94
Buy D.B., Kompan S.V., Polyakov S.A., Karam J. Linearization algorithms CLOS and LOO of the classes in programming languages the formal definitions	99
Buy D.B., Puzikova A.V. Mathematical theory of normalization: 2-4 normal forms	103
Vozniuk T.G. Building a classifier to solve pronominal anaphora based on the Tensor model	113
Galkin O.A. The research of recognition procedures based on the distributive location of multidimensional data and depth functions	117
Garkusha N.I. Obtaining the estimation of the stability area of equilibrium in the model of population with aftereffect	123
Hordiichuk O.V. Location oriented peer-to-peer networks	128
Donchenko V.S., Tarasova O. V Matrix multiple regression	133
Zavadskyi I.O. Prefix (Δ, k) -codes	139
Zavadskyi I.O., Petlova M.V. The quantum Toffoli gate, its implementation and application	143
Katerinich L.O., Kalashnikova A.A., Margolin A.G. Information analysis system for user authentication	147
Koval Iu. V. Multipoint stream protocol: definition and application for parallelizer and events analyzer mechanisms in gestures interface system creation	152
Kovaliov D.I. E-learning service in social network format	156
Marchenko O.O. An algorithm for extraction of explicit semantic relations from semantic-syntactic valence vectors of ontology concepts	160

Mashchenko S.O., Morenets V.I C- core of cooperative game with the fuzzy set of admissible coalitions	165
Nasirov E.M. Blocks approach of non-negative huge sparse matrix and tensors factorization	170
Panchenko T.V. Parallel Programs Correctness Proof Method Using Simplified State with Applications	174
Skotarenko F.M. Algorithm of linear discrimination problem for matrix spaces: base principals	178
Tereshchenko V.M., Osadchyi B.O., Petlovana M.V., Seleznov A.I., Tara A.M. Architecture of model of the united algorithmic environment (MUAE) for solving computational geometry problems	183
Khomicky V.V., Tereshchenko L.N., Fomin V.V., Abramova L. P. The use of numerical models SWAN spectral parameters for the calculation of wind waves in the area Curve spit	189
Tsaruk V. I. On transport flow optimization in Backman model	193
Sharapov M.M. Statistical correction of test results	199
Shlykov M. P. On the algorithm of Petunin confidence ellipsoid construction	203
Iakovenko O. Nonlinear programming software	209

RADIOPHYSICS

Veklich A.N., Kleshich M.M., Boretskij V.F., Fesenko S.O., Telega V.N. Plasma spectroscopy of electric arc between Cu-Ni electrodes in argon flow	215
Kozinetz A.V. Calculation of open circuit voltage for silicon solar cell with modified layer in space charge region	221
Liaskovskyi V. V., Bauzha O.S. Development of model unified mobile complex for registration of radioactive emissions	225
Manilov A.I. Improved mathematical processing of photoelectric response on modulated illumination	231
Pervak Yu.O. Features of the structure of dispersive mirrors with negative group delay	235
Slobodianiuk D. V., Popov M. O., Malyshev V. Yu., Moiseienko V. A. Signal profile distortion in multimode parametric amplifier	239

MODERN PHYSICS

Gavryushenko D.A., Dvoretzka V .I. The osmotic pressure of binary solution in the vicinity of the critical point of stratification	247
Gorkavenko T.V., Plyushchay I.V., Makara V.A. Ab initio calculation of equilibrium position of oxygen impurity in silicon	251
Gotsulskiy V.Ya., Svechnikova O.S., Stiranets M.V. The fine structure of low concentration peak light scattering aqueous solutions of ethanol	255
Lysov V.I., Tsaregradskaya T.L., Saenko G.V., Turkov O.V. The analysis possibilities practical using of conclusions high temperature stability amorphous alloys theory	259
Nikolaenko T. Yu., Pylypchuk Ie. V. Determination of physico-chemical properties of liquid nanosystems on the basis of magnetic permeability measurements	263

АЛГЕБРА, ГЕОМЕТРИЯ ТА ТЕОРИЯ ІМОВІРНОСТЕЙ

УДК 512.53+512.64

К.С. Алексеева, аспірант

Напівгрупа перетворень $\mathcal{IO}_{\mathbb{Z}}$

Київський національний університет імені
Тараса Шевченка, 01033, Київ, вул. Володи-
мирська, 64
e-mail: kseniia.tretiak@gmail.com

K.S. Alekseeva, Postgraduate Student

Semigroup of transformations $\mathcal{IO}_{\mathbb{Z}}$

Taras Shevchenko National University of
Kyiv, 01033, Kyiv, 64 Volodymyrska Str.
e-mail: kseniia.tretiak@gmail.com

У даній роботі вивчається інверсна напівгрупа монотонних перетворень $\mathcal{IO}_{\mathbb{Z}}$. Описані структура напівгрупи, деякі найпростіші властивості її елементів, ідеали та відношення Гріна напівгрупи. Показано, що існує така система твірних, що довільний елемент напівгрупи $\mathcal{IO}_{\mathbb{Z}}$ можна подати у вигляді добутку не більш ніж чотирьох елементів з цієї системи твірних. А також доведено, що група автоморфізмів напівгрупи $\mathcal{IO}_{\mathbb{Z}}$ ізоморфна нескінченній дієдральній групі $\mathbb{D}_{\infty}$.

Ключові слова: інверсна напівгрупа, напівгрупа перетворень, ідеал, відношення Гріна, автоморфізм.

The segroup theory is a relatively young but important direction in mathematics. It has been studied by many mathematicians during the last decades. Naturally there are issues related to the semigroups of transformations of some ordered set. We consider the so-called order-preserving transformations (a transformation $a : M \rightarrow M$ of poset M is called order-preserving, if for all $x < y$ from the domain of a we have $a(x) < a(y)$).

The present paper is dedicated to the inverse semigroup of order-preserving transformations $\mathcal{IO}_{\mathbb{Z}}$. In the first part of this paper we describe a structure of the semigroup $\mathcal{IO}_{\mathbb{Z}}$, properties of multiplication and we show, that there is such generating system, that any element of semigroup $\mathcal{IO}_{\mathbb{Z}}$ can be written as a product of at most four elements from that system. In the second part of this paper we describe ideals and Green's relations of the semigroup $\mathcal{IO}_{\mathbb{Z}}$. And in the end we prove that the group of automorphisms of $\mathcal{IO}_{\mathbb{Z}}$ is isomorphic to the infinite dihedral group $\mathbb{D}_{\infty}$.

Key Words: inverse semigroup, semigroup of transformations, ideal, Green's relation, automorphism.

Статтю представив доктор фізико-математичних наук професор Кириченко В.В.

1 Introduction

The important direction in the theory of semigroups of transformations of some set M is the study of semigroups of transformations that somehow consistent with some additional structure on the set M . Usually, this structure is a certain partial order on M . Consistency with the order can be understood in different ways. We consider the so-called order-preserving transformations. A transformation $a : M \rightarrow M$ of poset M is called *order-preserving*, if for all $x < y$ from the domain of a we have $a(x) < a(y)$. Many papers are dedicated to the study of semigroups of order-preserving transformations of the linearly ordered sets ([1], [2], [3], [4]). The semigroups of order-preserving transformations of the linearly ordered infinite sets have been studied in

papers ([5], [6]). However, at the moment there are many open questions for the case of infinite sets.

In this paper we consider the inverse semigroup $\mathcal{IO}_{\mathbb{Z}}$ of order-preserving transformations. In particular, we describe the ideals of this semigroup, its automorphisms, and show that there exists a system of generators such that every element of the semigroup $\mathcal{IO}_{\mathbb{Z}}$ can be written as a product of at most four elements from this system.

We keep the notations from [3].

2 Main definitions and properties

We denote by $\mathbb{Z}$ the set of all integers with the natural linear order.

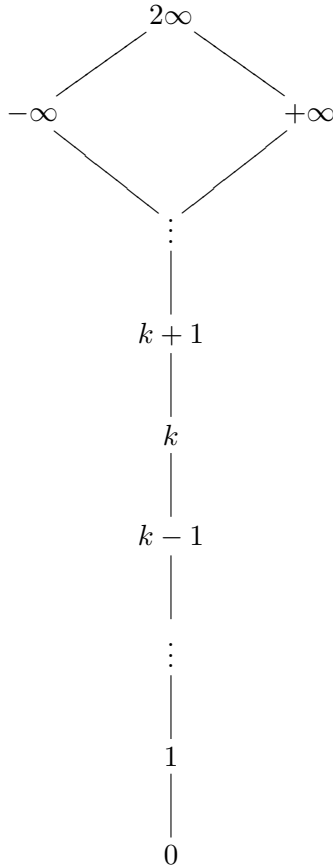
Definition 2.1. The semigroup of all partial injections $a : \mathbb{Z} \rightarrow \mathbb{Z}$, which preserve the natural order on $\mathbb{Z}$ (that is for every $x < y$ from the

domain of a we have $a(x) < a(y)$) is called an inverse semigroup, which preserves an order on the set $\mathbb{Z}$. We denote this semigroup by $\mathcal{IO}_{\mathbb{Z}}$.

We define the norm of a subset A on this set by the following rule:

$$\|A\| = \begin{cases} k, & \text{if } A \text{ is finite and contains } k \text{ elements,} \\ +\infty, & \text{if } A \text{ contains an infinite number of} \\ & \text{elements and is bounded from below,} \\ -\infty, & \text{if } A \text{ contains an infinite number of} \\ & \text{elements and is bounded from above,} \\ 2\infty, & \text{if } A \text{ contains an infinite number of} \\ & \text{elements and is not bounded.} \end{cases}$$

The norms take values in the set $\{0, 1, 2, \dots\} \cup \{+\infty, -\infty, 2\infty\}$, which is a poset with order determined by following Hasse diagram:



For a partial injection $a : \mathbb{Z} \rightarrow \mathbb{Z}$ we denote by $\text{dom}(a)$ and $\text{im}(a)$ the domain and the image of a respectively. If a is monotone, then $\|\text{dom}(a)\| = \|\text{im}(a)\|$. In this case the number $\text{nrnk}(a) = \|\text{dom}(a)\| = \|\text{im}(a)\|$ is called the normalized rank of a transformation a .

Proposition 1. For any $a, b \in \mathcal{IO}_{\mathbb{Z}}$

$$\text{nrnk}(ab) \leq \min(\text{nrnk}(a), \text{nrnk}(b)).$$

Proof. For every element $x \in \mathbb{Z}$ we have $(ab)(x) = b(a(x))$. This implies $\text{im}(ab) \subset \text{im}(b)$, so, clearly, we have $\|\text{im}(ab)\| \leq \|\text{im}(b)\|$. Hence $\text{nrnk}(ab) \leq \text{nrnk}(b)$. Let x be an arbitrary element from $\text{dom}(ab)$. Since $(ab)(x) = b(a(x))$, we have $x \in \text{dom}(a)$. Hence, $\text{dom}(ab) \subset \text{dom}(a)$, which implies inequality $\|\text{dom}(ab)\| \leq \|\text{dom}(a)\|$. Thus we have $\text{nrnk}(ab) \leq \text{nrnk}(a)$.

For any two subsets A and B of the set $\mathbb{Z}$ with equal norms less than 2∞ , there exists a unique element a of the semigroup $\mathcal{IO}_{\mathbb{Z}}$ such that $A = \text{dom}(a)$ and $B = \text{im}(a)$. If $\|A\| = \|B\| = 2\infty$, then there are infinitely many elements a of the set $\mathcal{IO}_{\mathbb{Z}}$ such that $A = \text{dom}(a)$ and $B = \text{im}(a)$. In the latter case in order to determine a one needs to specify its value at some point. For convenience, any element $a \in \mathcal{IO}_{\mathbb{Z}}$ with $A = \text{dom}(a)$, $B = \text{im}(a)$, and $m = a(n)$ will be denoted by $\pi_{A,B,n \rightarrow m}$, where $n \in A$, $m \in B$. The element $n \in A$ in this notation will be chosen as follows. It is the minimal element of the set A with respect to the order $0 < 1 < -1 < 2 < -2 < \dots < k < -k < k+1 < -k-1 < \dots$.

The semigroup $\mathcal{IO}_{\mathbb{Z}}$ contains an identity, which is the identical transformation $\text{id}_{\mathbb{Z}}$, and a zero, which is the nowhere defined transformation. The set of idempotents

$$E(\mathcal{IO}_{\mathbb{Z}}) = \{\text{id}_A, A \subset \mathbb{Z}\} := \{\pi_{A,A}, A \subset \mathbb{Z}\}$$

forms a semigroup, which is isomorphic to the semigroup $(2^{\mathbb{Z}}, \cap)$. The idempotent e is primitive element if and only if $e = \text{id}_{\{n\}}$ for some $n \in \mathbb{Z}$. The semigroup $\mathcal{IO}_{\mathbb{Z}}$ is an inverse semigroup since it is a regular semigroup in which idempotents commute. Inverse element of an element $\pi_{A,B,n \rightarrow m}$ is $\pi_{B,A,m \rightarrow n}$.

Proposition 2. Each element of the semigroup $\mathcal{IO}_{\mathbb{Z}}$ can be written as a product of at most four elements of the set

$$G := \{a \in \mathcal{IO}_{\mathbb{Z}} \mid \text{dom}(a) = \mathbb{Z} \text{ or } \text{im}(a) = \mathbb{Z}\}.$$

Proof. Every element of a normalized rank less than 2∞ can be written as a product of two elements of a normalized rank 2∞ : $\pi_{A,B,n \rightarrow m} = \text{id}_C \cdot b$, where $C \cap \text{dom}(b) = A$ and $b(A) = B$. Every element of a normalized rank 2∞ can be written as a product of two elements from the set G . Indeed, let $a \in \mathcal{IO}_{\mathbb{Z}}$ and $\text{nrnk}(a) = 2\infty$, moreover $a = \pi_{A,B,n \rightarrow m}$, then $a = \pi_{A,\mathbb{Z},n \rightarrow k} \pi_{\mathbb{Z},B,k \rightarrow m}$.

3 Ideals and Green's relations

Theorem 3.1. *Let $a \in \mathcal{IO}_{\mathbb{Z}}$. Then*

- 1) *a principal left ideal generated by an element a has the following form*

$$\mathcal{IO}_{\mathbb{Z}} \cdot a = \{b \in \mathcal{IO}_{\mathbb{Z}} : \text{im}(b) \subseteq \text{im}(a)\}$$

- 2) *a principal right ideal generated by an element a has the following form*

$$a \cdot \mathcal{IO}_{\mathbb{Z}} = \{b \in \mathcal{IO}_{\mathbb{Z}} : \text{dom}(b) \subseteq \text{dom}(a)\}$$

- 3) *a principal two-sided ideal generated by an element a has the following form*

$$\mathcal{IO}_{\mathbb{Z}} \cdot a \cdot \mathcal{IO}_{\mathbb{Z}} = \{b \in \mathcal{IO}_{\mathbb{Z}} : \text{nrank}(b) \leq \text{nrank}(a)\}$$

Proof.

- 1) Let $M_a = \{b \in \mathcal{IO}_{\mathbb{Z}} : \text{im}(b) \subseteq \text{im}(a)\}$. If $c \in \mathcal{IO}_{\mathbb{Z}}$ then $ca(x) = a(c(x))$ by definition. Thus $\text{im}(ca) \subseteq \text{im}(a)$ and $\mathcal{IO}_{\mathbb{Z}} \cdot a \subseteq M_a$. Now let $\text{im}(b) \subseteq \text{im}(a)$. Denote $A = \text{dom}(b)$, $B = \text{im}(b)$. Then $\|A\| = \|B\|$. Denote $C = a^{-1}(B)$. Since a^{-1} is partially defined injection, $\|C\| = \|B\|$. So $\|A\| = \|C\|$ and there exists a transformation $\alpha : A \rightarrow C$. Then $b = \alpha \cdot a$. Hence $M_a \subseteq \mathcal{IO}_{\mathbb{Z}} \cdot a$.
- 2) Analogous to the proof of statement 1.
- 3) Let $M_a = \{b \in \mathcal{IO}_{\mathbb{Z}} : \text{nrank}(b) \leq \text{nrank}(a)\}$. For arbitrary $c, d \in \mathcal{IO}_{\mathbb{Z}}$ we have the following inequalities:

$$\text{nrank}(ca) \leq \min\{\text{nrank}(c), \text{nrank}(a)\}$$

and

$$\text{nrank}(ad) \leq \min\{\text{nrank}(a), \text{nrank}(d)\}.$$

Let $b = c \cdot a \cdot d \in \mathcal{IO}_{\mathbb{Z}}$. Then we have

$$\begin{aligned} \text{nrank}(b) &\leq \\ &\leq \min\{\text{nrank}(c), \text{nrank}(a), \text{nrank}(d)\} \leq \\ &\leq \text{nrank}(a). \end{aligned}$$

Thus $\mathcal{IO}_{\mathbb{Z}} \cdot a \cdot \mathcal{IO}_{\mathbb{Z}} \subseteq M_a$. Let now $\text{nrank}(b) \leq \text{nrank}(a)$. Denote $A = \text{dom}(b)$, $B = \text{im}(b)$. Let $C \subseteq \text{dom}(a)$ be an arbitrary subset, such as $\|C\| = \text{nrank}(b) = \|A\|$, $\|D\| = \|a(C)\| = \|C\| = \text{nrank}(b) = \|B\|$. There exist some transformations $\alpha : C \rightarrow A$ and $\beta : D \rightarrow B$. Thus $b = \alpha^{-1} \cdot a \beta$. Hence $M_a \subseteq \mathcal{IO}_{\mathbb{Z}} \cdot a \cdot \mathcal{IO}_{\mathbb{Z}}$.

For every $k \in \mathbb{N}_0$ we denote by

$$I_k = \{b \in \mathcal{IO}_{\mathbb{Z}} : \text{nrank}(b) \leq k\},$$

$$I_F = \{b \in \mathcal{IO}_{\mathbb{Z}} : \text{nrank}(b) < \infty\},$$

$$I_{+\infty} = \{b \in \mathcal{IO}_{\mathbb{Z}} : \text{nrank}(b) \leq +\infty\},$$

$$I_{-\infty} = \{b \in \mathcal{IO}_{\mathbb{Z}} : \text{nrank}(b) \leq -\infty\}.$$

Set $F < \pm\infty$, $F > k$ for all $k \in \mathbb{N}_0$.

Proposition 3. *The sets I_F and I_k , $k \in \mathbb{N}_0$, are two-sided ideals.*

Proof. For the sets I_k , $k \in \mathbb{N}_0$, this statement follows from the definition of an ideal and proposition 1. For arbitrary $a \in I_F$ and $b \in \mathcal{IO}(\mathbb{N}, -\mathbb{N})$ we have $\text{nrank}(bab) \leq \text{nrank}(a) < \infty$. So $bab \in I_F$. Hence I_F is a two-sided ideal.

Proposition 4. *$I_{r_1} \subseteq I_{r_2}$ if and only if $r_1 \leq r_2$.*

Proof. Obviously, $r_1 \leq r_2$ implies $I_{r_1} \subseteq I_{r_2}$. We prove the converse statement by contradiction. Assume that $I_{+\infty} \subset I_{-\infty}$. In particular, it means that some element a with a normalized rank $+\infty$ belongs to the ideal $I_{-\infty}$. By definition of the ideal $I_{-\infty}$ we have that domain of every element of this ideal is bounded from above. However $\text{dom}(a)$ contains an infinite number of positives. We arrive to contradiction. Analogously one can show that $I_{-\infty} \not\subset I_{+\infty}$. This completes the proof.

Proposition 5. *Let $\mathcal{I}$ be the set of principal two-sided ideals. Then every two-sided ideal of the semigroup $\mathcal{IO}_{\mathbb{Z}}$ is a union of at most two sets from the set $\mathcal{I} \cup I_F$.*

Proof. Let $I \subset \mathcal{IO}_{\mathbb{Z}}$ be two-sided ideal. Every two-sided ideal is a union of principal two-sided ideals. Let the set of normalized ranks of elements of the ideal I contain 2∞ . Then decomposition of the ideal I into the union of principal ideals contains the principal ideal $\mathcal{IO}_{\mathbb{Z}}$. Therefore, $I = \mathcal{IO}_{\mathbb{Z}}$.

Let the set of normalized ranks of elements of the ideal I is bounded from above by an element l ($\pm\infty$). Then we take the ideal I_l ($I_{\pm\infty}$) in decomposition of the ideal I . If the set of normalized ranks of elements of the ideal I is not bounded from above, then we take the ideal I_F in decomposition of the ideal I .

Consider the case when the set of normalized ranks of elements of the ideal I is not bounded from above by elements $-\infty$ and $+\infty$. There exist elements $a, b \in I$ such that $\text{nrnk}(a) = -\infty$ and $\text{nrnk}(b) = +\infty$. So $I_{-\infty} \cup I_{+\infty} = I(a) \cup I(b) \subset I$. Suppose $I \not\subset I_{-\infty} \cup I_{+\infty}$. Then there exists an element $a \in I$ such that $a \in \mathcal{IO}_{\mathbb{Z}} \setminus (I_{-\infty} \cup I_{+\infty})$, that is normalized rank of this element is greater than $-\infty$ or $+\infty$. Therefore $\text{nrnk}(a) = 2\infty$. We arrive to contradiction. So $I \subset I_{-\infty} \cup I_{+\infty}$. Hence $I = I_{-\infty} \cup I_{+\infty}$. All other cases are similar.

The next theorem provides a description of Green's relations in $\mathcal{IO}_{\mathbb{Z}}$.

Theorem 3.2. *Let $a, b \in \mathcal{IO}_{\mathbb{Z}}$. Then*

- 1) $a\mathcal{L}b$ if and only if $\text{im}(a) = \text{im}(b)$;
- 2) $a\mathcal{R}b$ if and only if $\text{dom}(a) = \text{dom}(b)$;
- 3) $a\mathcal{H}b$ if and only if $\text{im}(a) = \text{im}(b)$ and $\text{dom}(a) = \text{dom}(b)$;
- 4) $a\mathcal{J}b$ if and only if $\text{nrnk}(a) = \text{nrnk}(b)$;
- 5) $\mathcal{D} = \mathcal{J}$.

Proof.

- 1) Let $a\mathcal{L}b$. Then $a \in L_b$, $b \in L_a$ and $\text{im}(a) \subseteq \text{im}(b)$, $\text{im}(b) \subseteq \text{im}(a)$. Therefore $\text{im}(a) = \text{im}(b)$. On the other hand, if $\text{im}(a) = \text{im}(b)$, then $b \in \mathcal{IO}_{\mathbb{Z}}a$ and $a \in \mathcal{IO}_{\mathbb{Z}}b$. This implies $\mathcal{IO}_{\mathbb{Z}}b \subseteq \mathcal{IO}_{\mathbb{Z}}a$ and $\mathcal{IO}_{\mathbb{Z}}a \subseteq \mathcal{IO}_{\mathbb{Z}}b$. Thus we have $\mathcal{IO}_{\mathbb{Z}}a = \mathcal{IO}_{\mathbb{Z}}b$, so $a\mathcal{L}b$.
- 2) Analogous to the proof of statement 1.
- 3) Let $a\mathcal{H}b$. Then $a\mathcal{L}b$ and $a\mathcal{R}b$ by definition, that is $\text{im}(a) = \text{im}(b)$ and $\text{dom}(a) = \text{dom}(b)$.
- 4) Let $a\mathcal{J}b$. Then $a \in J_b$, $b \in J_a$ and $\text{nrnk}(a) \leq \text{nrnk}(b)$, $\text{nrnk}(b) \leq \text{nrnk}(a)$. Therefore $\text{nrnk}(a) = \text{nrnk}(b)$. On the other hand if $\text{nrnk}(a) = \text{nrnk}(b)$, then $b \in \mathcal{IO}_{\mathbb{Z}}a\mathcal{IO}_{\mathbb{Z}}$ and $a \in \mathcal{IO}_{\mathbb{Z}}b\mathcal{IO}_{\mathbb{Z}}$. This implies $\mathcal{IO}_{\mathbb{Z}}a\mathcal{IO}_{\mathbb{Z}} \subseteq \mathcal{IO}_{\mathbb{Z}}b\mathcal{IO}_{\mathbb{Z}}$ and $\mathcal{IO}_{\mathbb{Z}}b\mathcal{IO}_{\mathbb{Z}} \subseteq \mathcal{IO}_{\mathbb{Z}}a\mathcal{IO}_{\mathbb{Z}}$. Thus we have $\mathcal{IO}_{\mathbb{Z}}a\mathcal{IO}_{\mathbb{Z}} = \mathcal{IO}_{\mathbb{Z}}b\mathcal{IO}_{\mathbb{Z}}$, so $a\mathcal{J}b$.
- 5) $\mathcal{D} \subseteq \mathcal{J}$ by definition. Let $a\mathcal{J}b$, then $\text{nrnk}(a) = \text{nrnk}(b)$ and $\|\text{dom}(a)\| = \|\text{im}(b)\|$. So there exists $c \in \mathcal{IO}_{\mathbb{Z}}$ such that $\text{dom}(c) = \text{dom}(a)$ and $\text{im}(c) = \text{im}(b)$. This implies $a\mathcal{R}c$ and $c\mathcal{L}b$. Then $a(\mathcal{R} \circ \mathcal{L})b$. Thus $a\mathcal{D}b$. Hence $\mathcal{J} \subseteq \mathcal{D}$.

4 Automorphisms

Recall that the infinite dihedral group $\mathbb{D}_{\infty}$ is the group of symmetries (both rotations and reflections) of the set of integers $\mathbb{Z}$ with respect to composition.

Theorem 4.1. $\text{Aut}(\mathcal{IO}_{\mathbb{Z}}) \simeq \mathbb{D}_{\infty}$.

Proof. Let f be an automorphism of $\mathcal{IO}_{\mathbb{Z}}$. An automorphism maps an ideal to an ideal, moreover the relation of inclusion is preserved. One can conclude from the diagram of ideals that $f(I_k) = I_k$ for every $k \in \mathbb{N}_0$, $f(I_F) = I_F$ and $f(I_{\pm\infty}) = I_{\pm\infty}$ or $f(I_{\pm\infty}) = I_{\mp\infty}$. Let $a \in \mathcal{IO}_{\mathbb{Z}}$ be an element of a normalized rank k for some $k \in \mathbb{N}_0$. Then $a \in I_k \setminus I_{k-1}$. This implies $f(a) \in I_k$ and $f(a) \notin I_{k-1}$. Thus $f(a) \in I_k \setminus I_{k-1}$ and $\text{nrnk}(f(a)) = k$. Similarly one can prove that $\text{nrnk}(f(a)) = \pm\infty$ for a of normalized rank $\pm\infty$ in case $f(I_{\pm\infty}) = I_{\pm\infty}$ and $\text{nrnk}(f(a)) = \mp\infty$ for a of normalized rank $\pm\infty$ in case $f(I_{\pm\infty}) = I_{\mp\infty}$. Let a be an element of normalized rank 2∞ . Then it does not belong to any ideal I_k , $I_{+\infty}$ and $I_{-\infty}$ for every $k \in \mathbb{N}_0$. Thus $f(a)$ does not belong to any of these ideals. Hence it has normalized rank 2∞ .

An automorphism f induces a bijection φ on the set of primitive elements. The map φ specifies a permutation τ_{φ} on the set $\mathbb{Z}$ by the following rule: $\tau_{\varphi}(x) = y$ if and only if $\varphi(\text{id}_{\{x\}}) = \text{id}_{\{y\}}$. So $\pi_{\{i\},\{j\}} = \text{id}_{\{i\}}\pi_{\{i\},\{j\}}\text{id}_{\{j\}}$. This implies $f(\pi_{\{i\},\{j\}}) = f(\text{id}_{\{i\}})f(\pi_{\{i\},\{j\}})f(\text{id}_{\{j\}})$ or $f(\pi_{\{i\},\{j\}}) = \text{id}_{\{\tau(i)\}}f(\pi_{\{i\},\{j\}})\text{id}_{\{\tau(j)\}}$. Since f preserves normalized ranks of the type k , where $k \in \mathbb{N}_0$, then $\text{nrnk} f(\pi_{\{i\},\{j\}}) = 1$. On the other hand $\text{dom}(\text{id}_{\{\tau(i)\}}f(\pi_{\{i\},\{j\}})\text{id}_{\{\tau(j)\}})$ does not contain any element $\tau(z)$, if $z \neq i$. Thus $\text{dom} f(\pi_{\{i\},\{j\}}) = \{\tau(i)\}$. Analogously $\text{im} f(\pi_{\{i\},\{j\}}) = \{\tau(j)\}$. Therefore $f(\pi_{\{i\},\{j\}}) = \pi_{\{\tau(i)\},\{\tau(j)\}}$.

Let $a \in \mathcal{IO}_{\mathbb{Z}}$ be an element such that $i \in \text{dom}(a)$ and $a(i) = j$. Then $\text{id}_{\{i\}} \cdot a \cdot \text{id}_{\{j\}} = \pi_{\{i\},\{j\}}$. Hence $\text{id}_{\{\tau(i)\}} f(a) \text{id}_{\{\tau(j)\}} = \pi_{\{\tau(i)\},\{\tau(j)\}}$. So $\tau(i) \in \text{dom}(f(a))$ and $f(a)(\tau(i)) = \tau(j)$. Since $i \in \text{dom}(a)$ is arbitrary, then $\text{dom}(f(a)) \supset \tau(\text{dom}(a))$. For every $i \notin \text{dom}(a)$ we have $\text{id}_{\{i\}} \cdot a = 0$. So $\text{id}_{\{\tau(i)\}} f(a) = 0$ and $\tau(i) \notin \text{dom} f(a)$. Thus $\text{dom} f(a) = \tau(\text{dom}(a))$ and $\varphi(a)(\tau(i)) = \tau(j)$ for every $i \in \text{dom}(a)$. Thus, an automorphism f is completely determined by permutation τ_f , which it induces on the set $\mathbb{Z}$.

If there exist pairs $i < j$, $t < s$ such that $\tau_f(i) < \tau_f(j)$, $\tau_f(s) < \tau_f(t)$, then we have $f(a) = \pi_{\{\tau_f(i), \tau_f(j)\}, \{\tau_f(t), \tau_f(s)\}} \notin \mathcal{IO}_{\mathbb{Z}}$ for $a = \pi_{\{i, j\}, \{t, s\}} \in \mathcal{IO}_{\mathbb{Z}}$. We arrive to contradiction. So for all i, j such that $i < j$ we have $\tau_f(i) < \tau_f(j)$ or for all i, j such that $i < j$ we have $\tau_f(i) > \tau_f(j)$.

If $\tau_f(i) < \tau_f(j)$ for every i, j such that $i < j$, then $\tau_f(i) = i + t$. In this case $f(a) = \pi_{\text{dom}(a), \text{im}(a)+t, n \rightarrow a(n)+t}$, $t \in \mathbb{Z}$. If $\tau_f(i) > \tau_f(j)$ for every i, j such that $i < j$, then $\tau_f(i) = -i - t$. In this case $f(a) = \pi_{\text{dom}(a), \text{im}(a)+t, n \rightarrow a(n)+t}$, $t \in \mathbb{Z}$.

Список використаних джерел

1. V. Fernandes The monoid of all injective order preserving partial transformations on a finite chain / V. Fernandes // Semigroup Forum – 2001. – Vol.62, Num.2. – P. 178 – 204.
2. O. Ganyushkin On the structure of $\mathcal{IO}_n$ / O. Ganyushkin, V. Mazorchuk // Semigroup Forum – 2003 – Vol.66, Num.3 – P. 455 – 483.
3. O. Ganyushkin Introduction to classical finite transformation semigroup / O. Ganyushkin, V. Mazorchuk // London:Springer – 2009. – 314 p.
4. A. Laradji On certain finite semigroups of order-decreasing transformations. I / A. Laradji, A. Umar // Semigroup Forum – 2004. – Vol.69, Num.2. – P. 184 – 200.
5. V. Pyekhtyeryev $\mathcal{H}$ –, $\mathcal{R}$ – and $\mathcal{L}$ –cross-sections of the infinite symmetric inverse semigroup $\mathcal{IS}_X$ / V. Pyekhtyeryev // Algebra Discrete Math –2005. –№1. – P. 92 – 104.
6. B.O. Пехтерев Конгруенції напівгрупи $\mathcal{IO}_{\mathbb{N}}$ / B.O. Пехтерев, К.С. Третьяк // Mat. Stud. – 2011. – №1. – С. 22 – 27.

References

1. FERNANDES V. (2001) *The monoid of all injective order preserving partial transformations on a finite chain* Semigroup Forum, Vol.62, Num.2, pp. 178 – 204.
2. GANYUSHKIN O., MAZORCHUK V. (2003) *On the structure of $\mathcal{IO}_n$* Semigroup Forum, Vol.66, Num.3, pp. 455 – 483.
3. GANYUSHKIN O., MAZORCHUK V. (2009) *Introduction to classical finite transformation semigroup*, London:Springer, 314 pp.
4. LARADJI A., UMAR A. (2004) *On certain finite semigroups of order-decreasing transformations. I*, Semigroup Forum, Vol.69, Num.2, pp. 184 – 200.
5. PYEKHTYERYEV V.O (2005) $\mathcal{H}$ –, $\mathcal{R}$ – and $\mathcal{L}$ –cross-sections of the infinite symmetric inverse semigroup $\mathcal{IS}_X$, Algebra Discrete Math, №1, pp. 92 – 104.
6. PYEKHTYERYEV V.O, TRETIK K.S. (2011) *Congruences of semigroup $\mathcal{IO}_{\mathbb{N}}$* , Mat. Stud., №1, pp. 22 – 27.

Received: 05.09.2014

УДК 517.91

Власенко І. Ю., к.ф.-м.н.

**Критерій топологічної спряженості
однорідних відображень $\mathbb{R}^n$**

Інститут математики НАНУ, м.
Київ, вул. Терещенківська, 3, e-mail:
vlasenko@imath.kiev.ua

I. Yu. Vlasenko, Ph.D. (Phys.-Math.)

**Criterion of topological conjugacy for
homogeneous mappings of $\mathbb{R}^n$**

Institute of Mathematics of the Na-
tional Academy of Sciences of Ukraine,
Kyiv, 3 Tereshchenkivska Street, e-mail:
vlasenko@imath.kiev.ua

У роботі розглянуто клас відображень n -вимірною евклідового простору, у якого координатні функції є однорідними. Описано ряд динамічних та топологічних властивостей цих відображень. Показано, що відображення цього класу мають 2 натуральні трансверсальні інваріантні шарування, де перше шарування це шарування на промені з початку координат, а друге шарування утворене динамічно заданими гіперсферами навколо початку координат. Для відображень даного класу побудовано набір топологічних інваріантів та дано критерій їх топологічної спряженості у термінах породженого відображення корозмірності 1, що задане на гіперсфері, що розділяє басейни притягання точок 0 та ∞ . Цей результат редукує проблему перевірки, чи є два відображення такого класу топологічно спряженими до задачі в на одиницю меншій розмірності.

Ключові слова: однорідні відображення, критерій спряженості, топологічні інваріанти.

We consider a class of maps of the n -dimensional Euclidean space that is a generalization of the class of maps having homogeneous polynomials as its coordinate functions. We describe some dynamical and topological properties of the maps of this class. It is shown that maps of this class has two natural transversal invariant foliations. The former is the foliation on the rays from the origin and the latter is on some hyperspheres around the origin generated by an invariant hypersphere that separates the basins of attraction of 0 and ∞ .

A criterion of their topological conjugacy is obtained. A homogeneous map induce a map on their invariant hypersphere that separates the basins of attraction. It is shown that two homogeneous maps are topologically conjugate if and only if the induced codimension 1 maps are topologically conjugate. This result reduce the problem of topological conjugacy for the maps of the considered class to a codimension 1 problem.

Key words: homogeneous mappings, conjugacy criterion, topological invariants.

Статтю представив доктор фізико-математичних наук, професор В. В. Кириченко.

Introduction. The non-invertible maps are much less studied in comparison to homeomorphisms with the exceptions of one-dimensional and holomorphic maps. Fine topological invariants of such maps are the invariants of their conjugacy class¹. The class of maps introduced here is a natural generalization of maps having homogeneous polynomials as coordinate functions. Note that homogeneous polynomials are already studied for the problem of their classification up to left-right topological equivalence, see [1]. But we consider these maps up to conjugacy which is more accurate equivalence that preserves fine details of

a map lost under left-right topological equivalence.

In the paper [6] similar results were obtained for the inner (open discrete) homogeneous mappings of $\mathbb{R}^2$. In this paper we describe some properties for these maps and prove a criterion of their topological conjugacy without assuming that the map is inner and in case of $\mathbb{R}^n$.

Preliminary information.

Definition 1. Call a continuous map $f: \mathbb{R}^n \rightarrow \mathbb{R}^n$ **homogeneous of a degree k** if $\forall t \geq 0 \forall \bar{x} \in \mathbb{R}^n$
 $f(t^k \bar{x}) = t^k f(\bar{x})$.

From the definition $f(\bar{0}) = \bar{0}$.

¹ f and g are topologically conjugate if there exists a homeomorphism h such that $fh = hg$.

Definition 2. A homogeneous map f is called **non-degenerate** if $f^{-1}(\bar{0}) = \bar{0}$.

Let $f : \mathbb{R}^n \rightarrow \mathbb{R}^n$ be a non-degenerate homogeneous map of a degree $k > 1$. Let a hypersphere $\hat{\mathbb{R}}^n$ be the one point compactification of $\mathbb{R}^n$ with the point ∞ . Then $f(\infty) = \infty$ and f also can be considered as the map $f : \hat{\mathbb{R}}^n \rightarrow \hat{\mathbb{R}}^n$. We assume that f can be a non-invertible and a non-surjective map. Note that if the homogeneous map is differentiable, then from the definition if $\bar{x}$ is critical, then the whole ray $\{t\bar{x} | t > 0\}$ is critical.

Some of the following lemmas were proved in the paper [6]. It was in case of $\mathbb{R}^2$, but the proof holds for $\mathbb{R}^n$ literally or with trivial obvious changes.

Lemma 1. The foliation of $\mathbb{R}^n \setminus \{\bar{0}\}$ of the rays from the origin is invariant under the action of f . Furthermore, any two pre-images of the same point can't belong to the same ray.

Proof. Let $x = ty \in \mathbb{R}^n \setminus \{\bar{0}\}$ and $x \neq y$. Then $x, y \neq \bar{0}$, $t \neq 0$, $t \neq 1$. Due to homogeneousness, $f(x) = f(ty) = t^k f(y)$, i.e. f maps a ray into a ray.

Show that any two pre-images of the same point can't belong to the same ray. Indeed, let $x = ty \in \mathbb{R}^n \setminus \{\bar{0}\}$ and $f(x) = f(y)$, $t \neq 0$, $t \neq 1$. Then $f(x) = t^k f(y) = t^k f(x)$ due to homogeneousness, which is not possible for $x \neq \bar{0}$, $t \neq 0$, $t \neq 1$ if f is not degenerate. $\square$

Lemma 2. For the map f each of the points $\bar{0}$ and ∞ has open basins of attraction.

Proof. Let D_r be a disc of radius r around the origin, and a hypersphere $S_r = \partial D_r$ be its boundary. Consider the image of D_r the set $f(D_r)$.

Note that the boundary $\partial f(D_r) \subset f(S_r)$. Consider $f(S_r)$. Let $d_1^{min} = \min \rho(O, f(S_1))$, $d_1^{max} = \max \rho(O, f(S_1))$, where ρ is the Euclidean metric. Then $d_1^{max} \geq d_1^{min} > 0$ because $f(S_1)$ is a compact hypersurface that does not intersect the point $\bar{0}$. Since $f(S_r)$ is homogeneous we have

$$d_r^{min} = \min \rho(O, f(S_r)) = r^k d_1^{min},$$

$$d_r^{max} = \max \rho(O, f(S_r)) = r^k d_1^{max}.$$

Thus for $r < 1/d_1^{max}$ $d_r^{max} < r$ and the interior $\text{Int}(D_r)$ of the disc D_r is the strictly invariant neighborhood containing the point $\bar{0}$, i.e. $f(\text{Int}(D_r)) \subset \text{Int}(D_r)$.

By definition, an attractor of the strictly invariant neighborhood $\text{Int}(D_r)$ is the set $\Lambda = \bigcap_{n \geq 0} f^n(\text{Int}(D_r))$. As $\bar{0} \in \Lambda$, $\Lambda \neq \emptyset$. And since $d_r^{max} = r^k d_1^{max}$, Λ is just the point $\bar{0}$.

Similarly, for $r > 1/d_1^{min}$ $d_r^{min} > r$ and the exterior $\hat{\mathbb{R}}^n \setminus D_r$ of the disc D_r is the strictly invariant neighborhood containing the point ∞ and the resulting attractor is the point ∞ itself.

The lemma is proven. $\square$

Lemma 3. On every ray from the origin there is exactly one point that is not attracted to the point $\bar{0}$ or ∞ .

Proof. Consider a ray from the origin. The basins of attraction of the points $\bar{0}$ and ∞ are open sets, the ray is a linearly connected set. Then the set N of points of the ray not belonging to the basins of attraction of the points $\bar{0}$ and ∞ is not empty.

Assume that the set N has two different points p_1 and $p_2 \neq p_1$. Since they are on the same ray their coordinates are proportional: $p_2 = tp_1$, $t > 0$, $t \neq 1$. Consider $f^n(p_1)$, $n > 0$. p_1 does not belong to the basins of attraction of the points $\bar{0}$ and ∞ . Therefore, the norm $\|f^n(p_1)\|$ is bounded: $\exists C_1, C_2 > 0$: $C_2 > \|f^n(p_1)\| > C_1$.

Consider $f^n(p_2)$, $n > 0$. $\|f^n(p_2)\| = \|t^{kn} f^n(p_1)\| = t^{kn} \|f^n(p_1)\|$. Then $t^{kn} C_2 > \|f^n(p_2)\| > t^{kn} C_1$. But $t \neq 1$. It follows that $f^n(p_2)$ tends either to $\bar{0}$ or to ∞ when $n \rightarrow \infty$ depending on whether $t > 1$ or $t < 1$. But p_2 belongs to N .

The obtained contradiction proves the lemma. $\square$

Lemma 4. Points that does not belong to the basins of attraction of the points $\bar{0}$ and ∞ create a hypersurface that is homeomorphic to the sphere S^{n-1} and separates the basins of attraction of the points $\bar{0}$ and ∞ .

Proof. Define the set of points that does not belong to the basins of attraction of the points $\bar{0}$ and ∞ by γ_1 . The basins of attraction of the points $\bar{0}$ and ∞ are open mutually exclusive sets by definition. As the set that divides them and their complement, the set γ_1 is closed. Also it is bounded, as follows from the proof of the lemma 2, then γ_1 is a compact set.

Note that the set of rays from the origin is homeomorphic to a hypersphere and is a compact

set. According to the lemma 3, there is a one-to-one map (bijection) between the rays from the origin and the points of γ_1 . Show that this bijection is continuous. Define a map $p: \mathbb{R}^n \setminus \{\bar{0}\} \rightarrow S^{n-1}$ as $p(\bar{x}) = \frac{\bar{x}}{\|\bar{x}\|}$. Then the restriction $q = p|_{\gamma_1}: \gamma_1 \rightarrow S^1$ by definition is continuous and is the bijection above. It is well-known (see, for example, [2]), that continuous bijection of compact sets is a homeomorphism. It follows that the one-to-one map between the rays from the origin and the points of γ_1 is indeed a homeomorphism and the lemma is proven. $\square$

Lemma 5. The hypersurface γ_1 that separates the basins of attraction of the points $\bar{0}$ and ∞ is totally invariant ($f^{-1}(\gamma_1) = \gamma_1$).

Proof. It follows from a more generic fact, that any union $\hat{A}$ of an attractor A and its basin of attraction can be represented as $\cup_n f^{-n}(U)$, where U is the strictly attracting neighborhood of the attractor A ($\overline{f(U)} \subset U$). As the representation of $\hat{A}$ is the union of all pre-images it is obvious that the result, the union $\hat{A}$ of an attractor A and its basin of attraction, is totally invariant. Since γ_1 is a complement to two such unions of $\bar{0}$ and its basin and ∞ and its basin which are totally invariant then γ_1 itself is totally invariant. $\square$

Consider a ray from the origin. Let it intersect the hypersurface γ_1 in a point p_1 . The points of the ray can be then represented as $p_t = tp_1$, $t > 0$. Using t as the coefficient of proportionality, let us build a family of hypersurfaces γ_t , $t > 0$, such that they are homothetic to γ_1 with respect to the origin. By construction it is a foliation of $\mathbb{R}^n \setminus \{\bar{0}\}$.

The set of rays from the origin is homeomorphic to the sphere S^{n-1} . Denote the set by S_ϕ^{n-1} . According to the lemma 1, f preserves the partitioning of $\mathbb{R}^n \setminus \{\bar{0}\}$ by the rays from the origin. Thus f induces a continuous map f_ϕ on S_ϕ^{n-1} .

Theorem 1 (Topological conjugacy criterion). Let f and g be homogeneous maps of a degree $k > 1$ and $f_\phi, g_\phi: S_\phi^{n-1} \rightarrow S_\phi^{n-1}$ be the continuous maps induced by f and g on the set S_ϕ^{n-1} . f and g are topologically conjugate $\iff f_\phi$ and g_ϕ are topologically conjugate.

Proof. According to the lemma 3, The hypersurface $\gamma_1(f)$ intersect each ray exactly in one point. Using this fact we can build a foliation

of $\mathbb{R}^n \setminus \{\bar{0}\}$ on homothetic images of $\gamma_1(f)$ with respect to the origin, the family $\gamma_t(f)$, $t > 0$. In the same way we can build the family $\gamma_t(g)$, $t > 0$.

Note that the homotheticity map from $\gamma_p(f)$ to $\gamma_q(f)$ is exactly the multiplication by the scalar $\frac{q}{p}$ for each coordinate. This map is also a homotheticity from $\gamma_p(g)$ to $\gamma_q(g)$. Denote this map by $H^{p,q}$. Note that because $\frac{p}{q} \frac{q}{p} = 1$ we have $H^{p,q} \circ H^{q,p} = \text{Id}$ (identity equation 1).

Since f_ϕ and g_ϕ are topologically conjugate there exists a homeomorphism $h_\phi: S_\phi^{n-1} \rightarrow S_\phi^{n-1}$ such that $h_\phi g_\phi = f_\phi h_\phi$. There is one-to-one correspondence between the rays from the set S_ϕ^{n-1} , the points of $\gamma_1(f)$ and the points of $\gamma_1(g)$ because of the fact that $\gamma_1(f)$ and $\gamma_1(g)$ intersect each ray in a single point. Then h_ϕ induces a homeomorphism of the hypersurfaces $h_1: \gamma_1(f) \rightarrow \gamma_1(g)$ such that on $\gamma_1(f)$ we have $h_1 g = f h_1$ (identity equation 2).

Define the homeomorphism h on $\mathbb{R}^n \setminus \{\bar{0}\}$. For that define $h_p: \gamma_p(f) \rightarrow \gamma_p(g)$ as $h_p = H^{1,p} \circ h_1 \circ H^{p,1}$. Define $h: \mathbb{R}^n \setminus \{\bar{0}\} \rightarrow \mathbb{R}^n \setminus \{\bar{0}\}$ as follows: if $\bar{x} \in \gamma_p(f)$ then $h(\bar{x}) = h_p(\bar{x})$. This definition is correct because the hypersurfaces belong to the foliation. The obtained map h is one-to-one and continuous by construction. It extends to the points $\bar{0}$ and ∞ by continuity: $h(\bar{0}) = \bar{0}$ and $h(\infty) = \infty$. A continuous one-to-one map of compacts is a homeomorphism. It follows that $h: \hat{\mathbb{R}}^n \rightarrow \hat{\mathbb{R}}^n$ constructed here is a homeomorphism of $\hat{\mathbb{R}}^n$.

Let us verify that h is required conjugacy homeomorphism. It is enough to check on $\mathbb{R}^n \setminus \{\bar{0}\}$.

Let $\bar{x} \in \gamma_p(f)$. Then $\exists \bar{x}_1 \in \gamma_1(f): \bar{x} = p\bar{x}_1 = H^{1,p}(\bar{x}_1)$. h commutes with homotheticity maps $H^{p,q}$ by construction, f and g commute with $H^{p,q}$ because they are homogeneous and $H^{p,q}$ is a multiplication by a scalar value. Using those commutations and identity equations (1) and (2) above, we get

$$\begin{aligned} h(f(\bar{x})) &= h_p(f(\bar{x})) = h_p(f(H^{1,p}(\bar{x}_1))) = \\ h_p(H^{1,p}(f(\bar{x}_1))) &= H^{1,p}(h_1(H^{p,1}(H^{1,p}(f(\bar{x}_1))))) = \\ H^{1,p}(h_1(f(\bar{x}_1))) &= H^{1,p}(g(h_1(\bar{x}_1))) = \\ g(H^{1,p}(h_1(H^{p,1}(\bar{x})))) &= g(h_p(\bar{x})) = g(h(\bar{x})). \end{aligned}$$

Thus, $h(f(\bar{x})) = g(h(\bar{x}))$ and f and g topologically conjugate with the help of h . The theorem is proven. $\square$

Corollary 1. If f is a homogeneous maps of a degree $k > 1$ then it induces a map f' on the set of rays from the origin such that f is topologically conjugate on $\mathbb{R}^n \setminus \{0\}$ to the map $S^{n-1} \times \mathbb{R} \ni (x, t) \mapsto (f'(x), t^k)$.

Conclusions

For a class of maps of $\mathbb{R}^n$ that is a generalization of the class of maps having homogeneous

polynomials as its coordinate functions we have described some topological invariants and have reduced the problem of their topological conjugacy to a codimension 1 problem.

The obtained results can be used for the global classification of the homogeneous mappings in dimension 2 and 3, as well as an initial step for the study of homogeneous mappings in greater dimensions.

All these results are applied in the study of the polynomial inner mappings.

Список використаних джерел

1. Арнольд В.И. Особенности дифференцируемых отображений. Т.1 / В.И.Арнольд, А.Н.Варченко, С.М.Гусейн-Заде // М., Наука. – 1982. – 303 с.
2. Власенко И. Топологические методы в изучении групп преобразований многообразий. / И.Власенко, С.Максименко, Е.Полулях // Праці Інституту математики НАН України. Математика та її застосування. Том 61 Інст. математики НАНУ. Київ, – 2006. – 363 с.
3. Власенко И. Внутренние отображения: топологические инварианты и их приложения. / И.Власенко // Праці Інституту математики НАН України. Математика та її застосування. Том 101 // Інст. математики НАНУ. Київ, – 2014. – 225 с.
4. Власенко И. Особенности динамики бесконечнократных внутренних по Трохимчуку эпиморфизмов. / И.Власенко // Труды Межд. конф. “Геометрия в Одессе – 2008” Інст. математики НАНУ. Київ, – 2008. – с. 373–389.
5. Власенко И. Динамика внутренних отображений. / И.Власенко // Нелінійні Коливання. – 2011. Vol.14, №2, – с. 181–186.
6. Власенко И. Критерий топологической сопряженности двумерных однородных внутренних отображений. / И.Власенко // Труды международного геометрического центра. – 2015. №1. – с. 6–15.

References

1. ARNOLD, V. I., GUSEIN-ZADE S. M. and VARCHENKO, A. N. (2012), “Singularities of differentiable maps. V. 1”, Springer. – xii+382 p.
2. VLASENKO I., MAKSIMENKO S. and POLULYAH E. (2006) “Topological methods in studying of the transformation groups of manifolds”, Kiev, Inst. of Math. – 363 p.
3. VLASENKO I. (2014) “Inner mappings: topological invariants and their applications”, Kiev, Inst. of Math. – 225 p.
4. VLASENKO I. (2008) “Osobennosti dinamiki beskonечноkратnyh vnutrennih po Trohimchuku epimorphismov”, Proceedings of the international conference “Geometry in Odessa – 2008” Kiev, Inst. of Math. – pp. 373–389.
5. VLASENKO I. (2011) “Dynamics of inner mappings”, Nonlinear Oscillations. - Vol.14, No.2, – pp. 181–186.
6. VLASENKO I. (2015) “On the topological conjugacy of two-dimensional homogeneous inner mappings”, Proc. Intern. Geom. Center, – No.1, – pp. 6–15.

Надійшла до редколегії 20.12.2014

УДК 517.91

О. М. Іванюк¹,
С. В. Білун², к. ф.-м. н., асистент

Деформації атомів степені 2 на поверхнях з краєм

¹Київський національний університет імені
Тараса Шевченка, м. Київ, просп. Глушкова, 4е

²Київський національний університет імені
Тараса Шевченка, м. Київ, просп. Глушкова, 4е
E-mail: ¹oxana_ivanyuk@ukr.net

O. M. Ivanyuk¹,
S. V. Bilun², Ph. D., Assistant Professor

Deformations of atoms of complexity 2 on surfaces with boundary

¹Taras Shevchenko National University of Kyiv,
Kyiv, Hlushkova prosp., 4e

²Taras Shevchenko National University of Kyiv,
Kyiv, Hlushkova prosp., 4e
E-mail: ¹oxana_ivanyuk@ukr.net

Розглядаються такі функції Морса на орієнтованих компактних поверхнях з межею, що їх обмеження на межу також є функцією Морса. Знайдено всі можливі деформації функцій в околах критичних рівнів з двома критичними точками до функцій, у яких по одній критичній точці на кожному критичному рівні. Показано, що таку деформацію в кожному з випадків можна здійснити двома способами.

Ключові слова: функція Морса, атом m -функції, топологічна класифікація.

We consider m -functions on a closed oriented two-dimensional manifold (surface) M with boundary ∂M that is smooth functions f such that: 1) f is a Morse function, 2) restriction $f|_{\partial M}$ is a Morse function. We investigate changing of neighborhood of a critical level (f -atom) points under a small deformation of the function. Near f -atom is considered up to homeomorphism that maps level component of function into level component and preserve the direction of function arising. An atom with one critical point (a simple atom) is stable under smooth deformation of the function. In one a one-parameter family of m -function, an atom with 2 critical points (complexity 2) can appear. Each atom of complexity 2 splits at a small movement in the one-parameter family of m -functions into two simple atoms by two ways. All possible deformations of atoms of complexity 2 of m -functions on surfaces with boundary to simple atoms are considered.

Key words: Morse function, atom of m -function, topological classification.

Статтю представив д. ф.-м. н. проф. Кириченко В.В.

Introduction. Let M be a closed oriented two-dimensional manifold (surface). Let f be a smooth function on M . Consider the Hamiltonian dynamical system given by the equation $\frac{dx}{dt} = \text{sgrad } f(x)$, $x \in M$. Then its trajectories lie on the components of the line-level function f . These components are called layers. Homeomorphism of surface reflecting layers on layers is called a layered equivalence. So the layered classification of functions sets the topological classification of Hamiltonian dynamical systems. On the set of all functions one can identify an open everywhere dense subset consisting of simple Morse functions. To study the functions A.Konrod [4] and H.Rib [7] introduced a graph, which is obtained from the surface after collecting each layer to the point. This graph is a complete topological invariant of simple Morse functions.

The analogue of Morse functions for manifolds with boundary are m -functions. The topological properties of m -functions were investigated in

[2,3,5,6,8–12]. In [3] all possible atoms of complexity 2 of m -functions on surfaces with boundary are considered. The purpose of the scientific work is to describe all possible deformations of atoms of complexity 2 to simple atoms.

m -function. Simple atoms. Let M is a smooth manifold of dimension n , $f : M \rightarrow \mathbb{R}$ is a smooth function.

The point $x \in M$ is a *critical* for function f if $df(x) = 0$. Herewith $f(x)$ is a *critical value* for function f .

The critical point is called *non-degenerate* if the second differential $d^2f = \sum \frac{\partial^2 f}{\partial x_i \partial x_j} dx_i dx_j$ is non-degenerate at this point.

Definition. A smooth function on a manifold M is a *Morse function* if it has no degenerate critical points.

Let M is a smooth compact manifold of dimension n with boundary.

Definition. The function $f : M \rightarrow \mathbb{R}$ is m -function if:

a) all its critical points are non-degenerate and do not situated on ∂M ;

б) the boundary of manifold can be represented as a union $\partial M = \partial M_- \cup \partial M_0 \cup \partial M_+$ so that the restriction f_∂ of function f on ∂M_0 is Morse function and when the set $\partial M_- \neq \emptyset$ ($\partial M_+ \neq \emptyset$) then function f takes minimum (maximum) value.

We will call m -functions f and g on the surfaces X^2 and Y^2 *layered equivalent* if there is a diffeomorphism $\lambda : X^2 \rightarrow Y^2$ which maps connected components line-level of function f into connected components line-level of function g . The pair (X^2, f) is layered equivalent to the pair (Y^2, g) .

We will explore the layered equivalence of m -functions in the neighborhoods of their critical values.

Definition. The neighborhood of the P^2 critical layer which is given by the inequality $c - \varepsilon \leq f \leq c + \varepsilon$ for sufficiently small ε which stratified on the line level of function f and which is considered accurate to the layered equivalence $P^2 = \{x : -\varepsilon \leq f(x) - c \leq \varepsilon\}$ is called an *atom*.

If the critical value c is local minimum or local maximum then an atom is called an *atom* A . If the critical value c is saddle then the corresponding atom is called a *saddle*.

Atom is called *simple* if m -function in the (P^2, f) is simple. Remaining atoms are called

complex.

Let c is the critical value of function f on X^2 and c' is the critical value of function g on Y^2 . We consider their specific layers: $f^{-1}(c)$ and $g^{-1}(c')$ and suppose that these layers are coherent.

m -functions f and g are called *layered equivalent* in the neighborhood of their particular layers $f^{-1}(c)$ and $g^{-1}(c')$ if there are two positive numbers ε i ε' and diffeomorphism $\lambda : f^{-1}(c - \varepsilon, c + \varepsilon) \rightarrow g^{-1}(c' - \varepsilon', c' + \varepsilon')$, which converts line-level of function f in line-level of function g and preserves the direction of functions growth that λ maps the range $(f > c)$ into the range $(g > c')$.

We consider the pair (P^2, f) where P^2 is a connected compact surface with non-empty edge ∂P^2 and f is m -function on it which has exactly one critical value C moreover $f^{-1}(c - \varepsilon) \cup f^{-1}(c + \varepsilon) = \partial P^2$. The class of equipped layered equivalence of this pair (P^2, f) is called *f-atom* or *equipped atom*.

Remark. Each atom corresponds to two f -atoms. Sometimes these two atoms can coincide, be equivalent.

Definition. *f-atom* is an atom with the previous definition for which the splitting of rings into positive and negative is observed.

The simplest examples of atoms is a neighborhood of maxima, minima and saddle critical points on surfaces. When passing through the critical level of m -function of the common situation levels adjustment may occur [2] (fig.1).

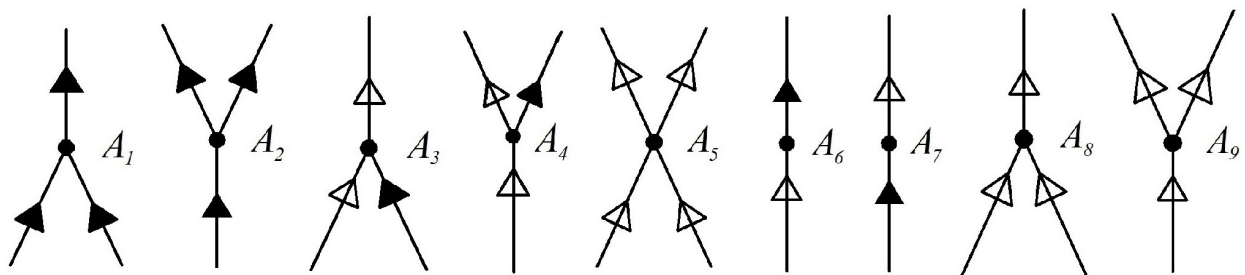


Fig. 1.

Atoms of complexity 2. All possible atoms of complexity 2 of m -functions on surfaces with boundary have been considered in the work [3].

Each atom of complexity 2 splits at a small movement in the one-parameter family of m -functions into two simple atoms by two ways (fig.2-11):

We stretch the points to different critical levels. Fig.2 shows that atom B_1 splits into two

simple atoms. Firstly we consider the case when the first critical point is above the second critical point. Having made a small movement we obtain two simple atoms: A_1 and A_6 . If we have a second critical point at a higher critical level than the first one we get other simple atoms: A_6 and A_3 . And similarly for B_2 , whereas B_1 and B_2 are the same as atoms although different as f -atoms.

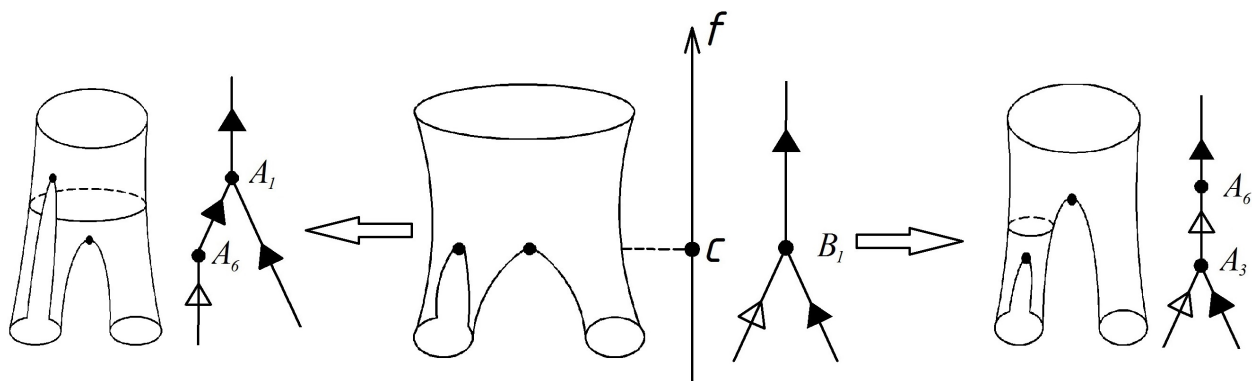


Fig. 2.

Fig.3 shows that atom B_3 splits into two simple atoms by two ways: A_3, A_7 and A_7, A_1 . Similarly for B_4 , whereas both atoms B_3 and B_4 are the same as atoms although different as f -atoms.

Atom B_5 splits into two simple atoms: A_8, A_5 and A_3, A_6 . And similarly for B_6 , whereas B_5 and B_6 are the same as atoms although different as f -atoms.

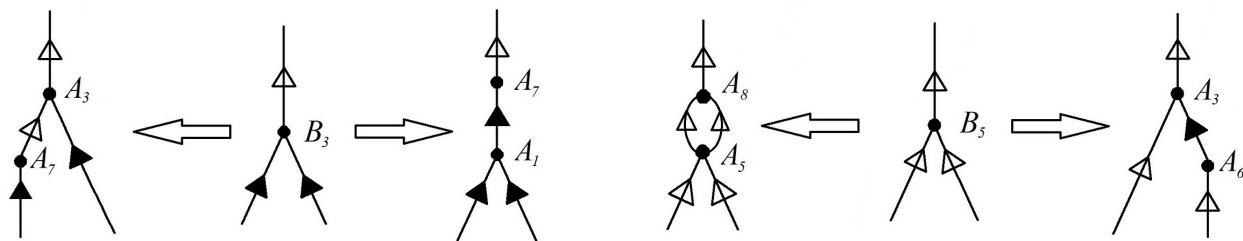


Fig. 3.

Fig.4 shows that atom B_7 splits into two simple atoms by two ways: A_3, A_8 and A_8, A_3 . Similarly for B_8 .

Atom B_9 splits into two simple atoms: A_5, A_9 and A_9, A_5 . And similarly for B_{10} .

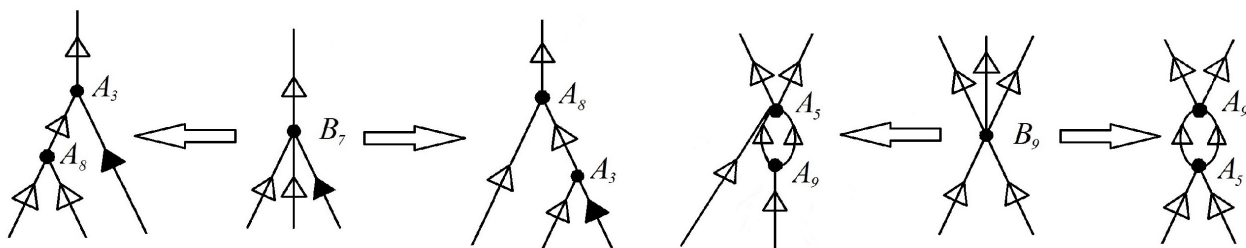


Fig. 4.

Fig.5 shows that atom B_{11} splits into two simple atoms by two ways: A_1, A_1 and A_1, A_1 . Similarly for B_{12} .

Atom B_{13} splits into two simple atoms: A_3, A_3 and A_3, A_1 . And similarly for B_{14} , whereas both atoms B_{13} and B_{14} are the same as atoms although different as f -atoms.

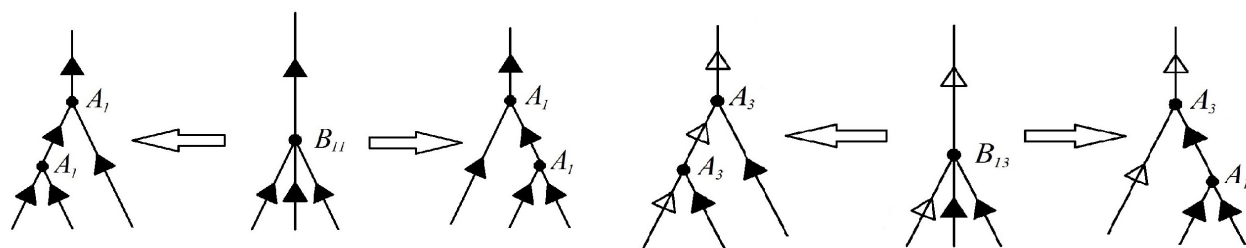


Fig. 5.

Fig.6 shows that atom B_{15} splits into two simple atoms by two ways: A_5, A_3 and A_5, A_4 . Similarly for B_{16} .

Atom B_{17} splits into two simple atoms: A_5, A_7 and A_9, A_3 . And similarly for B_{18} .

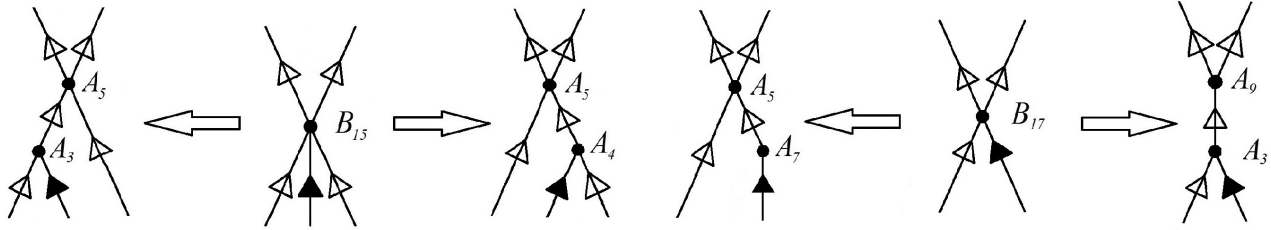


Fig. 6.

Fig.7 shows that atom B_{19} splits into two simple atoms by two ways: A_3, A_9 and A_9, A_3 . Similarly for B_{20} .

Atom B_{21} splits into such simple atoms: A_5, A_5 and A_5, A_5 .

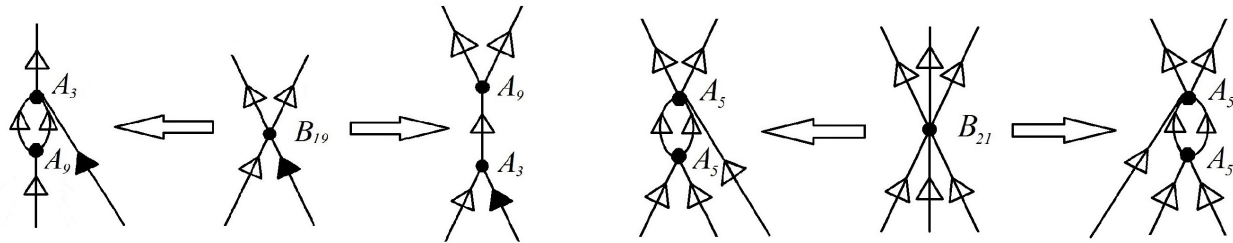


Fig. 7.

Fig.8 shows that atom B_{22} splits into two simple atoms by two ways: A_1, A_2 and A_2, A_1 .

And atom B_{23} splits into such simple atoms: A_4, A_4 and A_4, A_4 .

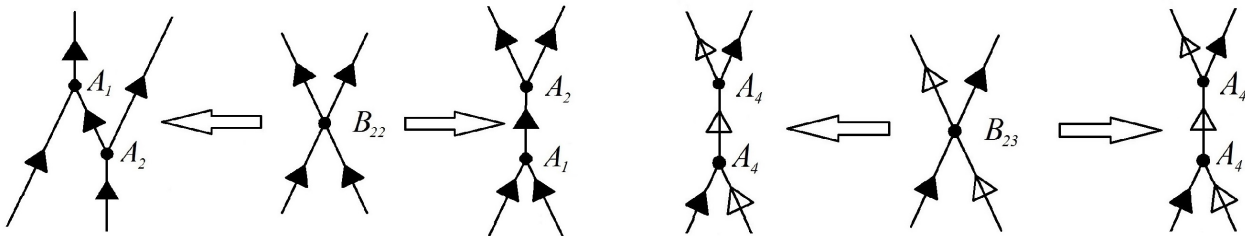


Fig. 8.

Fig.9 shows that atom B_{24} splits into two simple atoms by two ways: A_3, A_2 and A_4, A_3 . And in the same way for B_{25} .

Atom B_{26} splits into the following simple atoms: A_3, A_3 and A_5, A_5 .

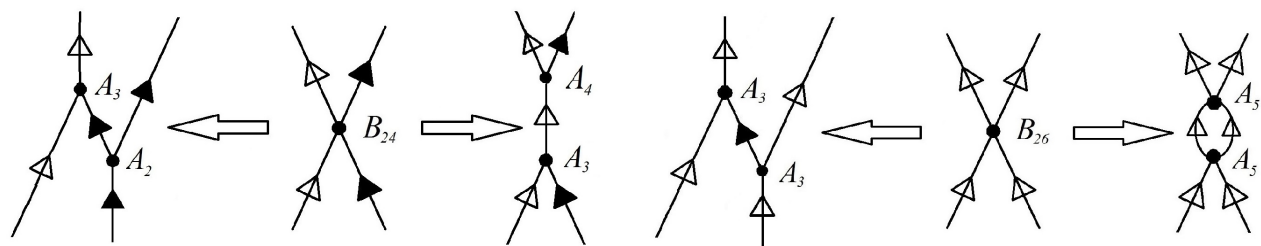


Fig. 9.

Fig.10 shows that atom B_{27} splits into two simple atoms by two ways: A_9, A_7 and A_9, A_7 . Similarly for B_{28} .

Atom B_{29} splits into the following simple atoms: A_8, A_9 and A_7, A_6 .

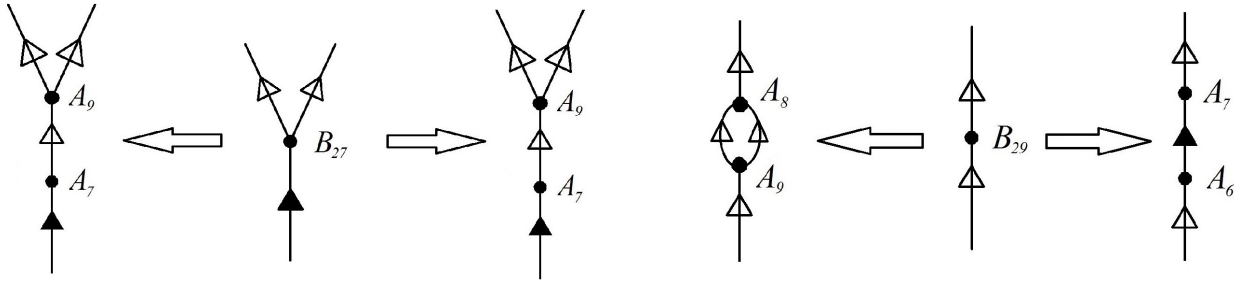


Fig. 10.

Fig.11 shows that atom B_{30} splits into two simple atoms: A_5 , A_5 and A_9 , A_8 .

Atom B_{31} splits into two simple atoms by two ways: A_9 , A_9 and A_9 , A_9 . And similarly for B_{32} .

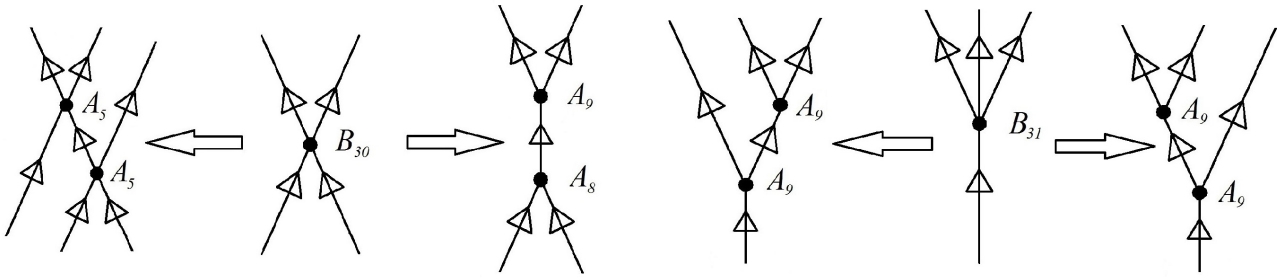


Fig. 11.

Summarizing all of the above the following theorem is just.

Theorem. All possible deformations of atoms of complexity 2 to simple atoms are set in the table:

B_1	B_2	B_3	B_4	B_5	B_6	B_7	B_8
A_1	A_7	A_3	A_6	A_8	A_5	A_3	A_9
A_6	A_2	A_7	A_4	A_5	A_9	A_8	A_4
A_6	A_4	A_7	A_2	A_3	A_7	A_8	A_4
A_3	A_7	A_1	A_6	A_6	A_4	A_3	A_9

B_9	B_{10}	B_{11}	B_{12}	B_{13}	B_{14}	B_{15}	B_{16}
A_5	A_8	A_1	A_2	A_3	A_4	A_5	A_4
A_9	A_5	A_1	A_2	A_3	A_4	A_3	A_5
A_9	A_5	A_1	A_2	A_3	A_2	A_5	A_3
A_5	A_8	A_1	A_2	A_1	A_4	A_4	A_5

B_{17}	B_{18}	B_{19}	B_{20}	B_{21}	B_{22}	B_{23}	B_{24}
A_5	A_6	A_3	A_8	A_5	A_1	A_4	A_3
A_7	A_5	A_9	A_4	A_5	A_2	A_4	A_2
A_9	A_4	A_9	A_4	A_5	A_2	A_4	A_4
A_3	A_8	A_3	A_8	A_5	A_1	A_4	A_3

B_{25}	B_{26}	B_{27}	B_{28}	B_{29}	B_{30}	B_{31}	B_{32}
A_1	A_3	A_9	A_6	A_8	A_5	A_9	A_8
A_3	A_3	A_7	A_8	A_9	A_5	A_9	A_8
A_3	A_5	A_9	A_6	A_7	A_9	A_9	A_8
A_4	A_5	A_7	A_8	A_6	A_8	A_9	A_8

Conclusions

We have considered all possible deformations of atoms of complexity 2 of m -functions on surfaces with boundary to simple atoms.

The results by O. V. Bolsinov and A. T. Fomenko [1] and O. O. Prishlyak [6] have been generalized in the work.

The obtained results can be used for the global classification of m -functions of complexity 2, as well as an initial step for the study of atoms of greater complexity.

All these results are applied in the study of vector fields of oblique gradient on the manifolds with boundary.

Список використаних джерел

1. Болсинов А. Интегрируемые гамильтоновы системы. Геометрия, топология, классификация. / А. Болсинов, А. Фоменко. – Киев: Изд-во Дом «Удмуртский университет». – 1999. – С. 66–99.
2. Пришляк О. О. Класифікація простих m -функцій на орієнтованих поверхнях / О. О. Пришляк, К. О. Пришляк, К. І. Міщенко, Н. В. Лукова // Журнал обчисл. та прикл. матем. – 2011. – №1(104). – С. 1–12.
3. Іванюк О. М. Атоми степені 2 на поверхнях з краєм / О. М. Іванюк, О. О. Пришляк // Proc. Intern. Geom. Center. – 2013. – Vol. 6. – №3. – С. 40–53.
4. Кронрод А. С. О функциях двух переменных / А. С. Кронрод // Успехи мат. наук. – 1950. – №5(35). – С. 24–134.
5. Максименко С. И. Классификация m -функций на поверхностях / С. И. Максименко // Укр. мат. журн. – 1999. – №8. – С. 1129–1135.
6. Prishlyak A. O. Morse functions with finite number of singularities on a plane / A. O. Prishlyak // Methods of func. and topology. – 2002. – Vol. 8. – №1. – P. 75–78.
7. Reeb G. Sur les points singuliers de une forme de pfaff completamente integrable ou de une function numerique / G. Reeb // Comptes Rendus Hebdomadaires des Seances de Academie des Sciences. – 1954. – Vol. 222. – P. 847–849.
8. Лукова Н. В. Функції загального положення на тривимірних многовидах з межею / Н. В. Лукова, О. О. Пришляк // Вісн. КНУ, Мат. Мех. – 2009. – №21. – С. 32–35.
9. Пришляк О. О. Топологічні властивості функцій на тривимірних тілах / О. О. Пришляк, К. О. Пришляк, О. Н. Вятчанінова // Журнал обчисл. та прикл. матем. – 2010. – №2(101). – С. 113–119.
10. Пришляк О. О. Класифікація простих m -функцій на орієнтованих поверхнях / О. О. Пришляк, К. О. Пришляк, К. І. Міщенко, Н. В. Лукова // Журнал обчисл. та прикл. матем. – 2011. – №1(104). – С. 116–127.
11. Пришляк О. О. m -функції на неорієнтованих поверхнях / О. О. Пришляк, К. О. Пришляк, Н. В. Лукова-Чуйко // Журнал обчисл. та прикл. матем. – 2012. – №2(108). – С. 176–185.
12. Іванюк О. М. Молекули m -функцій степені 2 на поверхнях з краєм / О. М. Іванюк, О. О. Пришляк // Proc. Intern. Geom. Center. – 2014. – Vol. 7. – №3. – С. 27–37.

References

1. BOLSINOV, A. V. and FOMENKO, A. T. (1999), “Integrable Hamiltonian systems”, *Geometry, topology, classification*, pp. 66–99.
2. PRISHLYAK, A. O., PRISHLYAK, K. O., MISCHENKO K. I. and LUKOVA N. V. (2011), “Classification of simple m -functions on oriented surfaces”, *Journal of calc. and appl. math*, No. 1 (104), pp. 1–12.
3. IVANYUK, O. M. and PRISHLYAK, A. O. (2013), “Atoms of complexity 2 on surfaces with boundary”, *Proc. Intern. Geom. Center*, v.6, No. 3, pp. 40–53.
4. KRONROD, A. S. (1950), “About functions of two variables”, *Russian Math. Surveys*, No. 5(35), pp. 24–134.
5. MAKSIMENKO, S. I. (1999), “Classification of m -functions on surfaces”, *Ukr. Math. Journal*, T. 51, No. 8, pp. 1129–1135.
6. PRISHLYAK, A. O. (2002), “Morse functions with finite number of singularities on a plane”, *Methods of func. and topology*, v. 8, No. 1, pp. 75–78.
7. REEB, G. (1954), “Sur les points singuliers de une forme de pfaff completamente integrable ou de une function numerique”, *Comptes Rendus Hebdomadaires des Seances de Academie des Sciences*, v. 222, pp. 847–849.
8. LUKOVA, N. V. and PRISHLYAK, A. O. (2009), “Function in general position on 3-manifolds with boundary”, *Bulletin of Taras Shevchenko National Univ. of Kyiv, Math. Mech.*, v. 21, pp. 32–35.
9. PRISHLYAK, A. O., PRISHLYAK, K. A. and VYATCHANINOVA, O. N. (2010), “Topological properties of functions on 3-dimensional bodies”, *J. Comput. and Applied Math*, No. 2 (101), pp. 113–119.
10. PRISHLYAK, A. O., PRISHLYAK, K. A., MISCHENKO, K. I. and LUKOVA, N. V. (2011), “Classification of simple m -function on oriented surfaces”, *J. Comput. and Applied Math*, No. 1 (104), pp. 116–127.
11. PRISHLYAK, A. O., PRISHLYAK, K. A. and LUKOVA-CHUIKO, N. V. (2012), “ m -functions on non-oriented surfaces”, *J. Comput. and Applied Math*, No. 2 (108), pp. 176–185.
12. IVANYUK, O. M. and PRISHLYAK, A. O. (2014), “Molecules of m -functions degree 2 on surfaces with boundary”, *Proc. Intern. Geom. Center*, v. 7, No. 3, pp. 27–37.

УДК 512.53

Є.А. Кочубінська¹, к.ф.-м.н., асистент
І.В. Семенюк², студент

Про узагальнений частковий вінцевий добуток напівгруп

¹Київський національний університет імені Тараса Шевченка, 01033, Київ, вул. Володимирська, 64.

²Київський національний університет імені Тараса Шевченка, 01033, Київ, вул. Володимирська, 64.

e-mail: ¹eugenia@univ.kiev.ua

E.A. Kochubinska¹, Ph.D., Assistant Professor
I.V. Semeniuk², Student

On a generalized partial wreath product of semigroups

¹Taras Shevchenko National University of Kyiv, 01033, Kyiv, 64 Volodymyrska st.

²Taras Shevchenko National University of Kyiv, 01033, Kyiv, 64 Volodymyrska st.,

e-mail: ¹eugenia@univ.kiev.ua

У статті запропоноване узагальнення часткового вінцевого добутку. Доведено, що узагальнений частковий вінцевий добуток напівгруп є напівгрупою. Показано, що узагальнений частковий вінцевий добуток інверсних напівгруп є інверсною напівгрупою. Знайдено вигляд ідемпотентів узагальненого часткового вінцевого добутку напівгруп.

Ключові слова: напівгрупа, частковий вінцевий добуток, узагальнений частковий вінцевий добуток, ідемпотент.

The construction of wreath product is well-known and widely used both for groups and semigroups. It gives a possibility to get new group (semigroups) from old. The semigroup of partial automorphisms of a rooted tree of a special kind is isomorphic to partial wreath product of inverse symmetric semigroups. However, sometimes it is necessary to consider a construction similar to wreath product, but which allows to get wreath product of a semigroup (group) with a set of semigroups (groups). It is shown In the present paper a generalization of partial wreath product of semigroups is introduced. It is proved that the generalized partial wreath product of semigroups is a semigroup. Besides, it is shown that the generalized wreath product of inverse semigroups is an inverse semigroup. It is also shown that under certain conditions the generalized wreath product of semigroup is isomorphic to the direct product of wreath products of semigroups. The description of idempotents of a generalized partial wreath product is provided.

Key Words: semigroup, partial wreath product, generalized partial wreath product, idempotent.

Communicated by Prof. Kirichenko V.V.

1 Introduction

The wreath product construction is a useful tool to obtain new groups or semigroups from old. However, it is interesting object of study. Depending on type of semigroups, slightly different modification of wreath product are considered. The most comprehensive survey is contained in [3]. In certain cases it is convenient to have a possibility construct a wreath product of a group (semigroup) with a set of groups (semigroups), for example, when describing closed subsemigroups of a semigroup of partial automorphisms of a rooted tree [1]. The generalized wreath product introduced in this paper provides such a possibility. We consider a generalized wreath product of semigroups and study its basic properties.

2 Partial wreath product of semigroups

We follow notations from [3].

Let P be a semigroup and (H, M) be a semigroup of partial transformations of a finite set M . Consider the set S^{PX} of partial functions from M to P :

$$S^{PX} = \{f: A \rightarrow P \mid \text{dom}(f) = A, A \subseteq M\}.$$

If $f, g \in S^{PX}$, then product fg is defined in a following way

$$\begin{aligned} \text{dom}(fg) &= \text{dom}(f) \cap \text{dom}(g), \\ (fg)(x) &= f(x)g(x) \text{ for each } x \in \text{dom}(fg). \end{aligned}$$

If $h \in H, f \in S^{PX}$, then f^h is defined as follows:

$$\begin{aligned} \text{dom}(f^h) &= \{x \in \text{dom}(h) : x^h \in \text{dom}(f)\} = \\ &= (\text{ran}(h) \cap \text{dom}(f))h^{-1}, (f^h)(x) = f(x^h). \end{aligned}$$

Definition 2.1. Partial wreath product of a semigroup P with a semigroup of partial transformations (H, M) of a finite set M is a set

$$\{(f, h) \in S^{PX} \times (H, M) \mid \text{dom}(f) = \text{dom}(h)\}$$

with multiplication defined by

$$(f, a)(g, b) = (fg^a, ab).$$

Partial wreath product of semigroups P and (H, M) is denoted by $P \wr_p H$.

Proposition 1. [2]

1. Partial wreath product of semigroup $P \wr_p H$ is a semigroup.
2. If semigroups P and (H, M) are inverse then the semigroup $P \wr_p H$ is inverse. The inverse element to (f, h) is the element $(f, h)^{-1} = ((f^{h^{-1}})^{-1}, h^{-1})$.

3 Generalized partial wreath product of semigroups

Let H be a semigroup which acts (possibly, partially) on a set M , $\{P_x\}_{x \in M}$ be any set of semigroups satisfying condition:

$$\text{if } y \in x^H \text{ (that is, } y = x^h \text{ for certain } h \in H), \\ \text{then } P_y \subset P_x.$$

Definition 3.1. Generalized partial wreath product of a group (semigroup) (H, M) with a set of semigroups $\{P_x\}_{x \in M}$ is a set

$$\begin{aligned} GWr(H, M, \{P_x\}) &= \\ &= \{(f, h) \mid h \in H, f: M \rightarrow \bigcup_x P_x, \\ &f(x) \in P_x, \text{dom}(f) = \text{dom}(h)\} \end{aligned}$$

with multiplication defined by

$$(f_1, h_1)(f_2, h_2) = (f_1 f_2^{h_1}, h_1 h_2), \quad (1)$$

where $(f_1, h_1), (f_2, h_2) \in GWr(H, M, \{P_x\})$.

We define the product of functions f_1 and f_2 in the same way as for usual wreath product:

$$(f_1 f_2)(x) = f_1(x) f_2(x),$$

and we also define

$$f^h(x) = f(x^h).$$

Theorem 3.1. The generalized partial wreath product $GWr(H, M, \{P_x\})$ of semigroup (H, M) with the set of semigroups $\{P_x\}_{x \in M}$ forms a semigroup under multiplication defined by (1).

Proof. We show that the set $GWr(H, M, \{P_x\})$ is closed under multiplication.

Let $(f_1, h_1), (f_2, h_2) \in GWr(H, M, \{P_x\})$. Note that

$$\text{dom}(f_1) = \text{dom}(h_1)$$

and

$$\text{dom}(f_2) = \text{dom}(h_2).$$

Then

$$(f_1, h_1)(f_2, h_2) = (f_1 f_2^{h_1}, h_1 h_2).$$

Clearly, $h_1 h_2 \in H$.

Since $P_{x^{h_1}} \subset P_x$, then $f_2(x^{h_1}) \in P_x$ and

$$f_1 f_2^{h_1}(x) = f_1(x) f_2(x^{h_1}) \in P_x.$$

And we need to check that equality

$$\text{dom}(f_1 f_2^{h_1}) = \text{dom}(h_1 h_2)$$

holds. We have

$$\text{dom}(h_1 h_2) = (\text{ran}(h_1) \cap \text{dom}(h_2))h_1^{-1},$$

$$\text{dom}(f_2^{h_1}) = (\text{ran}(h_1) \cap \text{dom}(f_2))h_1^{-1}.$$

Since

$$\text{dom}(f_1 f_2^{h_1}) = \text{dom}(f_1) \cap \text{dom}(f_2^{h_1})$$

and

$$\text{dom}(f_1) = \text{dom}(h_1), \quad \text{dom}(f_2) = \text{dom}(h_2),$$

then

$$\begin{aligned} \text{dom}(f_1 f_2^{h_1}) &= \\ &= \text{dom}(h_1) \cap (\text{ran}(h_1) \cap \text{dom}(f_2))h_1^{-1} = \\ &= (\text{ran}(h_1) \cap \text{dom}(h_2))h_1^{-1} = \text{dom}(f_2^{h_1}). \end{aligned}$$

To complete the proof, we note that every set Ah_1^{-1} is contained in $\text{dom}(h_1)$.

Thus, $(f_1, h_1)(f_2, h_2) \in GWr(H, M, \{P_x\})$, and generalize wreath product of semigroups is a semigroup. □

Consider now the semigroup $GW_r(H, M, \{P_x\})$. On the set M we introduce a relation $\sim$:

$$x \sim y \stackrel{\text{def}}{\iff} x \in y^H,$$

that is, there exists $h \in H$ such that $x = y^h$.

Lemma 1. *If H is an inverse group, then relation $\sim$ is the equivalence relation on the set M .*

Proof. Clearly, $x \in x^H$.

Since H is an inverse semigroup, then $x = y^h$ implies $y = x^{h^{-1}}$. Thus, $x \in y^H$ implies $y \in x^H$.

Let $x \sim y$ and $y \sim z$. Then there exists $h_1, h_2 \in H$ such that $x = y^{h_1}$, $y = z^{h_2}$. It follows $x = y^{h_1} = z^{h_2 h_1}$, and $x \sim z$. $\square$

Lemma 2. *For equivalence relation $\sim$ the following holds: if $x \sim y$, then $P_x = P_y$.*

Proof. It follows from the definition of the generalized wreath product that $x \in y^H$ implies $P_x \subset P_y$. Since $\sim$ is symmetric, then $P_y \subset P_x$. $\square$

Theorem 3.2. *If semigroup H and all the semigroups of the set $\{P_x\}_{x \in M}$ are inverse, then semigroup $GW_r(H, M, \{P_x\})$ is inverse.*

Proof. Let H and $\{P_x\}_{x \in M}$ be inverse semigroups. If $(f, h) \in GW_r(H, M, \{P_x\})$, then $h^{-1} \in H$ and $(f^{h^{-1}})^{-1} \in P_x$.

Thus, $((f^{h^{-1}})^{-1}, h^{-1}) \in GW_r(H, M, \{P_x\})$.

The element $((f^{h^{-1}})^{-1}, h^{-1})$ is inverse for (f, h) . Indeed,

$$\begin{aligned} (f, h)(f, h)^{-1}(f, h) &= (f, h)((f^{h^{-1}})^{-1}, h^{-1}) = \\ &= (f, h) = (f((f^{h^{-1}})^{-1})^h, hh^{-1})(f, h) = \\ &= (f((f^{h^{-1}})^{-1})^h f^{hh^{-1}}, hh^{-1}h) = (f, h). \end{aligned}$$

Similarly,

$$(f, h)^{-1}(f, h)(f, h)^{-1} = ((f^{h^{-1}})^{-1}, h^{-1}).$$

As H and $\{P_x\}_{x \in M}$ are inverse, the uniqueness of the element $(f, h)^{-1}$ follows from the uniqueness of the elements h^{-1} and $(f^{h^{-1}})^{-1}$. $\square$

Since $\sim$ is equivalence relation, the set M is partitioned into disjoint union of equivalence classes:

$$M = M_1 \cup M_2 \cup \dots \cup M_k.$$

Theorem 3.3. *If the semigroup H is inverse, then generalized partial wreath product is isomorphic to the direct product of partial wreath products:*

$$GW_r(H, M, \{P_{x_i}\}_{i=1}^k) \simeq \prod_{i=1}^k (P_{x_i} \wr_p (H, M_i))$$

where $x_i \in M_i$, $i = 1, 2, \dots, k$.

Proof. Let $(f, h) \in GW_r(H, M, \{P_{x_i}\}_{i=1}^k)$. Consider the map

$$\varphi : GW_r(H, M, \{P_{x_i}\}_{i=1}^k) \rightarrow \prod_{i=1}^k (P_{x_i} \wr_p (H, M_i)),$$

which acts as

$$\begin{aligned} (f, h) &\mapsto \\ &\mapsto ((f|_{M_1}, h|_{M_1}), (f|_{M_2}, h|_{M_2}), \dots, (f|_{M_k}, h|_{M_k})). \end{aligned}$$

First, we check that for all $i = 1, 2, \dots, k$ $(f|_{M_i}, h|_{M_i})$ is in $P_{x_i} \wr_p (H, M_i)$.

Let $x \in M_i$ for some $i = 1, \dots, k$. Since M_i is an equivalence class, then for every h from H we have $x^h|_{M_i} \in M_i$. Therefore, $h|_{M_i} \in (H, M_i)$.

From the definition of a generalized wreath product it follows that $f(x) \in P_{x_i}$ for any $x \in M_i$. Denote $\text{dom}(f_i) = B_i$. Clearly, $B_i \subseteq M_i$. From above we have

$$f_i : B_i \rightarrow P_{x_i}$$

is a function from $B_i \subset M_i$ to P_{x_i} as required. Therefore, $(f|_{M_i}, h|_{M_i}) \in P_{x_i} \wr_p (H, M_i)$.

Show that the map given above is a homomorphism. On one hand, we have

$$\begin{aligned} (fg^h, hb) &\mapsto \\ &\mapsto (((fg)^h|_{M_1}, hb|_{M_1}), \dots, ((fg)^h|_{M_k}, hb|_{M_k})). \end{aligned}$$

On the other hand,

$$\begin{aligned} &((f|_{M_1}, h|_{M_1}), \dots, (f|_{M_k}, h|_{M_k})) \times \\ &\times ((g|_{M_1}, b|_{M_1}), \dots, (g|_{M_k}, b|_{M_k})) = \\ &= (((fg)^h|_{M_1}), hb|_{M_1}), \dots, \\ &((fg)^h|_{M_k}), hb|_{M_k})) = \\ &= (((fg)^h|_{M_1}, hb|_{M_1}), \dots, ((fg)^h|_{M_k}, hb|_{M_k})). \end{aligned}$$

It is left to prove that this map is bijective. The injectivity follows from the definition of equivalence classes and from uniqueness of inverse element. Clearly, it is also surjective. So, given map is indeed an isomorphism. $\square$

4 Idempotents in generalized partial wreath product of semigroups

Theorem 4.1. *Element $(f, h) \in GWr(H, M, \{P_x\})$ is idempotent if and only if $h \in H$ is idempotent and f satisfies the condition: $f(x)f(x^h) = f(x)$ for each $x \in M$.*

Proof. An element $(f, h) \in GWr(H, M, \{P_x\})$ is an idempotent if and only if $(f, h)(f, h) = (f, h)$.

According to multiplication rule in the semigroup $GWr(H, M, \{P_x\})$ it means $(f, h)(f, h) = (ff^h, hh) = (f, h)$. Therefore, h is idempotent in H and $f(x)f(x^h) = f(x)$. $\square$

Theorem 4.2. *Let $(f, h) \in GWr(H, M, \{P_x\})$ be idempotent, then $f(y) \in P_x$ is an idempotent, where $y = x^h$ for some $x \in M$.*

Proof. Let $h \in H$ be idempotent and $x, y \in M$ such that $x^h = y$. Then $y = x^h = x^{h^2} = (x^h)^h = y^h$ and $f(y) = f(y^h) = f^h(y)$. Therefore from the previous theorem we have $f(x)f^h(x) = f(x)$ and

$$f(y)f^h(y) = f(y)f(y) = f(y),$$

so $f(y) \in P_x$ is an idempotent. $\square$

Список використаних джерел

1. Кочубінська Є.А Підстановкові зображення напівгрупи часткових автоморфізмів кореневого дерева/ Є.А. Кочубінська // Вісник Київського університету. Математика. Механіка. — 2007 — №17–18. — С. 33–37.
2. Kochubinska Ye. Combinatorics of partial wreath power of finite inverse symmetric semigroup $\mathcal{IS}_d$ / Є.А.Кочубінська // Algebra and Discrete Mathematics — 2007 — №1. — С. 49–61.
3. Meldrum J.D.P. Wreath product of groups and semigroups/ J.D.P. Meldrum. — Pitman Monographs and Surveys in Pure and Applied Mathematics, 74. Harlow: Longman Group Ltd — 1995.— xii, 324 c.

References

1. KOCHUBINSKA E.A. (2007), "Pidstanovkovi zobrazhennya napivgrupy chastkovykh avtomorfizmiv koreneвого dereva", *Visnyk Kyivskogo universytetu. Matematika. Mechanika.*, №17–18, pp. 33–37.
2. KOCHUBINSKA YE. (2007), "Combinatorics of partial wreath power of finite inverse symmetric semigroup $\mathcal{IS}_d$ ", *Algebra and Discrete Mathematics*, №1, pp. 49–61.
3. MELDRUM J.D.P., (1995) *Wreath product of groups and semigroups*. Pitman Monographs and Surveys in Pure and Applied Mathematics, 74. Harlow: Longman Group Ltd, 324 p.

Received: 14.12.2014

УДК 519.21

А.П. Петравчук¹, д.ф.-м.н., професор
В.В. Степух², аспірант

Базиси в алгебрах Лі диференціювань

¹Київський національний університет імені Тараса Шевченка, 01033, Київ, вул. Володимирська, 64.

²Київський національний університет імені Тараса Шевченка, 01033, Київ, вул. Володимирська, 64.

A.P. Petravchuk¹, doctor of science, professor
V.V. Stepukh², Ph.D student

On bases of Lie algebras of derivations

¹Taras Shevchenko National University of Kyiv, 01033, Kyiv, 64 Volodymyrska st., e-mail: ¹apetrav@gmail.com

²Taras Shevchenko National University of Kyiv, 01033, Kyiv, 64 Volodymyrska st., e-mail: ²svvhelios@gmail.com

Нехай K алгебраїчно замкнене поле характеристики 0 і R поле алгебраїчних функцій від двох змінних над K . Алгебра Лі всіх K -диференціювань поля R є двовимірним векторним простором над R (з стандартним базисом із частинних похідних). У роботі дано характеристизацію комутативних базисів цієї алгебри Лі і описано метод побудови деяких її комутативних базисів. Знайдено закон зміни дивергенції довільного диференціювання поля R при зміні базису простору диференціювань.

Ключові слова: многочлен, алгебра Лі, диференціювання, комутативний базис, дивергенція

Let K be an algebraically closed field of characteristic zero and R the field of algebraic functions in two variables over K containing the field $K(x, y)$. The set $\text{Der}(R)$ of all K -derivations of the field R is a Lie algebra. This Lie algebra also is a two-dimensional vector space over the field R (with the standard basis consisting of partial derivations). A basis of this vector space over R is called commutative if its elements commute in the Lie algebra $\text{Der}(R)$. A characterization of all commutative bases of $\text{Der}(R)$ is given in terms of divergence of their elements (the derivations from a basis of $\text{Der}(R)$ can have nonzero divergence in contrary to the Lie algebra of all derivations of the polynomial ring in two variables over K). A method of building some commutative bases of the vector space $\text{Der}(R)$ is pointed out, a formula for computing the divergence of a derivation from $\text{Der}(R)$ in any commutative basis is obtained. It turned out that the divergence of a derivation of the field R does not depend on choice of new basis consisting of coordinate polynomials.

Key Words: polynomial, Lie algebra, derivation, commutative basis, divergence.

Статтю представив доктор фізико-математичних наук, професор В.В.Кириченко

1 Introduction

Let $\mathbb{K}$ be an algebraically closed field of characteristic zero and $R \supseteq K(x_1, \dots, x_n)$ a field of algebraic functions in n variables, i.e. an algebraic extension of the field $K(x_1, \dots, x_n)$. The set of all $\mathbb{K}$ -derivations of the field R is a Lie algebra (with the Lie bracket $[D_1, D_2] = D_1D_2 - D_2D_1$) which will be denoted by $\widetilde{W}_n(R)$, this Lie algebra is also a vector space of dimension n over R (but not a Lie algebra over R). The basis $\{\frac{\partial}{\partial x_1}, \dots, \frac{\partial}{\partial x_n}\}$ is called standard, its elements satisfy obviously the relations $[\frac{\partial}{\partial x_i}, \frac{\partial}{\partial x_j}] = 0, i, j = 1, \dots, n$, i.e. this basis is commutative. All the commutative bases of the Lie algebra $W_n(\mathbb{K})$ (consisting of all $\mathbb{K}$ -derivations of the polynomial ring $\mathbb{K}[x_1, \dots, x_n]$) are characterized in [4] (see also, [5], [3]). In particular, the deri-

vations from such a basis are divergence-free [5] (see, for example [1], [9]). This characterization of commutative bases can not be transferred on the Lie algebra $\widetilde{W}_n(R)$ where $R = \mathbb{K}(x_1, \dots, x_n)$ since there are commutative bases of $\widetilde{W}_n(R)$ such that not all their elements have zero divergence. So, it is interesting to characterize commutative bases of $\widetilde{W}_n(R)$ in terms of divergence of their members.

Commutative bases of the Lie algebra $\widetilde{W}_2(R)$ are characterized (Theorem 1). A method of building some commutative bases in $\widetilde{W}_n(R)$ is pointed out, the law of changing of divergence when we change a basis of the field $\mathbb{K}(x, y)$ is found (Theorem 2).

We use standard notations in the paper (see, for example, [6]). Recall that every derivation of the field $\mathbb{K}(x_1, \dots, x_n)$ can be uniquely extended

to the field $R \supseteq \mathbb{K}(x_1, \dots, x_n)$ of algebraic functions in n variables (see, for example [1]). Therefore the standard derivations $\frac{\partial}{\partial x_i}$ of the subfield $\mathbb{K}(x_1, \dots, x_n)$ of the field R can be uniquely extended to the derivations of R and these derivations form a basis of $\widetilde{W}_n(R)$ (we save the same notations for such extended derivations). If D is a derivation of the field R written in the form $D = \sum_{i=1}^n a_i \frac{\partial}{\partial x_i}$, $a_i \in R$, then the divergence of D is $\text{div} D = \sum_{i=1}^n \frac{\partial a_i}{\partial x_i}$. Since D is a linear operator on the $\mathbb{K}$ -space R , we will write $\text{Ker} D$ for its kernel, this is subfield of R called field of constants of D in R .

2 Commutative bases of Lie algebras of derivations

Lemma 1. Let $D_1, D_2 \in \widetilde{W}_n(R)$ and $a, b \in R$. Then

- 1) $[aD_1, bD_2] = ab[D_1, D_2] + aD_1(b)D_2 - bD_2(a)D_1$;
- 2) if $a, b \in \text{Ker} D_1 \cap \text{Ker} D_2$, then $[aD_1, bD_2] = ab[D_1, D_2]$

Доведення. Straightforward check. $\square$

Lemma 2. Let $D_1 = a_{11} \frac{\partial}{\partial x_1} + a_{12} \frac{\partial}{\partial x_2}$, $D_2 = a_{21} \frac{\partial}{\partial x_1} + a_{22} \frac{\partial}{\partial x_2} \in \widetilde{W}_2(R)$ be nonzero derivations of the field $R \supseteq \mathbb{K}(x_1, x_2)$ such that $\Delta = \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} = 0$. The equality $[D_1, D_2] = 0$ holds if and only if $D_2 = aD_1$ for some $a \in \text{Ker} D_1$

Доведення. Since D_1, D_2 are linearly dependent over R , we have $D_2 = aD_1$ for some $a \in R$. The condition $[D_1, D_2] = 0$ is equivalent by Lemma 1 to the equality $D_1(a) = 0$, i.e. $a \in \text{Ker} D_1$. $\square$

Theorem 2.1. Let $D_1 = a_{11} \frac{\partial}{\partial x_1} + a_{12} \frac{\partial}{\partial x_2}$, $D_2 = a_{21} \frac{\partial}{\partial x_1} + a_{22} \frac{\partial}{\partial x_2}$ be derivations of a field $R \supseteq \mathbb{K}(x_1, x_2)$ such that $\Delta = \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} \neq 0$. The equality $[D_1, D_2] = aD_1 + bD_2$, for some $a, b \in R$ holds if and only if $\text{div} D_1 = \frac{D_1(\Delta)}{\Delta} - b$, $\text{div} D_2 = \frac{D_2(\Delta)}{\Delta} - a$.

Доведення. Denote by $B = (b_{ij})_{i,j=1}^2$ the matrix A^{-1} , where $A = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}$

Then

$$\frac{\partial}{\partial x_1} = b_{11}D_1 + b_{12}D_2, \quad \frac{\partial}{\partial x_2} = b_{21}D_1 + b_{22}D_2.$$

Using the approach from the paper [?] write down

$$\begin{aligned} 0 &= \frac{\partial}{\partial x_1} \frac{\partial}{\partial x_2} - \frac{\partial}{\partial x_2} \frac{\partial}{\partial x_1} = \frac{\partial}{\partial x_1} (b_{21}D_1 + b_{22}D_2) \\ &\quad - \frac{\partial}{\partial x_2} (b_{11}D_1 + b_{12}D_2) = \frac{\partial b_{21}}{\partial x_1} D_1 + b_{21} \frac{\partial}{\partial x_1} D_1 \\ &\quad + \frac{\partial b_{22}}{\partial x_1} D_2 + b_{22} \frac{\partial}{\partial x_1} D_2 - \left(\frac{\partial b_{11}}{\partial x_2} D_1 + b_{11} \frac{\partial}{\partial x_2} D_1 \right. \\ &\quad \left. + \frac{\partial b_{12}}{\partial x_2} D_2 + b_{12} \frac{\partial}{\partial x_2} D_2 \right). \end{aligned}$$

It follows from the last relation (after substitution $\frac{\partial}{\partial x_i} = b_{i1}D_1 + b_{i2}D_2$):

$$\begin{aligned} &\left(\frac{\partial b_{21}}{\partial x_1} - \frac{\partial b_{11}}{\partial x_2} \right) D_1 + \left(\frac{\partial b_{22}}{\partial x_1} - \frac{\partial b_{12}}{\partial x_2} \right) D_2 = \\ &b_{11}(b_{21}D_1 + b_{22}D_2)D_1 + b_{12}(b_{21}D_1 + b_{22}D_2)D_2 - \\ &b_{21}(b_{11}D_1 + b_{12}D_2)D_1 - b_{22}(b_{11}D_1 + b_{12}D_2)D_2. \end{aligned}$$

Therefore the next equalities hold:

$$\begin{aligned} &\left(\frac{\partial b_{21}}{\partial x_1} - \frac{\partial b_{11}}{\partial x_2} \right) D_1 + \left(\frac{\partial b_{22}}{\partial x_1} - \frac{\partial b_{12}}{\partial x_2} \right) D_2 = \\ &(b_{11}b_{22} - b_{21}b_{12})[D_2, D_1] = \det B[D_2, D_1] \quad (1) \end{aligned}$$

Taking into account the relations $\det B = \frac{1}{\Delta}$, $[D_2, D_1] = -aD_1 - bD_2$ we get from (1) that

$$\frac{\partial b_{12}}{\partial x_1} - \frac{\partial b_{21}}{\partial x_2} = -\frac{1}{\Delta}a, \quad \frac{\partial b_{22}}{\partial x_1} - \frac{\partial b_{12}}{\partial x_2} = -\frac{1}{\Delta}b$$

Since $B = A^{-1} = \frac{1}{\Delta} \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}$ we have

$$\begin{aligned} &\frac{\partial}{\partial x_1} \left(\frac{a_{21}}{\Delta} \right) + \frac{\partial}{\partial x_2} \left(\frac{a_{22}}{\Delta} \right) = \\ &-\frac{a}{\Delta}, \quad \frac{\partial}{\partial x_1} \left(\frac{a_{11}}{\Delta} \right) + \frac{\partial}{\partial x_2} \left(\frac{a_{12}}{\Delta} \right) = -\frac{b}{\Delta} \quad (2) \end{aligned}$$

The equation (2) implies

$$\text{div} D_1 = \frac{D_1(\Delta)}{\Delta} - b, \quad \text{div} D_2 = \frac{D_2(\Delta)}{\Delta} - a \quad (3)$$

Conversely, let the equalities (3) hold for some elements $a, b \in R$. As D_1, D_2 are linearly independent over R we get $[D_1, D_2] = \alpha D_1 + \beta D_2$ for some $\alpha, \beta \in R$. Repeating the considerations from the first part we get $\text{div} D_1 = \frac{D_1(\Delta)}{\Delta} - \beta$, $\text{div} D_2 = \frac{D_2(\Delta)}{\Delta} - \alpha$. But then $\alpha = a, \beta = b$. $\square$

Corollary 1. Let $D_1 = a_{11} \frac{\partial}{\partial x_1} + a_{12} \frac{\partial}{\partial x_2}$, $D_2 = a_{21} \frac{\partial}{\partial x_1} + a_{22} \frac{\partial}{\partial x_2} \in \widetilde{W}_2(\mathbb{K})$ be linearly independent over R . The derivations D_1 and D_2 commute if and only if $\text{div}(\frac{1}{\Delta} D_1) = 0$, $\text{div}(\frac{1}{\Delta} D_2) = 0$, where $\Delta = \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix}$.

While studying the Lie algebra $\widetilde{W}_2(R)$ a question arises about derivations $D_1, D_2 \in \widetilde{W}_2(R)$ such that $[D_1, D_2] = D_2$ (they generate in $\widetilde{W}_2(R)$ a two-dimensional over $\mathbb{K}$ nonabelian subalgebra). Such Lie subalgebras can also be characterized in terms of divergence:

Corollary 2. Let $D_1, D_2 \in \widetilde{W}_2(R)$ be linearly independent over R . The relation $[D_1, D_2] = D_2$ holds if and only if $\text{div}(\frac{1}{\Delta} D_1) = -\frac{1}{\Delta}$, $\text{div}(\frac{1}{\Delta} D_2) = 0$, in particular, if $\Delta = 1$ then $\text{div} D_2 = 0$, $\text{div} D_1 = -1$.

Note that a question was posed in [8] about structure of polynomials f, g satisfying the equality $\det J(f, g) = g$, i.e. about structure of Jacobian derivations satisfying the mentioned above conditions. Such polynomials in some cases were described in [2], see also [7].

Let us show a simple method how to build commutative bases in the Lie algebra $\widetilde{W}_n(R)$. Let $\{y_1, \dots, y_n\}$ be a transcendence basis of the field R over $\mathbb{K}$. Denote by $\frac{\partial}{\partial y_1}, \dots, \frac{\partial}{\partial y_n}$ the usual partial differentiations of the subfield $\mathbb{K}(y_1, \dots, y_n)$ of R and extend them (uniquely by Lemma 2) to derivations of the field R (we save the same notations). Since $[\frac{\partial}{\partial y_i}, \frac{\partial}{\partial y_j}] = 0$, $i, j = 1, \dots, n$ on the subfield $\mathbb{K}(y_1, \dots, y_n)$ of R and R is algebraic over $\mathbb{K}(y_1, \dots, y_n)$ then the commutativity preserves on the field R (by Lemma 2). Hence $\{\frac{\partial}{\partial y_1}, \dots, \frac{\partial}{\partial y_n}\}$ is a commutative basis for $\widetilde{W}_n(R)$. Write this basis in an explicit form. Denote $\Delta = \det J(y_1, \dots, y_n)$, the Jacobian of the algebraic functions $y_1, \dots, y_n$.

Consider the derivations $D_{y_1}, \dots, D_{y_n}$ of the field R defined by the rule:

$D_{y_i}(h) = \det J(y_1, \dots, y_{i-1}, h, y_{i+1}, \dots, y_n)$, where

$$J(y_1, \dots, y_{i-1}, h, y_{i+1}, \dots, y_n) = \begin{pmatrix} \frac{\partial y_1}{\partial x_1} & \dots & \frac{\partial y_1}{\partial x_n} \\ \dots & \dots & \dots \\ \frac{\partial h}{\partial x_1} & \dots & \frac{\partial h}{\partial x_n} \\ \dots & \dots & \dots \\ \frac{\partial y_n}{\partial x_1} & \dots & \frac{\partial y_n}{\partial x_n} \end{pmatrix}$$

is the Jacobi matrix of the functions $y_1, \dots, y_n$. Then $\frac{1}{\Delta} D_{y_i}(y_j) = \delta_{ij}$ and therefore $\frac{1}{\Delta} D_{y_i} = \frac{\partial}{\partial y_i}$

on the subfield $\mathbb{K}(y_1, \dots, y_n)$. But then $\frac{\partial}{\partial y_i} = \frac{1}{\Delta} D_{y_i}$ on the whole field R by Lemma 2. So $\{\frac{1}{\Delta} D_{y_i}\}$ is a commutative basis of the Lie algebra $\widetilde{W}_n(R)$. Note that all the derivations $\frac{1}{\Delta} D_{y_i}$ have zero divergence of and only if $\Delta \in \mathbb{K}^*$.

3 On divergence in different bases

Let $R \supseteq \mathbb{K}(x, y)$ be a field of algebraic functions in two variables. Then every derivation $D \in \widetilde{W}_2(R)$ has a unique expression $D = a_1 \frac{\partial}{\partial x} + a_2 \frac{\partial}{\partial y}$ for some $a_1, a_2 \in R$ and $\text{div} D = \frac{\partial a_1}{\partial x} + \frac{\partial a_2}{\partial y}$. If we take any elements f and $g \in R$ which algebraically independent over $\mathbb{K}$ then $D = b_1 \frac{\partial}{\partial f} + b_2 \frac{\partial}{\partial g}$ for some $b_1, b_2 \in R$ and a question arises how to compute $\text{div} D$ in the basis $\{\frac{\partial}{\partial f}, \frac{\partial}{\partial g}\}$ of the Lie algebra $\widetilde{W}_2(R)$ (we can consider $\text{div} D$ as a geometric characteristic of D). The answer is given by the next statement:

Theorem 3.1. Let $\{\frac{\partial}{\partial f}, \frac{\partial}{\partial g}\}$ be a commutative basis of the Lie algebra $\widetilde{W}_2(R) = \text{Der}_{\mathbb{K}} R$, where $R \supseteq \mathbb{K}(x, y)$ is a field of algebraic functions. If $D = P \frac{\partial}{\partial x} + Q \frac{\partial}{\partial y} \in \widetilde{W}_2(R)$ has an expression $D = R \frac{\partial}{\partial f} + S \frac{\partial}{\partial g}$ in the basis $\{\frac{\partial}{\partial f}, \frac{\partial}{\partial g}\}$ then $\text{div} D = \frac{\partial P}{\partial x} + \frac{\partial Q}{\partial y} = \frac{\partial R}{\partial f} + \frac{\partial S}{\partial g} - D(\Delta) \frac{1}{\Delta}$, where $\Delta = \det J(f, g)$ is the jacobian of f, g .

Доведення. Let us find a relation between coordinates (P, Q) and (R, S) of D . Note that

$$R = D(f) = P \frac{\partial f}{\partial x} + Q \frac{\partial f}{\partial y},$$

$$S = D(g) = P \frac{\partial g}{\partial x} + Q \frac{\partial g}{\partial y}. \quad (4)$$

Therefore $\begin{pmatrix} R \\ S \end{pmatrix} = \begin{pmatrix} \frac{\partial f}{\partial x} & \frac{\partial f}{\partial y} \\ \frac{\partial g}{\partial x} & \frac{\partial g}{\partial y} \end{pmatrix} \begin{pmatrix} P \\ Q \end{pmatrix}$, where $\begin{pmatrix} \frac{\partial f}{\partial x} & \frac{\partial f}{\partial y} \\ \frac{\partial g}{\partial x} & \frac{\partial g}{\partial y} \end{pmatrix} = J(f, g)$ is the Jacobi matrix of the algebraic functions f, g . Using the chain rule for differentiation we obtain

$$\frac{\partial R}{\partial f} = \frac{\partial R}{\partial x} \frac{\partial x}{\partial f} + \frac{\partial R}{\partial y} \frac{\partial y}{\partial f}, \quad \frac{\partial S}{\partial g} = \frac{\partial S}{\partial x} \frac{\partial x}{\partial g} + \frac{\partial S}{\partial y} \frac{\partial y}{\partial g}.$$

Since f, g are algebraically independent over $\mathbb{K}$ we have

$$F(x, f, g) = 0, \quad G(y, f, g) = 0 \quad (5)$$

for some nonzero polynomials $F(x, f, g), G(y, f, g) \in \mathbb{K}[u, v, w]$.

It is also obvious that $\frac{\partial F}{\partial x} \neq 0, \frac{\partial G}{\partial y} \neq 0$. From the equalities (5) it follows

$$\frac{\partial F}{\partial x} \frac{\partial x}{\partial f} + \frac{\partial F}{\partial f} = 0, \quad \frac{\partial F}{\partial x} \frac{\partial x}{\partial g} + \frac{\partial F}{\partial g} = 0$$

$$\frac{\partial G}{\partial y} \frac{\partial y}{\partial f} + \frac{\partial G}{\partial f} = 0, \quad \frac{\partial G}{\partial y} \frac{\partial y}{\partial g} + \frac{\partial G}{\partial g} = 0$$

and then

$$\frac{\partial x}{\partial f} = -\frac{\frac{\partial F}{\partial x}}{\frac{\partial F}{\partial f}}, \quad \frac{\partial x}{\partial g} = -\frac{\frac{\partial F}{\partial x}}{\frac{\partial F}{\partial g}} \quad (6)$$

$$\frac{\partial y}{\partial f} = -\frac{\frac{\partial G}{\partial y}}{\frac{\partial G}{\partial f}}, \quad \frac{\partial y}{\partial g} = -\frac{\frac{\partial G}{\partial y}}{\frac{\partial G}{\partial g}} \quad (7)$$

After differentiation the first equation in (5) by x and then by y we get

$$\begin{cases} \frac{\partial F}{\partial x} + \frac{\partial F}{\partial f} \frac{\partial f}{\partial x} + \frac{\partial F}{\partial g} \frac{\partial g}{\partial x} = 0 \\ \frac{\partial F}{\partial f} \frac{\partial f}{\partial y} + \frac{\partial F}{\partial g} \frac{\partial g}{\partial y} = 0 \end{cases} \quad (8)$$

Solving this system of linear equations we obtain

$$\frac{\partial F}{\partial f} = -\frac{\partial F}{\partial x} \frac{\partial g}{\partial y} \frac{1}{\Delta}, \quad \frac{\partial F}{\partial g} = \frac{\partial F}{\partial x} \frac{\partial f}{\partial y} \frac{1}{\Delta},$$

where $\Delta = \begin{vmatrix} \frac{\partial f}{\partial x} & \frac{\partial f}{\partial y} \\ \frac{\partial g}{\partial x} & \frac{\partial g}{\partial y} \end{vmatrix}$. Analogously it follows

from the second equality in (5) that

$$\begin{cases} \frac{\partial G}{\partial y} + \frac{\partial G}{\partial f} \frac{\partial f}{\partial y} + \frac{\partial G}{\partial g} \frac{\partial g}{\partial y} = 0 \\ \frac{\partial G}{\partial f} \frac{\partial f}{\partial x} + \frac{\partial G}{\partial g} \frac{\partial g}{\partial x} = 0 \end{cases} \quad (9)$$

that

$$\frac{\partial G}{\partial f} = \frac{\partial G}{\partial y} \frac{\partial g}{\partial x} \frac{1}{\Delta}, \quad \frac{\partial G}{\partial g} = -\frac{\partial G}{\partial y} \frac{\partial f}{\partial x} \frac{1}{\Delta}$$

Write now the divergence of D in the basis $\{\frac{\partial}{\partial f}, \frac{\partial}{\partial g}\}$

$$\frac{\partial R}{\partial f} + \frac{\partial S}{\partial g} = \frac{\partial R}{\partial x} \frac{\partial x}{\partial f} + \frac{\partial R}{\partial y} \frac{\partial y}{\partial f} + \frac{\partial S}{\partial x} \frac{\partial x}{\partial g} + \frac{\partial S}{\partial y} \frac{\partial y}{\partial g} =$$

$$(by(6), (7)) = \frac{\partial R}{\partial x} \left(-\frac{\frac{\partial F}{\partial x}}{\frac{\partial F}{\partial f}}\right) + \frac{\partial R}{\partial y} \left(-\frac{\frac{\partial G}{\partial y}}{\frac{\partial G}{\partial f}}\right) + \frac{\partial S}{\partial x} \left(-\frac{\frac{\partial F}{\partial x}}{\frac{\partial F}{\partial g}}\right) +$$

$$+ \frac{\partial S}{\partial y} \left(-\frac{\frac{\partial G}{\partial y}}{\frac{\partial G}{\partial g}}\right) = (by(8), (9)) = \frac{1}{\Delta} \left(\frac{\partial R}{\partial x} \frac{\partial g}{\partial y} + \frac{\partial R}{\partial y} -$$

$$\frac{\partial R}{\partial y} \frac{\partial g}{\partial x} - \frac{\partial S}{\partial x} \frac{\partial f}{\partial y} + \frac{\partial S}{\partial y} \frac{\partial f}{\partial x}\right)$$

But using (4) we obtain

$$\frac{\partial R}{\partial x} = \frac{\partial P}{\partial x} \frac{\partial f}{\partial x} + P \frac{\partial^2 f}{\partial x^2} + \frac{\partial Q}{\partial x} \frac{\partial f}{\partial y} + Q \frac{\partial^2 f}{\partial x \partial y}$$

$$\frac{\partial R}{\partial y} = \frac{\partial P}{\partial y} \frac{\partial f}{\partial x} + P \frac{\partial^2 f}{\partial x \partial y} + \frac{\partial Q}{\partial y} \frac{\partial f}{\partial y} + Q \frac{\partial^2 f}{\partial y^2}$$

$$\frac{\partial S}{\partial x} = \frac{\partial P}{\partial x} \frac{\partial g}{\partial x} + P \frac{\partial^2 g}{\partial x^2} + \frac{\partial Q}{\partial x} \frac{\partial g}{\partial y} + Q \frac{\partial^2 g}{\partial x \partial y}$$

$$\frac{\partial S}{\partial y} = \frac{\partial P}{\partial y} \frac{\partial g}{\partial x} + P \frac{\partial^2 g}{\partial x \partial y} + \frac{\partial Q}{\partial y} \frac{\partial g}{\partial y} + Q \frac{\partial^2 g}{\partial y^2}$$

Using the last equalities write down

$$\begin{aligned} \frac{\partial R}{\partial x} \frac{\partial g}{\partial y} - \frac{\partial S}{\partial x} \frac{\partial f}{\partial y} &= \left(\frac{\partial P}{\partial x} \frac{\partial f}{\partial x} + P \frac{\partial^2 f}{\partial x^2} + \frac{\partial Q}{\partial x} \frac{\partial f}{\partial y} + Q \frac{\partial^2 f}{\partial x \partial y}\right) \frac{\partial g}{\partial y} - \left(\frac{\partial P}{\partial x} \frac{\partial g}{\partial x} + P \frac{\partial^2 g}{\partial x^2} + \frac{\partial Q}{\partial x} \frac{\partial g}{\partial y} + Q \frac{\partial^2 g}{\partial x \partial y}\right) \frac{\partial f}{\partial y} \\ &= \frac{\partial P}{\partial x} \Delta + P \left(\frac{\partial^2 f}{\partial x^2} \frac{\partial g}{\partial y} - \frac{\partial^2 g}{\partial x^2} \frac{\partial f}{\partial y}\right) + Q \left(\frac{\partial^2 f}{\partial x \partial y} \frac{\partial g}{\partial y} - \frac{\partial^2 g}{\partial x \partial y} \frac{\partial f}{\partial y}\right) \end{aligned}$$

Analogously

$$\begin{aligned} \frac{\partial S}{\partial y} \frac{\partial f}{\partial x} - \frac{\partial R}{\partial y} \frac{\partial g}{\partial x} &= \left(\frac{\partial P}{\partial y} \frac{\partial f}{\partial x} + P \frac{\partial^2 f}{\partial x \partial y} + \frac{\partial Q}{\partial y} \frac{\partial f}{\partial y} + Q \frac{\partial^2 f}{\partial y^2}\right) \frac{\partial g}{\partial x} - \left(\frac{\partial P}{\partial y} \frac{\partial g}{\partial x} + P \frac{\partial^2 g}{\partial x \partial y} + \frac{\partial Q}{\partial y} \frac{\partial g}{\partial y} + Q \frac{\partial^2 g}{\partial y^2}\right) \frac{\partial f}{\partial x} \\ &= \frac{\partial Q}{\partial y} \Delta + P \left(\frac{\partial^2 g}{\partial x \partial y} \frac{\partial f}{\partial x} - \frac{\partial^2 f}{\partial x \partial y} \frac{\partial g}{\partial x}\right) + Q \left(\frac{\partial^2 g}{\partial y^2} \frac{\partial f}{\partial x} - \frac{\partial^2 f}{\partial y^2} \frac{\partial g}{\partial x}\right). \end{aligned}$$

It follows from the last equalities that

$$\begin{aligned} \frac{\partial R}{\partial f} + \frac{\partial S}{\partial g} &= \frac{1}{\Delta} \left(\left(\frac{\partial P}{\partial x} + \frac{\partial Q}{\partial y}\right) \Delta + P \frac{\partial \Delta}{\partial x} + Q \frac{\partial \Delta}{\partial y}\right) = \\ &= \frac{\partial P}{\partial x} + \frac{\partial Q}{\partial y} - D(\Delta) \frac{1}{\Delta} \end{aligned}$$

Therefore $\operatorname{div} D = \frac{\partial R}{\partial f} + \frac{\partial S}{\partial g} - D(\Delta) \frac{1}{\Delta}$ $\square$

Corollary 3. If $\Delta = \det J(f, g) \in \mathbb{K}^*$ then the divergence of a derivation $D \in \operatorname{Der} R$ does not depend on the choice of the commutative basis $\{\frac{\partial}{\partial f}, \frac{\partial}{\partial g}\}$. In particular, the divergence does not depend on the choice of the coordinate polynomials in the polynomial ring $\mathbb{K}[x, y]$.

Список використаних джерел

1. *Bavula V.V.* The groups of automorphisms of the Lie algebras of polynomial vector fields with zero or constant divergence / V.V. Bavula // arXiv:math.AG:1310.4741.
2. *Gavran V.S.* On weakly semisimple derivations of the polynomial ring in two variables / V.S. Gavran, V.V. Stepukh // Algebra and Discrete mathematics. - 2014. - №18. - P.50 - 58.
3. *Ievgen Makedonskyi*, On noncommutative bases of the free module $W_n(\mathbb{K})$ of all $\mathbb{K}$ -derivations of the polynomial ring in n variables. / Ievgen Makedonskyi // arXiv:math.RA:1105.4748.
4. *Nowicki A.* Commutative bases of derivations in polynomial and power series rings / A.Nowicki // J. Pure Appl. Algebra. - 1986. - 40. - P. 275-279.
5. *Nowicki A.* Polynomial derivations and their rings of constants / A. Nowicki // N.Copernicus University Press. - 1994. - Torun
6. *Nowicki A.* Rings of constants for k -derivations in $k[x_1, \dots, x_n]$ / A.Nowicki, M.Nagata // J. Math. Kyoto Univ. - 1988. - №28. - P.111 - 118.
7. *Petravchuk A.P.* On centralizers of elements in the Lie algebra of the special Cremona group $sa(2, k)$, / A.P. Petravchuk, O.G. Iena // Journal of Lie Theory. - 2006. - №3. - P.561 - 567.
8. *Stein Y.* Weakly nilpotent and weakly semisimple polynomials on the plane / Y. Stein // Int. Math. Research Notices. - 2000. - №13. - P. 681-698.
9. *Su Yucai* Structure of divergence free Lie algebras differential operators in two complex variables / Yucai Su, Xiaoping Xu // Journal of Algebra. - 2001. - 243. - № 2. - P. 557-595.

References

1. BAVULA, V.V. (2015), "The groups of automorphisms of the Lie algebras of polynomial vector fields with zero or constant divergence", *arXiv:math.AG:1310.4741*.
2. GAVRAN, V.S. and STEPUKH, V.V. (2014), "On weakly semisimple derivations of the polynomial ring in two variables", *Algebra and Discrete mathematics.*, no.18. - pp.50 - 58.
3. MAKEDONSKYI, IE. A. (2013), "On noncommutative bases of the free module $W_n(\mathbb{K})$ of all $\mathbb{K}$ -derivations of the polynomial ring in n variables", *arXiv:math.RA:1105.4748*.
4. NOWICKI, A. (1986), "Commutative bases of derivations in polynomial and power series rings", *J. Pure Appl. Algebra*, **40**, pp. 275-279.
5. NOWICKI, A. (1994), Polynomial derivations and their rings of constants, *N.Copernicus University Press.* - 1994. - Torun
6. NOWICKI, A. and NAGATA, M. (1988), Rings of constants for k -derivations in $k[x_1, \dots, x_n]$, *J. Math. Kyoto Univ.* - №28. - pp.111 - 118.
7. PETRAVCHUK, A.P. and IENA, O.G. (2006), On centralizers of elements in the Lie algebra of the special Cremona group $sa(2, k)$, *Journal of Lie Theory.* - №3. - pp.561 - 567.
8. STEIN, Y. (2000), Weakly nilpotent and weakly semisimple polynomials on the plane, *Int. Math. Research Notices.* - №13. - pp. 681-698.
9. SU, YUCAI and XU, XIAOPING (2001), Structure of divergence free Lie algebras differential operators in two complex variables, *Journal of Algebra.* - **243**. - № 2. - pp. 557-595.

Received: 08.12.2014

УДК 517.9

М.В. Плахотник¹, канд. ф.-м. н.

**Типи траєкторій періодичних точок
трикутних відображень, напівгрупа
ітерацій яких скінченна**

¹Київський національний університет
імені Тараса Шевченка,
83000, м. Київ, пр.-т. Глушкова, 4е,
e-mail: makar_plakhotnyk@ukr.net

M.V. Plakhotnyk¹, PhD in Math

**Types of periodical points trajectories of
triangular maps, whose semigroup of
iterations is finite**

¹Taras Shevchenko National University of Kyiv,
83000, Kyiv, Glushkova st., 4e,
e-mail: makar_plakhotnyk@ukr.net

Вивчаються неперервні трикутні відображення квадрату $[0, 1]^2$ в себе, ітерації яких утворюють скінченну групу. Доведено, що якщо напівгрупа ітерацій неперервного трикутного відображення квадрату в себе є скінченною циклічною групою, то воно або тривіальна, або складається з двох, або з чотирьох елементів. Введено поняття канонічного зображення циклічної групи неперервними відображеннями. Доведено, що групи, які зображаються відображеннями інтервалу, мають канонічне зображення, але група C_4 не має канонічного зображення трикутними відображеннями. Введено поняття типу траєкторії періодичної точки трикутного відображення квадрату та доведено, що обмеження та співіснування типів періодичних точок періоду 4 для трикутних відображень, що зображають групу C_4 , відсутні.

Ключові слова: Трикутне відображення, тип періодичної траєкторії, співіснування типів періодичних траєкторій.

So called triangular continuous square maps, whose iterations semigroup is finite are considered in the article. These are maps of square into itself, such that the first coordinate of the image is independent on the second coordinate of preimage. It is proved that this semigroup is a finite cyclic group only in the case if it is either trivial or has order 2 or 4. The notion of canonical representation of a cyclic group with triangular maps is introduced as such representations, where group identity corresponds to the identity map as a function, which does not move any point. It is proved that those groups, which can be exactly represented by interval maps have the canonical representation by triangular maps, but the group C_4 does not. The notion of a type of a cyclic trajectory of a point is introduced for triangular maps. All possible types of periodical trajectories for continuous interval maps are described. The coexistence of all types of periodical points for triangular map, which represents a group C_4 , is proved.

Key Words: Triangular map, periodical trajectory type, coexistence of periodical trajectories types.

Communicated by Prof. Kirichenko V.V.

Introduction

Triangular square I^2 into itself maps whose iterations semigroup is finite are considered in the article. Here we denote with I the interval $[0, 1]$.

A map $F : I^2 \rightarrow I^2$ are called continuous triangular, if it is of the form

$$F(x, y) = (f(x), g(x, y)) \quad (1)$$

for some continuous functions $f : I \rightarrow I$ and $g : I^2 \rightarrow I$. This means, that the first coordinate of the image of F is independent on the second coordinate of the argument.

Triangular maps form a semigroup with re-

spect to compositions. Remind, that compositions of a map F with itself is called iterations and the n -th iteration of F is denoted by F^n for $n \geq 1$.

Denote by F_{id} a triangular map, which does not move any elements of I^2 . This map is the identity element of semigroup of all triangular maps. We will differ representations of monoids (precisely, groups), where identity element is represented by F_{id} , or not.

For any point x_0 of the period n under the action of a map f of the interval I into itself, a permutation $\pi \in S_n$ corresponds to x_0 in the following rule (see. [5, p. 53], or a classical work [4]). Let $x_1 < \dots < x_n$ be an ordered orbit of x_0 . Then

permutation π of the set $\{1, \dots, n\}$ is defined by the rule $f(x_i) = x_{\pi(i)}$ for $1 \leq i \leq n$.

Definition 1. *The permutation π which is constructed above is called the **type** of periodical point x_0 in the action of the map f .*

The work [4] by O. Sharkovskii contains the world known theorem about the linear ordering of the sets of interval maps and about impossibility to specify this ordering. It is proved that sets B_n of those continuous interval I maps, which have the periodical point of a period n are linearly ordered by inclusion as follows.

$$\begin{aligned} B_1 \supset B_2 \supset B_{2^2} \supset B_{2^3} \supset \dots \\ \dots \supset B_{2^{2.5}} \supset B_{2^{2.3}} \supset \dots \supset B_{2^{.5}} \supset B_{2.3} \supset \dots \\ \dots B_9 \supset B_7 \supset B_5 \supset B_3, \end{aligned}$$

and all the inclusions are strict i.e. for all different $n_1, n_2 \in \mathbb{N}$ the inequality $B_{n_1} \neq B_{n_2}$ holds.

Remark 1 (see. [5], p. 58). *A cycle of period 3 may be only of the type $\pi_3 = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \end{pmatrix}$ or inverse. The cycle of a period 4 may be of some types: for example of $\pi_4^{(1)} = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 2 & 3 & 4 & 1 \end{pmatrix}$ or $\pi_4^{(2)} = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 3 & 4 & 2 & 1 \end{pmatrix}$. If the map f has a cycle of the type $\pi_4^{(1)}$ then it is easy to show that it also has a cycle of the period π_3 and whence the Sharkovskii theorem gives that it has points of all periods. In the same time interval map f may have a cyclic point with type $\pi_4^{(2)}$ and have no other point's periods except of those which yield from the Sharkovskii theorem i.e. 2 and 1.*

This remark explains the naturalness of considering of types of points trajectories in the study of properties of periodical points of maps.

It is proved at [7], that Sharkovskii theorem can be generalized to triangular maps.

In our work we will introduce the notion of canonical representation of a cyclic group by triangular map as such representation, where identity element of a group corresponds to F_{id} , mentioned above.

We mention in the Section 1, that only groups of 1, 2, or 4 elements can be exactly represented by triangular maps.

In Section 2 we prove, that there is no canonical representation of the group C_4 by continuous

triangular map, whence all its representations are non-canonical.

In Section 3 we generalize definition 1 for triangular maps and define the type of periodical point of triangular map. We call a type of the periodical trajectory of triangular map representable, if there exists a triangular map with a periodical point of such type. We describe all representable types and prove that any set of different representable types can be the whole set of all types of periodical trajectories of some triangular map.

1 Preliminaries

It is proved at [3] that if the semigroup of iterations of continuous map $f : I \rightarrow I$ is a finite group then this group consists of 1, or 2 elements. Also the following lemma is proved.

Lemma 1. *If the map $f \in C^0(I, I)$ has a periodical point with a period $m > 2$ then the semigroup of iterations of f is infinite.*

Simple algebraic reasonings give the following lemma (see. [3]).

Lemma 2. *For an arbitrary map H which maps some set M into its subset the uniform boundedness of points orbits is equivalent to the fact that iterations semigroup of this map is finite.*

Graphs of maps of I into itself, whose iterations semigroup is a finite group are described at [1]. Also the following theorem is proved.

Theorem 1. *If the map f of the interval I represents the finite group, then there exists an interval $[a, b] \subseteq I$, such that $f(I) = [a, b]$ and one of the following two conditions holds.*

1. *For any $x \in [a, b]$ the equality $f(x) = x$ holds and in this case the group of iterations of f is trivial;*
2. *The graph of f is symmetrical in the line $y = x$ for $x \in [a, b]$ but does not coincide with it. In this case the group of iterations of f is C_2 .*

All the semigroups which can be represented exactly with triangular maps of n -dimensional cube I^n are described at [2]. It is proved there that if cyclic group has exact representation of triangular maps of I^2 then its order is either 1, or 2, or 4.

Example 1. Let us give an example of a triangular map F of the square I^2 such that the equality $F = F^n$ holds for its iterations and the condition $F^n = F$ does not hold for $1 < n < 5$.

The deal of the example. Use the notation of definition 1 for the map F . Let the first coordinate f be defined with the formula $f(x) = 1 - x$ for all $x \in I$. Define the second coordinate g as follows.

Let $x < \frac{1}{2}$. Define $g(x, y) = x$ for $y \in [0, x]$; $g(x, y) = y$ for $y \in (x, 1 - x)$ and $g(x, y) = 1 - x$ for $y \in [1 - x, 1]$.

Also for $x \geq \frac{1}{2}$ define $g(x, y) = x$ for $y \in [0, 1 - x]$; $g(x, y) = 1 - y$ for $y \in (1 - x, x)$ and finally $g(x, y) = 1 - x$ for $y \in [x, 1]$.

The graphs of $g_{0,4}$ and $g_{0,8}$ are given at figure 1a). Now check the equality $F^5 = F$.

For $x \in (0, \frac{1}{2})$ and $y \in [0, x)$ we have $F(x, y) = (1 - x, x)$; $F^2(x, y) = F(1 - x, x) = (x, 1 - x)$; $F^3(x, y) = F(x, 1 - x) = (1 - x, 1 - x)$; $F^4(x, y) = F(1 - x, 1 - x) = (x, x)$; $F^5(x, y) = F(x, x) = (1 - x, x)$ whence $F(x, y) = F^5(x, y)$.

For $x \in (0, \frac{1}{2})$ and $y \in [x, 1 - x]$ we have $F(x, y) = (1 - x, y)$; $F^2(x, y) = F(1 - x, y) = (x, 1 - y)$; $F^3(x, y) = F(x, 1 - y) = (1 - x, 1 - y)$; $F^4(x, y) = F(1 - x, 1 - y) = (x, y)$ whence $F^5(x, y) = F(x, y)$.

For $x \in (0, \frac{1}{2})$ and $y \in (1 - x, 1]$ we have $F(x, y) = (1 - x, 1 - x)$; $F^2(x, y) = F(1 - x, 1 - x) = (x, x)$; $F^3(x, y) = F(x, x) = (1 - x, x)$; $F^4(x, y) = F(1 - x, x) = (x, 1 - x)$; $F^5(x, y) = F(x, 1 - x) = (1 - x, 1 - x)$ whence $F(x, y) = F^5(x, y)$.

The equality $F = F^5$ for $x \geq \frac{1}{2}$ is proved in the same way. $\square$

Definition 2. A *representation* of a cyclic group C_n by iterations of the map F is called **exact** if there is one-to-one correspondence between elements of C_n and iterations of F and this correspondence is concerted with the group operations in C_n .

Definition 3. An *exact representation* of a cyclic group C_n by iterations of the map F is called **canonical** representation if the identity of a group C_n corresponds to the neutral element of the semigroup of all maps which belongs to the set of iterations of F .

Remark 2. Theorem 1 yields that canonical representation of either trivial or the group C_2 by the maps of interval I is canonical if and only if f is

a bijection i.e. the equality $[a, b] = I$ holds where a and b are from the theorem formulation.

Remark 3. It is constructed at the example 1 the exact representation of the group C_4 but it is not canonical, because the group identity is represented by a map F^4 , which is not identity element of the semigroup of all square I^2 maps.

2 Canonical representations of cyclic groups with triangular maps

The main theorem of this section is the following.

Theorem 2. There is no canonical representation of the group C_4 with triangular maps of I^2 .

Remark 4. If a triangular map F represents the group C_4 canonically, then for every point $(x_1, y_1) \in I^2$ the condition $F^4(x_1, y_1) = (x_1, y_1)$ holds.

Lemma 3. If the semigroup of iterations of a triangular map F is finite, then trajectory of any periodical point of the period 4 is of the form

$$\begin{aligned} (x_1, y_1) &\rightarrow (x_2, y_2) \rightarrow (x_1, y_3) \rightarrow \\ &\rightarrow (x_2, y_4) \rightarrow (x_1, y_1), \end{aligned} \quad (2)$$

where $y_1 \neq y_3$.

Proof. Let the map F and point (x_1, y_1) be as in the condition of the lemma. Then x_1 is a periodical point of the map f .

If x_1 is a periodical point of the period 4 for the map f then Lemma 1 yields that semigroup of iterations of f is infinite whence semigroup of iterations of F is also infinite.

If x_1 is a fixed point of f then the point y_1 is a periodical point of the map $g_{x_1}(y) = g(x_1, y)$ and its period is 4 and trajectory of (x_1, y_1) under the action of F is looks as $F^n(x_1, y_1) = (x_1, g_{x_1}^n(y_1))$. Lemma 1 yields that semigroup of iterations of g_{x_1} is infinite which means that semigroup of iterations of F is also infinite.

So, the period of x_1 as a periodical point of f equals 2 which finishes the proof. $\square$

Lemma 4. If iterations of the triangular map F represent the group C_4 canonically then for every $x \in I$ the map $g_x(y) = g(x, y)$ is a homeomorphism of I with itself.

3 Non canonical representations of cyclic groups

We will consider in this section the question about possible types of periodical points of the period 4 under the action of the triangular map whose semigroup of iterations is finite.

Definition 4. We will call periodical sequences $\{(x_k, y_k)\}$ and $\{(\tilde{x}_k, \tilde{y}_k)\}$ to have the **same type** if for some fixed $t \in \mathbb{N}$ and all i, j inequalities $y_i \leq y_j$ and $\tilde{y}_{i+t} \leq \tilde{y}_{j+t}$ are equivalent and also inequalities $x_i \leq x_j$ and $\tilde{x}_{i+t} \leq \tilde{x}_{j+t}$ are equivalent.

Remark 5. Note that trajectories of periodical points (x_1, y_1) and $(\tilde{x}_1, \tilde{y}_1)$ under the action of triangular map F have the same type then points x_1 and $\tilde{x}_1$ as periodical points of the map f have the same period and the type of periodical trajectory in the sense of the Definition 1.

Definition 5. A **type** of a periodical sequence $\mathcal{P} = \{(x_k, y_k)\}$ is a finite sequence of points $\mathcal{T} = \{(a_k, b_k)\}$ of the set $\mathbb{N}^2$ which have the following properties:

1. The set of first values of the first coordinates of points of the set $\mathcal{T}$ together with the set of the second coordinated $\mathcal{T}$ are sets of some first natural numbers.

2. All elements of $\mathcal{T}$ are different.

3. Periodical sequences $\mathcal{P}$ and $\tilde{\mathcal{T}}$ have the same type where the infinite sequence $\tilde{\mathcal{T}}$ is obtained from the finite $\mathcal{T}$ one by its repeating infinite number of times.

Definition 6. We will a **type of periodical point** the type of its trajectory.

Remark 6. From Lemma 3 yields that a periodical point of the triangular map whose semigroup of iterations is finite may have either type

$$\mathcal{T}_1 = (1, y_1), (2, y_2), (1, y_3), (2, y_4), \quad (3)$$

of the type

$$\mathcal{T}_2 = (2, y_1), (1, y_2), (2, y_3), (1, y_4), \quad (4)$$

for some natural numbers y_1, y_2, y_3 and y_4 .

Definition 7. The finite sequence of natural numbers $\mathcal{T} = \{(a_k, b_k)\}$ is called an **admissible type** of periodical point if items 1 and 2 of Definition 5 hold.

Definition 8. The admissible type $\mathcal{T}$ of periodical point is called **representable** if there exists a continuous triangular map F whose semigroup of iterations is a finite group and which has a periodical point of the type $\mathcal{T}$.

The main result of this section is the following theorem.

Theorem 3. Each representable type of periodical point $\mathcal{T}$ of the period of the form (3) and (4) is representable.

Consider some lemmas which let decrease the number of types of periodical points whose representability should be studied.

Lemma 6. If for some natural numbers y_1, y_2, y_3 and y_4 the type of periodical points $\mathcal{T}_1 = [(1, y_1), (2, y_2), (1, y_3), (2, y_4)]$ is representable then type of periodical point $\mathcal{T}_2 = [(2, y_1), (1, y_2), (2, y_3), (1, y_4)]$ is also representable.

Proof. Let triangular map $F = (f, g)$ have a periodical point of the type $\mathcal{T}_1$. Consider the map $\tilde{F}$, defined as $\tilde{F}(x, y) = F(f(x), g(1 - x, y))$.

The periodical point which whose period type is $\mathcal{T}_1$ under the map F will be of the type $\mathcal{T}_2$ under the map $\tilde{F}$.

Note that all trajectories of $\tilde{F}$ are symmetrical in the line $x = 0,5$ to the correspond trajectories of F . $\square$

Lemma 6 lets to prove Theorem 3 only for admissible trajectories types $\mathcal{T}$ of period 4 of the form (3) and not to consider types of the form (4).

Lemma 7. If for some natural numbers y_1, y_2, y_3 and y_4 the type of trajectory $\mathcal{T}$ of the form (3) is representable the a type of trajectory $\tilde{\mathcal{T}} = [(1, y_1), (2, y_4), (1, y_3), (2, y_2)]$ is also representable.

Proof. If the map F has a point of the type $\mathcal{T}$ then this point has a type $\tilde{\mathcal{T}}$ under the action of the map $\tilde{F} = F^3$. $\square$

Lemma 8. If the type $\mathcal{T}$ of a periodical point of the form (3) is representable then the type

$$\tilde{\mathcal{T}} = (1, y_{\pi(1)}), (2, y_{\pi(2)}), (1, y_{\pi(3)}), (2, y_{\pi(4)})$$

of periodical point is also representable for any permutation $\pi \in S_4$.

Proof. Definition 5 of a periodical trajectory yields that type of a periodical point $\mathcal{T}$ of the form (3) will not change if we “permute cyclicly” points of the sequence $\mathcal{T}$ i.e. obtain

$$\mathcal{T}_1 = (2, y_2), (1, y_3), (2, y_4), (1, y_1),$$

$$\mathcal{T}_2 = (1, y_3), (2, y_4), (1, y_1), (2, y_2)$$

and

$$\mathcal{T}_3 = (2, y_4), (1, y_1), (2, y_2), (1, y_3).$$

Now lemma yields from Lemmas 6 and 7. $\square$

Lemma 9. *Type of trajectory $\mathcal{T}$ of the form (3) where $\{y_1, y_2, y_3, y_4\} = \{1, 2\}$ is representable.*

Proof. Lemma yields from the example 1 and Lemma 8. $\square$

Lemma 10. *Type of periodical point $\mathcal{T}$ of the form (3) where $\{y_1, y_2, y_3, y_4\} = \{1, 2, 3\}$ is representable.*

Proof. Taking into attention Lemma 8 it is enough to prove that there exists a triangular map F whose iterations form an exact representation of the group C_4 and which has a periodical point of the type

$$\mathcal{T} = (1, 1), (2, 1), (1, 2), (2, 3).$$

We will construct the map F with the periodical point of the type $\mathcal{T}$ in the same way as it was done at the Example 1.

Let the map f be defined with the formula $f(x) = 1 - x$ for all $x \in I$.

For $x \leq 1/2$ let g be defined as $g_x(y) = y + 1/2$ for $y \in [0, 0,5 - x]$ and let be $g_x(y) = 1 - x$ for $y \geq 0,5 - x$.

For $x > 1/2$ take $g_x(y) = 0,5$ for $y \in [0, 0,5]$; $g(y) = 1 - y$ for $y \in (0,5, x)$ and $g_x(y) = 1 - x$ for $y > x$.

Graphs of $g_0(y)$, $g_{0,4}(y)$, $g_{0,8}(y)$ and $g_1(y)$ are given at the figure 1b).

The fact that this map really has that properties as it is necessary can be proved with the direct checking in the same way as it was done at the Example 1. $\square$

Lemma 11. *The type of the sequence $\mathcal{T}$ of the form (3) where $\{y_1, y_2, y_3, y_4\} = \{1, 2, 3, 4\}$ is representable.*

Proof. The idea of the prove of this lemma is also analogical to the proof of the Lemma 9 and the Example 1.

According to Lemma 8 it is enough to prove that there exists a triangular map F whose iterations represent the group C_4 which have periodical point of the type

$$\mathcal{T} = (1, 1), (2, 2), (1, 3), (2, 4).$$

Take $f(x) = 1 - x$ for all $x \in I$.

For $x \leq 1/2$ take $g_x(y) = 1/3 + x/3$ for $y \leq x/3$; $g_x(y) = y + 1/3$ for $y \in (x/3, 2/3 - x)$ and $g_x(y) = 1 - x$ for $y \geq 2/3 - x$.

For $x > 1/2$ take $g_x(y) = 1/3 + x/3$ for $y \leq 2/3 - x/3$; $g_x(y) = 1 - y$ for $y \in (2/3 - x/3, x)$ and $g_x(y) = 1 - x$ for $y \geq x$.

Graphs of maps $g_0(y)$ and $g_4(y)$ are given at figure 1c). $\square$

Theorem 3 yields from Lemmas 9, 10 and 11.

4 Coexistence of types of periodical points

Remark 4 shows that the problem about coexistence of cycles of periodical points for interval maps was considered at first by A. Sharkovskii at his classical work [4]. In fact it is in this work the notion of type of periodical point of continuous interval map was given at first and the impotentness of some additional characteristic of a cycle except its period was explained.

Lemma 12. *If a triangular map F whose semi-group of iterations if finite has a periodical point of period 4 then it has a periodical point of the type $\mathcal{T} = (1, 1), (2, 1)$.*

Proof. Let (x_1, y_1) be a periodical point of the period 4. According to Lemma 3 its trajectory is of the form (2) i.e. x_1 is a periodical point of f and its period is 2.

Consider the map

$$\tilde{g}(y) = g_{x_2}(g_{x_1}(y)).$$

As a continuous interval I into I map it has a periodical point i.e. there exists a point y_0 such that trajectory of the point (x_1, y_0) looks as

$$(x_1, y_0) \rightarrow (x_2, y_0) \rightarrow (x_1, y_0).$$

This point has that type which is necessary. $\square$

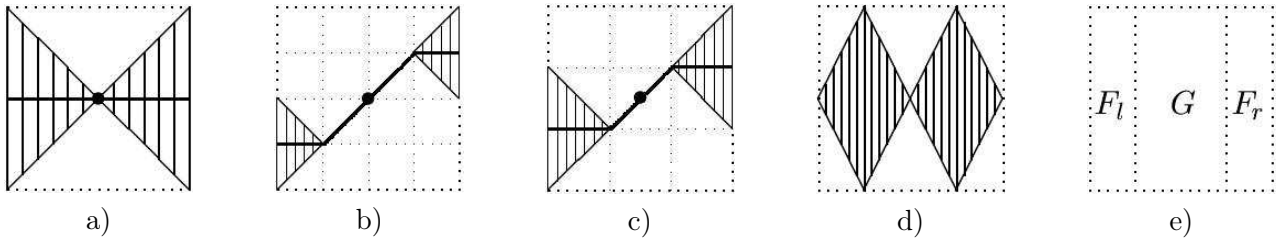


Figure 2:

Remark 7. Note that all triangular maps which were constructed in the process of proving of Lemmas 9, 10 and 11 have the following properties.

1. $g_{0,5}(y) = y$ for all $y \in I$ in particularly for $x = 0, 5$ the unique periodical point of F is its fixed point $(0, 5, 0, 5)$.

2. The points of period 4 have the same type for each of three constructed maps.

Taking into attention the Remark 7 we may obtain the following construction of triangular maps which have points of period 4 of different types.

Consider the transformation Ψ_1 which acts on the set of triangular maps and moves the map $F(x, y) = (f(x), g(x, y))$ to $\Psi_1(F)$ which is defines as $(\Psi_1(F))(x, y) = (\tilde{f}(x), g(x, y))$ where the function $\tilde{f}$ is defined as follows. If $x \leq 1/4$ then $\tilde{f}(x) = f(-2x + 1/2)$; if $1/4 < x \leq 3/4$ then $\tilde{f}(x) = f(2x - 1/2)$ and if $x \geq 3/4$ then $\tilde{f}(x) = f(-2x + 3/2)$.

For example the figure 2a) contains a shaded area of the square I^2 which is the periodical points set of the map F from the Example 1. Sets of periodical points of maps which were constructed at Lemmas 10 and 11 are given at figures 2b) and 2c). The bold curve at all three figures denotes the periodical points of period 2 and the point at the center of a rectangle is a fixed point.

The following lemma holds.

Lemma 13. Let the triangular map F is such that its the first coordinate is defined as $x \mapsto 1 - x$ and $F(0, 5, y) = 0, 5$ for all $y \in I$ then the following holds for $H = \Psi_1(F)$.

1. $H(0, 5, y) = H(0, y) = H(1, y) = 0, 5$ for all $y \in I$;

2. Sets of types of periodical points of maps F and H coincide.

3. Semigroups of iterations of the maps F and H coincide.

This lemma is proved with the direct checking.

For example the set of periodical points of map $\Psi_1(F)$ from the Example 1 if given at the figure 2d).

Define a transformation Ψ_2 which makes possible to construct the triangular map $\Psi_2(F, G)$ with two given triangular maps F and G and Ψ_2 will be such that the set of types of periodical points of the image map will be union of types of periodical points of F and G .

Additionally assume that as the triangular map F , as triangular map G are defined with the formula $x \rightarrow 1 - x$ for the first coordinate. Let F_l be the restriction of F to the set $[0, 1/2] \times I$ and let F_r be restriction of F to the set $[1/2, 1] \times I$. Then the transformation $\Psi_2(F, G)$ is schematically shown on the figure 2e) and it can be defined analytically as follows. $(\Psi_2(F, G))(x, y) = F(2x, y)$ for $x \leq 1/4$, $(\Psi_2(F, G))(x, y) = G(2x - 1/2, y)$ for $1/4 < x \leq 3/4$ and $(\Psi_2(F, G))(x, y) = F(2x - 1, y)$ for $x > 3/4$. The following lemma yields from the construction of Ψ_2 .

Lemma 14. Let the triangular map F and G are as in Lemma 13. Then the following holds for $H = \Psi_2(\Psi_1(F), \Psi_2(G))$.

1. $H(0, 5, y) = H(0, y) = H(1, y) = 0, 5$ for all $y \in I$;

2. Sets of types of periodical points of H is a union of sets of types of periodical points of F and types of periodical points of G ;

3. If semigroups of iterations of the maps F and G are finite groups the the semigroup of iterations of H is bigger of these two groups.

The reasonings above can be generalized in the following theorem.

Theorem 4. For arbitrary representable types $\mathcal{T}_1, \dots, \mathcal{T}_k$ of periodical points of period 4 there is a triangular map F whose iterations exactly represents a group C_4 and which has periodical points of each of types $\mathcal{T}_1, \dots, \mathcal{T}_k$ and has no periodical points of another period 4 types.

Proof. As is was noted in the Remark 7 the points of period 4 have the same type for each maps which were constructed at Lemmas 9, 10 and 11.

The constructions which were presented in the proof of further lemmas till Lemma 13 are

just transformations which construct new map all phase periodical period 4 points have the same type.

Now application of transformations Ψ_1 and Ψ_2 introduced above finishes the proof. $\square$

Список використаних джерел

References

- [1] *Плахотник М.В.* Зображення скінченної циклічної напівгрупи неперервними відображеннями інтервалу / М.В. Плахотник // Вісник Київського університету: серія Фізико-математичні науки - № 3, 2006р. - С.116-124.
- [2] *Плахотник М.В.* Конечные полугруппы, порожденные двумя идемпотентами, представимы треугольными отображениями / М.В. Плахотник // Труды ИПММ НАН Украины. - вып 15, 2007. - С. 151-159.
- [3] *Плахотник М.В.* Одновимірні динамічні системи з обмеженими у сукупності потужностями орбіт / М.В. Плахотник, В.В. Федоренко, Ю.В. Федоренко // Вісник Київського Університету; серія Фізико-математичні науки. - № 4, 2006. - С. 119-128.
- [4] *Шарковский А.Н.* Существование циклов непрерывного отображения прямой в себя / А.Н. Шарковский // Український математичний журнал. - 1964, том. 16, №1. - С. 61-71.
- [5] *Шарковский А.Н.* Введение в теорию функциональных уравнений / А.Н. Шарковский, Г.П. Пелюх // Киев: Наукова Думка. - 1974. - 119 с.
- [6] *Шарковский А.Н.* Динамика одномерных отображений / А.Н. Шарковский, С.Ф. Коляда, А.Г. Сивак, В.В. Федоренко // Киев: Наукова Думка. - 1989. - 216 с.
- [7] *Kloeden P.E.* On $\hat{S}$ harkovsky's cycle coexistence ordering/ P.E. Kloeden // Bull. Austral. Math soc. - 1979, Vol 20, No 2. - P. 171-178.
1. *PLAKHOTNYK M.V.* (2006) *Zobrazennia skinchennoi tsyklichnoji napivgrupy neperervnymy vidobrajenniamy intervalu.* Visnyk Kyjivskogo universytetu. Ser. Fyzyko-matematychni nauky.
2. *PLAKHOTNYK M.V.* (2007) *Konechnyje polugruppy, predstavimy dvumia treugolnymi otobrajenijami.* Trudy IPMM NAN Ukrainy, Vyp. 15
3. *PLAKHOTNYK M.V., FEDORENKO V.V. and FEDORENKO YU.V.* (2006) *Odnovymirni dynamichni systemy z обмеженими у сукупності потужностями орбіт.* Visnyk Kyjivskogo universytetu. Serija Fyzyko-matematychni nauky, No 3
4. *SHARKOVSKII A.N.* (1964) *Sushchestvovanie tsyklov nepreryvnogo otobrajenija priamoj v siebia* Ukrajskij matematychnyj jurnal., No 1.
5. *SHARKOVSKII A.N. and PELUH G.P.* (1974) *Vvedenije v teoriju funktsyonalnyh uravnenij.* Kyjiv, Naukova Dumka, 119 p.
6. *SHARKOVSKII A.N., KOLIADA S.F., SIVAK A.G. and FEDORENKO V.V.* (1989) *Dimanika odnomernyh otobrajenij.* Kyjiv, Naukova Dumka, 216 p.
7. *KLOEDEN P.E.* (1979) *On $\hat{S}$ harkovsky's cycle coexistence ordering* Bull. Austral. Math soc.

Received: 02.10.2014

УДК 512.53

Тетяна В. Турка, к. ф.-м. наук

Центр напівгрупи відповідностей

ДВНЗ «Донбаський державний педагогічний університет», 84116, Україна, Донецька обл., м. Слов'янськ, вул. Г. Батюка, 19,
e-mail: tvturka@gmail.com

Tetiana V. Turka, Ph. D.

The Center of Semigroup of Correspondences

Donbass State Pedagogical University,
84116, Ukraine, Donetsk region, Slovyansk,
Batykstr, 19.
e-mail: tvturka@gmail.com

У даній роботі описано центр напівгрупи відповідностей скінченної групи. Введено означення суперцентральності автоморфізму скінченної групи. Показано, що кожний автоморфізм циклічної групи є суперцентральним. Доведено, що центр напівгрупи відповідностей скінченної групи ізоморфний групі суперцентральної ізоморфізмів групи.

Ключові слова: центр напівгрупи, суперцентральний автоморфізм, напівгрупа відповідностей.

The problem of learning the semigroups of correspondences was set by Kurosh O. H. at his time. The work presents the centre of the semigroup of correspondences of the finite group. It is introduced the definition of the supercentral automorphism of the finite groups. We show that every automorphism of the cyclic group is supercentral and each homotopia of the vector space is the supercentral automorphism of the additive group of this space. It is proved that the set of all supercentral automorphisms form a subgroup of the group of the automorphisms of the finite group. It is also proved that the center the semigroup of the correspondences of the finite group is isomorphic to the group of the supercentral isomorphisms of the group. As a result, it is stated that the center of the semigroups of correspondences of the finite group is isomorphic to the group of the automorphisms of the finite group. It should be noted that if the finite group is a cyclic group or elementary abelian group, then the semigroup of correspondences has nontrivial center. If the group is generated by the element of order 2, then its semigroup of correspondences has the trivial center.

Key Words: center of a semigroup, supercentral automorphism, semigroup of correspondences

1 Вступ

Нехай G — універсальна алгебра. Якщо підалгебру з $G \times G$ розглядати як бінарне відношення на G , то множина $S(G)$ всіх підалгебр з $G \times G$ є напівгрупою відносно деморганівського добутку відношень. Напівгрупа $S(G)$ називається напівгрупою відповідностей алгебри G .

Задачу вивчення напівгруп відповідностей поставив ще в кінці 60-х років минулого століття відомий математик Курош О.Г. (див. [1]). Напівгрупи відповідностей довільних універсальних алгебр вивчав Іскандер [2]. Напівгрупи відповідностей скінченних груп вивчались, зокрема, в [3] і [4]. У даній роботі описується будова центра напівгрупи відповідностей скін-

ченної групи.

У роботі [3] показано, що коли G — група, то елементи напівгрупи $S(G)$ можна ототожнити з п'ятірками вигляду $(H_1, G_1, H_2, G_2, \varphi)$, де $H_1 \triangleleft G_1 < G$, $H_2 \triangleleft G_2 < G$, а φ — ізоморфізм факторгрупи G_1/H_1 на факторгрупу G_2/H_2 . При цьому відповідний елемент напівгрупи $S(G)$ — як підмножина із $G \times G$ — має вигляд

$$(H_1, G_1, H_2, G_2, \varphi) = \bigcup_{a \in G_1} (aH_1 \times \varphi(aH_1)).$$

Множини вигляду $aH_1 \times bH_2$, де $bH_2 = \varphi(aH_1)$, будемо називати блоками елемента $A = (H_1, G_1, H_2, G_2, \varphi)$.

2 Центр напівгрупи відповідностей

Нагадаємо, що центром $Z(T)$ напівгрупи T називається множина

$$Z(T) = \{a \in T \mid ax = xa \text{ для всіх } x \in T\}.$$

Зауважимо, що у випадку напівгрупи відповідностей $S(G)$ її центр $Z(S(G))$ завжди містить одиницю напівгрупи — діагональне відношення Δ , і нуль напівгрупи — пусте відношення (останнє — у випадку, коли пусте відношення включається в $S(G)$). Якщо центр містить лише ці елементи, будемо називати його *тривіальним*.

Твердження 1. *Нехай G — довільна універсальна алгебра. Якщо ненульовий елемент $A \in S(G)$ належить центру $Z(S(G))$ напівгрупи відповідностей $S(G)$, то проєкція елемента $A \subseteq G \times G$ на кожен множник збігається з G .*

Доведення. Для довільного бінарного відношення $\varphi \subseteq G \times G$ маємо рівності, які легко перевіряються безпосередньо:

$$\begin{aligned} \varphi \circ (G \times G) &= \text{pr}_1 \varphi \times G, \\ (G \times G) \circ \varphi &= G \times \text{pr}_2 \varphi. \end{aligned} \quad (1)$$

Повне відношення $G \times G$ є елементом напівгрупи відповідностей $S(G)$. Тому з рівностей (1) випливає, що для довільного ненульового елемента $A \in Z(S(G))$ виконується рівність

$$\text{pr}_1 A \times G = G \times \text{pr}_2 A.$$

Але тоді

$$\text{pr}_1 A = G, \quad G = \text{pr}_2 A. \quad \square$$

Нехай тепер G є групою,

$$A = (H_1, G_1, H_2, G_2, \varphi)$$

— елемент із центру $Z(S(G))$ її напівгрупи відповідностей $S(G)$. Із твердження 1 випливає, що $G_1 = G_2 = G$.

Лема 1. *Якщо елемент*

$$A = (H_1, G, H_2, G, \varphi)$$

напівгрупи відповідностей $S(G)$ належить центру, то $H_1 = H_2 = E$ і $\varphi \in \text{Aut } G$.

Доведення. Розглянемо елемент

$$B = (E, E, G, G, \varepsilon),$$

де ε — тривіальний автоморфізм одиничної групи. Тоді

$$B \circ A = (E, E, G, G, \varepsilon) = B,$$

$$A \circ B = (H_1, H_1, G, G, \varepsilon).$$

Отже, $H_1 = E$.

Аналогічно для елемента

$$C = (G, G, E, E, \varepsilon)$$

маємо:

$$A \circ C = (G, G, E, E, \varepsilon) = C,$$

$$C \circ A = (G, G, H_2, H_2, \varepsilon).$$

Тому $H_2 = E$.

Оскільки обидва фактори тепер дорівнюють $G/E \simeq G$, то $\varphi \in \text{Aut } G$. $\square$

Аutomорфізм φ групи G будемо називати *суперцентральним*, якщо він комутує з усіма ізоморфізмами факторів групи G у тому сенсі, що для довільних підгруп G_1, H_1, G_2, H_2 , таких, що $G_1 \triangleright H_1, G_2 \triangleright H_2$ і $G_1/H_1 \simeq G_2/H_2$, і для довільного ізоморфізму $\psi : G_1/H_1 \rightarrow G_2/H_2$ виконується рівність

$$\varphi(\psi(gH_1)) = \psi(\varphi(gH_1)) \quad (2)$$

для довільного $g \in G_1$.

Зрозуміло, що тотожний автоморфізм є суперцентральним. Розглянемо менш тривіальні приклади.

Твердження 2. *Кожен автоморфізм циклічної групи є суперцентральним.*

Доведення. Це очевидно для нескінченної циклічної групи Z , де є всього 2 автоморфізми: тотожний і $x \mapsto -x$.

Нехай тепер Z_n — циклічна група порядку n . Кожен автоморфізм групи Z_n має вигляд $\varphi_k : x \mapsto kx$, де елемент $k \in Z_n^*$ — фіксований. Для автоморфізма φ_k рівність (2) набуває вигляду

$$k \cdot \psi(g + N_1) = \psi(k \cdot (g + H_1))$$

і є правильною. $\square$

Елементарну абелеву групу Z_p^n зручно розглядати як адитивну групу n -вимірного векторного простору над полем Z_p . Тоді підгрупи збігатимуться з підпросторами, а автоморфізми — з невідродженими лінійними перетвореннями. Тому аналогічно попередньому доведенню для групи Z_n доводиться

Твердження 3. Для довільного елемента $k \in Z_p^*$ гомотетія $\varphi_k : x \mapsto kx$ є суперцентральним автоморфізмом адитивної групи векторного простору Z_p^n .

Твердження 4. Нехай G — скінченна група.

а) Якщо φ — суперцентральний автоморфізм, то для кожної підгрупи $H \leq G$ виконується рівність $\varphi(H) = H$.

б) Множина всіх суперцентральных автоморфізмів утворює підгрупу групи $\text{Aut } G$.

Доведення. а) Нехай $H \leq G$ — довільна підгрупа, φ — суперцентральний автоморфізм,

$$\varepsilon : H/E \rightarrow H/E$$

— тотожний ізоморфізм і g — довільний елемент із H . Тоді з рівності (2) випливає, що $\varphi(g) = \varepsilon(\varphi(g))$. Оскільки ε визначений лише на підгрупі H , то $\varphi(g) \in H$.

б) Позаяк група G — скінченна, то всі автоморфізми мають скінченний порядок. Тому досить перевірити лише замкненість множини суперцентральных автоморфізмів відносно множення. Нехай φ_1 і φ_2 — суперцентральні автоморфізми. Враховуючи а), замкненість відносно множення випливає з такого ланцюжка рівностей:

$$\begin{aligned} (\varphi_1\varphi_2)(\psi(gH_1)) &= \varphi_2(\varphi_1(\psi(gH_1))) = \\ &= \varphi_2(\psi(\varphi_1(gH_1))) = \\ &= \varphi_2(\psi(\varphi_1(g)H_1)) = \psi(\varphi_2(\varphi_1(g)H_1)) = \\ &= \psi(\varphi_2(\varphi_1(gH_1))) = \psi((\varphi_1\varphi_2)(gH_1)). \quad \square \end{aligned}$$

Теорема 1. Центр $Z(S(G))$ напівгрупи відповідностей $S(G)$ скінченної групи G збігається з множиною елементів вигляду $A = (E, G, E, G, \varphi)$, де φ — суперцентральний автоморфізм групи G .

Доведення. Із твердження 1 і леми 1 випливає, що кожен елемент $A \in Z(S(G))$ має вигляд $A = (E, G, E, G, \varphi)$, де $\varphi \in \text{Aut } G$. Покажемо спочатку, що $\varphi(H) = H$ для кожної підгрупи

$H \leq G$. Справді, якщо $H \leq G$, то $H \times H$ є елементом напівгрупи $S(G)$. Але

$$(H \times H) \circ A = H \times \varphi(H),$$

$$A \circ (H \times H) = (\varphi^{-1}(H) \times H).$$

Тому якщо $A \in Z(S(G))$, то $\varphi(H) = H$.

Розглянемо тепер довільний елемент $B = (H_1, G_1, H_2, G_2, \psi)$ напівгрупи $S(G)$. Запишемо B у вигляді

$$B = \bigcup_i (g_i H_1, g'_i H_2),$$

де $g'_i H_2 = \psi(g_i H_1)$. Безпосередньо перевіряється, що

$$\begin{aligned} B \circ A &= \bigcup_i (g_i H_1, \varphi(g'_i H_2)), \\ A \circ B &= \bigcup_i (\varphi^{-1}(g_i) H_1, g'_i H_2). \end{aligned} \quad (3)$$

Якщо $A \circ B = B \circ A$, то $\varphi^{-1}(g_i) H_1 = g_j H_1$ для деякого g_j . Тому $g_i H_1 = \varphi(g_j) H_1$. Крім того, має бути

$$(\varphi^{-1}(g_i) H_1, g'_i H_2) = (g_j H_1, \varphi(g'_j) H_2).$$

Але тоді

$$\begin{aligned} \psi(\varphi(g_j H_1)) &= \psi(\varphi(g_j) H_1) = \psi(g_i H_1) = \\ &= g'_i H_2 = \varphi(g'_j) H_2 = \varphi(g'_j H_2) = \\ &= \varphi(\psi(g_j H_1)). \end{aligned} \quad (4)$$

Оскільки елемент $g_j \in G_1$ може бути довільним, то це доводить виконання умови (2). Отже, автоморфізм φ є суперцентральним.

Навпаки, якщо φ — суперцентральний автоморфізм групи G , то для елемента $A = (G, E, G, E, \varphi)$ і довільного елемента $B = (G_1, H_1, G_2, H_2, \psi)$ із рівності $g_i H_1 = \varphi(g_j) H_1$ і рівностей (4) випливає, що праві частини в (3) будуть збігатися. А тому $A \circ B = B \circ A$ і A належить центру напівгрупи $S(G)$. $\square$

Наслідок 1. Центр $Z(S(G))$ напівгрупи відповідностей $S(G)$ скінченної групи G ізоморфний групі суперцентральных автоморфізмів групи G .

Доведення. Оскільки

$$(G, E, G, E, \varphi) \circ (G, E, G, E, \psi) = \\ = (G, E, G, E, \varphi\psi),$$

то це випливає з теореми 1 і твердження 4.b). $\square$

Наслідок 2. Якщо група автоморфізмів $\text{Aut } G$ групи G має тривіальний центр, то її напівгрупа відповідностей $S(G)$ має тривіальний центр.

Доведення. Це випливає з теореми 1 і того, що кожний суперцентральний автоморфізм належить центру групи $\text{Aut } G$. $\square$

Наслідок 3. Якщо G є циклічною групою Z_n (де $n > 2$) або елементарною абелевою групою Z_p^n (де $p > 2$ — просте), то напівгрупа відповідностей $S(G)$ має нетривіальний центр.

Доведення. Із тверджень 2 і 3 випливає, що в цих випадках існують нетривіальні суперцентральні автоморфізми. За теоремою 1 кожен із таких автоморфізмів дає нетривіальний елемент із центру напівгрупи $S(G)$. $\square$

Наслідок 4. Якщо група G породжується елементами порядку 2, то її напівгрупа відповідностей $S(G)$ має тривіальний центр.

Доведення. Із твердження 4.a) випливає, що кожний суперцентральний автоморфізм лишає нерухомими елементи порядку 2. А тому коли група породжується елементами порядку 2, то суперцентральним буде тільки тотожний автоморфізм. $\square$

Зауважимо, що умову наслідку 4 задовольняють, зокрема, елементарна абелева 2-група Z_2^n , дієдральна група D_n , симетрична група S_n і знакозмінна група A_n ($n > 4$).

Список використаних джерел

1. Курош А.Г. Общая алгебра (лекции 1969-70 учебного года) / А.Г. Курош // М.: Наука — 1974. — С. 160.
2. Искандер А.А. Структура соответствий универсальной алгебры / А.А. Искандер // Изв. АН СССР, серия матем. — 1965. — №29 — С.1357–1372.
3. Ганюшкін О.Г. Порядок напівгрупи відповідностей скінченної групи / О.Г. Ганюшкін, Т.В. Турка // Вісник Київського університету. Серія: фіз.-мат. науки. — 2009. — № 3. — С. 9–13.
4. Турка Т. Відношення Гріна на напівгрупі відповідностей скінченної групи / Т.В. Турка // Вісник Київського університету. Серія: фіз.-мат. науки. — 2010. — № 4. — С. 38–42.

References

1. KUROSH, A.G., (1974) *General algebra (lectures 1969-70 school years)* — М.: Nauka.—160 p.
2. ISKANDER A.A. (1965) *Structure of correspondence to a universal algebra* Izv.AN SSSR. Series: Mathematics, № 29. P. 1357–1372
3. GANYUSHKIN O., TURKA T. (2009) *The order of semigroups of correspondence of finite group* - The Bulletin of the University of Kiev. Series: Mechanics and mathematics, № 3. P. 9–13.
4. TURKA T. (2010) *Green's relations at semigroup of correspondences of finite group* The Bulletin of the University of Kiev. Series: Mechanics and mathematics, № 4. P. 38–42.

Received: 01.10.2014

УДК 519.21

Т.О. Яневич, к.ф.-м.н., с.н.с.

**Моделювання даних з великою
кількістю нулів при обстеженні
капітальних інвестицій України**

Київський національний університет імені
Тараса Шевченка, 01033, Київ, вул. Володи-
мирська, 64.

e-mail: yata452@univ.kiev.ua

T.O. Ianevych, PhD, Senior Researcher

**Modeling data with excess of zeros for
capital expenditure surveys in Ukraine**

Taras Shevchenko National University of
Kyiv, 01033, Kyiv, Volodymyrska st.64.

e-mail: yata452@univ.kiev.ua

У роботі вивчаються різні моделі та відповідні оціночні функції, що базуються на лінійних моделях, для даних, які містять велику кількість нулів. Досліджено ефективність таких оцінок при застосуванні вибіркового методу для обстеження капітальних інвестицій в Україні.

Ключові слова: Тобіт модель, Хекіт модель, загальна дворівнева модель, капітальні інвестиції, регресійні оцінки.

In the paper we examine the different models and corresponding Linear Model estimators for data containing many zero values. One of the models that widely used in such situations is Tobit model introduced by Tobin in 1958. It is developed for the censored data. For the data suffering from big number of zeros but not caused by censoring another models can be used: Heckit model or General Bivariate model. We examined the usefulness and efficiency of these models utilizing general linear regression estimator (GREG) in order to analyse the sample capital expenditure survey in Ukraine. For this we use Monte Carlo simulation method measuring the efficiency with the Absolute Relative Bias and the Relative Root Mean Square Error. It is appeared that GREG estimator leads to the biased but better results with regards to the accuracy comparing to the classical Horwitz-Thompson estimator. The usage of the Tobit and Heckit-based estimators fell short of expectations but still can be improved by changing or incorporating more independent variables.

Key Words: Tobit model, Heckit model, general bivariate model, capital expenditure, regression estimators.

Communicated by Prof. Kozachenko Yu.V.

1 Introduction

It is rather frequent situation when the economic data, especially microeconomic data, contain observations where some variable of interest is equal to zero for a number of the observations in the data set. Such data have excess of zero values and this can lead to a number of econometric problems when using Ordinary Least Squares (OLS) to estimate the unknown parameters of a regression model. We faced with this problem when start to work with Ukrainian capital expenditure survey.

The annual and quarterly surveys of capital expenditure collect important information for the Ukrainian National Accounts. This information is widely used by Ukrainian government and agencies, trade associations, universities and international organizations etc. for policy development and as a measure of regional activity.

The aim of the annual surveys is the structural detection whereas the quarterly surveys have the estimation of changes as an objective. All enterprises are to be observed in the annual survey. In the quarterly survey only the enterprises that have significant size of capital expenditure are usually surveyed. Before, only the enterprises, which take part in other quarter surveys (mostly in structural business survey), had a chance to be observed quarterly. Within the quarterly surveys the total of the capital expenditure was not adjusted for the unobserved enterprises. They contain many zero observations which have different nature. Some zeroes are obtained since the enterprises were not surveyed and some are real "genuine" zeroes corresponding to the situation when the enterprise doesn't have capital expenditure.

In order to avoid underestimation of the capital expenditure value in the quarterly surveys the probabilistic sampling should be implemented into

the investigation. Since the large and medium-sized enterprises are more valuable they are all observed every time. We survey by the means of the sample only the small and new enterprises. For new enterprises only the Horwitz-Thompson estimator can be used since we have little additional information for them. Thereby we focus on the investigation of small enterprises and perceive them as population.

2 Models for data with excess of zeros

The key decision facing any researcher working with a data set containing zeros is the choice of the appropriate model. The following summarizes the key elements of such a decision. Suppose that the variable of interest is y_i , and there are a large number of zero values for y in a given data set. The first step is to determine why the zeros are present in the data. There can be two alternatives: (1) the zeros appear as a result of censoring or (2) the zeros represent a decision that the researcher has no control over for some reason.

The first alternative usually corresponds to Tobit Model whereas the second one – to Heckit Model. The General Bivariate Model includes the features of both these models and can be considered as their generalization. Let us consider them in details.

2.1 Tobit Model

The Tobit Model was introduced by Tobin [5] in 1958. The Tobit Model, also called a censored regression model, is designed to estimate linear relationships between variables when there is either left- or right-censoring in the dependent variable (also known as censoring from below and above, respectively). Formally, it can be written as

$$\begin{aligned} y_i^* &= x_i' \beta + u_i, \quad i = 1, 2, \dots, n; \\ y_i &= y_i^*, \text{ if } y_i^* > 0; \\ y_i &= 0, \text{ if } y_i^* \leq 0. \end{aligned}$$

where u_i are i.i.d. drawings from $\mathcal{N}(0, \sigma_u)$, and y_i and x_i are observed in the sample, but y_i^* are unobserved if $y_i^* < 0$.

It is also worth to mention that there is sometimes confusion about the difference between truncated data and censored data. With censored variables, all of the observations are in the dataset, but we don't know the "true" values of some of

them. With truncation some of the observations are not included in the analysis because of the value of the variable. When a variable is censored, regression models for truncated data provide inconsistent estimates of the parameters.

The coefficient of such model can be calculated using the Maximum Likelihood method. In **R** there is a function `tobit()` in the **AER** package developed for this task. The Tobit Model allows different generalizations. For more information see [3].

2.2 Heckit Model

This type of model is appropriate when $y_i = 0$ because of a non-observable response. It means knowledge that $y_i = 0$ (as opposed to $d_i = 0$) is uninformative in estimating the determinants of the level of y_i . We can formulate it as follows

$$\begin{aligned} y_i &= d \cdot y_i^*, \quad i = 1, 2, \dots, n; \\ w_i &= \alpha' z_i + v_i; \\ \begin{cases} d_i = 0, & \text{if } w_i > 0; \\ d_i = 1, & \text{otherwise.} \end{cases} \\ y_i^* &= x_i' \beta + u_i. \end{aligned}$$

Here (v_i, u_i) are jointly normally distributed. If u and v are independent then we obtain so-called Two Part Model. The coefficient of the Heckit Model can be calculated using the Maximum Likelihood method or 2 step method, developed by Heckman [2] in 1976.

In **R** you can use the function **R** or `selection()` inside the `sampleSelection` package developed for this task. For details see [6].

2.3 General Bivariate Model

According to the Jones' [1] notation, the General Bivariate Model can be identified via the following three components. (In this paper Jones investigated the cigarette consumption and this caused the notations)

- Observed Consumption $\mathbf{y} = \mathbf{d} \cdot \mathbf{y}^{**}$
- Participation Equation $\mathbf{w} = \alpha' \mathbf{z} + \mathbf{v}$ and participation depends on the indicator variable so that

$$\begin{cases} d_i = 0, & \text{if } w_i > 0; \\ d_i = 1, & \text{otherwise.} \end{cases}$$

- Consumption Equation

$$\mathbf{y}^{**} = \max(0, \mathbf{y}^*)$$

$$\mathbf{y}^* = \mathbf{x}'\beta + \mathbf{u}.$$

So, in the Tobit Model the variables and the parameters in the Participation and Consumption Equations are identical. The Heckit model can be derived from the General Bivariate Model if there is no censoring in the Consumption Equation.

3 Analysing the data of expenditure for agricultural enterprises

In this section we plan to investigate the efficiency of the general regression estimator based on Tobit and Heckit models comparing to the classical Horvitz-Thompson estimator when analyzing capital expenditure of the Ukrainian agricultural enterprises.

For the population U consisting of N elements we denote as $(y_1, \dots, y_N)$ the vector of values for the variable y (here y_k are the values for the k -th element of the population. In our case we consider as a population U all the small agricultural enterprises of Ukraine. According to the State Register we had 37100 such enterprises in 2010, so $N = 37100$. The variable y that we are interested in is the capital expenditure of such enterprises in 2010. The parameter we investigate is the total capital expenditure $Y = \sum_{k \in U} y_k$. The underlying design is simple random sampling.

The main *relative* measure of efficiency for the estimator $\hat{\theta}$ of the parameter θ is the relative bias

$$\mathbb{B}(\hat{\theta}) = \frac{\mathbb{E}(\theta - \hat{\theta})}{\theta}$$

and the relative mean square error

$$\text{MSE}(\hat{\theta}) = \frac{\mathbb{E}(\theta - \hat{\theta})^2}{\theta^2}.$$

We compare the efficiency of two types of estimators for the total – classical Horvitz-Thompson estimator and the general regression estimator using the auxiliary information and assisted by different models:

$$\hat{Y}_{HT} = \sum_{k \in s} \frac{y_k}{\pi_k};$$

$$\hat{Y}_{GREG} = \sum_{k \in U} \hat{y}_k + \sum_{k \in s} \frac{y_k - \hat{y}_k}{\pi_k};$$

where π_k are the sample weights of the k -th element and $\hat{y}_k$ are the predictive values of the variable y for the element k obtained from some linear model. The additional information on these and other estimators can be found in the books [4], [7].

Since the explicit expression for the bias and mean square error is very hard to obtain so the Monte Carlo simulation method is usually used for estimation of such parameters. For the estimation of the relative bias we shall use the Absolute Relative Bias

$$ARB = \left| \frac{1}{K} \sum_{i=1}^K \hat{Y}(s_i) - Y \right| / Y,$$

and the Relative Root Mean Square Error for estimation relative mean square error

$$RRMSE = \sqrt{\frac{1}{K} \sum_{i=1}^K (\hat{Y}(s_i) - Y)^2 / Y^2}$$

Within the investigation we have made $K = 1000$ Monte Carlo simulation utilizing data from annual surveys of capital expenditure in 2009 and 2010. As far as these data appeared to be not very similar to the Normal we also investigated different estimator for log-transformed data.

So, we shall compare the following estimators

- Horvitz-Thompson estimator (HT);
- general regression estimator assisted by classical linear regression model with capital expenditure in 2009 as a regressor (GREG);
- general regression estimator assisted by classical linear regression model with log-transformed variables (LGREG);
- general regression estimator assisted by Tobit model with capital expenditure in 2009 as a regressor (TOB);
- general regression estimator assisted by Tobit model with log-transformed variables (LTOB);
- general regression estimator assisted by Heckit model with capital expenditure in 2009 as a regressor (HECK);

- general regression estimator assisted by Heckit model with log-transformed variables (LHECK);

The results of the simulation study are presented in the Tables:

%	HT	GREG	TOB	HECK
ARB	0.07	1.27	3.94	0.13
RRMSE	8.15	7.89	12.2	7.89

%	HT	LGREG	LTOB	LHECK
ARB	0.07	0.08	32.09	0.05
RRMSE	8.15	7.86	229.4	7.74

Список використаних джерел

1. Jones, A. M. A double-hurdle model of cigarette consumption / A. M. Jones // *Journal of Applied Econometrics* – 1989. – 4. – No. 1. – pp. 23–39.
2. Heckman, J. The Common Structure of Statistical Models of Truncation, Sample Selection and Limited Dependent Variables and a Simple Estimator for Such Models/ J. Heckman // *Annals of Economic and Social Measurement*. – 1976. – 5. – No. 4. – pp. 475–492.
3. Humphreys, B.R. Dealing with zeroes in economic data. / B.R. Humphreys // University of Alberta. – 2013. – www.ualberta.ca/ bhumphre/class/ (zeros_v1.pdf).
4. Särndal Model Assisted Survey Sampling / C.-E. Särndal, B. Swensson, J. Wretman // Springer-Verlag, New York. – 1992. – 694 p.
5. Tobin, J. Estimation of relationships for limited dependent variables / J. Tobin // *Econometrica*. – 1958. – 26. – pp. 24–36.
6. Toomet, O. Sample Selection Models in R: Package sampleSelection / O. Toomet, A. Henningsen // *Journal of Statist. Software*. – 2008. – 7. – www.jstatsoft.org/v27/i07
7. Василюк, О. Лекції з теорії і методів вибірових обстежень / О.І. Василюк, Т.О. Яковенко // ВПЦ "Київський університет", Київ, — 2010. – 208 С.

4 Conclusion

As we can see, usage of GREG estimator leads to biased but better results with regards to the accuracy. The usage of the Tobit and Heckit-based estimators fell short of expectations but still can be improved by changing or incorporating more independent variables. And the main useful thing is that all these estimators except HT can be used for the construction of small area estimators.

The small efficiency of the regression estimators based on the Tobit model can be explained by realizing that the process of zero appearing in this case is not connected to the censoring. The Heckit model explains this process much better and efficiency of the estimators is much better.

References

1. JONES, A. M. (1989), "A double-hurdle model of cigarette consumption", *Journal of Applied Econometrics*, v. 4, No. 1, pp. 23–39.
2. HECKMAN, J. (1976) "The Common Structure of Statistical Models of Truncation, Sample Selection and Limited Dependent Variables and a Simple Estimator for Such Models", *Annals of Economic and Social Measurement*, 5, No. 4, pp. 475–492.
3. HUMPHREYS, B.R. (2013) "Dealing with zeroes in economic data", *University of Alberta*, www.ualberta.ca/ bhumphre/class/ (zeros_v1.pdf).
4. SÄRNDAL, C.-E., SWENSSON, B. and WRETMAN, J. (1992) *Model Assisted Survey Sampling*, Springer-Verlag, New York, 694 P.
5. TOBIN, J. (1958) "Estimation of relationships for limited dependent variables", *Econometrica*, v. 26, pp. 24–36.
6. TOOMET, O. and HENNINGSEN, A. (2008) "Sample Selection Models in R: Package sampleSelection", *Journal of Statist. Software*, 7, www.jstatsoft.org/v27/i07
7. VASYLYK, O. and YAKOVENKO, T. (2010) *Lectures on theory and methods of survey sampling*, PC "Kyiv University", Kyiv, 208 P.

Received: 25.10.2014

**ДИФЕРЕНЦІАЛЬНІ
РІВНЯННЯ,
МАТЕМАТИЧНА ФІЗИКА
ТА МЕХАНІКА**

УДК 517.9

Асроров Ф. А.¹, к.ф.-м.н., н.с.

F. A. Asrorov¹, C.Sci (Phys-Math.).

Інтегральні множини розривних динамічних систем

Integral sets of discontinuous dynamical systems

¹ Київський національний університет імені Тараса Шевченка, 83000, м. Київ, пр-т. Глушкова 4д,
e-mail: farhod@univ.kiev.ua

¹ Taras Shevchenko National University of Kyiv, 83000, Kyiv, Glushkova st., 4d,
e-mail: farhod@univ.kiev.ua

Тематика роботи тісно пов'язана із застосуванням теорії багаточастотних коливань в системах звичайних диференціальних рівнянь з імпульсними збуреннями. Вивчається питання існування і властивостей інтегральних множин для класів лінійних та слабо нелінійних автономних імпульсно збурених систем. Крім того, досліджується якісна поведінка розв'язків в околі отриманої інтегральної множини. Для цього використовується ітераційний процес, який полягає в тому, що множина відшукується, як границя послідовності множин, кожна з яких є інтегральною множиною спеціально побудованих системи рівнянь. Однією з найважливіших достатніх умов для існування інтегральної множини є функція Гріна-Самойленка. Раніше це питання було вивчено для систем з імпульсним збуренням у фіксовані моменти часу. Новизна даної роботи полягає в тому, що тут розглядається розривна динамічна система, тобто коли фазова траєкторія зазнає імпульсного впливу при проходженні через фіксовану гіперповерхню фазового простору.

Ключові слова: функція Гріна-Самойленка, інтегральна множина, імпульсна дія.

Subject of the work is closely connected with applying the theory of multifrequency oscillations to the systems of ordinary differential equations with impulsive perturbations. We study the problem of existence and properties of integral sets for classes of linear and weakly non-linear autonomous systems with impulsive perturbations. Also we investigate the qualitative behavior of solution in a neighbourhood of obtained integral set. For this purpose we use iteration process that is we obtain integral set as a limit of sequence of integral sets for specially constructed systems of differential equations. One of the most important sufficient conditions for existence of integral set is Green-Samoilenko function. Previously, the issue has been studied for systems with impulsive perturbations at fixed moments of time. The novelty of this work lies in the fact that here we consider discontinuous dynamical system, that is a situation where the phase trajectory has impulsive perturbation when it passes through a fixed subset of the phase space.

Keywords: Green-Samoilenko function, integral set, impulsive action.

Статтю представив д.ф.-м.н., академік НАН України Перестюк М.О.

Досліджуємо питання існування інтегральних множин одного класу лінійних і слабо нелінійних систем диференціальних рівнянь, що піддаються імпульсному впливу в момент попадання зображаючої точки в задані множини фазового простору.

Розглянемо систему диференціальних рівнянь з імпульсним впливом

$$\frac{d\varphi}{dt} = \omega, \quad \frac{dx}{dt} = P(\varphi)x + f(\varphi, x), \quad \varphi \notin \Gamma, \quad (1)$$

в якій $\Delta x|_{\varphi \in \Gamma} = B(\varphi)x + I(\varphi, x)$,
 $x = (x_1, \dots, x_n)$, $\varphi = (\varphi_1, \dots, \varphi_m)$,
 $\omega = (\omega_1, \dots, \omega_m)$ - вектор з додатніми компонентами. Щодо множини Γ припускаємо,

що вона є підмножиною тора $\mathfrak{T}_m$ і являє собою
многовиди розмірності $m-1$, які можна
визначити рівнянням $\langle k, \varphi \rangle = 0 \pmod{2\pi}$, при
цьому $k = (k_1, \dots, k_m)$ – цілочисельний вектор,
 $\langle k, \omega \rangle \neq 0$. Остання умова забезпечує
приналежність кутової фазової змінної $\varphi(t)$
множині Γ в момент дії імпульсу як при $t=0$,
так і при $t \neq 0$. $P(\varphi)$ і $B(\varphi)$ – неперервні, 2π -
періодичні по кожній компоненті $\varphi_v, v = \overline{1, m}$
квадратні матриці, функції $f(\varphi, x)$ та $I(\varphi, x)$
визначено для всіх $\varphi \in \mathfrak{T}_m, x \in R^n$, неперервні
(кусково-неперервні з розривами першого роду
по φ), 2π -періодичні по φ і задовольняють
умові Ліпшиця по x рівномірно відносно
 $\varphi \in \mathfrak{T}_m$:

$$\begin{aligned} \|f(\varphi, x') - f(\varphi, x'')\| + \|I(\varphi, x') - I(\varphi, x'')\| \leq \\ \leq N \|x' - x''\| \end{aligned} \quad (2)$$

для всіх $x', x'' \in R^n$.

Допоміжна лема.

Лема. Для будь-якого розв'язку $t = t_i(\varphi)$
рівняння (3) справедлива рівність

$$t_i(\varphi - \omega t) - t_i(\varphi) = t \quad (3)$$

для всіх $\varphi \in \mathfrak{T}_m, t \in R^n$.

Будь яка неперервна траєкторія $\varphi = \omega t + \varphi_0$
перетинає многовид Γ через однакові часові
інтервали

$$\frac{2\pi}{\langle k, \omega \rangle}.$$

Позначимо через $t_i(\varphi)$ – моменти попадання
точки, що рухається по $\varphi + \omega t$ на Γ . Тоді для
будь-якого $i \in Z$ справедлива рівність

$$t_{i+1}(\varphi) - t_i(\varphi) = \frac{2\pi}{\langle k, \omega \rangle}.$$

Перш ніж сформулювати основний
результат, що відноситься до системи (1),
розглянемо наступну систему рівнянь:

$$\frac{d\varphi}{dt} = \omega, \quad \frac{dx}{dt} = P(\varphi)x + f(\varphi), \quad \varphi \notin \mathfrak{T}, \quad (4)$$

$$\Delta x|_{\varphi \in \Gamma} = B(\varphi)x + I(\varphi),$$

в якій $P(\varphi), B(\varphi), \omega, \Gamma$ – ті ж, що і в систем (1);
 $f(\varphi)$ і $I(\varphi)$ – неперервні (кусково-неперервні з
розривами першого роду по φ), 2π -періодичні
по φ функції.

Нехай $G(t, \tau, \varphi), G(\tau, \tau, \varphi) = E$ –
фундаментальна матриця системи рівнянь

$$\frac{dx}{dt} = P(\varphi_t(\varphi))x, \quad t \neq t_i(\varphi), \quad (5)$$

$$\Delta x|_{t=t_i(\varphi)} = B(\varphi_{t_i}(\varphi)).$$

Безпосередньо переконуємося, що $G(t, \tau, \varphi)$
задовольняє рівностям

$$\begin{aligned} G(t, \tau, \varphi + 2\pi) &= G(t, \tau, \varphi), \\ G(t, t + \tau, \varphi - \omega t) &= G(0, \tau, \varphi), \end{aligned} \quad (6)$$

для всіх $t, \tau \in R$ і $\varphi \in \mathfrak{T}_m$.

Нехай матриця $G(t, \tau, \varphi)$ і функції $t_i(\varphi)$ такі,
що функції

$$\begin{aligned} x_t(\varphi) &= \int_{-\infty}^t G(t, \tau, \varphi) f(\varphi + \omega \tau) d\tau + \\ &+ \sum_{t_i(\varphi) < t} G(t, t_i(\varphi), \varphi) I(\varphi + \omega t_i(\varphi)) \end{aligned} \quad (7)$$

являють собою сім'ю обмежених на всій осі
розв'язків рівнянь

$$\frac{dx}{dt} = P(\varphi + \omega \tau)x + f(\varphi + \omega \tau), \quad t \neq t_i(\varphi),$$

$$\Delta x|_{t=t_i(\varphi)} = B(\varphi + \omega t_i(\varphi)) + I(\varphi + \omega t_i(\varphi)).$$

Покладемо $x_t(\varphi) = u(\varphi + \omega t)$ і замінімо в (7)
 φ на $\varphi - \omega t$, отримаємо

$$\begin{aligned} u(\varphi) &= \int_{-\infty}^t G(t, \tau, \varphi - \omega t) f(\varphi + \omega(\tau - t)) d\tau + \\ &+ \sum_{t_i(\varphi - \omega t) < t} G(t, t_i(\varphi - \omega t), \varphi - \omega t) \times \\ &\times I(\varphi + \omega t_i(\varphi - \omega t) - t) \end{aligned}$$

Звідси, враховуючи, що

$$t_i(\varphi - \omega t) - t_i(\varphi) = t, \quad \varphi_{t+s}(\varphi) = \varphi_t(\varphi_s(\varphi)),$$

$$G(t, \tau, \varphi + 2\pi) = G(t, \tau, \varphi), \quad G(t, \tau, \varphi - \omega t) = G(0, \tau, \varphi),$$

маємо

$$\begin{aligned} u(\varphi) &= \int_{-\infty}^0 G(0, \tau, \varphi) f(\varphi + \omega \tau) d\tau + \\ &+ \sum_{t_i(\varphi) < 0} G(0, t_i(\varphi), \varphi) I(\varphi + \omega t_i(\varphi)). \end{aligned} \quad (8)$$

У припущенні, що інтеграл і сума в (7)
рівномірно збіжні, одержуємо, що функція $u(\varphi)$
визначає інтегральну множину системи рівнянь
(4).

Відзначимо, що для збіжності інтеграла і
суми з (8) достатньо, щоб функція $G(t, \tau, \varphi)$
задовольняла нерівність

$$\|G(t, \tau, \varphi)\| \leq K e^{-\gamma(t-\tau)}, \quad t \geq \tau \quad (10)$$

для всіх $t, \tau \in R$, $\varphi \in \mathfrak{Z}_m$ при деяких додатніх K і γ , а розв'язки рівняння (3) $t_i(\varphi)$ були такими, що

$$t_{i+1}(\varphi) - t_i(\varphi) \geq \theta > 0 \quad (11)$$

для всіх $i \in Z$, $\varphi \in \mathfrak{Z}_m$ і деякого $\theta > 0$.

При зазначених умовах з (8) маємо

$$\|x_i(\varphi)\| \leq \frac{K}{\gamma} \max_{\varphi \in \mathfrak{Z}_m} \|f(\varphi)\| + \frac{K}{1 - e^{-\gamma\theta}} \max_{\varphi \in \mathfrak{Z}_m} \|I(\varphi)\| \quad (12)$$

для всіх $t \in R$ і $\varphi \in \mathfrak{Z}_m$.

Асимптотичну стійкість цієї інтегральної множини $u(\varphi)$ забезпечує нерівність (10).

Таким чином, доведено наступна теорема.

Теорема 1. Нехай в системі рівнянь (5) функції $f(\varphi)$ і $I(\varphi)$ 2π -періодичні, неперервні на $\mathfrak{Z}_m$ (кусово-неперервні з розривами першого роду при $\varphi \in \Gamma$); $P(\varphi)$ і $B(\varphi)$ – неперервні на $\mathfrak{Z}_m$ 2π -періодичні матриці.

Якщо матриця $G(t, \tau, \varphi)$ задовольняє оцінці (10), а функції $t_i(\varphi)$ – нерівності (11), то система рівнянь (5) має асимптотично стійку інтегральну множину

$$x = u(\varphi), \quad u(\varphi + 2\pi) = u(\varphi),$$

де $u(\varphi)$ – кусково-неперервна з розривами першого роду на множині Γ функція. При цьому можна вказати таку постійну C , незалежну від функцій $f(\varphi)$ і $I(\varphi)$, що

$$\|u(\varphi)\| \leq C \max \left\{ \max_{\varphi \in \mathfrak{Z}_m} \|f(\varphi)\|, \max_{\varphi \in \mathfrak{Z}_m} \|I(\varphi)\| \right\}.$$

Відзначимо деякі окремі випадки систем (6), для яких нормована фундаментальна матриця $G(t, \tau, \varphi)$ допускає оцінку (10). Так, наприклад, якщо $B(\varphi) \equiv 0$, а траєкторії $\varphi + \omega\tau$ всюди щільно лежать на торі $\mathfrak{Z}_m$, то нерівність (10) виконується, якщо

$$\frac{1}{(2\pi)^m} \int_0^{2\pi} \dots \int_0^{2\pi} \Lambda(\varphi) d\varphi_1 \dots d\varphi_m < 0,$$

$\Lambda(\varphi)$ – найбільше з власних чисел симетричної матриці

$$P^*(\varphi) = \frac{1}{2} (P(\varphi) + P'(\varphi)),$$

де $P'(\varphi)$ – транспонована по відношенню до $P(\varphi)$ матриця.

Нерівність (10) виконується також у разі, якщо P і B – постійні матриці, комутуючі між собою, причому $E + B$ невиводжені і дійсні частини всіх власних чисел матриці

$$\Lambda = P + p \ln(E + B)$$

від'ємні.

Інтегральна множина $x = u(\varphi)$ системи рівнянь (1), будемо шукати як границю послідовності інтегральної множини $u_m(\varphi)$:

$$u(\varphi) = \lim_{m \rightarrow \infty} u_m(\varphi), \quad (13)$$

кожне з яких є інтегральною множиною системи рівнянь

$$\frac{dx}{dt} = \omega,$$

$$\frac{dx}{dt} = P(\varphi)x + f(\varphi, u_{m-1}(\varphi)), \quad \varphi \notin \mathfrak{Z}, \quad (14)$$

$$\Delta x|_{\varphi \in \Gamma} = B(\varphi)x + I(\varphi, u_{m-1}(\varphi)),$$

Згідно з теоремою 1, при кожному $m = 1, 2, \dots$ функція

$$\begin{aligned} u_m(\varphi) = & \int_{-\infty}^0 G(0, \tau, \varphi) \times \\ & \times f(\varphi + \omega\tau, u_{m-1}(\varphi + \omega\tau)) d\tau + \\ & + \sum_{t_i(\varphi) < 0} G(0, t_i(\varphi), \varphi) \times \\ & \times I(\varphi + \omega t_i(\varphi), u_{m-1}(\varphi + \omega t_i(\varphi))) \end{aligned} \quad (15)$$

визначає інтегральну множину системи (14). Рівномірну щодо $\varphi \in \mathfrak{Z}_m$ збіжність послідовності $u_m(\varphi)$ до граничної функції $u(\varphi)$ забезпечує нерівність (2) за умови малості константи Ліпшиця N .

Таким чином, справедливе наступне твердження.

Теорема 2. Нехай система рівнянь (1) така, що виконуються нерівності (2) і (10). Тоді при досить малій постійній Ліпшиця N система рівнянь (1) має асимптотично стійку інтегральну множину

$$x = u(\varphi), \quad u(\varphi + 2\pi) = u(\varphi),$$

де $u(\varphi)$ – кусково-неперервна з розривами першого роду на множині Γ функція така, що

$$\Delta u|_{\varphi \in \Gamma} = B(\varphi)u(\varphi) + I(\varphi, u(\varphi)).$$

Список використаних джерел

1. Асроров Ф.А. Обмежені розв'язки лінійних неоднорідних систем з імпульсною дією / Асроров Ф.А., Фекета П.В. // Наук. вісн. Ужгород. ун.-ту. Сер. матем. і інформ. 2010. – Вип.20. С. 4-12.
2. Асроров Ф.А. Функция Грина-Самойленко и существование интегральных множеств линейных расширений неавтономных систем / Асроров Ф.А., Перестюк Н.А. // Укр. мат. журн. 1994. т. 46, №8. С. 1067-1071.
3. Митропольский Ю.А. Исследование дихотомии линейных систем дифференциальных уравнений с помощью функций Ляпунова / Митропольский Ю.А., Самойленко А.М., Кулик В.Л. // – Киев: Наук. думка, 1990. 272 с.
4. Perestyuk N.A. Differential equations with impulsive effects: multivalued right-hand sides with discontinuous / Perestyuk N.A., Plotnikov V.A., Samoilenko A.M., Skripnik N.V. // De Gruyter 2011. – 307 p.
5. Самойленко А.М. Элементы математической теории многочастотных колебаний. Инвариантные торы / Самойленко А.М. // – М: Наука. Гл. ред. физ.-мат. лит., 1987. – 304 с.
6. Самойленко А.М. Дифференциальные уравнения с импульсным воздействием / Самойленко А.М., Перестюк Н.А. // К.: Вища шк. Головное изд-во, 1987. – 288 с.
7. Фекета П.В. Інтегральні множини розширень неавтономних систем на торі з імпульсними збуреннями / Фекета П.В., Асроров Ф.А. // Наук. вісн. Ужгород. ун.-ту. Сер. матем. і інформ. 2012. – Вип. 23. С. 125-132.
8. Samoilenko A.M. Impulsive differential equations / Samoilenko A.M., Perestyuk N.A. // Singapore-New Jersey-London-Hong Kong: World Scientific 1995. – 462 p.

References

1. ASROROV F.A., FEKETA P.V. (2010) *Bounded solutions of linear inhomogeneous systems with impulsive action* // *Bull. Uzhhorod University. Math. and Informatics Series*, 2010. – vol.20. p. 4-12.
2. ASROROV F.A., PERESTYUK N.A. (1994) *Green-Samoilenko function and existence of integral sets of linear extensions of non-autonomous systems* // *Ukr. Mat. Zhurn.* 1994. vol. 46, No. 8. p. 1067-1071.
3. MITROPOLSKY YU.A., SAMOILENKO A.M., KULYK V.L. (1990) *Investigation of dichotomy of linear systems of differential equations via Lyapunov functions.* – Kyiv: Nauk. Dumka.
4. PERESTYUK N.A., PLOTNIKOV N.A., SAMOILENKO A.M., SKRIPNIK N.V. (2011) *Differential equations with impulsive effects: multivalued right-hand sides with discontinuous.* De Gruyter.
5. SAMOILENKO A.M. (1987) *Elements of mathematical theory of multi-frequency oscillations. Invariant tori.* – M: Nauka..
6. SAMOILENKO A.M., PERESTYUK N.A. (1987) *Differential equations with impulsive actions.* – K.: Vyscha shkola.
7. FEKETA P.V., ASROROV F.A. (2012) *Integral sets of extensions of non-autonomous systems on torus with impulsive perturbations*// *Bull. Uzhhorod University. Math. and Informatics Series*, 2012. – vol.23. p. 125-132.
8. SAMOILENKO A.M., PERESTYUK N.A. (1995) *Impulsive differential equations.* Singapore-New Jersey-London-Hong Kong: World Scientific.

Надійшла до редколегії 11.06.2015

УДК 532.5

О. Л. Голиченко¹, аспірант
В. С. Малуґа², к. ф.-м. н., с. н. с.

Особливості перемішування рідини в скінченному циліндрі.

¹ Київський національний університет імені
Тараса Шевченка, 03680, м. Київ, пр-т. Глушко-
ва 4е, e-mail: Golichenko_O_L@mail.ru

² Інститут гідромеханіки НАН України, 03057,
м. Київ, вул. Желябова, 8/4,
e-mail: v_s_malyuga@ukr.net

O. L. Golichenko¹, graduate student,
V. S. Malyuga², PhD (Phys.-Math.), sen. res.

Mixing characteristics in a finite cylindrical container.

¹ Taras Shevchenko National University of Kyiv,
03680, Kyiv, Glushkova st., 4e,
e-mail: Golichenko_O_L@mail.ru

² Institute of hydromechanics, 03057, Kyiv,
Zhelyabova str., 8/4,
e-mail: v_s_malyuga@ukr.net

Розглядаються особливості перемішування рідини в тривимірних періодичних течіях. Показано, що природа перемішування рідини у тривимірних течіях має значно складніший характер порівняно з перемішуванням у двовимірних течіях. Проведено класифікацію періодичних точок. Показано, що якісне перемішування рідини можливе лише в течіях, що мають виключно ізольовані періодичні точки. Перемішування в течіях, що мають періодичні лінії, завжди має двовимірний характер, а отже отримання якісної мікстури в таких течіях неможливе. Перемішування в течіях з періодичними лініями і ізольованими періодичними точками ілюструються на прикладі трьох періодичних течій Стокса в скінченному циліндрі. Наведено перетин Пуанкаре для цих течій. Для моделювання перемішування пасивної домішки використовувалась чисельна техніка, що базується на трекінгу матеріальних поверхонь.

Ключові слова: течії Стокса, ламінарне перемішування, хаотична адвекція, періодичні лінії

The paper addresses the kinematic features of three-dimensional mixing in periodic flows. It is shown that the fluid mixing in three-dimensional flows has considerably more difficult characteristics as compared to the mixing in two-dimensional flows. A classification of periodic points, providing important information about the mixing properties of a flow, is presented. It is shown that qualitative three-dimensional mixing takes place only in flows that contain only isolated periodic points. Flows with periodic lines possess only quasi-two-dimensional mechanisms of chaotic advection and therefore the high quality of mixture cannot be reached in such the flows. Three examples of Stokes flow in a finite cylindrical cavity with discontinuous periodic motion of its end walls are used to illustrate the apparent differences between the mixing in flows with periodic lines and in flows with only isolated periodic points. The dynamics of the Poincare map is analysed for all the identified types of periodic points. It is shown that the stable and unstable manifolds of points on the periodic lines create two surfaces in the flow which intersect infinitely often. The Poincare map is presented for the flows under consideration. A numerical technique based on tracking of a material surface is presented to study the mixing process.

Key Words: Stokes flow; finite cylinder; laminar mixing; periodic lines.

Статтю представив д.ф.-м.н., проф. Жук Я.О.

Вступ

На відміну від двовимірних течій нестисливої рідини, для яких рівняння руху індивідуальної пасивної частинки рідини може завжди бути записане у гамільтоновій формі, а отже при описанні такої системи можуть бути застосовані добре відомі методи гамільтонової механіки, дослідження перемішування рідини у

тривимірних течіях все ще представляє значні труднощі. Наприклад, важливою інформацією, що говорить про розташування областей регулярного руху рідини, які є бар'єрами для процесу хаотичного перемішування, є інформація про розташування і класифікацію періодичних точок. У двовимірних течіях періодичні точки розділяють на три класи: еліптичні, гіперболічні і параболічні. Відомо, що острови регулярного

руху виникають лише довкола еліптичних точок. У тривимірних течіях класифікація періодичних точок є значно складнішою. Більше того, у тривимірних течіях періодичні точки можуть бути не лише ізольованими, а й об'єднуватись у періодичні лінії. В даній роботі проведено повну класифікацію періодичних точок у тривимірних течіях, виділено дев'ять типів періодичних точок і показано поведінку відображення Пуанкаре поблизу кожного типу точок. Показано, що острови регулярного руху можуть виникати лише довкола періодичних точок, об'єднаних в періодичні лінії. В течіях, де є лише ізольовані періодичні точки, островів регулярного руху бути не може, а отже процес хаотичного перемішування відбувається по всій області.

В даній статті ми розглядаємо процес перемішування в'язкої нестисливої рідини в скінченному циліндричному контейнері, якщо рух рідини викликано періодичним рухом торцевих стінок циліндра. На цьому прикладі ми наглядно демонструємо особливості перемішування рідини у тривимірних течіях, що суттєво відрізняються від детально вивчених особливостей перемішування у двовимірних течіях. З цією метою ми розробили техніку визначення періодичних структур, провели детальну класифікацію періодичних точок і визначили їх многовиди, побудували перетин Пуанкаре. Розрахунки проводились для трьох типових протоколів руху торцевих стінок циліндра. Дві течії мали періодичні лінії, а третя мала лише ізольовані періодичні точки. Наведені результати дуже наглядно демонструють, що лише в третій течії перемішування буде ефективним.

Класифікація періодичних точок та структура многовидів

Згідно з теоремою Брауера про нерухому точку будь-яке неперервне нестисливе відображення замкнутої області в себе в скінченновимірному евклідовому просторі має нерухому точку. Очевидно, що така нерухома точка відображення буде періодичною точкою течії. Розташування і тип періодичних точок в течії визначають характер, інтенсивність і якість перемішування рідини в течії.

Нехай в початковий момент часу $t=0$ пасивна матеріальна точка мала ейлерові координати $\mathbf{x}_0$. А в деякий момент часу t ця матеріальна точка вже має ейлерові координати $\mathbf{x}$. Такий процес перенесення пасивної

матеріальної точки можна представити оператором $\mathbf{x} = \Phi_t(\mathbf{x}_0)$. Тоді точка $\mathbf{P}$ є періодичною точкою потоку з періодом T , якщо $\mathbf{P} = \Phi_T(\mathbf{P})$. В термінах відображення Пуанкаре періодична точка з періодом T визначається як нерухома точка. Дійсно, якщо відображення Пуанкаре визначено як $\mathbf{x}_{n+1} = f(\mathbf{x}_n)$, де $f(\mathbf{x}) \equiv \Phi_T(\mathbf{x})$. То для періодичної точки з періодом T маємо $\mathbf{P} = f(\mathbf{P})$. Аналогічно, періодичні точки течії з періодом nT визначаються як нерухомі точки відображення Пуанкаре порядку n , тобто $\mathbf{P} = f^n(\mathbf{P})$. Класифікація періодичних точок подібна до класифікації особливих точок поля течії [1, 2]. У двовимірних течіях періодична точка може бути класифікована як еліптична, гіперболічна або параболічна в залежності від структури поля в околі періодичної точки [3]. У тривимірних течіях класифікація періодичних точок значно складніша. Опишемо класифікацію періодичних точок у тривимірних течіях більш детально. Нехай у початковий момент часу деяка матеріальна частинка близька до $\mathbf{P}$ знаходилась у позиції $\mathbf{P} + d\mathbf{x}_0$. Через період ця матеріальна частинка переходить у позицію $\mathbf{P} + d\mathbf{x}_1 = f(\mathbf{P} + d\mathbf{x}_0)$. Після лінеаризації отримуємо $d\mathbf{x}_1 = \mathbf{F}d\mathbf{x}_0$, де $\mathbf{F} = \partial f(\mathbf{x}) / \partial \mathbf{x} |_{\mathbf{x}=\mathbf{P}}$ – матриця Якобі в точці $\mathbf{P}$. Властивості стійкості відображення f визначаються властивостями матриці $\mathbf{F}$, які в свою чергу визначаються значеннями власних чисел матриці. Власні числа матриці є коренями рівняння $|\mathbf{F} - \lambda \mathbf{E}| = 0$, яке у тривимірному випадку приймає наступну форму:

$$\lambda^3 - J_1\lambda^2 + J_2\lambda - J_3 = 0, \quad (1)$$

де три інваріанти матриці визначені наступним чином $J_1 = \text{tr}(\mathbf{F})$, $2J_2 = \text{tr}^2(\mathbf{F}) - \text{tr}(\mathbf{F}^2)$, $J_3 = \det(\mathbf{F})$. Всі елементи матриці $\mathbf{F}$ обраховувались чисельно. При цьому умова нестисливості рідини $J_3 = 1$ використовувалась для контролю точності чисельних розрахунків. Чисельна похибка $|J_3 - 1|$ не перевищувала 10^{-3} поблизу поверхні циліндра і набувала значно менших значень всередині розрахункової області.

Очевидно, що рівняння (1) завжди має один дійсний додатний корінь. Нехай це буде λ_1 . З урахуванням цього факту рівняння (1) може бути представлене у вигляді

$$(\lambda - \lambda_1) \left(\lambda^2 + (\lambda_1 - J_1)\lambda + \frac{1}{\lambda_1} \right) = 0. \quad (2)$$

Два інших власних числа набувають значень:

$$\lambda_{2,3} = \frac{(J_1 - \lambda_1)}{2} \pm \sqrt{D}, \quad (3)$$

де дискримінант

$$D = \frac{(J_1 - \lambda_1)^2}{4} - \frac{1}{\lambda_1}. \quad (4)$$

Власні числа λ_2, λ_3 можуть набувати як дійсні так і комплексно-спряжені значення. Причому перехід відбувається при $D = 0$, тобто при $\lambda_2 = \lambda_3 = \pm 1/\sqrt{\lambda_1}$. Тепер очевидно, що у тривимірних течіях можна виділити дев'ять класів періодичних точок, залежно від значень λ_1 і D . Розглянемо класифікацію періодичних точок більш детально.

Якщо $\lambda_1 = 1$, то це означає, що в напрямку відповідного власного вектора $\mathbf{n}_1$ не відбувається розтягування або стискання. У канонічній системі координат $(\eta^{(1)}, \eta^{(2)}, \eta^{(3)})$, осі якої визначаються власними векторами, умова $\lambda_1 = 1$ означає, що $\eta^{(1)} = f(\eta^{(1)})$. Тоді для двох інших власних значень виконується умова $\lambda_2 \lambda_3 = 1$, тобто в канонічній площині $(\eta^{(2)}, \eta^{(3)})$ для відображення f виконується умова нестисливості. Таким чином, випадок $\lambda_1 = 1$ є по суті двовимірним. Такі періодичні точки можуть бути розбиті на три класи: еліптичні, гіперболічні та параболічні. Умова $\lambda_1 = 1$ виконується в точках періодичних ліній. Дійсно, на періодичній лінії в напрямку дотичної не відбувається розтягування або стискання, тобто напрямком дотичної до періодичної лінії відповідає напрямку власного вектора $\mathbf{n}_1$. Таким чином, класифікація точок періодичної лінії еквівалентна класифікації періодичних точок у двовимірних течіях. Тобто, періодичні лінії можна розбити на еліптичні і гіперболічні відрізки, що з'єднуються в параболічних точках. З (4) випливає, що точка періодичної лінії є еліптичною, якщо $-1 < J_1 < 3$. У цьому випадку $D < 0$, а отже власні числа $\lambda_{2,3}$ є комплексно-спряжені. Якщо $J_1 < -1$ або $J_1 > 3$, то $D > 0$, а отже така точка періодичної лінії є гіперболічною. У двовимірних системах еліптичні точки оточені островами регулярного руху. Відображення Пуанкаре в таких островах

представляє собою обертання довкола еліптичної періодичної точки. Таким чином, рідина всередині еліптичної області не покидає цю область і виявляється ізольованою від решти рідини в області перемішування. У тривимірних областях еліптичні ділянки періодичних ліній також оточені островами, що мають форму труб. З гіперболічними точками у двовимірних системах пов'язані стійкий і нестійкий многовиди, що представляють собою лінії, вздовж яких відбувається стискання та розтягування рідини. У тривимірному випадку стійкий і нестійкий многовиди, пов'язані з гіперболічними ділянками періодичної лінії, об'єднуються у поверхні. Таким чином, відображення Пуанкаре довкола періодичної лінії має по суті двовимірну природу.

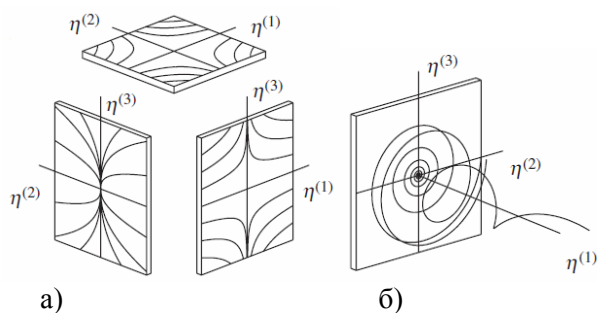


Рис. 1. Фазовий портрет періодичної точки: а) типу гіперболічний вузол ($D > 0$), б) типу гіперболічний фокус ($D < 0$).

Якщо $\lambda_1 \neq 1$, то така періодична точка є ізольованою. Природа відображення Пуанкаре в околі такої періодичної точки є суттєво тривимірною. Класифікація ізольованих точок буде значно складнішою, порівняно з класифікацією періодичних ліній. На рис. 1 а) показано фазовий портрет в канонічній системі координат $(\eta^{(1)}, \eta^{(2)}, \eta^{(3)})$ при умові $D > 0$. У цьому випадку всі три власні числа λ_i набувають дійсні значення. Фазові портрети в координатних площинах представляють собою два сідла і вузол. На рис. 1 б) показано фазовий портрет відображення в околі періодичної точки при умові $D < 0$. У цьому випадку власні числа $\lambda_{2,3}$ набувають комплексно-спряжені значення, що відповідають за обертання у площині $(\eta^{(2)}, \eta^{(3)})$. У виродженому випадку $D = 0$ маємо два рівні власні числа $\lambda_2 = \lambda_3 = \pm 1/\sqrt{\lambda_1}$. А отже у площині $(\eta^{(2)}, \eta^{(3)})$ фазовий портрет представляє собою дикритичний або вироджений вузол. Таким чином, при $\lambda_1 \neq 1$ періодична точка є

ізолюваною і може бути класифікована як гіперболічний вузол або гіперболічний фокус.

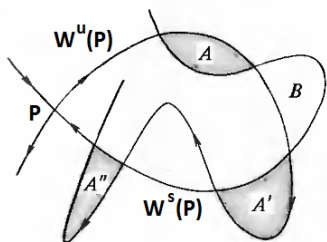


Рис. 2. Стійкий та нестійкий многовиди і гомоклінічні точки їх перетину (площі $A = A' = A'' = \dots$).

Очевидно, що фазовий портрет ізолюваних періодичних точок не має замкнутих ліній або поверхонь, а отже довкола таких точок не може утворюватись ізолюваних островів регулярного руху подібних до тих, що утворюються довкола еліптичних точок періодичних ліній. У течії, що має лише ізолювані періодичні точки і не має періодичних ліній, рідина повинна перемішуватись по всій області течії.

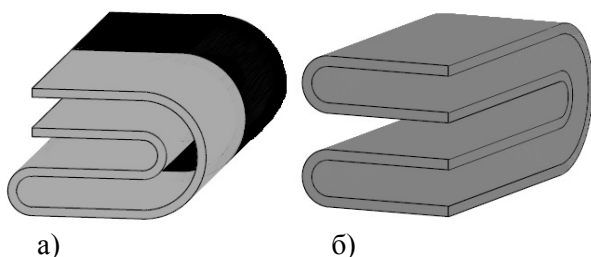


Рис. 3. Схематичне зображення відображення підкови: а) у квазі-двовимірній течії, що має періодичні лінії, б) у реально тривимірній течії.

Будь-яка ізолювана періодична гіперболічна точка пов'язана у просторі з поверхнею і лінією, що називаються многовидами. Стійкий многовид $W^s(\mathbf{P})$ – це множина точок, які в процесі відображення Пуанкаре прямують до періодичної точки $\mathbf{P}$. Тобто для будь-якої точки $\mathbf{x} \in W^s(\mathbf{P})$ маємо $f^n(\mathbf{x}) \rightarrow \mathbf{P}$ при $n \rightarrow \infty$. Нестійкий многовид $W^u(\mathbf{P})$ – це множина точок, які прямують до $\mathbf{P}$ при зворотньому процесі, тобто для довільної точки $\mathbf{x} \in W^u(\mathbf{P})$ маємо $f^{-n}(\mathbf{x}) \rightarrow \mathbf{P}$ при $n \rightarrow \infty$. У двовимірному випадку стійкий та нестійкий многовиди – це дві лінії, що перетинаються безліч разів (рис. 2). Точки перетину називаються гомоклінічними (при перетині стійкого і нестійкого многовидів однієї періодичної точки) або гетероклінічною (при перетині стійкого і нестійкого многовидів

двох різних періодичних точок). Оскільки точок перетину стійкого і нестійкого многовидів безліч, вони ущільнюються при наближенні до періодичної точки. Таким чином, многовиди формують складну шарову структуру. Завдяки цій властивості перемішування рідини представляє собою видовження рідкої плями вздовж нестійкого многовиду і багаторазове складання при переході через гомоклінічні точки. Таке видовження і складання рідкої плями називається відображенням підкови. Очевидно, що у тривимірних течіях, що мають періодичні лінії, механізм перемішування подібний до того, що має місце у двовимірних течіях. Відображення підкови такої течії схематично показано на рис. 3 а). У такому квазі-двовимірному відображенні темна рідина не змішується зі світлою. Видовження і складання рідини відбувається лише в одному напрямку. Хоча таке відображення підкови також приводить до складної шарової структури, однак наявність лише двовимірного механізму перемішування не дозволяє одержати якісну однорідну мікстуру в усій тривимірній області течії. На рис. 3 б) показано відображення підкови для реально тривимірної течії. Видовження і складання рідини по черзі в двох напрямках веде до ефективного перемішування в усій області течії.

Чисельні результати

Для розрахунків було використано розв'язок задачі Стокса, представлений в [4]. Наведені вище теоретичні міркування проілюструємо на прикладі трьох можливих періодичних течій, що генеруються у скінченному циліндрі. В усіх трьох випадках течії в циліндрі генеруються періодичним рухом торцевих поверхонь циліндра. Нехай циліндр займає область $0 < r < 1$, $0 < \theta < 2\pi$, $-h < z < h$. Бічна поверхня циліндра $r=1$ нерухома, а торцеві сторони рухаються періодично. У загальному випадку рух торцевих сторін можна записати у наступній формі:

$$\mathbf{u} = u_-(t)\mathbf{e}_x + v_-(t)\mathbf{e}_y, \quad z = -h, \quad (5)$$

$$\mathbf{u} = u_+(t)\mathbf{e}_x + v_+(t)\mathbf{e}_y, \quad z = h. \quad (6)$$

Розглянемо три типові протоколи руху торцевих сторін:

- Задача 1. Рухається лише верхня торцева стінка. Нижня залишається нерухомою. Рух верхньої стінки є зигзагоподібним: $u_+ = V$, $v_+ = u_- = v_- = 0$ при $0 < t < T/2$ і $v_+ = V$,

$u_+ = u_- = v_- = 0$, при $T/2 < t < T$. Тобто, протягом першої половини періоду верхня торцева стінка рухається зі сталою швидкістю в напрямку осі x , протягом другої половини періоду – в напрямку осі y .

- Задача 2. Верхня і нижня стінки рухаються по черзі в протилежних напрямках паралельно осі x . Тобто, $u_+ = V$, $v_+ = u_- = v_- = 0$ при $0 < t < T/2$ і $u_- = -V$, $v_- = u_+ = v_+ = 0$ при $T/2 < t < T$.
- Задача 3. Протокол руху стінок в цій задачі складається з трьох кроків: $u_+ = V$, $v_+ = u_- = v_- = 0$ при $0 < t < T/3$, $u_- = -V$, $v_- = u_+ = v_+ = 0$ при $T/3 < t < 2T/3$, $v_+ = V$, $u_+ = u_- = v_- = 0$ при $2T/3 < t < T$.

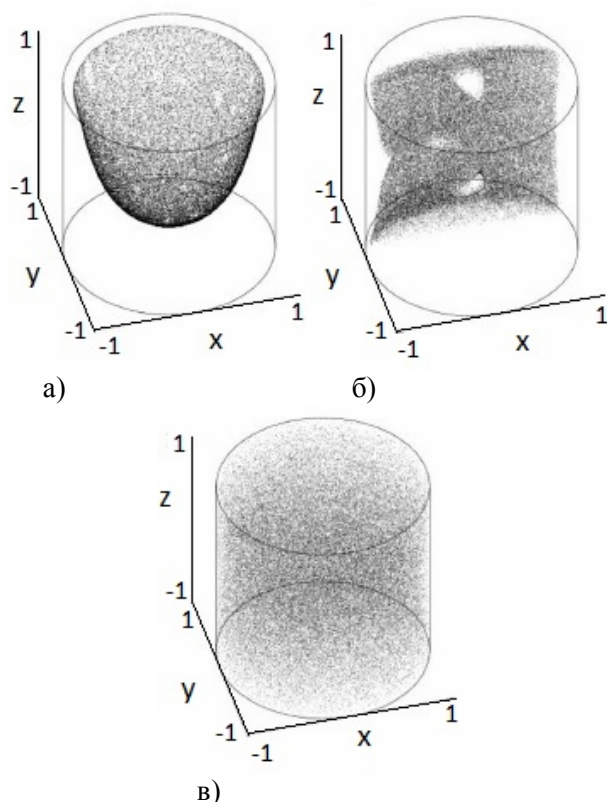


Рис. 4. Перетин Пуанкаре: а) задача 1, б) задача 2, в) задача 3.

Течії у перших двох задачах мають періодичні лінії, а течія задачі 3 має лише ізольовані періодичні точки (див. [2]). Отже, згідно викладеної у попередньому розділі теорії перемішування у таких течіях має двовимірний характер. Хоча режим перемішування рідини може бути хаотичним, в силу його двовимірної природи пляма пасивної домішки не буде перемішуватись по всій області. Крім того, в таких течіях можуть виникати і зони

регулярного перемішування, що утворюють острови довкола еліптичних ділянок періодичних ліній, що також заважають перемішуванню рідини. Течія задачі 3 має лише ізольовані періодичні точки, а отже довкола таких точок не можуть утворюватись острови регулярного руху. Пляма пасивної домішки повинна перемішуватись по всій області течії. Дійсно, наведені теоретичні висновки підтверджено на рис. 4, на якому наведено перетин Пуанкаре для всіх трьох задач. Очевидно, що лише в третій задачі матеріальна точка блукає по всій області течії. В перших двох задачах наявність періодичних ліній спричинює двовимірний характер перемішування.

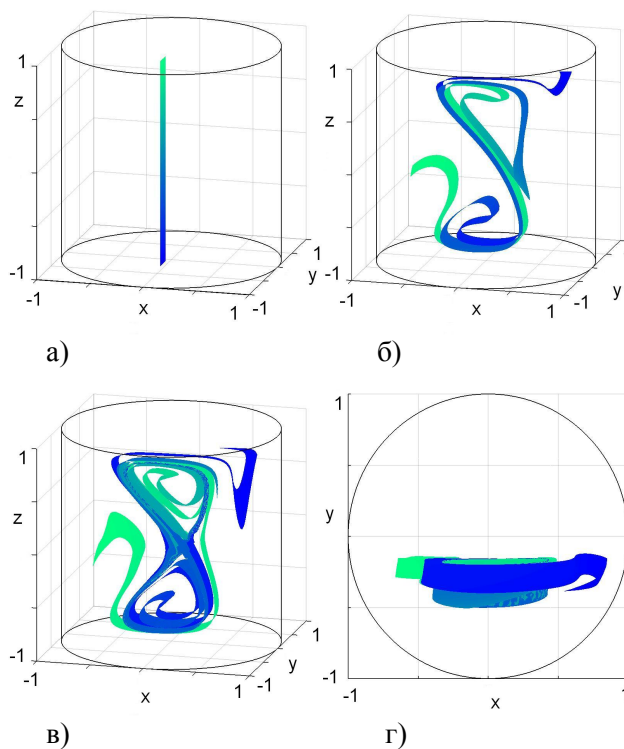


Рис. 5. Трекінг поверхні пасивної домішки (задача 2): а) початкова позиція $t = 0$, б) $t = 4T$, в) $t = 10T$, г) вигляд згори.

Для відстеження руху пасивної домішки було використано техніку, що базується на трекінгу матеріальних поверхонь. Рідка поверхня представлялась неструктурованою триангулярною сіткою. Результати трекінгу поверхні в Задачі 2 показано на рис. 5. Видно, що рідка поверхня після 10 періодів представляє собою складну видовжену складчасту структуру, як і повинно бути при хаотичному перемішуванні рідини. Однак, оскільки така течія має періодичну лінію, пасивна домішка не

розмивається по всій області течії, а залишається зосередженою в досить вузькій області паралельній осі X . На рис. 6 представлено результати трекінгу поверхні в Задачі 1. Видно, що пасивна домішка також не розмивається по всій області течії, а утворює складчасту структуру в деякій локальній області, обмеженій поверхнею обертання. На рис. 7 представлено результати трекінгу поверхні в Задачі 3. Початкова поверхня

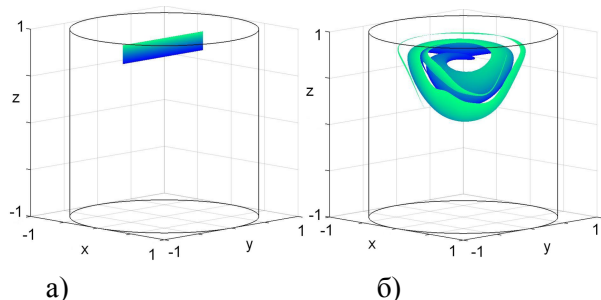


Рис. 6. Трекінг поверхні пасивної домішки (задача 1): а) $t = 0$, б) $t = 17T$.

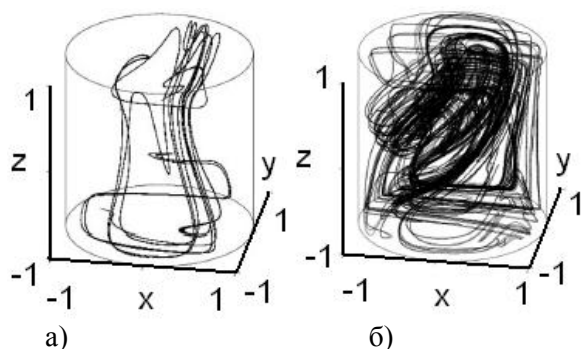


Рис. 7. Трекінг поверхні пасивної домішки (задача 3): а) $t = 6T$, б) $t = 10T$.

представляла собою сферу з центром в ізольованій періодичній точці. Лише у цьому випадку пасивна домішка досить швидко заповнює всю область течії.

Висновки

У тривимірних течіях періодичні точки проявляють значно більшу різноманітність порівняно з двовимірними течіями. Більше того, у тривимірних течіях періодичні точки можуть об'єднуватись у періодичні лінії. В роботі наведено загальну класифікацію періодичних точок у тривимірних течіях, проаналізовано їх фазові портрети та вплив на процес перемішування рідини. Показано, що класифікація точок періодичних ліній подібна до класифікації періодичних точок у двовимірних течіях. Тому процес перемішування в таких течіях має таку ж саму природу, що і процес перемішування у двовимірних течіях. Будь-який матеріальний об'єм зазнає видовження, згинання і складення лише в одному напрямку. Утворена таким чином складчаста структура подібна до тих, що утворюються при перемішуванні у двовимірних течіях. Пляма пасивної домішки не розмивається по всій області течії, а отже отримати якісну мікстуру в таких течіях неможливо. Ефективний міксинг можливий лише у тривимірних течіях, що мають лише ізольовані періодичні точки. Лише у таких течіях пляма пасивної домішки розмивається по всій області течії.

References

1. Chong M.S. A general classification of three-dimensional flow fields / M.S. Chong, A.E. Perry, B.J. Cantwell // *Phys. Fluids A*. – 1990. – 2. – P. 765-776.
2. Malyuga V.S. Mixing in the Stokes flow in a cylindrical container / V.S. Malyuga, V.V. Meleshko, M.F.M. Speetjens, H.J.H. Clercx, G.J.F. van Heijst // *Proc. R. Soc. Lond. A*. – 2002. – 458. – P. 1867-1885.
3. Ottino J.M. The kinematics of mixing: stretching, chaos, and transport / Cambridge University Press, 1989. – p. 364.
4. Meleshko V.V. Steady Stokes flow in a finite cylinder / V.S. Malyuga, A.M. Gomilko // *Proc. Roy. Soc. Lond.* – 2000. – 456. – P. 1741-1758.
5. CHONG, M.S., PERRY, A.E., CANTWELL, B.J. (1990) A general classification of three-dimensional flow fields. *Phys. Fluids A*. 2. p. 765-776.
6. Malyuga, V.S., Meleshko, V.V., Speetjens, M.F.M., Clercx, H.J.H., van Heijst, G.J.F. (2002) Mixing in the Stokes flow in a cylindrical container. *Proc. R. Soc. Lond. A*. 458. p. 1867-1885.
7. OTTINO, J.M. (1989) *The kinematics of mixing: stretching, chaos, and transport*. Cambridge University Press.
8. MELESHKO, V.V. MALYUGA, V.S., GOMILKO, A.M. (2000) Steady Stokes flow in a finite cylinder. 456. p. 1741-1758.

Надійшла до редколегії 03.06.15

УДК 539.3

Горошко О. О.¹, д.ф.-м.н., проф.,
Кикоть С. В.², к.ф.-м.н., доц.

Коливання гнучкого трубопроводу на пружній основі при закритичних швидкостях транспортування рідини

¹ Київський національний університет імені
Тараса Шевченка, 83000, м. Київ, пр-т.
Глушкова 4д.

² Національний транспортний університет,
01010, м. Київ, вул. Суворова, 1,
e-mail: a.goroshko@gmail.com,
s.v.kykot@gmail.com

Oleg O. Goroshko¹, Dr.Sc. (Phys.-Math.), Prof.,
Sergiy V. Kykot², Ph.D., (Phys.-Math.), Ass. Prof.

Oscillations of a flexible conduit on an elastic foundation at overcritical values of transported fluid

¹ Taras Shevchenko National University of Kyiv,
83000, Kyiv, Glushkova st., 4d,

² National Transport University, 01010, Kyiv,
Suvorova 1,
e-mail: a.goroshko@gmail.com,
s.v.kykot@gmail.com

Досліджуються коливання гнучкого трубопроводу на пружній основі з шарнірним закріпленням кінців під впливом потоку рідини, що рухається з швидкістю значення якої перевищують значення другої критичної швидкості. На основі двохвильового представлення розв'язку, побудовано рівняння для визначення власних частот та ексцесів, якими визначаються деформації форм коливань при закритичних швидкостях руху потоку рідини. Встановлено, що в закритичній області транспортування рідини має місце скінчене число тонів власних коливань.

Ключові слова: гнучкий трубопровід, рухоме навантаження, двохвильовий коливний процес, частота коливань, ексцес, закритична швидкість.

The paper considers oscillations of a flexible conduit with fluid flow, which is located on elastic foundation with pin jointed ends. The fluid flow velocity is more than the second critical velocity value. The inertia forces in full of the moving load are taken into consideration in the mathematical model construction.

Object of study – a flexible conduit with fluid flow, which is located on an elastic foundation.

Purpose – investigation of influence of the fluid flow velocity on the natural-vibration frequencies and determination the conditions of loss of stability of the conduit on an elastic foundation.

Method of study is based on extraction and approximate investigation of one-frequency two-wave oscillation processes in mechanical structures with moving loads.

Based on two-wave representation of the solution, equations for determining the natural frequencies and excesses which define deformation of the oscillations forms depending on the overcritical values of the fluid flow velocity are constructed.

The results of the paper can be used for further developments of direct methods for constructing approximate solutions of complicated dynamic systems.

Forecast assumptions about the object of study – taking into consideration of the inertia forces of moving load in full gives an opportunity qualitatively reveal the basic characteristics of dynamics of a structure with a moving load.

Key words: conduit, moving loading, double wave oscillation process, oscillation frequency, excesses, overcritical velocity.

Статтю представив д.ф.-м.н., проф. Жук Я. О.

У науковій монографії [1] та статті [3] доведено, що коливання балок та гнучких об'єктів скінченної довжини, які лежать на пружній вінкерівській основі і відбуваються під дією рухомих рівномірно розподілених інерційних навантажень (потoku рідини), що рухаються з постійною швидкістю, моделюється

суперпозицією двох груп стоячих хвиль з однаковими частотами, але з різними формами і фазами.

Дослідження динаміки трубопроводу при закритичних швидкостях транспортування рідини розглянуто у роботі [4].

Метою даної роботи є вивчення динаміки коливань гнучкого трубопроводу на пружній основі при закритичних швидкостях руху потоку рідини.

Постановка задачі та методика розв'язання. Розглянемо особливості коливань одновимірного гнучкого трубопроводу на пружній основі під впливом потоку рідини, що рухається з швидкістю $v > v_{kp2}$, значення якої перевищують критичні [3].

Диференціальне рівняння поперечних коливань гнучкого трубопроводу на пружній основі з рухомою рідиною при повному врахуванні сил інерції рухомого навантаження в даному випадку набуває вигляду

$$\rho \frac{\partial^2 u}{\partial t^2} + 2\rho_2 v \frac{\partial^2 u}{\partial x \partial t} - (T - \rho_2 v^2) \frac{\partial^2 u}{\partial x^2} + ku = 0, \quad (1)$$

де $u(x, t)$ – поперечні відхилення труби;

T – її поздовжній натяг; $\rho = \rho_1 + \rho_2$ – сума розподілених мас труби і рухомої рідини;

v – швидкість транспортування потоку рідини;

k – коефіцієнт жорсткості опорної основи.

Граничні умови, які не залежать від того, контактує трубопровід з поверхнею чи ні, а визначаються способами закріплення та навантаження кінців. Наприклад, розглянемо випадок шарнірного закріплення на краях гнучкого трубопроводу

$$u(0, t) = 0, \quad u(l, t) = 0. \quad (2)$$

Наявність гіроскопічних сил, представлених членами зі змішаними похідними в рівняннях динаміки трубопроводів з рухомою рідиною, не дають змоги застосувати методи розділення змінних та прямі методи математичної фізики в класичній формі до дослідження коливного руху.

При докритичних швидкостях руху рідини рівняння (1) є рівнянням гіперболічного типу, а при закритичних швидкостях руху рідини рівняння (1) перетворюється в рівняння еліптичного типу.

У роботі [3] для коливання гнучкого трубопроводу на пружній основі під впливом потоку рідини, що рухається з постійною докритичною швидкістю $v < v_{kp2}$, встановлено значення трьох критичних швидкостей руху рідини

$$v_{kp1} = \sqrt{\frac{T}{\rho_2}}, \quad v_{kp2} = \sqrt{\frac{T(\rho_1 + \rho_2)}{\rho_1 \rho_2}} = v_{kp1} \sqrt{1 + \frac{\rho_2}{\rho_1}},$$

$$v_{kp3} = \sqrt{\frac{T}{\rho_2} + \frac{k}{\rho_2 \lambda_n^2}} = v_{kp1} \sqrt{1 + \frac{k}{T} \left(\frac{l}{\pi n} \right)^2};$$

а також умови $v_{kp1} < v_{kp3} < v_{kp2}$, при

$$k < \frac{\rho_2}{\rho_1} \left(\frac{n\pi}{l} \right)^2 T, \quad n = 1, 2, 3, \dots, \infty.$$

Для розв'язання крайової задачі (1), (2) будемо шукати розв'язок рівняння (1) в комплексному вигляді [1,2]

$$u(x, t) = aX(x)e^{i(\omega t + \gamma x)}. \quad (3)$$

Підставивши розв'язок у формі (3) в рівняння (1) та, відокремивши дійсну й уявну частини, отримаємо систему рівнянь:

$$\begin{aligned} [(T - \rho_2 v^2)]X''(x) + \\ + [(\rho_1 + \rho_2)\omega^2 + 2\rho_2 v \omega \gamma - (T - \rho_2 v^2)\gamma^2 - k]X(x) = 0, \\ [\rho_2 v \omega - (T - \rho_2 v^2)\gamma]X'(x) = 0. \end{aligned} \quad (4)$$

Рівняння (4) дають змогу побудувати власні функції $X_n(x)$, знайти частоти власних коливань ω_n та ексцеси γ_n , що характеризують зміну форм власних коливань під дією рухомого навантаження.

Із другого рівняння системи (4) знаходимо значення параметра γ за формулою

$$\gamma = \frac{\rho_2 v \omega}{T - \rho_2 v^2} = -\frac{\rho_2 v \omega}{\rho_2 v^2 - T}. \quad (5)$$

Виключивши з першого рівняння системи (4) параметр γ за допомогою формули (5), отримаємо звичайне диференціальне рівняння для визначення власних чисел та форм

$$X''(x) + \lambda^2 X(x) = 0, \quad (6)$$

де

$$\lambda^2 = \frac{k}{\rho_2 v^2 - T} - \omega^2 \frac{\rho_1 \rho_2 v^2 - T(\rho_1 + \rho_2)}{(\rho_2 v^2 - T)^2}.$$

При нульових граничних умовах із (6) знаходимо

$$X_n(x) = \sin \lambda_n x, \quad \text{де } \lambda_n = \frac{\pi n}{l} \quad (n = 1, 2, 3, \dots, \infty).$$

Цим параметрам відповідає частинний комплексний розв'язок на частоті n -го тону

$$u_n(x, t) = a_n \sin \lambda_n x [\cos \omega_n t \cos \gamma_n x - \sin \omega_n t \sin \gamma_n x + i(\sin \omega_n t \cos \gamma_n x + \cos \omega_n t \sin \gamma_n x)] \quad (7)$$

У відповідності з (5) – (6) визначаємо

$$\omega_n = \frac{n\pi}{l} \sqrt{\frac{(\rho_2 v^2 - T) \left(T + k \left(\frac{l}{n\pi} \right)^2 - \rho_2 v^2 \right)}{\rho_1 \rho_2 v^2 - T(\rho_1 + \rho_2)}},$$

$$\gamma_n = -\frac{n\pi}{l} \rho_2 v \sqrt{\frac{T + k \left(\frac{l}{n\pi} \right)^2 - \rho_2 v^2}{(\rho_2 v^2 - T)(\rho_1 \rho_2 v^2 - T(\rho_1 + \rho_2))}},$$

$$n = 1, 2, 3, \dots, m$$

Формули (8) встановлюють залежність частоти власних коливань ω та параметра γ від швидкості руху інерційного навантаження.

Дійсних значень частоти набувають в закритичній області при $v_{kp2} < v < v_{kp3}$, або

$$v_{kp1} \sqrt{1 + \frac{\rho_2}{\rho_1}} < v < v_{kp1} \sqrt{1 + \frac{k}{T} \left(\frac{l}{\pi n} \right)^2},$$

звідки одержуємо умови: $\frac{\rho_2}{\rho_1} < \frac{k}{T} \left(\frac{l}{\pi n} \right)^2$, та

$$k > \frac{\rho_2}{\rho_1} \left(\frac{n\pi}{l} \right)^2 T, \quad n = 1, 2, 3, \dots, m.$$

Внаслідок лінійності задачі частинними розв'язками рівняння (1) будуть як дійсна, так і уявна частини (7)

$$u_{n1}(x, t) = a_n \sin \lambda_n x (\cos \omega_n t \cos \gamma_n x - \sin \omega_n t \sin \gamma_n x),$$

$$u_{n2}(x, t) = a_n \sin \lambda_n x (\sin \omega_n t \cos \gamma_n x + \cos \omega_n t \sin \gamma_n x). \quad (9)$$

Щодо формул (9) можемо зауважити, що розв'язок $u_{n2}(x, t)$ відрізняється від $u_{n1}(x, t)$

лише запізненням на $\Delta t = \frac{2\pi}{\omega_n}$.

Для побудови графічних залежностей, як приклад, приймалися такі параметри системи:

$$\rho_1 = \rho_2 = \rho/2, \quad l = 5 \text{ м}, \quad T/\rho = \Omega^2 (l/\pi)^2, \quad \Omega = \sqrt{2\pi} \text{ с}^{-1}.$$

На рисунках 1 та 2 показано залежності частоти власних коливань ω_n та параметра γ_n від швидкості руху потоку рідини v , що визначаються формулами (8).

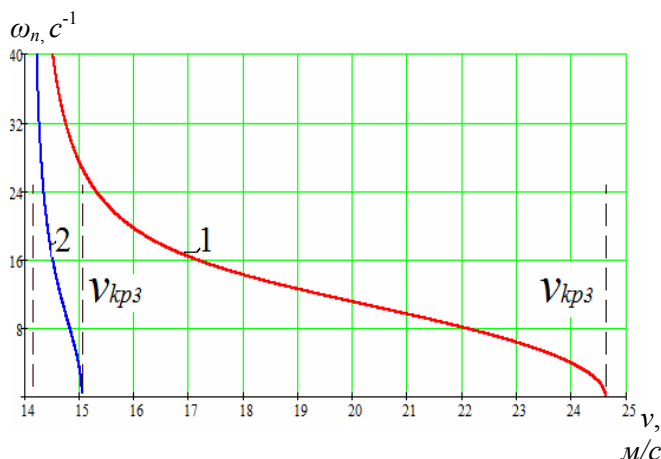


Рис. 1. Залежність частоти від значення закритичної швидкості руху рідини

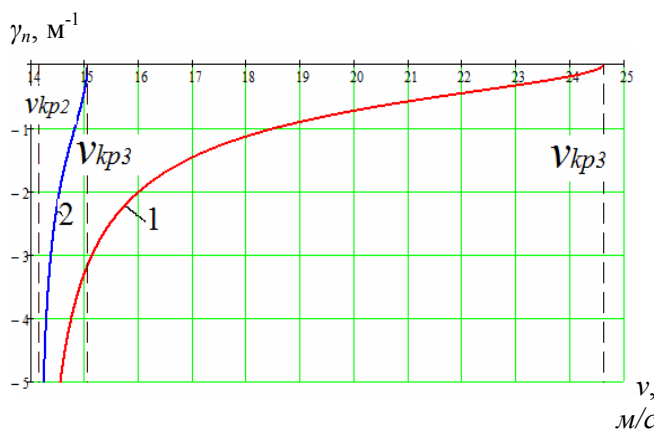


Рис. 2. Залежність експесу від значення закритичної швидкості руху рідини

Загальний розв'язок рівняння коливання трубопроводу (1), який будується у формі

$$u(x, t) = \sum_{n=1}^m u_n(x, t) \quad \text{для закритичних значень}$$

швидкості руху потоку рідини $v_{kp2} < v < v_{kp3}$ набуває вигляду

$$u(x, t) = \sum_{n=1}^m a_n \left(\sin \frac{n\pi x}{l} \cos \gamma_n x \cos(\omega_n t + \varphi_n) - \sin \frac{n\pi x}{l} \sin \gamma_n x \sin(\omega_n t + \varphi_n) \right). \quad (10)$$

Загальний розв'язок (10) перепишемо у двоххвильовій формі

$$u(x, t) = \sum_{n=1}^m a_n (\Phi_n \cos(\omega_n t + \varphi_n) - \Psi_n \sin(\omega_n t + \varphi_n)), \quad (11)$$

де

$$\begin{aligned} \Phi_n(x) &= \sin \frac{n\pi x}{l} \cos \gamma_n x, \\ \Psi_n(x) &= \sin \frac{n\pi x}{l} \sin \gamma_n x. \end{aligned} \quad (12)$$

Маємо дві різні форми стоячих хвиль, утворені деформуванням форм власних коливань $\sin \frac{n\pi x}{l}$, зсувами, що визначаються параметром γ_n , але однаковими частотами. Параметр γ_n це своєрідний ексцес (переки́с) форм коливань, які задовольняють одні й ті ж граничні та початкові умови, але зміщені за фазою на $\frac{\pi}{2}$. При нульовій швидкості руху навантаження ці форми збігаються, тобто має місце класичний підхід до розділення змінних.

Після нескладних тригонометричних перетворень розв'язок (11) зводиться до вигляду

$$u(x, t) = \sum_{n=1}^m a_n \sin \frac{n\pi x}{l} \cos(\omega_n t + \gamma_n x + \varphi_n). \quad (13)$$

Висновки. Отримані розрахункові формули та графічні залежності дають змогу зробити певні висновки щодо коливного процесу гнучкого трубопроводу при закритичних швидкостях транспортування рідини.

У закритичній області швидкостей руху потоку рідини коливання гнучкого трубопроводу мають двоххвильовий характер.

Залежно від співвідношень величин: швидкості транспортування рідини v , поздовжнього натягу T , та коефіцієнта жорсткості основи k , – при закритичних швидкостях транспортування рідини має місце скінчене число тонів власних коливань.

Список використаних джерел

1. Горошко О. О. Двоххвильові процеси в механічних системах / Горошко О. О., Дем'яненко А. Г., Кіба С. П. – К.: Либідь, 1991. – 188 с.
2. Горошко О. О. Дослідження одночастотних режимів в двоххвильових системах (в системах з рухомим навантаженням) / О. О. Горошко // Вісник Київського національного університету імені Тараса Шевченка. Серія : фіз.-мат. науки. – 2002. – № 2. – С. 93–98.
3. Кикоть С. В. Коливання гнучкого трубопроводу на пружній основі з рухомою рідиною. / С. В. Кикоть // Вісник НТУ: В 2-х частинах: Ч. 2 – К.: НТУ – 2012. – Вип. 26 – С. 559 – 563.
4. Лимарченко О. С. Поведінка трубопроводу при закритичних швидкостях течії рідини / О. С. Лимарченко, О. П. Тімохін // Вісник Київського національного університету імені Тараса Шевченка. Серія : фіз.-мат. науки. – 2013. – № 3. – С. 173–175.

References

1. GOROSHKO O., KIBA S., DEM'YANENKO A. (1991) Two-waves processes in mechanical systems. Kyiv: Lybid.
2. GOROSHKO O. (2002) An investigation of the one-frequency regimes in two-waves systems (in the systems with a moving loading). Bulletin of Taras Shevchenko National University of Kyiv. Series: Physics & Mathematics.(2). p. 93–98.
3. KYKOT' S. (2012) Oscillations of a flexible conduit on an elastic foundation with moving fluid *Visnyk NTU: In two parts*. Kyiv: NTU. 26(2). p. 559–563.
4. LIMARCHENKO O. (2013) Behavior of pipeline for supercritical velocities of liquid flow. Bulletin of Taras Shevchenko National University of Kyiv. Series: Physics & Mathematics.(3). p. 173–175.

Надійшла до редколегії 29.05.2015

УДК 517.9

Н.Л. Денисенко, к.ф.-м.н.

**Про періодичні розв'язки систем
нелінійних диференціально-
функціональних рівнянь з малим
параметром**

Національний Технічний Університет України
"Київський політехнічний інститут", Україна, 03056,
Київ, пр. Перемоги, 37,
e-mail: natalia_den@bigmir.net

N.L. Denysenko, Ph.D.

**About periodic solutions of systems of
nonlinear differential-functional equations
with a small parameter**

National Technical University of Ukraine "Kyiv
Polytechnic Institute 37, Prospect Peremohy, 03056, Kyiv-
56, Ukraine,
e-mail: natalia_den@bigmir.net

В даній роботі отримано умови існування та єдиності періодичних розв'язків широкого класу систем нелінійних диференціально-функціональних рівнянь з відхиленнями аргументу і малим параметром, запропоновано метод побудови таких розв'язків та досліджено їх властивості.

Ключові слова: диференціально-функціональні рівняння, відхилення аргумента.

Differential-functional equations with deviations of argument were the main subject of research of many mathematicians, and they are widely used in various fields of science and technology. Extensive application, as well as the diversity and complexity of the problems that arise in their research, are stimulating the research of various aspects the theory of such equations. There are currently a large number of publications in which a number of actual problems the theory of differential-functional equations with deviations of argument have been investigated, in particular for differential-functional equations with linearly transformed argument. Among them a special place is occupied by papers in which the questions the existence of continuously differentiable, bounded and periodic solutions such equations have been studied and their properties have been investigated. In this paper we obtain the sufficient conditions for existence and uniqueness of continuous periodic solutions of systems of nonlinear differential-functional equations with deviations of argument and a small parameter. We propose the method of construction such solutions for the given systems in case when the parameter is equal to zero and when it is different from zero. Furthermore, we investigated properties of such solutions when parameter tends to zero.

Key Words: differential-functional equations, deviations of argument.

Статтю представив академік НАНУ, доктор фіз.-мат. наук, проф. Перестюк М. О.

1. Постановка задачі. Дана стаття присвячена дослідженню питання існування та властивостей періодичних розв'язків систем нелінійних диференціально-функціональних рівнянь вигляду

$$\begin{aligned} \dot{x}(t) = & Ax(t) + \\ & + f(t, x(t), x(\lambda_1 t + \psi_1(t, \varepsilon)), \dots, x(\lambda_k t + \psi_k(t, \varepsilon))) + \\ & + \varepsilon F(t, x(t), x(\lambda_1 t + \psi_1(t, \varepsilon)), \dots, \\ & \dots, x(\lambda_k t + \psi_k(t, \varepsilon)), \varepsilon), \end{aligned} \quad (1)$$

де $\lambda_i \in \mathbb{N}$, $i = \overline{1, k}$; ε — достатньо малий невід'ємний скалярний параметр ($\varepsilon \in I = [0, \varepsilon_0]$, ε_0 — достатньо мале); $t \in \mathbb{R} = (-\infty, +\infty)$; A — дійсна стала $(n \times n)$ -матриця; $\psi_i : \mathbb{R} \times I \rightarrow \mathbb{R}$, $i = \overline{1, k}$, $f : \mathbb{R} \times \mathbb{R}^n \times \dots \times \mathbb{R}^n \rightarrow \mathbb{R}^n$,

$$F : \mathbb{R} \times \mathbb{R}^n \times \dots \times \mathbb{R}^n \times I \rightarrow \mathbb{R}^n.$$

Різні випадки таких рівнянь досліджувались багатьма математиками і на сьогодні існує велика кількість робіт, які одержані при вивченні питань їх теорії [1–4]. У даній роботі узагальнюються результати роботи [4] для систем рівнянь (1).

2. Основний результат. Нехай виконуються умови:

1*) числа λ_i , $i = \overline{1, k}$, є цілі додатні;

2*) всі компоненти вектор-функцій $f(t, x_0, x_1, \dots, x_k)$, $F(t, x_0, x_1, \dots, x_k, \varepsilon)$ та функції $\psi_i(t, \varepsilon)$, $i = \overline{1, k}$, є неперервними за всіма змінними і T -періодичними за t функціями, тобто

$$\begin{aligned} f(t+T, x(t), x(\lambda_1 t + \psi_1(t, \varepsilon)), \dots, x(\lambda_k t + \psi_k(t, \varepsilon))) &\equiv \\ \equiv f(t, x(t), x(\lambda_1 t + \psi_1(t, \varepsilon)), \dots, x(\lambda_k t + \psi_k(t, \varepsilon))), \\ F(t+T, x(t), x(\lambda_1 t + \psi_1(t, \varepsilon)), \dots, x(\lambda_k t + \psi_k(t, \varepsilon)), \varepsilon) &\equiv \\ \equiv F(t, x(t), x(\lambda_1 t + \psi_1(t, \varepsilon)), \dots, x(\lambda_k t + \psi_k(t, \varepsilon)), \varepsilon), \end{aligned}$$

$$\psi_i(t+T, \varepsilon) \equiv \psi_i(t, \varepsilon), \quad i = \overline{1, k};$$

$$3^*) |f(t, \tilde{x}_0, \tilde{x}_1, \dots, \tilde{x}_k) - f(t, \tilde{\tilde{x}}_0, \tilde{\tilde{x}}_1, \dots, \tilde{\tilde{x}}_k)| \leq l_1 \sum_{i=0}^k |\tilde{x}_i - \tilde{\tilde{x}}_i|,$$

$$|F(t, \tilde{x}_0, \tilde{x}_1, \dots, \tilde{x}_k, \varepsilon) - F(t, \tilde{\tilde{x}}_0, \tilde{\tilde{x}}_1, \dots, \tilde{\tilde{x}}_k, \varepsilon)| \leq l_2 \sum_{i=0}^k |\tilde{x}_i - \tilde{\tilde{x}}_i|,$$

де $t \in \mathbb{R}$, $\tilde{x}_i, \tilde{\tilde{x}}_i \in \mathbb{R}^n$, $i = \overline{0, k}$, l_1, l_2 — довільні додатні сталі.

Припустимо, що власні значення a_j , $j = \overline{1, n}$, матриці A задовольняють умову $\operatorname{Re} a_j(A) \neq 0$, $j = \overline{1, n}$. У цьому випадку, як відомо, існує неособлива матриця C , яка приводить матрицю A до вигляду $A = C^{-1} \operatorname{diag}(A_1, A_2) C$, де A_1, A_2 — деякі сталі матриці розмірності $p \times p$ і $(n-p) \times (n-p)$, власні значення яких задовольняють умови

$$\begin{aligned} \operatorname{Re} a_j(A_1) < 0, j = 1, \dots, p, (0 < p \leq n), \\ \operatorname{Re} a_j(A_2) > 0, j = p+1, \dots, n. \end{aligned} \quad (2)$$

Дослідимо спочатку питання про існування T -періодичних розв'язків системи рівнянь (1) при $\varepsilon = 0$, тобто системи рівнянь вигляду

$$\dot{x}(t) = Ax(t) + \quad (3)$$

$$+ f(t, x(t), x(\lambda_1 t + \psi_1(t, 0)), \dots, x(\lambda_k t + \psi_k(t, 0))).$$

Для цього виконаємо перетворення

$$\dot{x}(t) = Ax(t) + y(t), \quad (4)$$

де $y(t) \in C^0$, де C^0 — простір неперервних T -періодичних вектор-функцій з нормою $\|y(t)\| = \max_t |y(t)|$. Тоді із (4) безпосередньо випливає, що $x(t)$ визначається єдиним чином за допомогою співвідношення

$$x(t) = \int_{-\infty}^{+\infty} G(t-\tau)y(\tau)d\tau, \quad (5)$$

де

$$G(t) = \begin{cases} C^{-1} \operatorname{diag}(e^{A_1 t}, 0) C, & \text{при } t > 0; \\ -C^{-1} \operatorname{diag}(0, e^{A_2 t}) C, & \text{при } t < 0. \end{cases} \quad (6)$$

При цьому для матричної функції $G(t) = (g_{ij}(t))$ виконуються такі умови:

а) $G(+0) - G(-0) = E$, де E — одинична матриця розмірності $n \times n$;

б) $|G(t)| \leq K e^{-\alpha|t|}$ при всіх $t \neq 0$, де $K > 0$, $\alpha > 0$ і $|G| = \max_{1 \leq i \leq n} \sum_{j=1}^n |g_{ij}|$;

в) $\dot{G} = AG$, $t \neq 0$.

Внаслідок перетворення (4) система рівнянь (3) набуває вигляду

$$\begin{aligned} y(t) &= f\left(t, \int_{-\infty}^{+\infty} G(t-\tau)y(\tau)d\tau, \right. \\ &\lambda_1 \int_{-\infty}^{+\infty} G(\lambda_1(t-\tau))y(\lambda_1\tau + \psi_1(t, 0))d\tau, \dots \\ &\left. \dots, \lambda_k \int_{-\infty}^{+\infty} G(\lambda_k(t-\tau))y(\lambda_k\tau + \psi_k(t, 0))d\tau\right), \quad (7) \end{aligned}$$

де $G(t)$ визначається за допомогою співвідношення (6). Для системи рівнянь (7) має місце наступна теорема.

Теорема 1. Нехай виконуються умови $1^* - 3^*$ і

4) всі власні значення a_j , $j = \overline{1, n}$, матриці A такі, що має місце (2), тобто $\exists K > 0$, $\alpha > 0$: $|G(t)| \leq K e^{-\alpha|t|}$ при всіх $t \neq 0$;

5) $\max_{t \in \mathbb{R}} |f(t, 0, \dots, 0)| \leq f^* < +\infty$;

6) $\frac{2Kl_1(k+1)}{\alpha} = \Theta < 1$.

Тоді існує єдиний неперервний T -періодичний розв'язок $\gamma = \gamma(t)$ системи рівнянь (7).

Доведення. Побудуємо розв'язок системи рівнянь (7) за допомогою методу послідовних наближень, які визначимо формулами

$$y_0(t) \equiv 0,$$

$$y_m(t) = f\left(t, \int_{-\infty}^{+\infty} G(t-\tau)y_{m-1}(\tau)d\tau, \right.$$

$$\left. \lambda_1 \int_{-\infty}^{+\infty} G(\lambda_1(t-\tau))y_{m-1}(\lambda_1\tau + \psi_1(t, 0))d\tau, \dots \right.$$

$$\dots, \lambda_k \int_{-\infty}^{+\infty} G(\lambda_k(t-\tau)) y_{m-1}(\lambda_k \tau + \psi_k(t, 0)) d\tau, \quad (8)$$

$$m = 1, 2, \dots$$

Покажемо, що при всіх $m = 1, 2, \dots$, $t \in \mathbb{R}$ виконуються співвідношення

$$|y_m(t) - y_{m-1}(t)| \leq f^* \Theta^{m-1}. \quad (9)$$

Справді, зважаючи на умови теореми, маємо

$$|y_1(t) - y_0(t)| = |f(t, 0, \dots, 0)| \leq f^*,$$

тобто при $m = 1$ оцінка (9) має місце. Розмірковуючи за індукцією, припустимо, що оцінка (9) доведена уже для деякого $m \geq 1$ і покажемо, що вона не зміниться при переході від m до $m + 1$. Дійсно, із (8) одержуємо

$$\begin{aligned} & |y_{m+1}(t) - y_m(t)| \leq \\ & \leq l_1 \left(\int_{-\infty}^{+\infty} |G(t-\tau)| |y_m(\tau) - y_{m-1}(\tau)| d\tau + \right. \\ & \quad \left. + \sum_{i=1}^k \lambda_i \int_{-\infty}^{+\infty} |G(\lambda_i(t-\tau))| \times \right. \\ & \quad \left. \times |y_m(\lambda_i \tau + \psi_i(t, 0)) - y_{m-1}(\lambda_i \tau + \psi_i(t, 0))| d\tau \right) \leq \\ & \leq l_1 \left(\int_{-\infty}^{+\infty} K e^{-\alpha|t-\tau|} f^* \Theta^{m-1} d\tau + \right. \\ & \quad \left. + \sum_{i=1}^k \lambda_i \int_{-\infty}^{+\infty} K e^{-\alpha \lambda_i |t-\tau|} f^* \Theta^{m-1} d\tau \right) \leq \\ & \leq f^* \Theta^{m-1} K l_1 \left(\int_{-\infty}^{+\infty} e^{-\alpha|t-\tau|} d\tau + \sum_{i=1}^k \lambda_i \times \right. \\ & \quad \left. \times \int_{-\infty}^{+\infty} e^{-\alpha \lambda_i |t-\tau|} d\tau \right) \leq f^* \Theta^{m-1} K l_1 \left(\frac{2}{\alpha} + \right. \\ & \quad \left. + \sum_{i=1}^k \lambda_i \frac{2}{\alpha \lambda_i} \right) = f^* \Theta^{m-1} K l_1 \frac{2(k+1)}{\alpha} = f^* \Theta^m. \end{aligned}$$

Цим доведено, що оцінка (9) має місце для довільного $m \geq 1$.

По індукції покажемо, що наближення $y_m(t)$, $m = 0, 1, 2, \dots$, є T -періодичними вектор-функціями по t , тобто

$$y_m(t+T) = y_m(t), \quad m = 0, 1, \dots \quad (10)$$

В силу умов теореми, маємо $y_0(t+T) = 0 = y_0(t)$, тобто співвідношення (10) виконується при $m = 0$. Припустимо, що співвідношення (10) доведене уже для деякого $m \geq 0$ і покажемо його справедливості для $m + 1$. Із (8) маємо

$$\begin{aligned} y_{m+1}(t+T) &= f\left(t+T, \int_{-\infty}^{+\infty} G(t+T-\tau) y_m(\tau) d\tau, \right. \\ & \quad \lambda_1 \int_{-\infty}^{+\infty} G(\lambda_1(t+T-\tau)) y_m(\lambda_1 \tau + \psi_1(t+T, 0)) d\tau, \dots \\ & \quad \left. \dots, \lambda_k \int_{-\infty}^{+\infty} G(\lambda_k(t+T-\tau)) y_m(\lambda_k \tau + \psi_k(t+T, 0)) d\tau \right) = \\ &= f\left(t, \int_{-\infty}^{+\infty} G(t-s) y_m(s) ds, \right. \\ & \quad \lambda_1 \int_{-\infty}^{+\infty} G(\lambda_1(t-s)) y_m(\lambda_1 s + \psi_1(t, 0)) ds, \dots \\ & \quad \left. \dots, \lambda_k \int_{-\infty}^{+\infty} G(\lambda_k(t-s)) y_m(\lambda_k s + \psi_k(t, 0)) ds \right) = \\ &= y_{m+1}(t). \end{aligned}$$

Отже, співвідношення (10) виконуються при всіх $m \geq 0$.

Таким чином, всі наближення $y_m(t)$, $m = 0, 1, 2, \dots$, є неперервними T -періодичними вектор-функціями, для яких справедливі оцінки (9). Враховуючи умову 6 теореми і (9), приходимо до висновку, що ряд $\sum_{m=1}^{+\infty} (y_m(t) - y_{m-1}(t))$ рівномірно збігається для довільного $t \in \mathbb{R}$ до деякої неперервної T -періодичної вектор-функції $\gamma(t) = \lim_{m \rightarrow +\infty} y_m(t)$, яка є розв'язком системи рівнянь (7) (це легко показати, якщо в (8) перейти до границі при $m \rightarrow +\infty$). При цьому для довільного $t \in \mathbb{R}$ має місце нерівність

$$|\gamma(t)| \leq \sum_{m=1}^{+\infty} |y_m(t) - y_{m-1}(t)| \leq \sum_{m=1}^{+\infty} f^* \Theta^{m-1} = \frac{f^*}{1-\Theta}.$$

Для завершення доведення теореми покажемо, що розв'язок $\gamma(t)$ системи (7) — єдиний.

Дійсно, нехай існує ще один неперервний T -періодичний $\eta(t)$ системи рівнянь (7) такий, що $\gamma(t) \neq \eta(t)$. Тоді одержуємо

$$\begin{aligned} |\gamma(t) - \eta(t)| &\leq l_1 \left(\int_{-\infty}^{+\infty} |G(t - \tau)| |\gamma(\tau) - \eta(\tau)| d\tau + \right. \\ &\quad \left. + \sum_{i=1}^k \lambda_i \int_{-\infty}^{+\infty} |G(\lambda_i(t - \tau))| \times \right. \\ &\quad \left. \times |\gamma(\lambda_i \tau + \psi_i(t, 0)) - \eta(\lambda_i \tau + \psi_i(t, 0))| d\tau \right) \leq \\ &\leq l_1 \left(\int_{-\infty}^{+\infty} K e^{-\alpha|t-\tau|} |\gamma(\tau) - \eta(\tau)| d\tau + \right. \\ &\quad \left. + \sum_{i=1}^k \lambda_i \int_{-\infty}^{+\infty} K e^{-\alpha\lambda_i|t-\tau|} \times \right. \\ &\quad \left. \times |\gamma(\lambda_i \tau + \psi_i(t, 0)) - \eta(\lambda_i \tau + \psi_i(t, 0))| d\tau \right) \leq \\ &\leq l_1 \frac{2K(k+1)}{\alpha} \max_t |\gamma(t) - \eta(t)| \leq \Theta \|\gamma(t) - \eta(t)\|, \end{aligned}$$

де $\|\gamma(t) - \eta(t)\| = \max_t |\gamma(t) - \eta(t)|$. Отже, одержимо співвідношення

$$\|\gamma(t) - \eta(t)\| \leq \Theta \|\gamma(t) - \eta(t)\|,$$

яке може мати місце лише у випадку, коли $\Theta \geq 1$, що суперечить умові 6 даної теореми. Цим доведено, що вектор-функція $\gamma(t)$ є єдиним неперервним T -періодичним розв'язком системи рівнянь (7). Теорему 1 доведено.

Враховуючи теорему 1 і співвідношення (5), приходимо до висновку, що вектор-функція

$$\bar{x}(t) = \int_{-\infty}^{+\infty} G(t - \tau) \gamma(\tau) d\tau$$

є єдиним неперервним T -періодичним розв'язком системи (3).

Перейдемо тепер до дослідження T -періодичних розв'язків системи рівнянь (1) при $\varepsilon \neq 0$. Виконаємо в системі рівнянь (1) взаємно-однозначну заміну змінних

$$x(t) = y(t) + \bar{x}(t), \quad (11)$$

де $\bar{x}(t)$ — T -періодичний розв'язок системи рівнянь (1) при $\varepsilon = 0$. В результаті отримаємо систему рівнянь вигляду

$$\dot{y}(t) = Ay(t) +$$

$$\begin{aligned} &+ \varphi(t, y(t), y(\lambda_1 t + \psi_1(t, \varepsilon)), \dots, y(\lambda_k t + \psi_k(t, \varepsilon))) + \\ &+ \varepsilon \Phi(t, y(t), y(\lambda_1 t + \psi_1(t, \varepsilon)), \dots, y(\lambda_k t + \psi_k(t, \varepsilon)), \varepsilon), \end{aligned} \quad (12)$$

де

$$\begin{aligned} &\varphi(t, y(t), y(\lambda_1 t + \psi_1(t, \varepsilon)), \dots, y(\lambda_k t + \psi_k(t, \varepsilon))) = \\ &= f(t, y(t) + \bar{x}(t), y(\lambda_1 t + \psi_1(t, \varepsilon)) + \bar{x}(\lambda_1 t + \psi_1(t, \varepsilon)), \dots, \\ &\quad \dots, y(\lambda_k t + \psi_k(t, \varepsilon)) + \bar{x}(\lambda_k t + \psi_k(t, \varepsilon))) - \\ &\quad - f(t, \bar{x}(t), \bar{x}(\lambda_1 t + \psi_1(t, 0)), \dots, \bar{x}(\lambda_k t + \psi_k(t, 0))); \\ &\Phi(t, y(t), y(\lambda_1 t + \psi_1(t, \varepsilon)), \dots, y(\lambda_k t + \psi_k(t, \varepsilon)), \varepsilon) = \\ &= F(t, y(t) + \bar{x}(t), y(\lambda_1 t + \psi_1(t, \varepsilon)) + \bar{x}(\lambda_1 t + \psi_1(t, \varepsilon)), \dots, \\ &\quad \dots, y(\lambda_k t + \psi_k(t, \varepsilon)) + \bar{x}(\lambda_k t + \psi_k(t, \varepsilon)), \varepsilon). \end{aligned}$$

Легко переконатися, що вектор-функції $\varphi(t, y(t), y(\lambda_1 t + \psi_1(t, \varepsilon)), \dots, y(\lambda_k t + \psi_k(t, \varepsilon)))$, $\Phi(t, y(t), y(\lambda_1 t + \psi_1(t, \varepsilon)), \dots, y(\lambda_k t + \psi_k(t, \varepsilon)), \varepsilon)$ задовольняють умови $2^* - 3^*$ і $\varphi(t, 0, \dots, 0) \equiv 0$.

Якщо виконаємо перетворення

$$\dot{y}(t) = Ay(t) + z(t), \quad (13)$$

де $z(t) \in C^0$, де C^0 — простір неперервних на $\mathbb{R}$ T -періодичних вектор-функцій з нормою $\|z(t)\| = \max_t |z(t)|$, то одержимо, що $y(t)$ визначається єдиним чином за допомогою рівності

$$y(t) = \int_{-\infty}^{+\infty} G(t - \tau) z(\tau) d\tau, \quad (14)$$

де $G(t)$ визначається співвідношенням (6). В результаті перетворення (13) система рівнянь (12) набуває вигляду

$$\begin{aligned} z(t) &= \varphi\left(t, \int_{-\infty}^{+\infty} G(t - \tau) z(\tau) d\tau, \right. \\ &\quad \lambda_1 \int_{-\infty}^{+\infty} G(\lambda_1(t - \tau)) z(\lambda_1 \tau + \psi_1(t, \varepsilon)) d\tau, \dots \\ &\quad \dots, \lambda_k \int_{-\infty}^{+\infty} G(\lambda_k(t - \tau)) z(\lambda_k \tau + \psi_k(t, \varepsilon)) d\tau \Big) + \\ &\quad + \varepsilon \Phi\left(t, \int_{-\infty}^{+\infty} G(t - \tau) z(\tau) d\tau, \right. \\ &\quad \left. \lambda_1 \int_{-\infty}^{+\infty} G(\lambda_1(t - \tau)) z(\lambda_1 \tau + \psi_1(t, \varepsilon)) d\tau, \dots \right) \end{aligned}$$

$$\dots, \lambda_k \int_{-\infty}^{+\infty} G(\lambda_k(t-\tau))z(\lambda_k\tau + \psi_k(t, \varepsilon))d\tau, \varepsilon). \quad (15)$$

Для системи рівнянь (15) має місце наступна теорема.

Теорема 2. *Нехай виконуються умови $1^* - 3^*$*

- 4) всі власні значення a_j , $j = \overline{1, n}$, матриці A такі, що має місце (2), тобто $\exists K > 0$, $\alpha > 0$: $|G(t)| \leq Ke^{-\alpha|t|}$ при всіх $t \neq 0$;

- 5) $\varphi(t, 0, \dots, 0) \equiv 0$,

$$\max_{t \in \mathbb{R}} |\Phi(t, 0, \dots, 0, \varepsilon)| \leq M < +\infty;$$

$$6) \quad \frac{2K(k+1)(l_1 + \varepsilon l_2)}{\alpha} = \theta < 1. \quad (16)$$

Тоді існує єдиний неперервний T -періодичний розв'язок $\tilde{\gamma} = \tilde{\gamma}(t, \varepsilon)$ системи рівнянь (15) такий, що $\lim_{\varepsilon \rightarrow 0} \tilde{\gamma}(t, \varepsilon) = 0$.

Доведення проводиться за тією ж схемою, що і доведення теореми 1. При цьому послідовні наближення визначаються наступними формулами

$$\begin{aligned}
z_0(t, \varepsilon) &\equiv 0, \\
z_m(t, \varepsilon) &= \varphi\left(t, \int_{-\infty}^{+\infty} G(t - \tau) z_{m-1}(\tau, \varepsilon) d\tau, \right. \\
&\quad \lambda_1 \int_{-\infty}^{+\infty} G(\lambda_1(t - \tau)) z_{m-1}(\lambda_1 \tau + \psi_1(t, \varepsilon), \varepsilon) d\tau, \dots \\
&\quad \dots, \lambda_k \int_{-\infty}^{+\infty} G(\lambda_k(t - \tau)) z_{m-1}(\lambda_k \tau + \psi_k(t, \varepsilon), \varepsilon) d\tau) + \\
&\quad \left. + \varepsilon \Phi\left(t, \int_{-\infty}^{+\infty} G(t - \tau) z_{m-1}(\tau, \varepsilon) d\tau, \right. \right. \\
&\quad \left. \lambda_1 \int_{-\infty}^{+\infty} G(\lambda_1(t - \tau)) z_{m-1}(\lambda_1 \tau + \psi_1(t, \varepsilon), \varepsilon) d\tau, \dots \right. \\
&\quad \left. \dots, \lambda_k \int_{-\infty}^{+\infty} G(\lambda_k(t - \tau)) z_{m-1}(\lambda_k \tau + \psi_k(t, \varepsilon), \varepsilon) d\tau, \varepsilon\right),
\end{aligned}
\tag{17}$$

Розмірковуючи за індукцією можна показати, що при всіх $t \in \mathbb{R}$ виконуються співвідношення

$$z_m(t+T, \varepsilon) = z_m(t, \varepsilon), \quad m = 0, 1, \dots;$$

$$|z_m(t, \varepsilon) - z_{m-1}(t, \varepsilon)| \leq \tilde{M}\theta^{m-1}, \quad m = 1, 2, \dots, \quad (18)$$

де $\tilde{M} = \tilde{M}(\varepsilon) > M$ — деяка додатна стала, що залежить від ε .

Звідси слідує, що послідовність T -періодичних вектор-функцій $z_m(t, \varepsilon)$, $m = 0, 1, 2, \dots$, рівномірно збігається для довільного $t \in \mathbb{R}$ до деякої неперервної T -періодичної вектор-функції $\tilde{\gamma} = \tilde{\gamma}(t, \varepsilon)$, яка є розв'язком системи рівнянь (15) (в цьому легко переконатися, якщо в (17) перейти до границі при $m \rightarrow +\infty$).

Покажемо, що розв'язок $\tilde{\gamma}(t, \varepsilon)$ системи рівнянь (15) задовольняє умову $\lim_{\varepsilon \rightarrow 0} \tilde{\gamma}(t, \varepsilon) = 0$. Для цього доведемо, що при всіх $t \in \mathbb{R}$, $m = 0, 1, 2, \dots$, виконуються співвідношення

$$\lim_{\varepsilon \rightarrow 0} z_m(t, \varepsilon) = 0. \quad (19)$$

Розглядаючи послідовно (17), $m = 0, 1, 2, \dots$,
отримаємо

$$\lim_{\varepsilon \rightarrow 0} z_0(t, \varepsilon) = 0,$$

$$\begin{aligned} \lim_{\varepsilon \rightarrow 0} z_1(t, \varepsilon) &= \lim_{\varepsilon \rightarrow 0} \left[\varphi \left(t, \int_{-\infty}^{+\infty} G(t - \tau) z_0(\tau, \varepsilon) d\tau, \right. \right. \\ &\quad \lambda_1 \int_{-\infty}^{+\infty} G(\lambda_1(t - \tau)) z_0(\lambda_1 \tau + \psi_1(t, \varepsilon), \varepsilon) d\tau, \dots \\ &\quad \dots, \lambda_k \int_{-\infty}^{+\infty} G(\lambda_k(t - \tau)) z_0(\lambda_k \tau + \psi_k(t, \varepsilon), \varepsilon) d\tau \Big) + \\ &\quad \left. + \varepsilon \Phi \left(t, \int_{-\infty}^{+\infty} G(t - \tau) z_0(\tau, \varepsilon) d\tau, \right. \right. \\ &\quad \lambda_1 \int_{-\infty}^{+\infty} G(\lambda_1(t - \tau)) z_0(\lambda_1 \tau + \psi_1(t, \varepsilon), \varepsilon) d\tau, \dots \\ &\quad \left. \dots, \lambda_k \int_{-\infty}^{+\infty} G(\lambda_k(t - \tau)) z_0(\lambda_k \tau + \psi_k(t, \varepsilon), \varepsilon) d\tau, \varepsilon \right) \Big] = \\ &= \varphi(t, 0, \dots, 0) = 0; \end{aligned}$$

$$\lim_{\varepsilon \rightarrow 0} z_m(t, \varepsilon) = \lim_{\varepsilon \rightarrow 0} \left[\varphi \left(t, \int_{-\infty}^{+\infty} G(t - \tau) z_{m-1}(\tau, \varepsilon) d\tau, \right. \right.$$

$$\begin{aligned} & \lambda_1 \int_{-\infty}^{+\infty} G(\lambda_1(t-\tau)) z_{m-1}(\lambda_1\tau + \psi_1(t, \varepsilon), \varepsilon) d\tau, \dots \\ & \dots, \lambda_k \int_{-\infty}^{+\infty} G(\lambda_k(t-\tau)) z_{m-1}(\lambda_k\tau + \psi_k(t, \varepsilon), \varepsilon) d\tau + \\ & + \varepsilon \Phi\left(t, \int_{-\infty}^{+\infty} G(t-\tau) z_{m-1}(\tau, \varepsilon) d\tau, \right. \\ & \left. \lambda_1 \int_{-\infty}^{+\infty} G(\lambda_1(t-\tau)) z_{m-1}(\lambda_1\tau + \psi_1(t, \varepsilon), \varepsilon) d\tau, \dots \right. \\ & \left. \dots, \lambda_k \int_{-\infty}^{+\infty} G(\lambda_k(t-\tau)) z_{m-1}(\lambda_k\tau + \psi_k(t, \varepsilon), \varepsilon) d\tau, \varepsilon\right) = \\ & = \varphi(t, 0, \dots, 0) = 0, \quad m \geq 2, \\ & \dots \end{aligned}$$

що і необхідно було довести.

Доведення єдиності неперервного T -періодичного розв'язку $\tilde{\gamma}(t, \varepsilon)$ системи рівнянь (15) проводимо аналогічно до того, як це було зроблено у теоремі 1. Отже, вектор-функція $\tilde{\gamma}(t, \varepsilon)$

є єдиним неперервним T -періодичним розв'язком системи рівнянь (15), що задовольняє умову $\lim_{\varepsilon \rightarrow 0} \tilde{\gamma}(t, \varepsilon) = 0$. Теорему 2 доведено.

Враховуючи теорему 2 і співвідношення (14), знаходимо, що вектор-функція

$$\bar{y}(t, \varepsilon) = \int_{-\infty}^{+\infty} G(t-\tau) \tilde{\gamma}(\tau, \varepsilon) d\tau$$

є єдиним неперервним T -періодичним розв'язком системи рівнянь (1) при $\varepsilon \neq 0$, для якого $\lim_{\varepsilon \rightarrow 0} \bar{y}(t, \varepsilon) = 0$.

Таким чином, в силу (11) і теорем 1, 2 приходимо до висновку, що система рівнянь (1) має єдиний неперервний T -періодичний розв'язок $\hat{x}(t, \varepsilon) = \bar{y}(t, \varepsilon) + \bar{x}(t)$ такий, що $\lim_{\varepsilon \rightarrow 0} \hat{x}(t, \varepsilon) = \bar{x}(t)$, де $\bar{x}(t)$ — неперервний T -періодичний розв'язок системи рівнянь (1) при $\varepsilon = 0$, а $\bar{y}(t, \varepsilon)$ — неперервний T -періодичний розв'язок системи рівнянь (1) при $\varepsilon \neq 0$.

З а у в а ж е н н я. При деяких додаткових припущеннях теореми 1, 2 мають місце також у випадку, коли функції $\psi_i(t, \varepsilon)$, $i = \overline{1, k}$, залежать від невідомої функції $x(t)$.

Список використаних джерел

1. Kato T. The functional-differential equation $y'(x) = ay(\lambda x) + by(x)$ / Т. Kato, J.B. McLeod // Bull. Amer. Math. Soc. — 1971. — **77**. — P.891–937.
2. Самойленко А.М. Ограниченные на всей вещественной оси решения систем нелинейных дифференциально-функциональных уравнений и их свойства / А.М. Самойленко, Г.П. Пелюх // Укр. мат. журн. — 1994. — **46**, №6. — С. 737–747.
3. Денисенко Н.Л. Асимптотичні властивості неперервних розв'язків систем дифференціально-функціональних рівнянь з лінійними перетвореннями аргументу / Н.Л. Денисенко // Наукові вісті НТУУ "КПІ". — 2008. — №3. — С. 135–141.
4. Денисенко Н.Л. Про властивості неперервних періодичних розв'язків систем дифференціально-функціональних рівнянь із малим параметром / Н.Л. Денисенко // Нелінійні коливання. — 2014. — **17**, №3. — С. 332–340.

References

1. KATO, T. and MCLEOD, J.B. (1971) The functional-differential equation $y'(x) = ay(\lambda x) + by(x)$. *Bull. Amer. Math. Soc.* **77**. p. 891–937.
2. SAMOILENKO, A.M. and PELIUKH, G.P. (1994) Ogranichennyie na vsej veschestvennoy osi resheniya sistem nelineynykh differentsialno-funktsionalnykh uravneniy i ih svoystva. *Ukrainskyi matematychnyi zhurnal*. **46**(6). p. 737–747.
3. DENYSENKO, N.L. (2008) Asymptotychni vlastyvoli neperervnykh rozviazkiv system dyferentsialno-funktsionalnykh rivnians z liniinymy peretvorenniamy arhumentu. *Naukovi visti NTUU "KPI"*. (3). p. 135–141.
4. DENYSENKO, N.L. (2014) Pro vlastyvoli neperervnykh periodychnykh rozviazkiv system dyferentsialno-funktsionalnykh rivnians iz malym parametrom. *Nelineini kolyvannia*. **17**(3). p. 332–340.

Надійшла до редколегії 20.05.2015

УДК 517.929

Т.О. Єрєміна, старший викладач

**Про побудову неперервних розв'язків
систем нелінійних
різницево-функціональних рівнянь**

Національний технічний університет України «Київ-
ський політехнічний інститут», 03056, м. Київ, пр. Пе-
ремоги, 37,
e-mail: ierominat@ukr.net

T.O. Yeromina, senior teacher

**About construction of continuous
solutions of nonlinear
difference-functional equations**

National Technical University of Ukraine «Kyiv Polytech-
nic Institute», 03056, Kyiv-56, Prospect Peremohy, 37,
e-mail: ierominat@ukr.net

Розроблено метод побудови неперервних розв'язків одного класу систем нелінійних різницево-функціональних рівнянь і досліджено їх властивості.

Ключові слова: різницеві рівняння, функціональні рівняння, різницево-функціональні рівняння.

It is considered one class of systems of nonlinear difference-functional equations which includes properties of difference and q-difference (functional) equations what is important tool in solving many problems in applied sciences. Therefore, under various assumptions, the difference equations have been the object of research of many mathematicians, the development of their theory acquired a number of currents. Today this theory is widely branching and already has a lot of tangible results. The main task of this work is a study of the structure of the set of continuous solutions of one class of systems of nonlinear difference-functional equations. Determined sufficient conditions of the existence of continuous solutions, developed a method of constructing such solutions of systems of nonlinear difference-functional equations and also investigated the behavior of these solutions at t goes to infinity. It should be noted that the method of proof also gives possibility to construct a periodic solution and investigate its properties. Performed researches add to already existent works of other mathematicians and facilitate further study of the continuous solutions of wider classes of difference-functional equations.

Key Words: difference equations, functional equations, difference-functional equations.

Статтю представив академік НАНУ, доктор фіз.-мат. наук, проф. Перестюк М. О.

Окремі класи систем нелінійних різницево-функціональних рівнянь при різних припущеннях були об'єктом дослідження багатьох математиків [1-4]. Особлива увага приділялася вивченню питань існування неперервних розв'язків, дослідженню структури їх множини та поведінці при $t \rightarrow \infty$ [3-4]). Продовжуючи ці дослідження виникла задача про отримання аналогічних результатів для рівнянь вигляду

$$x(qt) = \Lambda x(t) + f(t, x(t+1)), \quad (1)$$

де $t \in \mathbb{R}$, Λ – дійсна $(n \times n)$ - матриця вигляду $\Lambda = \text{diag}(\lambda_1, \dots, \lambda_n)$, $f: \mathbb{R} \times \mathbb{R}^n \rightarrow \mathbb{R}^n$, q - деяка дійсна стала, при певних припущеннях відносно λ_i , $i = \overline{1, n}$ та q .

Розглянемо систему нелінійних різницево-функціональних рівнянь вигляду (1). При цьому припустимо, що виконуються наступні умови:

1. $|\lambda_i| \neq 0$, $i = 1, 2, \dots, n$, $q > 0$;

2. вектор-функція $f(t, x)$ є неперервною обмеженою при всіх $t \in \mathbb{R}$, $x \in \mathbb{R}^n$ і $f(t, 0) = 0$;

3. для довільних $t \in \mathbb{R}$, $x, y \in \mathbb{R}^n$ виконується співвідношення

$$|f(t, x) - f(t, y)| \leq l|x - y|, \quad (2)$$

де l - деяка додатна стала.

Теорема 1. Нехай виконуються умови 1.-

3. і умови:

4. $0 < \lambda_i < 1$, $i = 1, \dots, n$, $q > 1$;

5. $\Delta = \frac{l}{1-\lambda^*} < 1$,

де $1 > \lambda^* > \max\{\lambda_i, i = 1, \dots, n\}$.

Тоді система рівнянь (1) має сім'ю неперервних обмежених при $t \geq T > 0$ (T - деяка достатньо велика додатна стала) розв'язків у вигляді ряду

$$x(t) = \sum_{i=0}^{\infty} x_i(t), \quad (3)$$

де $x_i(t)$, $i = 0, 1, \dots$ - деякі неперервні обмежені при $t \geq T > 0$ вектор-функції, які задовольняють умові $\lim_{t \rightarrow +\infty} x(t) = 0$.

Доведення. Розглянемо послідовність систем рівнянь вигляду

$$x_0(qt) = \Lambda x_0(t), \quad (4_0)$$

$$x_1(qt) = \Lambda x_1(t) + f(t, x_0(t+1)), \quad (4_1)$$

$$x_i(qt) = \Lambda x_i(t) + f\left(t, \sum_{l=0}^{i-1} x_l(t+1)\right) - f\left(t, \sum_{l=0}^{i-2} x_l(t+1)\right), \quad (4_i)$$

$i = 2, 3, \dots$, і покажемо, що вони мають сім'ї неперервних обмежених при $t \geq T > 0$ розв'язків, які задовольняють умови

$$|x_i(t)| \leq M \Delta^i, i = 1, 2, \dots, \quad (5_i)$$

де M - деяка додатна стала.

Система рівнянь (4₀) має множину неперервних при $t \geq T > 0$ розв'язків вигляду

$$x_0(t) = t^\nu \omega\left(\frac{\ln t}{\ln q}\right), \quad (6_0)$$

де $\omega(\tau) = (\omega_1(\tau), \omega_2(\tau), \dots, \omega_n(\tau))$, $\omega_i(\tau)$, $i = 1, \dots, n$ - довільні неперервні 1-періодичні функції,

$$t^\nu = \text{diag}\left(t^{\frac{\ln \lambda_1}{\ln q}}, t^{\frac{\ln \lambda_2}{\ln q}}, \dots, t^{\frac{\ln \lambda_n}{\ln q}}\right),$$

які задовольняють умові

$$|x_0(t)| \leq |t^\nu| |\omega(\tau)| \leq t^{\frac{\ln \lambda^*}{\ln q}} |\omega(\tau)| \leq t^{\frac{\ln \lambda^*}{\ln q}} M, \quad (5_0)$$

де $M = \max_{\tau} |\omega(\tau)|$.

Підставляючи в (4₁) ряд

$$x_1(t) = \sum_{j=0}^{\infty} \Lambda^j f\left(q^{-(j+1)}t, x_0(q^{-(j+1)}t+1)\right), \quad (6_1)$$

можна переконатися, що він є її формальним розв'язком. Більше цього, в силу (2), (5₀) і $\frac{\ln \lambda^*}{\ln q} < 0$, отримаємо

$$\begin{aligned} |x_1(t)| &\leq \sum_{j=0}^{\infty} |\Lambda|^j \left| f\left(q^{-(j+1)}t, x_0(q^{-(j+1)}t+1)\right) \right| \leq \\ &\leq \sum_{j=0}^{\infty} (\lambda^*)^j l \left| x_0(q^{-(j+1)}t+1) \right| \leq \\ &\leq \sum_{j=0}^{\infty} (\lambda^*)^j l M \left(\frac{1}{q^{j+1}}t+1 \right)^{\frac{\ln \lambda^*}{\ln q}} \leq \\ &\leq l M \sum_{j=0}^{\infty} (\lambda^*)^j \leq M \frac{l}{1-\lambda^*} \leq M \Delta. \end{aligned}$$

Отже, системи рівнянь (4₀), (4₁) мають сім'ї неперервних при $t \geq T > 0$ розв'язків, для кожного з яких виконуються співвідношення (5₀), (5₁).

Враховуючи умови теореми і оцінки (5₀), (5₁), можна послідовно показати, що ряди

$$x_i(t) =$$

$$\begin{aligned} &= \sum_{j=0}^{\infty} \Lambda^j \left[f\left(q^{-(j+1)}t, \sum_{l=0}^{i-1} x_l\left(q^{-(j+1)}t+1\right)\right) - \right. \\ &\quad \left. - f\left(q^{-(j+1)}t, \sum_{l=0}^{i-2} x_l\left(q^{-(j+1)}t+1\right)\right) \right], \quad (6_i) \end{aligned}$$

$i = 2, 3, \dots$ рівномірно збігаються при всіх $t \geq T > 0$ і є розв'язками відповідних систем рівнянь (4_i), $i = 2, 3, \dots$. Дійсно, легко переконатися, що ряди (6_i), $i = 2, 3, \dots$ є формальними розв'язками систем рівнянь (4_i), $i = 2, 3, \dots$. Доведемо їх збіжність.

Оскільки ряд (6₁) рівномірно збігається при $t \geq T > 0$, то розмірковуючи за індукцією, припустимо, що збіжність ряду (6_i) доведена уже для деякого $i \geq 1$ і доведемо, що ряд (6_{i+1}) також рівномірно збігається при $t \geq T > 0$ і виконується оцінка (5_{i+1}). В силу (2), (6_{i+1}) та (5_i) отримаємо

$$\begin{aligned} |x_{i+1}(t)| &\leq \sum_{j=0}^{\infty} |\Lambda|^j \left| f\left(q^{-(j+1)}t, \sum_{l=0}^i x_l\left(q^{-(j+1)}t+1\right)\right) - \right. \\ &\quad \left. - f\left(q^{-(j+1)}t, \sum_{l=0}^{i-1} x_l\left(q^{-(j+1)}t+1\right)\right) \right| \leq \\ &\leq \sum_{j=0}^{\infty} |\Lambda|^j l \left| \sum_{l=0}^i x_l\left(q^{-(j+1)}t+1\right) - \sum_{l=0}^{i-1} x_l\left(q^{-(j+1)}t+1\right) \right| \leq \\ &\leq \sum_{j=0}^{\infty} |\Lambda|^j l |x_i\left(q^{-(j+1)}t+1\right)| \leq \\ &\leq l M \Delta^i \sum_{j=0}^{\infty} (\lambda^*)^j \leq M \Delta^i \frac{l}{1-\lambda^*} = M \Delta^{i+1}. \end{aligned}$$

Отже, системи рівнянь (4_i), $i=1, 2, \dots$ мають розв'язки у вигляді рядів (6_i), $i=1, 2, \dots$, що рівномірно збігаються при $t \geq T > 0$ до деяких неперервних вектор-функцій $x_i(t)$, $i=1, 2, \dots$, які задовольняють умови (5_i). Із (5_i) безпосередньо випливає, що ряд (3) рівномірно збігається при всіх $t \geq T > 0$ до деякої неперервної вектор-функції $x(t)$, яка задовольняє умові

$$|x(t)| \leq \frac{M}{1-\Delta}.$$

Для завершення доведення теореми достатньо, очевидно, показати, що ряд (3) задовольняє систему рівнянь (1), тобто виконується тожність

$$\sum_{i=0}^{\infty} x_i(qt) \equiv \Lambda \sum_{i=0}^{\infty} x_i(t) + f\left(t, \sum_{i=0}^{\infty} x_i(t+1)\right).$$

Для цього потрібно довести, що ряд

$$f\left(t, x_0(t+1)\right) + \sum_{i=2}^{\infty} \left[f\left(t, \sum_{l=0}^{i-1} x_l(t+1)\right) - f\left(t, \sum_{l=0}^{i-2} x_l(t+1)\right) \right] f\left(t, \sum_{l=0}^{i-1} x_l(t+1)\right) - f\left(t, \sum_{l=0}^{i-2} x_l(t+1)\right), \quad (9_i)$$

рівномірно збігається при всіх $t \geq T > 0$ і його сумою є $f\left(t, \sum_{l=0}^{\infty} x_l(t+1)\right)$.

А це в свою чергу, буде доведено, якщо вдасться показати, що при всіх $t \geq T > 0$ має місце співвідношення

$$\lim_{i \rightarrow \infty} f\left(t, \sum_{l=0}^i x_l(t+1)\right) = f\left(t, \sum_{l=0}^{\infty} x_l(t+1)\right). \quad (7)$$

Нехай ε - як завгодно мале додатне число. Тоді в силу умов теореми існує такий номер N , що виконується умова

$$M \frac{l}{1-\Delta} \Delta^{N+1} < \varepsilon.$$

Далі, приймаючи до уваги умови теореми і оцінки (5_i), $i=0,1, \dots$, знаходимо

$$\left| f\left(t, \sum_{l=0}^{\infty} x_l(t+1)\right) - f\left(t, \sum_{l=0}^i x_l(t+1)\right) \right| \leq l \left| \sum_{l=i+1}^{\infty} x_l(t+1) \right| \leq lM \frac{\Delta^{i+1}}{1-\Delta} < \varepsilon$$

при всіх $i \geq N$. Цим самим співвідношення (7) і теорема доведена.

Теорема 2. Нехай виконуються умови 1.-3. і умови:

6. $\lambda_i > 1, i = 1, \dots, n, 0 < q < 1$;
7. $\Delta = \frac{l}{\lambda_* - 1} < 1$,
- де $1 < \lambda_* < \min \{\lambda_i, i = 1, \dots, n\}$.

Тоді система рівнянь (1) має сім'ю неперервних обмежених при $t \geq T > 0$ (T - деяка достатньо велика додатна стала) розв'язків у вигляді ряду

$$x(t) = \sum_{i=0}^{\infty} x_i(t), \quad (8)$$

де $x_i(t), i = 0, 1, \dots$ - деякі неперервні обмежені при $t \geq T > 0$ вектор-функції, які

задовольняють умові $\lim_{t \rightarrow +\infty} x(t) = 0$.

Доведення: Розглянемо послідовність систем рівнянь вигляду

$$x_0(qt) = \Lambda x_0(t), \quad (9_0)$$

$$x_1(qt) = \Lambda x_1(t) + f\left(t, x_0(t+1)\right), \quad (9_1)$$

$$x_i(qt) = \Lambda x_i(t) +$$

$i = 2, 3, \dots$ Покажемо, що вони мають сім'ю неперервних обмежених при $t \geq T > 0$ розв'язків, які задовольняють умови

$$|x_i(t)| \leq M \Delta^i, i = 1, 2, \dots \quad (10_i)$$

де M - деяка додатна стала.

Дійсно, система рівнянь (9₀) має сім'ю неперервних при $t \geq T > 0$ розв'язків, які задовольняють умові

$$|x_0(t)| \leq \frac{M}{t^{\left| \frac{\ln \lambda_*}{\ln q} \right|}}, \quad (10_0)$$

де $1 < \lambda_* < \min \{\lambda_i, i = 1, \dots, n\}$. Далі, підставляючи в (9₁) ряд

$$x_1(t) = - \sum_{j=0}^{\infty} \Lambda^{-(j+1)} f\left(q^j t, x_0(q^j t + 1)\right), \quad (11_1)$$

можна переконатися, що він є її формальним розв'язком. Більше того, в силу (2), (10₀), отримуємо

$$\begin{aligned} |x_1(t)| &\leq \sum_{j=0}^{\infty} |\Lambda^{-1}|^{j+1} |f(q^j t, x_0(q^j t + 1))| \leq \\ &\leq \sum_{j=0}^{\infty} \left(\frac{1}{\lambda_*}\right)^{j+1} l |x_0(q^j t + 1)| \leq \\ &\leq \frac{1}{\lambda_*} \sum_{j=0}^{\infty} \left(\frac{1}{\lambda_*}\right)^j l \frac{M}{(q^j t + 1)^{\left| \frac{\ln \lambda_*}{\ln q} \right|}} \leq \\ &\leq \frac{1}{\lambda_*} \sum_{j=0}^{\infty} \left(\frac{1}{\lambda_*}\right)^j l M \leq \\ &\leq \frac{M}{\lambda_*} \frac{l}{1 - \frac{1}{\lambda_*}} \leq M \frac{l}{\lambda_* - 1} \leq M \Delta. \end{aligned}$$

Отже, системи рівнянь (9₀), (9₁) мають сім'ю неперервних при $t \geq T > 0$ розв'язків, для кожного з яких виконуються співвідношення (10₀), (10₁).

Враховуючи умови теореми і оцінки (10₀), (10₁), можна послідовно показати, що ряди

$$x_i(t) = - \sum_{j=0}^{\infty} \Lambda^{-(j+1)} \left[f\left(q^j t, \sum_{l=0}^{i-1} x_l(q^j t + 1)\right) - \right.$$

$$-f\left(q^j t, \sum_{l=0}^{i-2} x_l(q^j t + 1)\right)\Bigg], \quad (11_i)$$

$i = 2, 3, \dots$ рівномірно збігаються при всіх $t \geq T > 0$ і є розв'язками відповідних систем рівнянь (9_i) , $i = 2, 3, \dots$. Дійсно, легко переко-
нати, що ряди (11_i) , $i = 2, 3, \dots$ є формальни-
ми розв'язками систем рівнянь (9_i) , $i = 2, 3, \dots$.
Доведемо їх збіжність.

Справді, оскільки ряд (11_1) рівномірно збі-
гається при $t \geq T > 0$, то розмірковуючи за ін-
дукцією, припустимо, що збіжність ряду (11_i)
доведена уже для деякого $i \geq 1$ і доведемо,
що ряд (11_{i+1}) також рівномірно збігається при
 $t \geq T > 0$ і виконується оцінка (10_{i+1}) . В силу
(2), (11_{i+1}) та (10_i) отримаємо

$$\begin{aligned} |x_{i+1}(t)| &\leq \sum_{j=0}^{\infty} |\Lambda^{-1}|^{j+1} \left| f\left(q^j t, \sum_{l=0}^i x_l(q^j t + 1)\right) - \right. \\ &\quad \left. - f\left(q^j t, \sum_{l=0}^{i-1} x_l(q^j t + 1)\right) \right| \leq \\ &\leq \sum_{j=0}^{\infty} |\Lambda^{-1}|^{j+1} l \left| \sum_{l=0}^i x_l(q^j t + 1) - \right. \\ &\quad \left. - \sum_{l=0}^{i-1} x_l(q^j t + 1) \right| \leq \\ &\leq \sum_{j=0}^{\infty} |\Lambda^{-1}|^{j+1} l |x_i(q^j t + 1)| \leq \\ &\leq l M \Delta^i \frac{1}{\lambda_*} \sum_{j=0}^{\infty} \left(\frac{1}{\lambda_*}\right)^j \leq \\ &\leq M \Delta^i \frac{l}{\lambda_* - 1} = M \Delta^{i+1}. \end{aligned}$$

Список використаних джерел

1. Birkhoff G.D. General theory of linear difference equations. / Birkhoff G.D. — Trans. Amer. Math. Soc. — 1911. — 12. — pp. 243 – 284.
2. Митропольский Ю.А. Системы эволюционных уравнений с периодическими и условно-периодическими коэффициентами. / Митропольский Ю.А., Самойленко А.М., Мартынюк Д.И. — К.: Наукова думка. — 1985. — 216 с.
3. Пелюх Г.П. К теории систем линейных разностных уравнений с непрерывным аргументом / Пелюх Г.П. — ДАН. — 2006. — 73. — № 2. — С. 269–272.
4. Єрьоміна Т.О. Дослідження структури множини неперервних розв'язків систем лінійних різницево-функціональних рівнянь / Єрьоміна Т.О. — // Нелінійні коливання. — 2014. — 17, № 3. — С. 341–350.

Отже, системи рівнянь (9_i) , $i=1,2,\dots$ мають розв'язки у вигляді рядів (11_i) , $i=1,2,\dots$, що рівномірно збігаються при $t \geq T > 0$ до деяких неперервних вектор-функцій $x_i(t)$, $i=1,2,\dots$, які задовольняють умови (10_i) . Із (10_i) безпосередньо випливає, що ряд (8) рівномірно збігається при всіх $t \geq T > 0$ до деякої неперервної вектор-функції $x(t)$, яка задовольняє умови

$$|x(t)| \leq \frac{M}{1 - \Delta}.$$

Далі, діючи аналогічно тому, як і при доведенні теореми 1, можна показати, що ряд (8) задовольняє систему рівнянь (1), тобто виконується тотожність

$$\sum_{i=0}^{\infty} x_i(qt) \equiv \Lambda \sum_{i=0}^{\infty} x_i(t) + f\left(t, \sum_{i=0}^{\infty} x_i(t+1)\right),$$

що і завершує доведення теореми.

References

1. BYRKHOV, G. (1911) General theory of linear difference equations. — Trans. Amer. Math. Soc., 12, pp. 243 – 284.
2. MITROPOLSKY, Y., SAMOILENKO, A., MARTINUK, D. (1985) Sistemy evolyutsionnyh uravneniy s periodicheskimi i uslovno-periodicheskimi koefitsientami. — Kiev: Naukova Dumka.
3. PELYUKH, G. (2006) K teorii sistem lineynyh raznostnyh uravneniy s nepreryvnym argumentom. DAN. — 73, № 2. — pp. 269–272.
4. YEROMINA, T. (2014) Doslidzhennya struktury mnozhyny neperervnyh rozvyazkiv sistem liniynyh riznytsevo-funktsionalnyh rivnyan. Neliniyni kolyvannya. — 17, № 3. — pp. 341–350.

Надійшла до редколегії 05.05.2015

УДК 539.3

Клімчук Т. В.¹, аспірант
Острик В. І.¹, д. ф.-м. н., проф.

Ковзний контакт напівнескінченного штампа з пружною смугою

¹ Київський національний університет імені
Тараса Шевченка, 03680,
м. Київ, просп. Глушкова, 4 е,
e-mail: tarasyniv@ukr.net

T. V. Klimchuk¹, graduate student
V. I. Ostryk¹, Dr. Sci. (Phys.-Math.), Prof.

Sliding contact of the semi-infinite stamp with an elastic strip

¹ Taras Shevchenko National University of Kyiv,
03680, Kyiv, Glushkova st., 4 e,
e-mail: tarasyniv@ukr.net

Розглянуто плоску контактну задачу теорії пружності про вдавлювання напівнескінченного штампa в пружну смугу, нижня межа якої жорстко закріплена, у випадку повного проковзування між підшовою штампa і межею смуги з урахуванням сил тертя в області контакту. Із застосуванням методу Вінера – Гопфа отримано точний аналітичний розв'язок задачі. Визначено розподіли напружень в області контакту та на закріпленій грані смуги, а також нормальні переміщення вільної від навантаження частини межі смуги.

Ключові слова: ковзний контакт, пружна смуга, штамп, метод Вінера – Гопфа, напруження.

In present paper the sliding contact problem of the semi-infinite stamp with rectilinear basis and an elastic strip with rigidly fixed bottom edge is considered. Corresponding mixed boundary-value problem of theory of elasticity was reduced to single integral equation by means of using the Fourier integral transform. The functional equation was obtained after expanding the integral equation on whole real axis and settling the functions which are analytical in upper and lower half-planes of the complex plane. The solution of the last one was realized by the Wiener – Hopf method. The factorization of the coefficient of functional equation was implemented in the form of infinite products which contain roots of transcendental equations. Obtained exact analytical solution for stress and displacement includes the power series with exponential convergence. The results of the calculation of contact stress and normal displacement on the top edge and also normal and tangential stresses on the bottom edge are presented. The calculation was performed for different values of friction coefficient and compared with results of smooth contact.

Key words: sliding contact, elastic strip, stamp, Wiener – Hopf method, stress.

Статтю представив д. ф.-м. н., проф. Жук Я. О.

Вступ

У контактній механіці окремий розділ складають задачі для пружної смуги, що пов'язано з необхідністю досліджень властивостей конструкційних матеріалів з покриттями. Отримані, в основному, наближені асимптотичні розв'язки контактних задач для вузької або широкої (в порівнянні з розміром штампa) смуги у випадку гладкого контакту, тобто без врахування сил тертя в області контакту [1, 2].

Дана робота присвячена дослідженню напружено-деформованого стану в пружній смузі при фрикційному контакті з напівнескінченим штампом. Раніше ця задача без врахування сил тертя в області контакту розв'язана в [3].

Постановка задачі

Розглянемо пружну смугу $-\infty < x < \infty$, $-h \leq y \leq h$ з модулем зсуву G та коефіцієнтом Пуассона ν

в умовах плоскої деформації (рис. 1). Нехай нижня грань $y = -h$ ($-\infty < x < \infty$) смуги жорстко закріплена. У верхню грань $y = h$ смуги на проміжку $0 \leq x < \infty$ вдавлюється штамп з прямолінійною основою, який під дією рівномірно розподілених нормального p і дотичного q навантажень рухається рівномірно з малою швидкістю так, що динамічними ефектами можна знехтувати. Інша частина $-\infty < x < 0$ верхньої межі $y = h$ вільна від навантаження. В області контакту $0 \leq x < \infty$, $y = h$ підшови штампa з частиною верхньої грані смуги нормальні та дотичні напруження підпорядковані закону тертя Амонтона (Кулона). Для забезпечення проковзування штампa вздовж смуги на інтенсивності p і q навантаження необхідно накласти умову $q = \mu_0 p$, де $|\mu_0|$ – коефіцієнт тертя; випадку руху штампa вліво відповідають значення $\mu_0 > 0$, $q > 0$, вправо – значення $\mu_0 < 0$, $q < 0$ ($p > 0$).

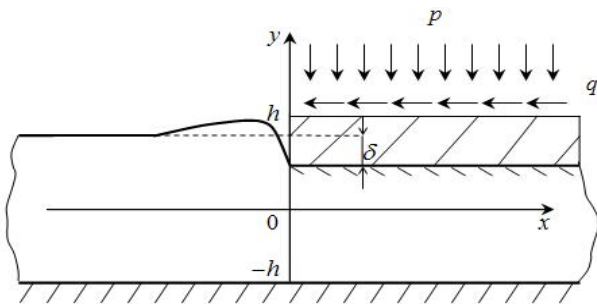


Рис. 1. Взаємодія штампа з пружною смугою

Крайові умови задачі наступні:

$$\begin{aligned} u_y|_{y=h} &= -\delta, \quad \tau_{yx}|_{y=h} = \mu_0 \sigma_y|_{y=h} \quad (0 < x < \infty), \\ \sigma_y|_{y=h} &= 0, \quad \tau_{yx}|_{y=h} = 0 \quad (-\infty < x < 0), \\ u_x|_{y=-h} &= 0, \quad u_y|_{y=-h} = 0 \quad (-\infty < x < \infty), \end{aligned} \quad (1)$$

де δ – просідання штампа. Величина δ визначається з умов рівноваги елемента $x_0 \leq x \leq x_0 + \Delta x$, $-h \leq y \leq h$ смуги на нескінченності ($x_0 \rightarrow \infty$), який знаходиться в умовах однорідного напружено-деформованого стану: $\sigma_y = -p$, $\varepsilon_x = 0$, $\varepsilon_y = -\delta / (2h)$. Із закону Гука отримаємо

$$\delta = \frac{1-2\nu}{1-\nu} \frac{ph}{G}.$$

Інтегральне рівняння задачі та його розв'язання

Введемо невідому функцію нормальних контактних напружень

$$\sigma(x) = \frac{1}{2G} \sigma_y|_{y=h} \quad (0 < x < \infty) \quad (2)$$

та її трансформанту Фур'є

$$\tilde{\sigma}(\mu) = \frac{1}{2\pi} \int_0^\infty \sigma(r) e^{i\mu r} dr.$$

Розглянемо основну змішану крайову задачу для смуги, коли на грані $y = h$ задано нормальні $2G\sigma(x)$ та дотичні $2G\mu_0\sigma(x)$ напруження, а на грані $y = -h$ – нульові переміщення. Розв'язок шукатимемо за допомогою інтегрального перетворення Фур'є. На грані $y = h$ смуги матимемо:

$$u_y|_{y=h} = 2h \int_{-\infty}^\infty \frac{\lambda(2\mu h)}{2\mu h \Delta(2\mu h)} \tilde{\sigma}(\mu) e^{-i\mu x} d\mu, \quad (3)$$

$$\begin{aligned} \lambda(z) &= 2(1-\nu) \left[(3-4\nu) \operatorname{sh} 2z - 2z \right] - \\ &\quad - 2i\mu_0 \left[(3-4\nu)(1-2\nu) \operatorname{sh}^2 z - z^2 \right], \\ \Delta(z) &= (3-4\nu)(\operatorname{ch} 2z - 1) + 8(1-\nu)^2 + 2z^2. \end{aligned}$$

Розв'язок (3) задовольняє всі крайові умови (1), окрім першої. Підставивши вираз для нормальних переміщень з (3) в першу крайову умову (1), отримаємо інтегральне рівняння

$$2h \int_{-\infty}^\infty K(2\mu h) \tilde{\sigma}(\mu) e^{-i\mu x} d\mu = -\delta \quad (0 \leq x < \infty). \quad (4)$$

Функцію контактних напружень подамо у вигляді

$$\sigma(x) = -\frac{p}{2G} + \sigma_1(x), \quad \sigma_1(\infty) = 0.$$

Виконавши заміни змінних

$$x = 2h\xi, \quad r = 2h\eta, \quad \tau = 2\mu h$$

та ввівши нову невідому функцію

$$\varphi(\xi) = -\sqrt{2\pi} 2h\delta^{-1} \sigma_1(2h\xi) \quad (0 < \xi < \infty), \quad (5)$$

інтегральне рівняння (4) перепишемо у формі

$$\int_0^\infty k(\xi - \eta) \varphi(\eta) d\eta = f(\xi) \quad (0 \leq \xi < \infty), \quad (6)$$

$$k(\xi - \eta) = \frac{1}{2\pi} \int_{-\infty}^\infty K(\tau) e^{-i\tau(\xi - \eta)} d\tau,$$

$$f(\xi) = \frac{\sqrt{2\pi} hp}{\delta} \sum_{k=1}^\infty \frac{\lambda(-is_k)}{s_k^2 \Delta'(-is_k)} e^{-s_k \xi},$$

де s_k ($k = 1, 2, \dots$) – корені рівняння $\Delta(is) = 0$ з півплощини $\operatorname{Re} s > 0$.

Рівняння (6) є класичним інтегральним рівнянням Вінера – Гопфа. Для його розв'язання застосуємо метод факторизації [4].

Вважаючи, що $\varphi(\xi) = 0$ при $\xi < 0$, розповсюдимо інтегральне рівняння (6) на всю числову вісь і застосуємо до нього інтегральне перетворення Фур'є. Увівши невідомі функції комплексної змінної

$$\Phi^+(z) = \frac{1}{\sqrt{2\pi}} \int_0^\infty \varphi(\xi) e^{iz\xi} d\xi, \quad (7)$$

$$\Phi^-(z) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^0 e^{iz\xi} d\xi \int_0^\infty k(\xi - \eta) \varphi(\eta) d\eta,$$

аналітичні відповідно у півплощинах $\operatorname{Im} z > c^+$ і $\operatorname{Im} z < c^-$ ($c^+ < 0$, $c^- > 0$), отримаємо функціональне рівняння

$$K(z) \Phi^+(z) - \Phi^-(z) = F^+(z) \quad (c^+ < \operatorname{Im} z < c^-), \quad (8)$$

$$\begin{aligned} F^+(z) &= \frac{1}{\sqrt{2\pi}} \int_0^\infty f(\xi) e^{iz\xi} d\xi = \\ &= -\frac{1}{\delta} \frac{hp}{G} \sum_{k=1}^\infty \frac{\lambda(-is_k)}{s_k^2 \Delta'(-is_k)} \frac{1}{iz - s_k}. \end{aligned}$$

Факторизацію коефіцієнта $K(z)$ рівняння (8) виконаємо у вигляді нескінченних добутків:

$$K(z) = K(0)K^+(z)K^-(z), \quad K(0) = \frac{1-2\nu}{1-\nu},$$

$$K^+(z) = \prod_{k=1}^{\infty} \left(1 + \frac{iz}{\zeta'_k}\right) \left(1 - \frac{iz}{s_k}\right)^{-1},$$

$$K^-(z) = \prod_{k=1}^{\infty} \left(1 + \frac{iz}{\zeta_k}\right) \left(1 + \frac{iz}{s_k}\right)^{-1},$$

де ζ'_k і ζ_k ($k=1, 2, \dots$) – корені рівняння $\lambda(is) = 0$ із півплощин $\text{Re } s < 0$ і $\text{Re } s > 0$ відповідно.

Функцію $F^+(z)/K^-(z)$ подамо як різницю аналітичних у півплощинах $\text{Im } z > c^+$ і $\text{Im } z < c^-$ функцій:

$$\frac{F^+(z)}{K^-(z)} = f^+(z) - f^-(z),$$

$$f^+(z) = \frac{i}{\delta} \frac{hp}{G} \sum_{k=1}^{\infty} \frac{\lambda(-is_k)}{s_k^2 \Delta'(-is_k)} \frac{1}{K^-(is_k)(is_k + z)}.$$

В результаті згідно з [4] отримаємо розв'язок функціонального рівняння (8):

$$\Phi^+(z) = f^+(z) / [K(0)K^+(z)],$$

$$\Phi^-(z) = K^-(z)f^-(z).$$

Застосувавши обернене перетворення Фур'є до першого співвідношення (7), отримаємо розв'язок інтегрального рівняння (6):

$$\varphi(\xi) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \Phi^+(t) e^{-i\xi t} dt \quad (\xi > 0). \quad (9)$$

Визначення напружень та переміщень

Обчисливши інтеграл із (9) за допомогою теорії лишків та врахувавши заміни (2), (5), отримаємо контактні напруження ($\xi > 0$):

$$\frac{\sigma_y}{p} \Big|_{y=h} = -1 - K(0) \sum_{k=1}^{\infty} \frac{\Delta(i\zeta'_k) K^-(i\zeta'_k)}{\lambda'(i\zeta'_k)} e^{\zeta_k \xi}. \quad (10)$$

Аналогічно, обрахувавши інтеграл із (3) для $x < 0$, знайдемо нормальні переміщення ненавантаженої частини верхньої грані смуги ($\xi < 0$):

$$\frac{u_y}{\delta} \Big|_{y=h} = -\frac{1}{K(0)} \sum_{k=1}^{\infty} \frac{1}{K^+(is_k)} \frac{\lambda(is_k)}{s_k^2 \Delta'(is_k)} e^{s_k \xi}. \quad (11)$$

Поблизу краю штамп контактні напруження та нормальні переміщення мають характерну для задач ковзного контакту поведінку

$$\sigma_y \Big|_{y=h} = O(x^{-1/2+\gamma}), \quad x \rightarrow +0;$$

$$u_y \Big|_{y=h} = O((-x)^{1/2+\gamma}), \quad x \rightarrow -0,$$

$$\gamma = \frac{1}{\pi} \arctg \frac{(1-2\nu)\mu_0}{2(1-\nu)}.$$

яку визначаємо з формул (3), (7) із застосуванням тауберових теорем [4].

Напруження на нижній грані смуги дістанемо у вигляді комплексної комбінації

$$\frac{1}{p} (\sigma_y + i\tau_{yx}) \Big|_{y=-h} = -\frac{1}{2} (1 + i\mu_0) + 2(1-\nu) \frac{1}{\pi i} \int_{-\infty}^{\infty} \frac{\tilde{\lambda}(\tau)}{\tau \Delta(\tau) K^+(\tau)} e^{-i\tau \xi} d\tau, \quad (12)$$

$$\tilde{\lambda}(\tau) = [\tau + 2(1-\nu)] \text{ch } \tau + [\tau - (1-2\nu)] \text{sh } \tau -$$

$$-i\mu_0 \{ [\tau - 2(1-\nu)] \text{ch } \tau + [\tau + (1-2\nu)] \text{sh } \tau \}.$$

Обчислення напружень і переміщень, віднесених до величин p і δ відповідно, за формулами (10)–(12) проведено для коефіцієнта Пуассона $\nu = 1/3$ і різних значень коефіцієнта тертя $|\mu_0|$ ($\mu_0 = 0.25, 0.5, 1$ при проковзуванні штамп вліво; $\mu_0 = 0$ при гладкому контакті; $\mu_0 = -0.25, -0.5, -1$ при проковзуванні штамп вправо).

На рис. 2 наведено розподіли нормальних контактних напружень $\sigma_y/p \Big|_{y=h}$. Видно, що зі збільшенням коефіцієнта тертя у випадку руху штамп вправо ($\mu_0 < 0$) спостерігається зменшення контактного тиску в середній частині області контакту ($0.3 < x/(2h) < 1$). На краю області контакту ($x \rightarrow +0$) контактний тиск необмежено зростає. На відміну від контакту штамп з пружною півплощиною, де вплив дотичних зусиль, які виникають в області контакту за наявності тертя, на розподіл контактного тиску є незначним [1], у даній задачі цей вплив є суттєвим.

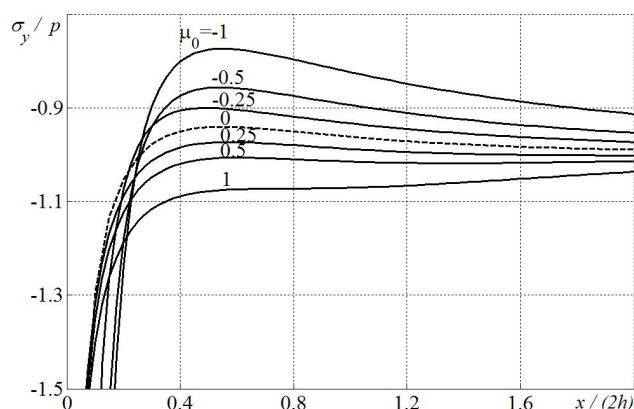


Рис. 2. Розподіли контактних напружень ($y = h$)

На рис. 3 зображено нормальні переміщення вільної від навантаження частини верхньої грані смуги. На відстані від краю штамп, яка приблизно дорівнює ширині смуги, її верхня межа піднімається, утворюючи пагорбок, висота якого зростає при збільшенні параметра μ_0 ; для від'ємних значень $\mu_0 < -0.5$ пагорбок зникає і вся верхня межа смуги просідає разом зі штампом.

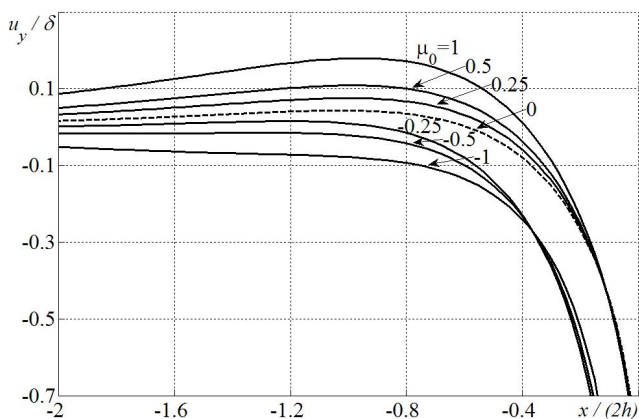


Рис. 3. Розподіли нормальних переміщень ($y = h$)

Рис. 4 і 5 ілюструють розподіли нормальних $\sigma_y/p|_{y=-h}$ і дотичних $\tau_{yx}/p|_{y=-h}$ напружень на закріпленій нижній грані смуги. Привертає увагу той факт, що у випадку проковзування штамп вправо ($\mu_0 < 0$) у певній зоні нижньої грані смуги поза штампом нормальні напруження переходять у напруження розтягу. При збільшенні навантаження це може призвести до відшарування смуги від жорсткої основи. Зсувні напруження на нижній грані смуги зменшуються поблизу краю штамп-

па внаслідок видавлювання пружного матеріалу з під штамп.

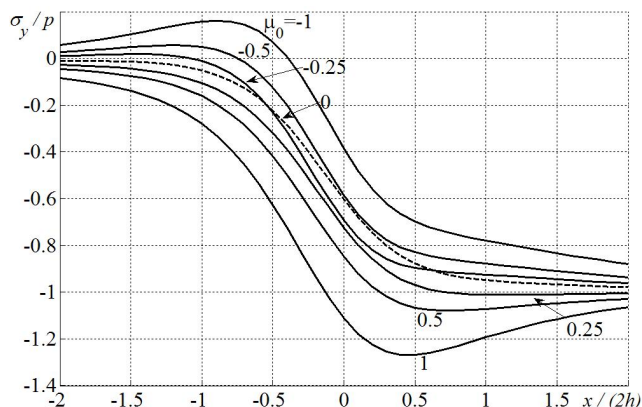


Рис. 4. Розподіли нормальних напружень ($y = -h$)

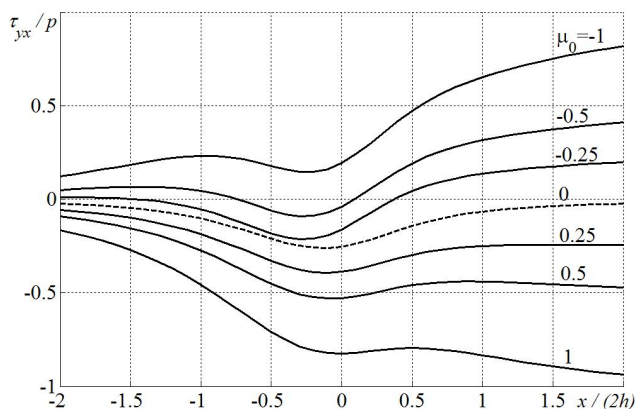


Рис. 5. Розподіли дотичних напружень ($y = -h$)

Список використаних джерел

1. Джонсон К. Механика контактного взаимодействия / К. Джонсон. – Москва: Мир, 1989. – 510 с.
2. Механика контактных взаимодействий / Под ред. И. И. Воровича, В. М. Александрова. – Москва, 2001. – 672 с.
3. Улитко А. Ф. Смешанная задача теории упругости для полосы на жестком основании / А. Ф. Улитко, В. И. Острик // Тр. III Всерос. конф. по теории упругости с междунар. участием. – Ростов-на-Дону: Новая книга, 2004. – С. 372–375.
4. Нобл Б. Метод Винера – Хопфа / Б. Нобл. – Москва: Изд-во иностр. лит., 1962. – 280 с.

References

1. DZHONSON, K. (1989) *Mehanika kontaktноgo vzaimodeistvija*. Moskva: Mir.
2. VOROVICH, I. and ALEKSANDROV, V. (editors) (2001) *Mehanika kontaktnih vzaimodeistvii*. Moskva.
3. ULITKO, A. and OSTRIK, V. (2004) Smeshannaja zadacha teorii uprugosti dlja polosity na zestkom osnovanii. *Trudy III Vserosijskoj konferentsii po teorii uprugosti*. Rostov-na-Donu: Novaja kniga. P. 372 – 375.
4. NOBL, B. (1958) *Methods based on the Wiener – Hopf technique*. London – New York – Paris – Los Angeles.

Надійшла до редколегії 09.06.15

УДК 539.595

Лимарченко О.С.¹, д.тех.н., проф.
Маджид М.¹, студент

Динаміка трубопроводу, що обертається, з рідиною

¹ Київський національний університет імені Тараса Шевченка, 03680, м. Київ, пр-т. Глушкова 4е,
e-mail: olelim2010@yahoo.com

Limarchenko O.S.¹, doctor of engineering, professor
Majid M.¹, student

Dynamics of rotating pipeline with liquid

¹ Taras Shevchenko National University of Kyiv,
03680, Kyiv, Glushkova ave., 4e,
e-mail: olelim2010@yahoo.com

Досліджується сумісний рух пружного трубопроводу з рідиною, що обертається з заданою кутовою швидкістю. Вивчаються нелінійні коливання системи для різних значень кутової швидкості. Показано, що обертання зменшує запас стійкості трубопроводу і призводить до прояву біфуркаційних режимів.

Ключові слова: пружний трубопровід з рідиною, обертання, режими втрати стійкості, альтернативне положення рівноваги.

We consider combined motion of rotating elastic pipeline with liquid. Pipeline is rotated with given angular velocity. It was shown that the presence of rotation does not change normal modes of oscillations, however, normal frequencies and, as consequence, critical velocities of liquid flow decrease. Rotation is reduced in the motion equations as supplementary force, which acts similar to centrifugal forces and weakens elastic forces. Therefore, under the presence of rotation stability loss is attained for lesser values of liquid velocities. Different modes of behavior of rotating pipeline were considered in a vicinity of liquid flow with half of critical velocity for pipeline without rotation. Numerical results showed that there are three typical modes of the system behavior. Increase of angular velocity results in normal frequency lowering. After reaching certain value of angular velocity, which corresponds to critical velocity of liquid flow in rotating reservoir, pipeline begins to perform oscillations in a vicinity of alternative position, normal frequencies increase. Further increase of rotation lead to growth of amplitudes of oscillations and finally this causes transition to oscillations with larger amplitudes and loss of stability.

Key Words: elastic pipeline with liquid, rotation, mode of loss of stability, alternative equilibrium position.

Статтю представив доктор фіз.-мат. наук, професор Я.О. Жук

Вступ

В техніці широко використовуються трубопроводи з рідиною, що обертаються. Насамперед такі інженерні системи пов'язані з використанням різних технологій буріння, коли по трубопроводу тече рідина, яка або змиває частину зруйнованої породи, або розкручує турбінний бур. При моделюванні трубопроводів переважна більшість задач, що розглядаються, досліджується без врахування поступального чи обертального руху основи. Розв'язання таких задач найбільш успішно проводиться на основі наближених методів, основаних на варіаційних алгоритмах. Додаткову складність ця задача містить через потребу описувати вихідну механічну систему в мішаному

вигляді – лагранжевий опис для пружної труби, ейлерів опис для рідини. В цілому розроблений раніше алгоритм для випадку поступального руху основи, на якій встановлено трубопровід, дозволив виконати модифікацію для випадку трубопроводу, який обертається [1].

Досліджується рух пружного трубопроводу з рідиною, що тече, на основі, яка виконує заданий обертальний рух навколо вісі, яка збігається з поздовжньою віссю трубопроводу в незбуреному прямолінійному стані. З одного боку трубопровід консольно защемлений, другий край вільний. Розглянуто декілька характерних діапазонів обертання трубопроводу. Вважається, що в початковий момент часу система знаходиться у спокої, проте кінематично збурена перша форма

коливань з амплітудою $c_1(0) = 0,01$. Вважаємо, що рідина однорідна, нестислива, ідеальна. Її поздовжній рух заданий.

Метою роботи є вивчення динамічних процесів руху системи трубопроводу – рідина при обертанні навколо поздовжньої вісі.

Математична модель

Математична модель системи будується повністю аналогічно випадку поступального руху основи трубопроводу, тому приведемо лише відмінності і додаткові члени в моделі. На відміну від випадку без обертання в системі додається потенціальна енергія, обумовлена дією відцентрових сил, що виникають через обертання

$$P_{rot} = -\frac{1}{2}(\mu + \rho)\omega^2 \int_0^l u^2 dx.$$

Тут μ і ρ є лінійними густинами труби і рідини, ω – задана кутова швидкість обертання трубопроводу, u – поперечна деформація трубопроводу, l – довжина трубопроводу.

В підсумку наявність цього додаткового члена після застосування методу модальної декомпозиції в формі розкладу рухів трубопроводу по формах коливань A_i з амплітудними параметрами

$u(x, t) = \sum_i c_i(t) A_i(x)$
 c_i () приведе до таких рівнянь руху

$$\begin{aligned} \ddot{c}_r = & -\frac{EJ}{\rho + \mu} \kappa_r^4 c_r + \frac{7}{2} \frac{\rho V^2}{(\rho + \mu) N_r} \sum_i c_i \beta_{ir}^2 + \\ & + \frac{2\rho V}{(\rho + \mu) N_r} \sum_i \dot{c}_i (\beta_{ri}^1 - \beta_{ir}^1) - \frac{PF}{(\rho + \mu) N_r} \sum_i c_i \beta_{ir}^2 + \\ & + (\rho + \mu) \omega^2 c_r - \frac{2\rho \dot{V}}{(\rho + \mu) N_r} \sum_i c_i \beta_{ir}^2 - \\ & - \frac{1}{2N_r} \sum_{ijk} \ddot{c}_i c_j c_k d_{jki}^2 - \sum_{ijk} \dot{c}_i \dot{c}_j c_k \frac{1}{N_r} \left(d_{jki}^2 - \frac{1}{2} d_{krij}^2 \right) - \\ & - \frac{EJ}{(\rho + \mu) N_r} \sum_{ijk} c_i c_j c_k d_{ijkl}^6 - \frac{EF}{2(\rho + \mu) N_r} \sum_{ijk} c_i c_j c_k d_{ijk}^4 - \\ & - \frac{13}{4} \frac{\rho V^2}{(\rho + \mu) N_r} \sum_{ijk} c_i c_j c_k d_{ijk}^4. \end{aligned}$$

Зміст коефіцієнтів рівнянь залишається таким де як і у випадку трубопроводу без обертання [1]. Звернемо увагу, що в наслідок обертання в рів-

няннях руху додається один член, який підкреслений двома рисками. Відразу зазначимо, що його дія є протилежною до дії оновлюючої сили. В той же час структура рівнянь в порівнянні з випадком без обертання не змінилася. Форми коливань трубопроводу і хвильові числа k_i у випадку обертання залишаються такими ж як і у випадку нерухомої основи трубопроводу, проте частотні параметри p_i змінюються

$$p_i^2 = k_i^4 a^2 - \omega^2, \quad \text{де} \quad a^2 = \frac{EJ}{\mu + \rho}.$$

Тут EJ – згинна жорсткість трубопроводу, a – швидкість поширення збурень.

Для оцінки як обертання впливає на критичну швидкість течії рідини випишемо лінеаризоване рівняння руху системи в дискретних параметрах для першої форми коливань

$$\ddot{c}_1 = \Omega^2 c_1.$$

де Ω – частота коливань трубопроводу.

$$\Omega^2 = -\frac{EJ}{(\rho + \mu)} \chi_1^4 + \frac{7\rho V^2}{2(\rho + \mu) N_1} \beta_{11}^2 - \frac{PF}{(\rho + \mu) N_1} \beta_{11}^2 - \omega^2.$$

Критичну швидкість V_{KP}^1 одержимо з умови рівності нулю частотного параметра ω . Тоді

$$V_{KP}^1 = \pm \sqrt{\frac{2}{7\rho} \left(\frac{EJ N_1 \chi_1^4}{\beta_{11}^2} + PF - \omega^2 \frac{(\rho + \mu) N_1}{\beta_{11}^2} \right)}.$$

Із виразу для критичної швидкості видно, що обертання сприяє зменшенню критичної швидкості.

Чисельний приклад

Розглянемо задачу про рух трубопроводу з консольним закріпленням з одного боку і вільним краєм з іншого. Трубопровід може виконувати обертання з заданою кутовою швидкістю. При числовій реалізації введемо наступні фізичні і геометричні параметри трубопроводу при розгляді впливу кутового руху. Вважаємо трубопровід циліндричним, виробленим з алюмінію, довжиною 1 м, внутрішній радіус 2 см, товщина труби – 2 мм. Фізичні параметри для рідини приймаються як для води, що тече в трубопроводі. Розглядається випадок початкового відхилення трубопроводу по першій формі з амплітудою $c_1(0) = 0,01$.

Розглянемо варіанти обертання трубопроводу з кутовими швидкостями до $\omega = 20\pi$. При цьому швидкість течії рідини $V = 0,5V_{KP}^1$, де критична швидкість обирається для випадку, коли обертання

трубопроводу відсутнє. Перша критична швидкість для трубопроводу без обертання є характерною величиною, яка для обраних фізичних і геометричних параметрів рівна 20,94 м/с. Будемо розглядати поведінку трубопроводу при швидкості течії рідини $V = 0,5V_{KP}^1$, що приблизно відповідає 10 м/с. Період першої форми коливань для такого трубопроводу буде $T = 0,1567$ с. При обчисленнях вводиться безрозмірний час як тривалість періоду коливань за першою формою. Далі в результатах час фактично відповідає кількості періодів коливань по першій формі. Ця власна частота коливань буде $\Omega_1 = \frac{2\pi}{T} \approx 40,1 \frac{1}{c}$.

Чисельні результати одержувалися на основі нелінійної моделі динаміки трубопроводу. При

цьому приймалося до уваги 7 форм коливань трубопроводу.

Для різних частот обертання можна на основі прийнятої моделі дослідити поведінку системи і тепер вже за рахунок зміни швидкості обертання змінювати ступінь наближення системи до критичної швидкості, тобто до внутрішнього резонансу. Було проаналізовано поведінку системи на інтервалі зміни швидкостей обертання від $\omega = 0$ до $\omega = 20\pi$. Найбільш характерні випадки поведінки системи спостерігаються при $\omega_1 = 0$, $\omega_2 = 10\pi$, $\omega_3 = 16\pi$, $\omega_4 = 20\pi$. Нумерація кривих виконана у відповідності з нумерацією частот обертання трубопроводу.

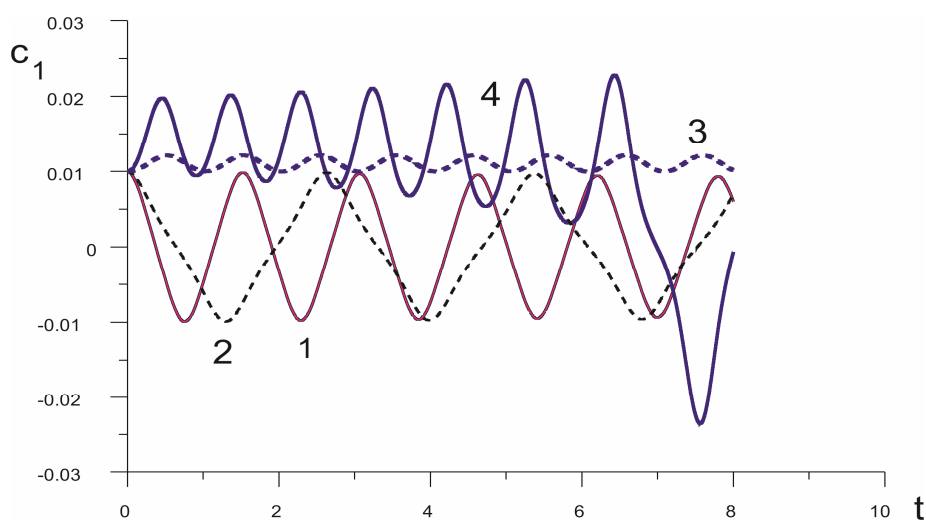


Рис 1. Зміна в часі амплітуди першої форми коливань трубопроводу для різних варіантів кутової швидкості

На Рис. 1 показано зміну в часі амплітуди першої форми коливань трубопроводу для різних варіантів кутової швидкості обертання. Для $\omega_1 = 0$ в системі відбуваються стійкі коливання. При збільшенні швидкості обертання $\omega_2 = 10\pi$ стійкість процесів не порушується, проте у відповідності із співвідношенням для залежності частоти коливань від кутової швидкості частота коливань спадає. При подальшому збільшенні швидкості обертання до $\omega_3 = 16\pi$ в системі відбувається наближення до області, де швидкість течії стає близькою до критичної. Тепер коливання відбуваються вже в околі альтернативного положення динамічної рівноваги. При

цьому частота коливань значно зростає і стає навіть вищою за випадок коливань трубопроводу без обертання. Тобто обертання в цьому діапазоні зміни кутових швидкостей має стабілізуючу роль. При подальшому зростанні частоти обертання $\omega_4 = 20\pi$ частота коливань змінюється несуттєво, однак зростає амплітуда коливань і в кінці інтервалу часу, що розглядається, коливання навколо альтернативного положення рівноваги втрачають стійкість.

Таким чином обертання приводить до виникнення трьох режимів динамічної поведінки трубопроводу.

До швидкостей обертання приблизно рівних $\omega_2 = 15\pi$ в системі відбуваються стійкі коливання

в околі початкового прямолінійного положення трубопроводу; при цьому із збільшенням швидкості обертання частота коливань падає.

В діапазоні $15\pi < \omega < 18\pi$ коливання трубопроводу відносно прямолінійного положення можуть втратити стійкість і будуть відбуватися відносно альтернативного положення рівноваги; при цьому частота коливань зростає і стає більшою ніж у випадку коливань трубопроводу без обертання.

При коливаннях трубопроводу в діапазоні більшому за $18\pi < \omega$ коливання відбуваються із зростаючою амплітудою; можлива втрата стійкості коливань відносно альтернативного положення рівноваги і перехід до коливань відносно початкового прямолінійного положення трубопроводу; частота коливань змінюється несуттєво в порівнянні з коливаннями на другому діапазоні динамічної поведінки трубопроводу.

Висновки

На основі варіаційного підходу поширено застосування моделі [1, 2] на випадок обертального руху трубопровода з рідиною.

Чисельно проаналізовано групу прикладів по коливаннях трубопроводу, що обертається, з рідиною. Досліджується розвинення перехідних

процесів на інтервалі перших 8–10 періодів коливань по першій формі.

Показано, що обертання трубопроводу призводить до появи трьох діапазонів поведінки трубопроводу з рідиною, що тече. Вони обумовлені тим, що обертання зменшує значення критичних швидкостей течії рідини.

В докритичному діапазоні розвиток коливань відбувається стійко відносно прямолінійного положення рівноваги трубопроводу з частотою, яка зменшується із зростанням частоти обертання.

В білякритичному діапазоні коливання відбуваються в околі альтернативного положення рівноваги з частотою, що незначно перевершує частоту коливань трубопроводу без обертання. При зростанні швидкості обертання трубопроводу частота коливань змінюється незначно, амплітуда зростає і можлива втрата стійкості коливань трубопроводу.

Таким чином в цілому обертання трубопроводу спричиняє дестабілізуючу роль для коливань. В цьому випадку прояви режимів втрати стійкості, прояви біфуркацій цих режимів можливі навіть при швидкостях течії, які набагато менші за критичні значення швидкостей для випадку трубопроводу без обертання.

Список використаних джерел

1. Лимарченко В.О. Варіаційний метод дослідження нелінійних задач динаміки трубопроводів з рідиною / Лимарченко В.О., Лимарченко О.С., Тімохін О.П. // Збірник праць Інституту математики НАН України. – 2012. – Т. 9. – №2. – С. 224-239
2. Limarchenko O., Rotational motion of structures with tanks partially filled by liquid / O. Limarchenko, G. Matarazzo, V. Yasinsky – Kyiv: FADA, 2003. – 286 p.

References

1. LIMARCHENKO, V.O., LIMARCHENKO O.S., TIMOKHIN A.P. (2012) *Variational method for investigation of nonlinear problems of dynamics of pipelines with liquid*. – Proceedings of Institute of Mathematics NAS of Ukraine. – 9. – № 2, P. 224-239
2. LIMARCHENKO, O., MATARAZZO, G., YASINSKY, V. (2003) *Rotational motion of structures with tanks partially filled by liquid*. Kyiv.

Надійшла до редколегії 19.03.15

УДК 539.3

Остос О. Х.¹, студент
Острик В. І.¹, д. ф.-м. н., проф.

Згин пружної смуги зосередженими силами

¹ Київський національний університет імені
Тараса Шевченка, 03680,
м. Київ, просп. Глушкова, 4 е,
e-mail: ostos.alexander1994@yandex.ua

Ostos O. H.¹, student
V. I. Ostryk¹, Dr. Sci. (Phys.-Math.), Prof.

Bending of elastic strip by concentrated forces

¹ Taras Shevchenko National University of Kyiv,
03680, Kyiv, Glushkova st., 4 e,
e-mail: ostos.alexander1994@yandex.ua

Розглянуто згин пружної смуги трьома зосередженими нормальними силами, одна з яких прикладена до однієї грані смуги, а дві інших сили – до іншої грані. Із застосуванням інтегрального перетворення Фур'є знайдено точний аналітичний розв'язок задачі. Подано результати розрахунків розподілів напружень в різних поперечних перерізах смуги та нормальних переміщень на середній лінії смуги. Проведено порівняння отриманих результатів з результатами наближеного розв'язку задачі у межах теорії балок Кірхгофа.

Ключові слова: згин, пружна смуга, напруження, переміщення, балка.

It was considered the bend of elastic strip by three normal concentrated forces, which one of them was put on the one side of the strip, and two others forces were put on the another side. With the application of integral transformation of Fourier it was found the exact analytic solution of the problem, according to which the stresses were expressed by convergent integrals of Fourier, and displacements by divergent integrals. To define the latter it was used the apparatus of general functions. It was given the results of calculations of the stress distribution in the different across sections of the bend and the normal displacements on the middle line of the bend. It was compared the obtained results with the results of the approximate solution of the problem in the bound of the theory of Kirchhoff's beams. In the last case it was considered the bend of the beam with the hinge fixing edges by central concentrated force. It was elucidated, that in case of relatively thin beams the theory of Kirchhoff give good results for crooks and stresses on some distance from the points of putting the concentrated forces. It was defined the zones of great concentration of tangent stresses in the across sections of the strip near the lines where concentrated forces had effect. Stress components exceed more than several times the stress components, which were obtained by the beam's theory in the zones.

Key Words: bend, elastic strip, stress, displacement, beam.

Статтю представив д. ф.-м. н., проф. Жук Я. О.

Вступ

Як відомо [1], використання прикладних теорій згину тонких пластин, балок та оболонок, які базуються на певних фізичних гіпотезах, призводить до значних похибок при визначенні напружено-деформованого стану вказаних тіл поблизу їх меж та в околі концентраторів напружень.

З метою визначення меж застосування теорії Кірхгофа згину балок та уточнення отриманих за цією теорією результатів нижче розглядається задача про згин балки, яка шарнірно опирається своїми кінцями на дві опори, центральною зосередженою нормальною силою. Ця задача еквівалентна задачі циліндричного згину смугової шарнірно опертої пластини рівномірно розподіленими уздовж лінії симетрії пластини поперечними

силами. Порівнюються два розв'язки: точний розв'язок рівнянь теорії пружності для пружної смуги, тобто для пластини, яка знаходиться у стані плоскої деформації, та наближений розв'язок за теорією балок Кірхгофа з використанням гіпотези прямої нормалі.

Підхід, який полягає в комбінуванні розв'язку рівнянь теорії пружності в околі зосередженої нормальної сили (задачі Фламана для півплощини [2]) та розв'язку за теорією Кірхгофа поза вказаним околom, розвинуто в роботах [3, 4].

Постановка задачі

Пружна смуга, яка займає область $-\infty < x < \infty$, $-h \leq y \leq h$ і має пружні характеристики G і ν (G – модуль зсуву, ν – коефіцієнт Пуассона),

знаходиться у стані плоскої деформації. Смуга згинається трьома зосередженими нормальними силами: сила P прикладена до верхньої грані $y = h$ смуги у точці $x = 0$, сили $P/2$ – до нижньої грані $y = -h$ у точках $x = a$, $x = -a$.

Відповідна крайова задача для смуги складається із двох рівнянь рівноваги Ламе відносно компонент u_x , u_y вектора переміщень [2]:

$$\begin{aligned} 2(1-\nu)\frac{\partial^2 u_x}{\partial x^2} + (1-2\nu)\frac{\partial^2 u_x}{\partial y^2} + \frac{\partial^2 u_y}{\partial x \partial y} &= 0, \\ \frac{\partial^2 u_x}{\partial x \partial y} + (1-2\nu)\frac{\partial^2 u_y}{\partial x^2} + 2(1-\nu)\frac{\partial^2 u_y}{\partial y^2} &= 0 \end{aligned} \quad (1)$$

і крайових умов відносно нормальних та дотичних напружень

$$\begin{aligned} \sigma_y|_{y=h} &= -P\delta(x), \\ \sigma_y|_{y=-h} &= -\frac{P}{2}[\delta(x-a) + \delta(x+a)], \\ \tau_{xy}|_{y=\pm h} &= 0 \quad (-\infty < x < \infty), \end{aligned} \quad (2)$$

де $\delta(x)$ – дельта-функція Дірака.

Розв'язання крайової задачі

Для розв'язання крайової задачі (1), (2) застосуємо інтегральне перетворення Фур'є за змінною x . Відносно трансформант переміщень

$$\tilde{u}_x = \int_{-\infty}^{\infty} u_x e^{ipx} dx, \quad \tilde{u}_y = \int_{-\infty}^{\infty} u_y e^{ipx} dx \quad (3)$$

із (1) отримаємо систему звичайних диференціальних рівнянь

$$\begin{aligned} (1-2\nu)\frac{d^2 \tilde{u}_x}{dy^2} - 2(1-\nu)p^2 \tilde{u}_x - ip\frac{d\tilde{u}_y}{dy} &= 0, \\ -ip\frac{d\tilde{u}_x}{dy} + 2(1-\nu)\frac{d^2 \tilde{u}_y}{dy^2} - (1-2\nu)p^2 \tilde{u}_y &= 0. \end{aligned} \quad (4)$$

Увівши розв'язувальну функцію $\phi(y)$ за допомогою рівностей

$$\begin{aligned} \tilde{u}_x &= 2(1-\nu)\phi''(y) - (1-2\nu)p^2\phi(y), \\ \tilde{u}_y &= ip\phi'(y) \end{aligned} \quad (5)$$

так, щоб задовольнити тотожно друге рівняння (4), із першого рівняння (4) отримаємо диференціальне рівняння

$$\phi^{IV}(y) - p^2\phi''(y) + p^4\phi(y) = 0, \quad (6)$$

загальним розв'язком якого є функція

$$\phi(y) = A_0 \operatorname{ch} py + B_0 \operatorname{sh} py + C_0 y \operatorname{ch} py + D_0 y \operatorname{sh} py \quad (7)$$

з довільними сталими A_0 , B_0 , C_0 , D_0 .

Із (5), (7) знаходимо

$$\begin{aligned} \tilde{u}_x &= A \operatorname{ch} py + B \operatorname{sh} py + C[(3-4\nu)p^{-1} \operatorname{sh} py + y \operatorname{ch} py] + \\ &+ D[(3-4\nu)p^{-1} \operatorname{ch} py + y \operatorname{sh} py], \end{aligned} \quad (8)$$

$$-i\tilde{u}_y = A \operatorname{sh} py + B \operatorname{ch} py + C y \operatorname{sh} py + D y \operatorname{ch} py,$$

$$A = (A_0 p + D_0) p, \quad B = (B_0 p + C_0) p, \quad C = C_0 p^2, \quad D = D_0 p^2.$$

Трансформанти напружень дістаємо із (8) за допомогою закону Гука і співвідношень Коші [2]:

$$\begin{aligned} i(2G)^{-1} \tilde{\sigma}_x &= A p \operatorname{ch} py + B p \operatorname{sh} py + C[(3-2\nu) \operatorname{sh} py + \\ &+ p y \operatorname{ch} py] + D[(3-2\nu) \operatorname{ch} py + p y \operatorname{sh} py], \\ -i(2G)^{-1} \tilde{\sigma}_y &= A p \operatorname{ch} py + B p \operatorname{sh} py + C[(1-2\nu) \operatorname{sh} py + \\ &+ p y \operatorname{ch} py] + D[(1-2\nu) \operatorname{ch} py + p y \operatorname{sh} py], \\ (2G)^{-1} \tilde{\tau}_{xy} &= A p \operatorname{sh} py + B p \operatorname{ch} py + C[2(1-\nu) \operatorname{ch} py + \\ &+ p y \operatorname{sh} py] + D[2(1-\nu) \operatorname{sh} py + p y \operatorname{ch} py]. \end{aligned} \quad (9)$$

Крайові умови (2) відносно трансформант напружень набувають вигляду

$$\begin{aligned} \tilde{\sigma}_y|_{y=h} &= -\frac{P}{2G}, \quad \tilde{\sigma}_y|_{y=-h} = -\frac{P}{2G} \cos ap, \\ \tilde{\tau}_{xy}|_{y=\pm h} &= 0. \end{aligned} \quad (10)$$

Задовольнивши виразами (9) систему рівностей (10), отримаємо

$$\begin{aligned} A &= \frac{iP}{G} \frac{2(1-\nu) \operatorname{sh} ph + ph \operatorname{ch} ph}{p(\operatorname{sh} 2ph + 2ph)} \cos^2 ap, \\ B &= \frac{iP}{G} \frac{2(1-\nu) \operatorname{ch} ph + ph \operatorname{sh} ph}{p(\operatorname{sh} 2ph - 2ph)} \sin^2 ap, \end{aligned} \quad (11)$$

$$C = -\frac{iP}{G} \frac{\operatorname{ch} ph \sin^2 ap}{\operatorname{sh} 2ph - 2ph}, \quad D = -\frac{iP}{G} \frac{\operatorname{sh} ph \cos^2 ap}{\operatorname{sh} 2ph + 2ph}.$$

Оберненим перетворенням Фур'є із (8), (9), (11) знаходимо компоненти переміщень та напружень у пружній смузі, що відповідають розв'язку крайової задачі (1), (2). Зокрема, вертикальні переміщення на середній лінії смуги набувають вигляду

$$u_y|_{y=0} = -\frac{P}{\pi G} \int_{-\infty}^{\infty} \frac{2(1-\nu) \operatorname{ch} \tau + \tau \operatorname{sh} \tau}{\tau(\operatorname{sh} 2\tau - 2\tau)} \sin^2 \frac{a\tau}{2h} e^{-ix\tau/h} d\tau. \quad (12)$$

В останньому виразі підінтегральна функція має порядок величини τ^{-2} при $\tau \rightarrow 0$, а інтеграл є розбіжним і його слід розуміти в узагальненому сенсі. З використанням узагальненого значення інтеграла

$$\int_{-\infty}^{\infty} \frac{e^{-ix\tau/h}}{\tau^2} d\tau = -\pi \frac{|x|}{h}, \quad (13)$$

яке можна отримати подвійним інтегруванням за x інтеграла

$$\int_{-\infty}^{\infty} e^{-ix\tau/h} d\tau = 2\pi h \delta(x), \quad (14)$$

із (12) отримаємо

$$u_y|_{y=0} = 3(1-\nu) \frac{Pa^2}{8Gh^3} |x| + \frac{2P}{\pi G} \times \\ \times \int_0^{\infty} \left(\frac{2(1-\nu)ch\tau + \tau sh\tau}{\tau(2\tau - sh\tau)} \sin^2 \frac{a\tau}{2h} + \frac{3(1-\nu)a^2}{8h^2\tau^2} \right) \cos \frac{x\tau}{h} d\tau. \quad (15)$$

Наближений розв'язок за теорією Кірхгофа

У теорії Кірхгофа циліндричного згину тонких пластин переміщення u_y наближено дорівнюють функції прогину $-w(x)$ (з протилежним знаком), яка вважається незмінною по товщині пластини і задовольняє диференціальне рівняння

$$w^{IV}(x) = \frac{P}{D} \delta(x) \quad (D = \frac{4}{3(1-\nu)} Gh^3, -a < x < a) \quad (16)$$

та крайові умови шарнірного опирання

$$w(\pm a) = 0, \quad w''(\pm a) = 0. \quad (17)$$

Розв'язком крайової задачі (16), (17) є функція

$$w(x) = \frac{P}{12D} (|x|^3 - 3ax^2 + 2a^3), \quad (18)$$

а згинальний момент і узагальнена перерізуювальна сила виражаються через прогин так:

$$M_x = -D \frac{d^2 w}{dx^2}, \quad Q_x = -D \frac{d^3 w}{dx^3}. \quad (19)$$

Крім того, горизонтальні переміщення та напруження приймаються такими:

$$u_x = -y \frac{dw}{dx}, \quad \sigma_x = \frac{3y}{2h^3} M_x, \quad \sigma_y = 0, \\ \tau_{xy} = \frac{3}{4h^3} (h^2 - y^2) Q_x. \quad (20)$$

Результати обчислень

Безрозмірні вертикальні переміщення $\bar{u}_y = \frac{Gh^3}{Pa^3} (u_y|_{y=0} - u_y|_{y=0, x=a})$ середньої лінії смуги у точці $x = 0$, які обчислені за формулою (15) при $\nu = 0,3$, подано у табл. 1 для різних значень відносної ширини смуги h/a . Із зменшенням останнього параметра абсолютні значення $|\bar{u}_y|$ наближаються до значення 0,0875 безрозмірного прогину $\bar{w} = \frac{Gh^3}{Pa^3} w(x)$ у центрі балки ($x = 0$), розрахованого за формулою (18), а для $\bar{a} = a/h = 8$ від-

різняються від нього на 5 %. Графічно переміщення $\bar{u}_y$ (криві 1, 2, 3 для $\bar{a} = 2, 4, 8$ відповідно) і прогин $-\bar{w}$ (пунктир) показано на рис. 1.

Таблиця 1

a/h	$-\bar{u}_y (x=0)$	a/h	$-\bar{u}_y (x=0)$
2	0,1416	12	0,0896
3	0,1141	15	0,0889
4	0,1033	20	0,0883
6	0,0950	50	0,0877
8	0,0919	80	0,0876
10	0,0904	100	0,0875

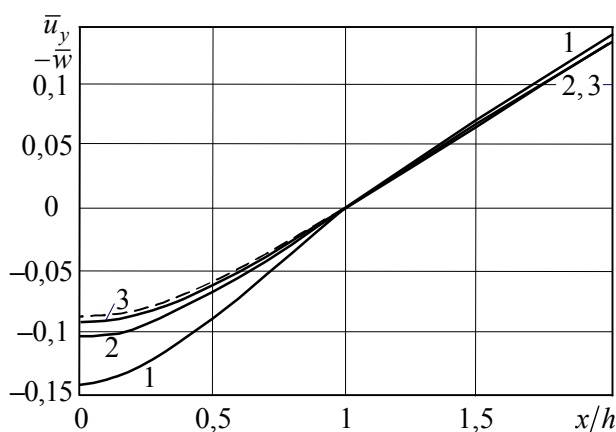


Рис. 1

Всі інші розрахунки проведено для значення $\bar{a} = 8$ і коефіцієнта Пуассона $\nu = 0,3$.

Розподіли безрозмірних нормальних напружень $\bar{\sigma}_x = (h/P)\sigma_x$ уздовж поперечних перерізів смуги $x = 0, 25a; 0,5a; 0,7a; a$ зображені кривими 1, 2, 3, 4 відповідно на рис. 2. Пунктирні лінії відповідають теорії Кірхгофа. Поодаль від ліній дії сил теорія Кірхгофа дає задовільні результати (розподіли 2, 3), а поблизу від однієї із вказаних ліній (розподіл 1) або безпосередньо на лінії $x = a$ дії сили $P/2$ (розподіл 4) напруження σ_x за точним розв'язком мають нелінійний розподіл, особливо в околі точки прикладання сили.

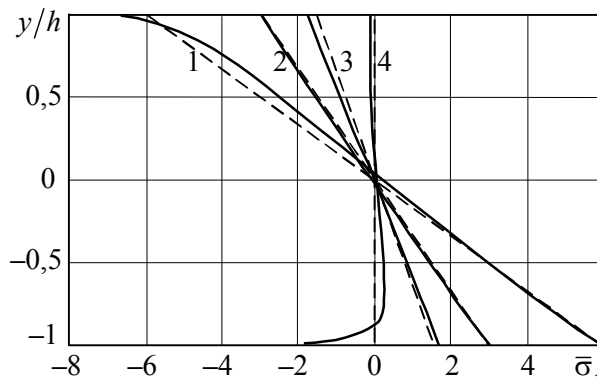


Рис. 2

На рис. 3 показано розподіли безрозмірних дотичних напружень $\bar{\tau}_{xy} = (h/P)\tau_{xy}$ по ширині смуги вздовж прямих $x = 0,9a; 0,95a; 0,97a; a$ (криві 1–4). За теорією Кірхгофа усі ці розподіли збігаються між собою і показані пунктирною лінією. Вони практично збігаються з розподілами дотичних напружень за точним розв'язком на певній відстані від ліній дій сил ($0,1a < |x| < 0,9a$).

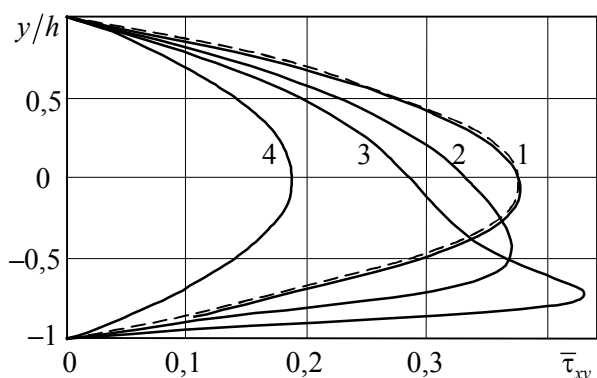


Рис. 3

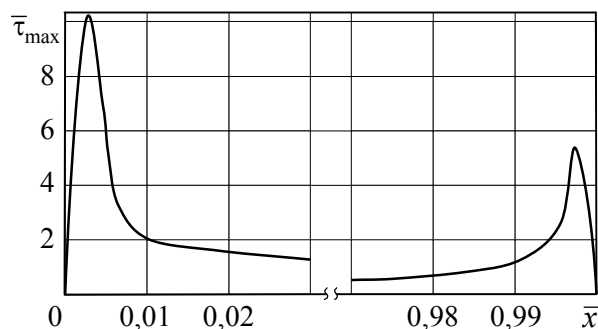


Рис. 4

У перерізах смуги, близьких до лінії $x = a$ дії сили максимум безрозмірних дотичних напружень $\bar{\tau}_{max}$ помітно зростає, а точка, у якій досягається цей максимум зсувається із середньої лінії смуги і підходить близько до межі смуги і точки прикладання сили. Як показано на рис. 4, значення $\bar{\tau}_{max}$ у поперечних перерізах смуги поблизу ліній $x = 0$, $x = a$ дій сил перевищують відповідне значення за теорією Кірхгофа ($\bar{\tau}_{max} = 0,37$) у декілька разів.

На рис. 5 показано вигляд нормалей $x = 0,125a; 0,25a; 0,5a; 0,75a$ (криві 1–4) після деформації пластини. Пунктирні лінії відповідають теорії Кірхгофа. Тут $\tilde{u}_x = k(G/P)u_x$, $\tilde{u}_y = y/h + k(G/P) \times (u_y - u_y|_{y=0})$, $k = 0,1149$. Незважаючи на те, що

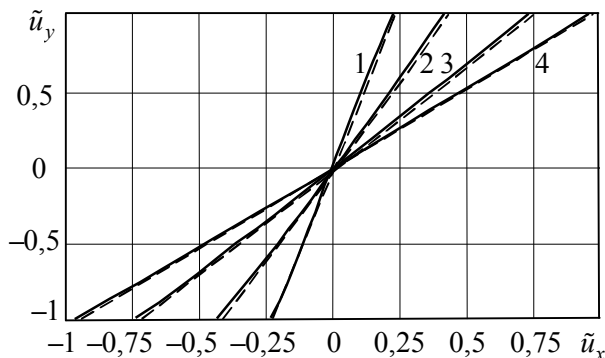


Рис. 5

за точним розв'язком нормаль пластини поза малими околами точок прикладання сил залишається майже прямою, все ж таки використання теорії Кірхгофа може давати значні похибки при обчисленні напружень та переміщень у пластині.

Список використаних джерел

1. Тимошенко С. П. Пластинки и оболочки / С. П. Тимошенко, С. Войновский-Кригер. – Москва: Наука, 1966. – 636 с.
2. Тимошенко С. П. Теория упругости / С. П. Тимошенко, Дж. Гудьер. – Москва: Наука, 1979. – 560 с.
3. Мазур С. Г. Уточненная теория изгиба тонких пластин сосредоточенными силами / С. Г. Мазур // Прикл. механика. – 1979. – 15, № 4. – С. 86-89.
4. Моргунов М. О. Границі застосування теорії Кірхгофа при згині пружного шару зосередженими силами / М. О. Моргунов // Вісник Київського університету. Математика. Механіка. – 2000. – № 5. – С. 72-80.

References

1. TIMOSHENKO, S. and Voinovskiy-Kriger, S. (1966) *Plastinki i obolochki*. Moskva: Nauka.
2. TIMOSHENKO, S. and Gud'jer, D. (1979) *Teoriya uprugosti*. Moskva: Nauka.
3. MAZUR, S. G. (1979) Utochnennaja teoriya izgiba tonkih plastin sosredotochennymi silami. *Prikladnaja mehanika*. 15, N 4. P. 86-89.
4. MORGUNOV, M. O. (2000) Granytsi zastosuвання teorii Kirhgofa pri zgyni pruzhnogo sharu zoseredzhenymy sylamy. *Visnyk Kyjv's'kogo universytetu. Matematika. Mehanika*. N 5. P. 72-80.

Надійшла до редколегії 08.06.15

КОМП'ЮТЕРНІ НАУКИ ТА ІНФОРМАТИКА

УДК 517.956

Анікушин А. В., к.ф.-м.н., асистент
Номіровський Д. А., д.ф.-м.н., професор

**Якісний аналіз лінійних
диференціальних моделей
псевдогіперболічного типу в деяких
класах узагальнених функцій**

Київський національний університет імені Та-
раса Шевченка, 03680, м. Київ, пр-т. Глушко-
ва 4д,
e-mail: anik_andrii@ukr.net
Київський національний університет імені Та-
раса Шевченка, 03680, м. Київ, пр-т. Глушко-
ва 4д,
e-mail: kashpir@mail.ru

A. V. Anikushyn, Ph.D.
D. A. Nomirovskii, Ph.D.

**Qualitative analysis of linear
differential pseudohyperbolic
equations in some classes
of distributions**

Department of Computational Mathematics of
Taras Shevchenko National University of Kyiv,
03680, Kyiv, Glushkova st., 4d,
e-mail: anik_andrii@ukr.net
Department of Computational Mathematics of
Taras Shevchenko National University of Kyiv,
03680, Kyiv, Glushkova st., 4d,
e-mail: kashpir@mail.ru

Вивчається крайова задача для лінійного диференціального псевдогіперболічного оператора третього порядку. Диференціальний оператор діє у просторах узагальнених функцій скінченного порядку типу просторів Соболева. Розглядається узагальнена постановка. Доведено, що псевдогіперболічний оператор задовольняє ланцюг апріорних нерівностей в різних класах узагальнених функцій. Доведено теореми існування та єдиності розв'язку псевдогіперболічного рівняння.

Ключові слова: псевдогіперболічне рівняння, якісний аналіз, узагальнені функції, розв'язок.

The paper deals with linear third order differential pseudo-hyperbolic equations defined in a cylindrical domain. We require that the state function of the equations satisfies the homogeneous initial and boundary conditions of the first, second, or third types. The main operator of the pseudo-hyperbolic equation is required to be uniformly elliptic with respect to the space variables.

The chief result of the article is the existence and uniqueness theorems for the pseudo-hyperbolic equations in the terms of theory of distributions (generalized functions) of finite order. Since the problem is considered in the weak formulations, we continuously extend the classical pseudo-hyperbolic operator and its adjoint operator to operators defined in certain spaces of distributions. This allows us to prove an infinite chain of a priori inequalities for the extended pseudo-hyperbolic operators and their adjoint operators. The proof is based on using specific auxiliary integral and differential operators with respect to the time variable. The existence and uniqueness theorems for both the pseudo-hyperbolic equations and their adjoint ones proceed from these inequalities.

Key Words: pseudo-hyperbolic equation, qualitative analysis, distributions, solution.

Статтю представив д.т.н., с.н.с. Кудін В.І.

У циліндричній області $(t, \xi) \in Q = (0, T) \times \Omega$, де $\Omega \subset \mathbb{R}^n$ — обмежена однозв'язна область з регулярною межею $\partial\Omega$, розглянемо лінійний диференціальний псевдогіперболічний оператор:

$$\mathcal{L}u = \frac{\partial^2 u}{\partial t^2} + A \left(\frac{\partial u}{\partial t} \right) + B(u). \quad (1)$$

Диференціальні оператори A та B не зале-

жать від часової змінної t і задаються стандартними загальними диференціальними виразами другого порядку за просторовими змінними $\xi = (\xi_1, \dots, \xi_n)$. Наприклад,

$$A(u) = - \sum_{i,j=1}^n (a_{ij} u_{\xi_i})_{\xi_j} + \sum_{i=1}^n a_i u_{\xi_i} + au.$$

Вимагається, щоб оператор A був рівномірно

еліптичним в області $\bar{\Omega}$.

Межа $\partial\Omega$ області Ω складається з трьох регулярних частин $\gamma_1, \gamma_2, \gamma_3$. Покладемо $\Gamma_i = (0, T) \times \gamma_i$, де $i \in \{1, 2, 3\}$. Функція стану $u(t, \xi)$ задовольняє однорідні крайові умови:

$$u|_{\Gamma_1} = \frac{\partial \mathcal{L}u}{\partial \vec{n}}|_{\Gamma_2} = \left(a_0(x)u + \frac{\partial \mathcal{L}u}{\partial \vec{n}}\right)|_{\Gamma_3} = 0 \quad (2)$$

та початкові умови:

$$u|_{t=0} = \frac{\partial u}{\partial t}|_{t=0} = 0, \quad (3)$$

де $\partial \mathcal{L}u / \partial \vec{n} = (\text{grad}_{\mathcal{L}} u, \vec{n})_{\mathbb{R}^n}$, а $\mathcal{L}$ -градієнт $\text{grad}_{\mathcal{L}}$ визначено в [1].

В статті [1] для оператора $\mathcal{L}$ доведено апіорні нерівності в деяких просторах узагальнених функцій скінченного порядку:

$$\begin{aligned} \|u\|_{W_0^0} &\leq c_1 \|\mathcal{L}u\|_{(W_T^1)^*} \leq c_2 \|u\|_{H_0^1} \leq \\ &\leq c_3 \|\mathcal{L}u\|_{(H_T^1)^*} \leq c_4 \|u\|_{W_0^1}, \end{aligned} \quad (4)$$

де $u(t, \xi)$ — довільна функція простору W_0^1 , c_i — додатні сталі, що не залежать від $u(t, \xi)$.

Допоміжні інтегро-диференціальні оператори, які використано у доведенні нерівностей (4), дозволяють встановити подібні співвідношення для спряженого псевдогіперболічного оператора. Крім цього, на основі отриманих апіорних нерівностей можна вивчати лінійний псевдогіперболічний оператор, що діє в інших парах просторів. Зокрема, нерівності (4) можна "зсунути" на диференціальний оператор d^k/dt^k при довільних цілих значеннях k .

Розповсюдимо позначення $u^{(k)}$ для k -ої похідної за змінною t для функції $u \in L_0$ на випадок цілих від'ємних значень k . У випадку $k \in \mathbb{Z}$, $k < 0$ під $u^{(k)}$ будемо розуміти функцію, яку одержано після $|k|$ -кратного застосування до функції $u(t, \xi)$ оператора інтегрування

$$u(t, \xi) \rightarrow \int_0^t u(\tau, \xi) d\tau.$$

Використовуючи інтегральний оператор

$$v(t, \xi) \rightarrow \int_T^t v(\tau, \xi) d\tau,$$

введемо для кожної функції $v \in L_T$ подібне позначення: $v^{[k]}$. Зауважимо, що функція $v^{[k]}$ для

$k \geq 0$ визначається як звичайна k -та похідна функції $v \in L_T$ за змінною t .

Використовуючи наведені позначення, уведемо до розгляду простори W_0^k , H_0^k , W_T^k і H_T^k для випадку цілих значень k . Наприклад, під W_T^k будемо розуміти простір, побудований як поповнення множини функцій L_T за нормою:

$$\|v\|_{W_T^k}^2 = \sum_{i=1}^n \int_Q \left(v_{\xi_i}^{[k]}\right)^2 dQ.$$

Зрозуміло, що природне розширення лінійного відображення $u \rightarrow u^{(k)}$, де $k \in \mathbb{Z}$, з множини L_0 на весь простір W_0^l , де $l \in \mathbb{Z}$, задає ізометрію між простором W_0^l (або відповідно H_0^l) і простором W_0^{l-k} (відповідно H_0^{l-k}).

З аналогічних причин розширення відображення $v \rightarrow v^{[k]}$, де $k \in \mathbb{Z}$, задає ізометрію між простором W_T^l (або відповідно H_T^l) і простором W_T^{l-k} (відповідно H_T^{l-k}).

Для всіх u і $k, m \in \mathbb{Z}$ виконується рівність:

$$(u^{(k)})^{(m)} = u^{(k+m)}.$$

Зокрема, рівність $(u^{(1)})^{(-1)} = u$, яку формально можна подати у вигляді:

$$\int_0^t u(\tau, \xi) d\tau = u(t, \xi),$$

справджується для всіх W_0^0 , незважаючи на те, що простір W_0^0 містить і такі нескінченно диференційовні функції $u(t, \xi)$, які не обертаються в нуль при $t = 0$.

Аналогічне твердження має місце і для відображення: $v \rightarrow v^{[k]}$.

Використовуючи формулу інтегрування частинами і початкові умови на функції u і v , можна показати, що для довільних функцій $u \in L_0$ і $v \in L_T$ має місце рівність:

$$\begin{aligned} (\mathcal{L}u, v)_{L_2(Q)} &= (u, v_{tt} - av_t + bv)_{L_2(Q)} + \\ &+ \sum_{i,j=1}^n \left(\int_0^t u_{\xi_i} d\tau, a_{ij} v_{\xi_j t} - b_{ij} v_{\xi_j t} \right)_{L_2(Q)} + \quad (5) \\ &+ \sum_{i=1}^n \left(\int_0^t u_{\xi_i} d\tau, a_i v_{tt} - b_i v_t \right)_{L_2(Q)} - \\ &- \left(\int_0^t a_0 u d\tau, v_t \right)_{L_2(\Gamma_3)}. \end{aligned}$$

Праву частину останньої рівності (5) можна розглядати як означення розширення оператора $\mathcal{L}$ до оператора (за яким збережемо

попереднє позначення), що діє в просторах $\mathcal{L}: H_0^0 \rightarrow (W_T^2)^*$ і визначений на всьому просторі H_0^0 .

Неважко довести, що розширений оператор $\mathcal{L}: H_0^0 \rightarrow (W_T^2)^*$ є лінійним і неперервним. Спряжений лінійний неперервний оператор $\mathcal{L}^*: W_T^2 \rightarrow (H_0^0)^*$ також визначається правою частиною рівності (5).

Лема 1. Нехай $\mathcal{L}: H_0^0 \rightarrow (W_T^2)^*$. Для довільної функції $u \in H_0^0$ мають місце нерівності

$$\|u\|_{W_0^{-1}} \leq c_1 \|\mathcal{L}u\|_{(W_T^2)^*} \leq c_2 \|u\|_{H_0^0}, \quad (6)$$

де c_1, c_2 — додатні сталі, що не залежать від функції $u(t, \xi)$.

Доведення. Якщо до виразу, визначеному правою частиною рівності (5), застосувати нерівність Коші-Буняковського в інтегральній формі і нерівність Фрідрікса, то отримаємо доведення правої частини (6).

Для доведення нерівності $\|u\|_{W_0^{-1}} \leq c_1 \|\mathcal{L}u\|_{(W_T^2)^*}$ розглянемо значення функціонала $\mathcal{L}u \in (W_T^2)^*$ на елементі

$$v^{[-1]}(t, \xi) = \int_T^t v(\tau, \xi) d\tau,$$

де функція $v(t, \xi)$ — розв'язок задачі Коші для звичайного диференціального рівняння першого порядку:

$$e^{Mt}(v_t - v) = \int_0^t u(\tau, \xi) d\tau, \quad v|_{t=T} = 0. \quad (7)$$

Відзначимо, що легко встановити належність функції $v^{[-1]}(t, \xi)$ простору W_T^2 . Справді, оскільки функція $u(t, \xi)$ належить простору H_0^0 , то функція $u^{(-1)}(t, \xi)$ належить простору H_0^1 . З леми 1 роботи [1] випливає, що функція $e^{-Mt}u^{(-1)}(t, \xi)$ є елементом простору W_0^0 . Використовуючи лему 3 роботи [1] і посилаючись на рівність (7), маємо, що $v \in W_T^1$. Тому $v^{[-1]}(t, \xi)$ належить простору W_T^2 .

Використовуючи формулу інтегрування частинами, можна показати, що має місце рівність:

$$\langle \mathcal{L}u, v^{[-1]} \rangle_{W_T^2} = \langle \mathcal{L}(-u^{(-1)}), v \rangle_{W_T^1}.$$

Ураховуючи те, що функції $u^{(-1)}$ і v зв'язані між собою рівністю (7), для оцінки знизу виразу

$\langle \mathcal{L}(-u^{(-1)}), v \rangle_{W_T^1}$ можна повторити міркування леми 4 статті [1]. Це дозволяє стверджувати, що для деякої досить великої додатної сталої c мають місце співвідношення:

$$\begin{aligned} c \langle \mathcal{L}(-u^{(-1)}), v \rangle_{W_T^1} &\geq \|u^{(-1)}\|_{W_0^0}^2 + \|v\|_{W_T^1}^2 = \\ &= \|u\|_{W_0^{-1}}^2 + \|v^{[-1]}\|_{W_T^2}^2 \geq 2\|u\|_{W_0^{-1}}\|v^{[-1]}\|_{W_T^2}. \end{aligned}$$

Застосовуючи до виразу $\langle \mathcal{L}u, v^{[-1]} \rangle_{W_T^2}$ нерівність Шварца, отримуємо:

$$\begin{aligned} c\|\mathcal{L}u\|_{(W_T^2)^*}\|v^{[-1]}\|_{W_T^2} &\geq c \langle \mathcal{L}u, v^{[-1]} \rangle_{W_T^2} \geq \\ &\geq 2\|u\|_{W_0^{-1}}\|v^{[-1]}\|_{W_T^2}. \end{aligned}$$

Якщо в останній нерівності ліву і праву частини скоротити на $\|v^{[-1]}\|_{W_T^2}$, то отримаємо ліву частину нерівностей (6). $\square$

Використовуючи формулу інтегрування частинами, праву частину рівності (5) для всіх $u \in L_0$ і $v \in L_T$ можна переписати у вигляді:

$$\begin{aligned} (u, v_{tt} - av_t + bv)_{L_2(Q)} &+ (a_0u, v)_{L_2(\Gamma_3)} - \\ &- \sum_{i,j=1}^n (u_{\xi_i}, a_{ij}v_{\xi_j t} - b_{ij}v_{\xi_j})_{L_2(Q)} - \\ &- \sum_{i=1}^n (u_{\xi_i}, a_i v_t - b_i v)_{L_2(Q)}. \end{aligned} \quad (8)$$

Отриманий вираз задає звуження оператора $\mathcal{L}: H_0^0 \rightarrow (W_T^2)^*$. Такий звужений оператор буде визначений на всій множині W_0^0 та буде діяти у просторах $\mathcal{L}: W_0^0 \rightarrow (H_T^2)^*$.

Лема 2. Нехай $\mathcal{L}: W_0^0 \rightarrow (H_T^2)^*$. Для довільної функції $u \in W_0^0$ має місце нерівність

$$\|u\|_{H_0^0} \leq c_1 \|\mathcal{L}u\|_{(H_T^2)^*} \leq c_2 \|u\|_{W_0^0}, \quad (9)$$

де c_1, c_2 — додатні сталі, що не залежать від функції $u(t, \xi)$.

Доведення. Якщо до виразу (8) застосувати нерівність Коші-Буняковського в інтегральній формі і нерівність Фрідрікса, то отримаємо доведення правої частини (9).

Для доведення нерівності $\|u\|_{H_0^0} \leq c_1 \|\mathcal{L}u\|_{(H_T^2)^*}$ розглянемо значення функціонала $\mathcal{L}u$ на елементі $v(t, \xi)$, де функція $v(t, \xi)$ —

розв'язок задачі Коші для звичайного диференціального рівняння другого порядку:

$$e^{Mt}(v - v_t + \beta v_{tt}) = u, \quad v|_{t=T} = v_t|_{t=T} = 0.$$

Відзначимо, що з леми 2 статті [1] випливає, що функція $v(t, \xi)$ належить простору W_T^2 , а отже і простору H_T^2 . Використовуючи формулу інтегрування частинами, можна показати, що має місце рівність:

$$\langle \mathcal{L}u, v \rangle_{H_T^2} = \langle \mathcal{L}(-u^{(-1)}), v_t \rangle_{H_T^1}.$$

Ураховуючи, що функції $u^{(-1)}$ і v зв'язані між собою співвідношенням

$$e^{Mt}(v - v_t + \beta v_{tt}) = (u^{(-1)})_t,$$

для оцінки знизу виразу $\langle \mathcal{L}(-u^{(-1)}), v_t \rangle_{H_T^1}$ можна повторити міркування леми 5 статті [1]. Це дозволяє стверджувати, що для деякої досить великої додатної сталої c мають місце співвідношення:

$$\begin{aligned} c \langle \mathcal{L}(-u^{(-1)}), v_t \rangle_{H_T^1} &\geq \|u^{(-1)}\|_{H_0^1}^2 + \|v_t\|_{H_T^1}^2 = \\ &= \|u\|_{H_0^0}^2 + \|v\|_{H_T^2}^2 \geq 2\|u\|_{H_0^0}\|v\|_{H_T^2}. \end{aligned}$$

Застосовуючи до виразу $\langle \mathcal{L}u, v \rangle_{H_T^2}$ нерівність Шварца:

$$c\|\mathcal{L}u\|_{(H_T^2)^*}\|v\|_{H_T^2} \geq c\langle \mathcal{L}u, v \rangle_{H_T^2} \geq 2\|u\|_{H_0^0}\|v\|_{H_T^2}$$

і скорочуючи на $\|v\|_{H_T^2}$, завершуємо доведення леми: $\square$

Міркуючи аналогічним чином, доведені апріорні нерівності можна "зсунути" на оператор $u \rightarrow u^{(k)}$ для довільного цілого значення k . Таким чином, має місце теорема.

Теорема 1. Для довільного числа $k \in \mathbb{Z}$ і довільної функції $u \in W_0^k$ мають місце нерівності

$$\begin{aligned} \|u\|_{W_0^{k-1}} &\leq c_1 \|\mathcal{L}u\|_{(W_T^{2-k})^*} \leq c_2 \|u\|_{H_0^k} \leq \\ &\leq c_3 \|\mathcal{L}u\|_{(H_T^{2-k})^*} \leq c_4 \|u\|_{W_0^k}, \end{aligned}$$

де c_i — додатні сталі, що не залежать від функції $u(t, \xi)$.

Використовуючи міркування двоїстості, аналогічні апріорні нерівності можна довести і для спряженого оператора $\mathcal{L}^*$.

Теорема 2. Для довільного числа $k \in \mathbb{Z}$ і довільної функції $v \in W_T^k$ мають місце нерівності

$$\begin{aligned} \|v\|_{W_T^{k-1}} &\leq c_1 \|\mathcal{L}^*v\|_{(W_0^{2-k})^*} \leq c_2 \|v\|_{H_T^k} \leq \\ &\leq c_3 \|\mathcal{L}^*v\|_{(H_0^{2-k})^*} \leq c_4 \|v\|_{W_T^k}, \end{aligned}$$

де c_i — додатні сталі, що не залежать від функції $v(t, \xi)$.

З тверджень теорем 1, 2 випливає ін'єктивність операторів $\mathcal{L}$ і $\mathcal{L}^*$.

Вивчимо питання існування та єдиності розв'язку псевдогіперболічного операторного рівняння $\mathcal{L}u = f$. У випадку, коли права частина f є узагальненою функцією певного скінченного порядку, деякі результати щодо існування та єдиності узагальненого розв'язку лінійного псевдогіперболічного рівняння отримано в та інших роботах [2-7].

Теорема 3. Нехай $\mathcal{L}$ — псевдогіперболічний оператор.

1. Для довільної правої частини $f \in (W_T^k)^*$, де $k \in \mathbb{Z}$, існує єдиний розв'язок $u \in W_0^{1-k}$ рівняння $\mathcal{L}u = f$.
2. Для довільної правої частини $f \in (H_T^k)^*$, де $k \in \mathbb{Z}$, існує єдиний розв'язок $u \in H_0^{2-k}$ рівняння $\mathcal{L}u = f$.

Доведення. Доведемо лише першу частину теореми. Друга частина доводиться аналогічно.

Маємо $f \in (W_T^k)^* \subset (W_T^{k+1})^*$. Використовуючи нерівності теореми 2, маємо, що для всіх функцій $v \in W_T^{k+1} \subset H_T^{k+1} \subset W_T^k$ виконуються нерівності:

$$\begin{aligned} |\langle f, v \rangle_{W_T^{k+1}}| &= |\langle f, v \rangle_{W_T^k}| \leq \\ &\leq \|f\|_{(W_T^k)^*} \|v\|_{W_T^k} \leq c_1 \|f\|_{(W_T^k)^*} \|\mathcal{L}^*v\|_{(W_0^{1-k})^*}. \end{aligned}$$

Зважаючи на ін'єктивність оператора $\mathcal{L}^*$, вираз $\langle f, v \rangle_{W_T^{k+1}}$ можна розглядати як означення деякого лінійного неперервного функціонала $l(\omega)$ від $\omega = \mathcal{L}^*v$, визначеного на множині $\mathcal{L}^*(W_T^{k+1}) \subset (H_0^{1-k})^*$ простору $(W_0^{1-k})^*$.

За теоремою Хана-Банаха цей функціонал може бути неперервно розширений на весь простір $(W_0^{1-k})^*$. З рефлексивності простору W_0^{1-k} випливає існування такого елемента $u \in W_0^{1-k}$,

що функціонал l допускає представлення у вигляді $l(\omega) = \langle \omega, u \rangle_{W_0^{1-k}}$. Підставляючи $\omega = \mathcal{L}^*v$ в рівність $l(\omega) = \langle \omega, u \rangle_{W_0^{1-k}}$, отримуємо

$$\begin{aligned}\langle f, v \rangle_{W_T^{k+1}} &= \langle \mathcal{L}^*v, u \rangle_{W_0^{1-k}} = \\ &= \langle \mathcal{L}u, v \rangle_{H_T^{k+1}} = \langle \mathcal{L}u, v \rangle_{W_T^{k+1}}.\end{aligned}$$

Звідки, зважаючи на довільність функції $v \in W_T^{k+1}$, маємо $\mathcal{L}u = f$.

Єдиність розв'язку випливає з ін'єктивності оператора $\mathcal{L}$. $\square$

Аналогічна теорема має місце і для спряженого оператора $\mathcal{L}^*$.

Список використаних джерел

1. Номіровський Д. А. Априорні нерівності для лінійних диференціальних псевдогіперболічних моделей / Д. А. Номіровський // Вісник Київського національного університету імені Тараса Шевченка. Серія: фізико-математичні науки. — 2013. — №1. — С. 215-220.
2. Lyashko S. I. Generalized Optimal Control of Linear Systems with Distributed Parameters / S. I. Lyashko. — London : Kluwer Academic Publishers, 2002. — 466 p.
3. Ляшко С. И. Обобщенное управление линейными системами / С. И. Ляшко. — К. : Наукова думка, 1998. — 472 с.
4. Палієнко Л.І. Моделювання та узагальнення оптимізація у псевдогіперболічних системах : дис. ... канд. фіз.-мат. наук: 01.05.02 / Л. І. Палієнко — К., 1999. — 122 с.
5. Семенов В.В. Модели і методи узагальненої оптимізації лінійних систем з розподіленими параметрами : дис. ... канд. фіз.-мат. наук: 01.05.02 / В. В. Семенов. — К., 2002. — 153 с.
6. Номіровський Д. А. О единственной разрешимости псевдогиперболических уравнений с сингулярными правыми частями / Д. А. Номіровський // Математические заметки. — 2006. — Т. 80, В. 4. — С. 582-595.
7. Аникушин А. В. Обобщенное оптимальное управление системами, которые описываются линейными интегро-дифференциальными уравнениями с неотрицательно-определенными интегральными операторами / А. В. Аникушин // Проблемы управления и информатики. — 2014. — №1. — С. 33-41.

Теореми 1-3 можуть бути використані для розв'язання задачі оптимального керування та дослідження керованості псевдогіперболічної системи, зокрема, для встановлення траєкторної та фінальної керованостей в різних класах керуючих впливів, доведення існування оптимального програмного керування системою через праву частину рівняння, вивчення диференціальних властивостей функціонала якості, побудови чисельних методів наближеного розв'язання задачі оптимального керування, побудови чисельних методів наближеного розв'язання псевдогіперболічного рівняння та дослідження збіжності запропонованих чисельних методів [2,3].

References

1. NOMIROVSKII, D. (2013) A priori inequalities for linear differential pseudohyperbolic equations. *Bulletin of Taras Shevchenko National University of Kyiv. Series: Physics & Mathematics*. 1. p.215-220.
2. LYASHKO, S. (2002) *Generalized Optimal Control of Linear Systems with Distributed Parameters*. London: Kluwer Academic Publishers.
3. LYASHKO, S. (1998) *Generalized Optimal Control of Linear Systems*. Kiev: Naukova dumka.
4. PALIENKO, L. (1999) *Simulation and generalized optimization of pseudohyperbolic systems* (Candidate's dissertation).
5. SEMENOV, V. (2002) *Models and methods of generalized optimization for linear systems with distributed parameters* (Candidate's dissertation).
6. NOMIROVSKII, D. (2006) Unique solvability of pseudohyperbolic equations with singular right-hand sides. *Mathematical Notes*. 80 (4). p.550-562.
7. ANIKUSHYN, A. (2014) Generalized optimal control for systems described by linear integro-differential equations with nonnegative definite integral operators. *Journal of Automation and Information Sciences*. 46 (6). p.58-67.

Надійшла до редакції 22.05.2014

УДК 004.655

Буй Д.Б., д.ф.-м.н., проф.,
Канарська І.С., аспірантка,
Сенченко О.С., к.ф.-м.н., доц.

D.B. Buy, Dr. Sc., professor,
I.S. Kanarska, postgraduate student,
A.S. Senchenko, PhD

Критерії збереження активного домену сигнатурними операціями табличних алгебр

The criteria for preservation of active domain by signature operations in the table algebra

Київський національний університет імені
Тараса Шевченка, 03187, м. Київ, пр-т Глушкова
4д,
e-mail: buy@unicyb.kiev.ua, iren_kiss@mail.ru,
senchenko_as@mail.ru

Taras Shevchenko National University of Kyiv,
03187, Kyiv, Glushkova av., 4d,
e-mail: buy@unicyb.kiev.ua, iren_kiss@mail.ru,
senchenko_as@mail.ru

В роботі знайдено необхідні та достатні умови, за яких сигнатурні операції табличних алгебр зберігають активний домен. Результати роботи можуть бути використаними в теорії табличних алгебр для дослідження збереження денотатів, та для оптимізації запитів в реляційних базах даних.

Ключові слова: бази даних, таблична алгебра, активний домен.

In this paper is investigated the properties of the active domain in the table algebras, which are modern analogue of Codd's well-known relational algebra and formed the theoretical foundation of modern query language databases. Elements of the carrier of table algebra specify relational data structures, and signature operations are based on the basic table manipulation in relational algebra and SQL-like languages. Active domain of attribute in the table is the set of all values of attribute in the rows of the table. The necessary and sufficient conditions under which the active domain preserved signature operations table algebra (intersection, difference, selection, projection, active complement, join, division) are found. Also is found a criterion of distribution of active domain with respect to the difference between the tables. The results can be used in the theory of table algebras to the study table denotation preservation, equivalent transformations of expressions and in a relational database to ensure data integrity, and, in our view, to optimize queries.

Key Words: database, table algebra, active domain.

Статтю представив д.ф.-м.н., проф. Анісімов А.В.

Вступ. Системи керування базами даних широко використовуються у діяльності людини. Найбільш розповсюдженою є реляційна модель даних, математична модель якої вперше була запропонована Е. Коддом [1, 2]. Реляційна база даних є скінченною множиною скінченних відношень між заздалегідь визначеними множинами даних. Табличні алгебри, що були введені В.Н. Редьком і Д.Б. Буєм [3], побудовані на основі реляційних алгебр Е. Кодда, суттєво їх уточнюють та складають теоретичний фундамент мов запитів сучасних табличних баз даних. Елементи їх носія уточнюють реляційні структури даних, а операції побудовані на базі основних маніпуляцій у SQL-подібних мовах.

У табличних алгебрах важливу роль відіграє активний домен атрибуту таблиці – множина значень, яка складається з усіх можливих значень

цього атрибуту в цій таблиці. Наприклад, завдяки дослідженню збереження активного домену в [4] вдалося встановити непохідність деяких операцій табличних алгебр відносно решти.

В монографії [4] знайдено взаємозв'язки між активними доменами вихідних таблиць та таблиці, що є результатом виконання операцій табличних алгебр, більшість з цих взаємозв'язків подано у вигляді включень. У даній роботі знайдено та доведено критерії, за виконання яких ці включення перетворюються в рівності.

Основні визначення

Зафіксуємо деяку непорожню множину атрибутів $\mathfrak{R} = \{A_1, \dots, A_n\}$. Довільну скінченну підмножину множини $\mathfrak{R}$ назовемо схемою, причому схема може бути порожньою множиною. Рядком s схеми R називається

множина пар $s = \{(A'_1, d_1), \dots, (A'_k, d_k)\}$, проекція якої за першою компонентою дорівнює R , причому атрибути $A'_1, \dots, A'_k$ є попарно різними, тобто рядок є функціональним бінарним відношенням. Таблиця T схеми R – скінченна множина рядків схеми R .

Активним доменом атрибуту A відносно таблиці T називається множина $D_{A,T} = \{d \mid \exists s \in T \wedge (A, d) \in s\}$, яка, змістовно кажучи, складається з усіх значень атрибуту A в рядках таблиці T .

На множині таблиць однакової схеми введено операції об'єднання, перетину та різниці таблиць як обмеження одноіменних теоретико-множинних операцій. За аналогією з арифметичною операцією різниці для таблиці $T_1 - T_2$ перший аргумент назвемо таблицею-зменшуваним, а другий – таблицею-від'ємником.

Проекцією за множиною атрибутів $X \subseteq R$ називається унарна параметрична операція π_X , значенням якої є таблиця, що складається з обмежень по X усіх рядків вихідної таблиці: $\pi_X(T) = \{s \mid X \mid s \in T\}$ (тут обмеження розуміємо стандартно: $s \mid X = s \cap (X \times pr_2 s)$, де $pr_i s$ – проекція рядка s за i -ю компонентою).

Селекцією за предикатом $P: S \rightarrow \{true, false\}$, де S – множина усіх рядків, називається унарна параметрична операція σ_P , яка співставляє таблиці її підтаблицю, що містить рядки, на яких предикат P набуває значення «істина»: $\sigma_P(T) = \{s \mid s \in T \wedge P(s) = true\}$.

Для введення операції з'єднання визначимо допоміжне поняття. Бінарні відношення ρ та τ називаються сумісними (позначається $\rho \approx \tau$), якщо $\rho \mid X = \tau \mid X$, де $X = pr_1 \rho \cap pr_1 \tau$.

З'єднання – бінарна операція $\otimes$, значенням якої є таблиця, що складається з усіх можливих об'єднань сумісних рядків вихідних таблиць: $T_1 \otimes T_2 = \{s_1 \cup s_2 \mid s_1 \in T_1 \wedge s_2 \in T_2 \wedge s_1 \approx s_2\}$.

Насиченням $C(T)$ називається таблиця $\prod_{A \in R} D_{A,T}$, де R – схема таблиці T , $\prod$ – оператор узагальненого декартового добутку, що відповідає індексуванню за всіма атрибутами A схеми T ($A \mapsto D_{A,T}, A \in R$). Активним доповненням таблиці T є таблиця $\tilde{T} = C(T) - T$.

Нехай R_1 – схема таблиці T_1 , R_2 – схема таблиці T_2 та $R_2 \subseteq R_1$. Ділення T_1 на T_2 –

таблиця схеми $R_1 - R_2$, така, що:

$$T_1 \div_{R_2}^{R_1} T_2 = \{s \in \pi_{R_1-R_2}(T_1) \mid \{s\} \otimes T_2 \subseteq T_1\}.$$

Іншими словами, таблиця $T_1 \div_{R_2}^{R_1} T_2$ складається з усіх

таких рядків $s \in \pi_{R_1-R_2}(T_1)$, що для будь-якого рядка $s' \in T_2$, рядок $s \cup s'$ належить таблиці T_1 .

Також на множині таблиць введено операцію перейменування, яка не буде використана в роботі, тому її означення не наводимо. Табличною алгеброю назвемо часткову алгебру з носієм – множиною всіх таблиць (довільних схем) – та наведеними вище операціями.

Виділяють дві особливі таблиці: таблицю $T_\varepsilon = \{\varepsilon\}$, де ε – порожній рядок, при цьому схема таблиці T_ε є порожньою множиною, та таблицю $T_\emptyset = \emptyset$ – порожню множину рядків довільної схеми. Зауважимо, що для цих таблиць значення активного домену є порожньою множиною.

Основні результати

В монографії [4] у підрозділі про взаємну похідність операцій табличних алгебр знайдено взаємозв'язки між активними доменами вихідних таблиць та таблиці, що є результатом виконання цих операцій; у загальному випадку для перетину, різниці, селекції, проекції, з'єднання, ділення та активного доповнення ці взаємозв'язки задаються у вигляді включень. У даній роботі знайдені критерії, за яких ці включення перетворюються в рівності.

Розглянемо збереження активного домену операцією перетину. Справедливе твердження:

Теорема 1. Нехай T_1 та T_2 – таблиці схеми R . Для всіх атрибутів $A_i \in R$ рівність $D_{A_i, T_1 \cap T_2} = D_{A_i, T_1} \cap D_{A_i, T_2}$ виконується тоді та тільки тоді, коли для будь-якого $x \in D_{A_i, T_1}$ виконується імплікація $x \in D_{A_i, T_2} \rightarrow \exists s(s \in T_1 \cap T_2 \wedge (A_i, x) \in s)$ (1).

Доведення.

Необхідність. Нехай виконується (1). Доведемо рівність $D_{A_i, T_1 \cap T_2} = D_{A_i, T_1} \cap D_{A_i, T_2}$. Від супротивного, припустимо, що ця рівність не виконується. Тоді, з урахуванням доведеного в [4] включення $D_{A_i, T_1 \cap T_2} \subseteq D_{A_i, T_1} \cap D_{A_i, T_2}$, це можливо лише у тому випадку, коли існує такий $x \in D_{A_i, T_1} \cap D_{A_i, T_2}$, що $x \notin D_{A_i, T_1 \cap T_2}$. З приналежності $x \in D_{A_i, T_1} \cap D_{A_i, T_2}$ та (1) випливає

існування такого рядка s , що $s \in T_1$, $s \in T_2$ та $(A_i, x) \in s$, тому, $x \in D_{A_i, T_1 \cap T_2}$. Отримана суперечність доводить необхідність.

Достатність. Нехай виконується рівність $D_{A_i, T_1 \cap T_2} = D_{A_i, T_1} \cap D_{A_i, T_2}$. Від супротивного, припустимо, що не виконується умова (1), тобто, існує такий $x \in D_{A_i, T_1}$, що $x \notin D_{A_i, T_2}$, та не існує такого рядка s , що $s \in T_1 \wedge s \in T_2 \wedge (A_i, x) \in s$. Тоді з рівності $D_{A_i, T_1 \cap T_2} = D_{A_i, T_1} \cap D_{A_i, T_2}$ випливає, що $x \in D_{A_i, T_1 \cap T_2}$, тому за визначенням операції перетину таблиць існує такий рядок z , що $z \in T_1$, $z \in T_2$ та $(A_i, x) \in z$; суперечність доводить достатність теореми.

Розглянемо збереження активного домену операцією різниці. Спочатку знайдемо критерій збереження активного домену таблиці-зменшеного таблицею-різницею.

Теорема 2. Нехай T_1 та T_2 – таблиці схеми R . Для всіх атрибутів $A_i \in R$ рівність $D_{A_i, T_1 - T_2} = D_{A_i, T_1}$ виконується тоді та тільки тоді, коли для будь-якого $x \in D_{A_i, T_1}$ існує такий рядок $s \in T_1$, що $(A_i, x) \in s$ та $s \notin T_2$.

Доведення.

Необхідність. Нехай виконується умова теореми. Доведемо рівність $D_{A_i, T_1 - T_2} = D_{A_i, T_1}$. Від супротивного, припустимо, що ця рівність не виконується. Тоді, з урахуванням доведеного в [4] включення $D_{A_i, T_1 - T_2} \subseteq D_{A_i, T_1}$, це можливо лише у тому випадку, коли існує такий $x \in D_{A_i, T_1}$, що $x \notin D_{A_i, T_1 - T_2}$. З умов $s \in T_1$ та $s \notin T_2$ за визначенням різниці випливає, що $s \in T_1 - T_2$, а з $(A_i, x) \in s$ випливає, що $x \in D_{A_i, T_1 - T_2}$. Отримана суперечність доводить необхідність.

Достатність. Нехай виконується рівність $D_{A_i, T_1 - T_2} = D_{A_i, T_1}$. Від супротивного, припустимо, що умова теореми не виконується, тобто існує такий $x \in D_{A_i, T_1}$, що для всіх таких рядків $s \in T_1$, для яких $(A_i, x) \in s$ виконується приналежність $s \in T_2$. Тоді, за визначенням операції різниці, всі такі рядки s , що $(A_i, x) \in s$ не належать таблиці $T_1 - T_2$, тому $x \notin D_{A_i, T_1 - T_2}$, отже в цьому випадку $D_{A_i, T_1 - T_2} \neq D_{A_i, T_1}$. Одержали суперечність, що доводить достатність теореми.

Знайдемо критерій дистрибутивності активного домену відносно різниці таблиць.

Теорема 3. Нехай T_1 та T_2 – таблиці схеми R . Для всіх атрибутів $A_i \in R$ рівність $D_{A_i, T_1 - T_2} = D_{A_i, T_1} - D_{A_i, T_2}$ виконується тоді та тільки тоді, коли для будь-якого $x \in D_{A_i, T_1}$ виконується імплікація $x \in D_{A_i, T_2} \rightarrow (\forall s \in T_1 ((A_i, x) \in s \rightarrow s \in T_2))$ (2).

Доведення.

Необхідність. Нехай виконується (2). Доведемо рівність $D_{A_i, T_1 - T_2} = D_{A_i, T_1} - D_{A_i, T_2}$. Спочатку доведемо, що $D_{A_i, T_1 - T_2} \subseteq D_{A_i, T_1} - D_{A_i, T_2}$. Від супротивного, припустимо, що це включення не виконується, тоді існує такий $x \in D_{A_i, T_1 - T_2}$, що $x \notin (D_{A_i, T_1} - D_{A_i, T_2})$. З приналежності $x \in D_{A_i, T_1 - T_2}$ випливає, що існує такий рядок $s \in T_1 - T_2$, що $(A_i, x) \in s$, тому $s \in T_1$ та $s \notin T_2$. З приналежності $s \in T_1$ випливає, що $x \in D_{A_i, T_1}$, а з урахуванням того, що $x \notin D_{A_i, T_1} - D_{A_i, T_2}$ випливає, що $x \in D_{A_i, T_2}$. Тоді з виконання (2) одержимо, що $s \in T_2$. Суперечність доводить включення $D_{A_i, T_1 - T_2} \subseteq D_{A_i, T_1} - D_{A_i, T_2}$.

Доведемо обернене включення. Від супротивного, припустимо, що воно не виконується, тому існує такий $x \in D_{A_i, T_1} - D_{A_i, T_2}$, що $x \notin D_{A_i, T_1 - T_2}$. Тоді з приналежності $x \in D_{A_i, T_1} - D_{A_i, T_2}$ випливає, що $x \in D_{A_i, T_1}$ та $x \notin D_{A_i, T_2}$, тому існує такий рядок $s \in T_1$, що $(A_i, x) \in s$. З $x \notin D_{A_i, T_2}$ випливає, що $s \notin T_2$, тоді за визначенням операції різниці таблиць $s \in T_1 - T_2$, з чого випливає приналежність $x \in D_{A_i, T_1 - T_2}$. Суперечність доводить включення $D_{A_i, T_1} - D_{A_i, T_2} \subseteq D_{A_i, T_1 - T_2}$ та необхідність теореми.

Достатність. Нехай виконується рівність $D_{A_i, T_1 - T_2} = D_{A_i, T_1} - D_{A_i, T_2}$. Від супротивного, припустимо, що не виконується умова (2), тобто, існує такий $x \in D_{A_i, T_1}$, що $x \in D_{A_i, T_2}$, та існує такий рядок $s \in T_1$, що $(A_i, x) \in s$ та $s \notin T_2$. Тоді з $x \in D_{A_i, T_1}$ та $x \in D_{A_i, T_2}$ випливає, що $x \notin D_{A_i, T_1} - D_{A_i, T_2}$, а з $s \in T_1$ та $s \notin T_2$ за визначенням різниці таблиць випливає, що $s \in T_1 - T_2$, тому $x \in D_{A_i, T_1 - T_2}$, що порушує

рівність $D_{A_i, T_1 - T_2} = D_{A_i, T_1} - D_{A_i, T_2}$. Одержали суперечність, що доводить достатність теореми.

Розглянемо збереження активного домену операцією проекції. Справедливе твердження:

Твердження 1. Нехай T – таблиця схеми R . Для всіх атрибутів $A_i \in R$ рівність $D_{A_i, \pi_X(T)} = D_{A_i, T}$ виконується тоді та тільки тоді, коли $A_i \in X$.

Розглянемо збереження активного домену операцією селекції. Справедливе твердження:

Теорема 4. Нехай T – таблиця схеми R . Для всіх атрибутів $A_i \in R$ рівність $D_{A_i, \sigma_P(T)} = D_{A_i, T}$ виконується тоді та тільки тоді, коли для будь-якого $x \in D_{A_i, T}$ існує такий рядок $s \in T$, що $(A_i, x) \in s$ та $P(s) = \text{true}$.

Доведення.

Необхідність. Нехай виконується умова теореми. Доведемо рівність $D_{A_i, \sigma_P(T)} = D_{A_i, T}$. Від супротивного, припустимо, що ця рівність не виконується. Тоді, з урахуванням доведеного в [4] включення $D_{A_i, \sigma_P(T)} \subseteq D_{A_i, T}$, це можливо лише у тому випадку, коли існує такий $x \in D_{A_i, T}$, що $x \notin D_{A_i, \sigma_P(T)}$. Тоді за визначенням селекції виконується $s \in \sigma_P(T)$, тому, $x \in D_{A_i, \sigma_P(T)}$; суперечність доводить необхідність.

Достатність. Нехай виконується рівність $D_{A_i, \sigma_P(T)} = D_{A_i, T}$. Від супротивного, припустимо, що умова (2) не виконується, тобто, існує такий $x \in D_{A_i, T}$, що для всіх рядків s , для яких $(A_i, x) \in s$ виконується $P(s) = \text{false}$. Тоді за означенням селекції $x \notin D_{A_i, \sigma_P(T)}$, тобто рівність $D_{A_i, \sigma_P(T)} = D_{A_i, T}$ не виконується. Одержали суперечність, що доводить достатність теореми.

Розглянемо збереження активного домену операцією з'єднання. Справедливе твердження:

Теорема 5. Нехай T_1 – таблиця схеми R_1 . Для всіх атрибутів $A_i \in R_1$ рівність $D_{A_i, T_1 \otimes T_2} = D_{A_i, T_1}$ виконується тоді та тільки тоді, коли для будь-якого $x \in D_{A_i, T_1}$ існують такі рядки $s \in T_1$, $s' \in T_2$, що $(A_i, x) \in s$ та $s \approx s'$.

Доведення.

Необхідність. Нехай виконується умова теореми. Доведемо рівність $D_{A_i, T_1 \otimes T_2} = D_{A_i, T_1}$. Від супротивного, припустимо, що ця рівність не виконується. Тоді, з урахуванням доведеного в

[4] включення $D_{A_i, T_1 \otimes T_2} \subseteq D_{A_i, T_1}$, це можливо лише у тому випадку, коли існує такий $x \in D_{A_i, T_1}$, що $x \notin D_{A_i, T_1 \otimes T_2}$. З приналежності $x \in D_{A_i, T_1}$ та виконання умови теореми випливає існування в таблиці T_1 такого рядка s , що $(A_i, x) \in s$ та рядка $s' \in T_2$, причому $s \approx s'$. Тоді за визначенням з'єднання в таблиці $T_1 \otimes T_2$ існує рядок $s \cup s'$, для цього рядка виконується $(A_i, x) \in s \cup s'$, тому $x \in D_{A_i, T_1 \otimes T_2}$, суперечність доводить необхідність.

Достатність. Нехай виконується рівність $D_{A_i, T_1 \otimes T_2} = D_{A_i, T_1}$. Від супротивного, припустимо, що умова теореми не виконується, тобто, існує такий $x \in D_{A_i, T_1}$, що для будь-якого рядка $s \in T_1$, для якого $(A_i, x) \in s$, не існує сумісного з ним рядка у таблиці T_2 . Тоді, за визначенням з'єднання всі такі рядки s , що $(A_i, x) \in s$, не мають участь у з'єднанні, тому $x \notin D_{A_i, T_1 \otimes T_2}$, отже $D_{A_i, T_1 \otimes T_2} \neq D_{A_i, T_1}$; суперечність доводить достатність теореми.

Розглянемо збереження активного домену операцією активного доповнення.

Теорема 6. Нехай T – таблиця схеми R . Для всіх атрибутів $A_i \in R$ рівність $D_{A_i, \tilde{T}} = D_{A_i, T}$ виконується тоді та тільки тоді, коли для будь-якого $x \in D_{A_i, T}$ існують такі значення $d_1 \in D_{A_1, T}, \dots, d_n \in D_{A_n, T}$, що рядок $\{(A_1, d_1), \dots, (A_i, x), \dots, (A_n, d_n)\} \notin T$. (3)

Доведення.

Необхідність. Нехай виконується умова (3). Доведемо рівність $D_{A_i, \tilde{T}} = D_{A_i, T}$. Від супротивного, припустимо, що ця рівність не виконується. Тоді, з урахуванням доведеного в [4] включення $D_{A_i, \tilde{T}} \subseteq D_{A_i, T}$, це можливо лише у випадку, коли існує такий $x \in D_{A_i, T}$, що $x \notin D_{A_i, \tilde{T}}$. З умови (3) випливає, що $s = \{(A_1, d_1), \dots, (A_i, x), \dots, (A_n, d_n)\} \notin T$ та $s \in C(T)$, тому за визначенням активного доповнення $s \in \tilde{T}$, отже $x \in D_{A_i, \tilde{T}}$. Отримана суперечність доводить необхідність.

Достатність. Нехай виконується рівність $D_{A_i, \tilde{T}} = D_{A_i, T}$. Від супротивного, припустимо, що умова (3) не виконується, тобто, існує такий $x \in D_{A_i, T}$, що для всіх значень $d_1 \in D_{A_1, T}, \dots,$

$d_n \in D_{A_n, T}$ рядок $\{(A_1, d_1), \dots, (A_i, x), \dots, (A_n, d_n)\}$ належить таблиці T . Тоді з рівності $D_{A_i, \tilde{T}} = D_{A_i, T}$ випливає, що $x \in D_{A_i, \tilde{T}}$, тому у $\tilde{T}$ існує рядок $s = \{(A_1, d'_1), \dots, (A_i, x), \dots, (A_n, d'_n)\}$. З визначення активного доповнення випливає, що $s \in C(T)$ та $s \notin T$. З приналежності $s \in C(T)$ та доведеної в [4] рівності $D_{A_i, C(T)} = D_{A_i, T}$ випливають приналежності $d'_1 \in D_{A_1, T}, \dots, d'_n \in D_{A_n, T}$. Одержали суперечність, що доводить достатність.

Розглянемо збереження активного домену операцією ділення. Справедливе твердження:

Теорема 7.

Нехай T_1 – таблиця схеми R_1 та T_2 – таблиця схеми R_2 . Для всіх атрибутів $A_i \in R_1 - R_2$ рівність $D_{A_i, T_1 \div_{R_2}^{R_1}} = D_{A_i, T_1}$ виконується тоді та тільки тоді, коли для будь-якого $x \in D_{A_i, T_1}$ існує такий рядок $s \in \pi_{R_1 - R_2}(T_1)$, що $(A_i, x) \in s$, та для всіх рядків $s' \in T_2$, рядок $s \cup s'$ належить T_1 .

Доведення.

Необхідність. Нехай виконується умова теореми. Доведемо рівність $D_{A_i, T_1 \div_{R_2}^{R_1}} = D_{A_i, T_1}$. З урахуванням доведеного в [4] включення $D_{A_i, T_1 \div_{R_2}^{R_1}} \subseteq D_{A_i, T_1}$, треба показати, що $D_{A_i, T_1} \subseteq D_{A_i, T_1 \div_{R_2}^{R_1}}$. Нехай $x \in D_{A_i, T_1}$. За визначенням операції ділення таблиць рядок s

належить $T_1 \div_{R_2}^{R_1} T_2$, тому з $(A_i, x) \in s$ випливає

$x \in D_{A_i, T_1 \div_{R_2}^{R_1} T_2}$, що доводить необхідність.

Достатність. Нехай виконується рівність $D_{A_i, T_1 \div_{R_2}^{R_1} T_2} = D_{A_i, T_1}$. Від супротивного, припустимо, що умова теореми не виконується, тобто, існує такий $x \in D_{A_i, T_1}$, що для всіх таких рядків $s \in \pi_{R_1 - R_2}(T_1)$, для яких $(A_i, x) \in s$, існує такий рядок $s' \in T_2$, що рядок $s \cup s'$ не належить таблиці T_1 . Тоді за визначенням операції ділення, всі такі рядки s , що $(A_i, x) \in s$ не належать таблиці $T_1 \div_{R_2}^{R_1} T_2$, тому $x \notin D_{A_i, T_1 \div_{R_2}^{R_1} T_2}$, отже в цьому випадку $D_{A_i, T_1 \div_{R_2}^{R_1} T_2} \neq D_{A_i, T_1}$. Одержали суперечність, що доводить достатність теореми.

Висновки

В роботі знайдено критерії, за яких у табличних алгебрах операції перетину, різниці, селекції, проєкції, з'єднання, активне доповнення та ділення зберігають активний домен, та критерій дистрибутивності активного домену відносно різниці. Результати роботи можуть бути використаними в теорії табличних алгебр для дослідження збереження денотатів та для оптимізації запитів у реляційних базах даних.

Список використаних джерел

1. Codd E.F. A Relational Model of Data for Large Shared Data Banks / E.F. Codd // Communications of the ACM. – 1970. – 13 (6). – P. 377–387.
2. Codd E.F. The Relational Model for Database Management: version 2 / E.F. Codd. – Addison-Wesley, 1990. – 538 p.
3. Редько В.Н. К основаниям теории реляционных моделей баз данных / В.Н. Редько, Д.Б. Буй // Кибернетика и системный анализ. – 1996. – №4. – С. 3–12.
4. Реляційні бази даних: табличні алгебри та SQL-подібні мови / В.Н. Редько, Ю.Й. Брона, Д.Б. Буй, С.А. Поляков. – Київ: «Академперіодика», 2001. – 198 с.

References

1. CODD E.F. (1970) A Relational Model of Data for Large Shared Data Banks. *Communications of the ACM*. 13 (6). p. 377–387.
2. CODD E.F. (1990) *The Relational Model for Database Management: version 2*. Addison-Wesley.
3. REDKO V.N. and BUY D.B. (1996) Foundations of the theory of relational database models. *Cybernetics and Systems Analysis*. 32(4). p. 3 – 12.
4. REDKO V.N. et al. (2001) *Relational data base: table algebras and family of the SQL languages*. Kyiv: Academperiodica.

Надійшла до редколегії 27.05.15

УДК 004.655

Буй Д.Б.¹ д.-р. фіз.-мат. н., професор,
Компан С.В.², Поляков С.А.³ канд. фіз.-мат. н.,
Карам Д.⁴ аспірант

Алгоритми CLOS та LOOPS лінеаризації класів в мовах програмування: формальна побудова

Київський національний університет імені Тараса
Шевченка, 03680, м. Київ, пр-т. Глушкова 4д

e-mail: ¹ buy@unicyb.kiev.ua

e-mail: ² robin_2005@mail.ru

e-mail: ³ sergey.a.polyakov@gmail.com

e-mail: ⁴ karamjasim1978@yahoo.com

D.B. Buy¹ professor, S.V. Kompan²,
S.A. Polyakov³ PhD, J. Karam⁴ postgraduate

Linearization algorithms CLOS and LOOPS of the classes in programming languages: the formal definitions

¹Taras Shevchenko National University of Kyiv,
03680, Kyiv, Glushkova st., 4d

e-mail: ¹ buy@unicyb.kiev.ua

e-mail: ² robin_2005@mail.ru

e-mail: ³ sergey.a.polyakov@gmail.com

e-mail: ⁴ karamjasim1978@yahoo.com

В статті досліджуються методи вирішення конфліктів імен в мовах програмування CLOS та LOOPS, які дозволяють використовувати множинне успадкування. Ці методи засновані на лінеаризації класів, яка полягає в трансформації ієрархії класів з множинним успадкуванням в ієрархію одиночного успадкування, при якому проблема вибору відповідного атрибута вирішується тривіально. Вводиться визначення ієрархії класів.

Ключові слова: множинне успадкування, CLOS, LOOPS, алгоритми лінеаризації

The methods of conflict resolution of names in programming languages CLOS and LOOPS that allow multiple inheritance were researched. Those methods are based on the linearization of the classes. The idea of linearization is to reduce multiple inheritance to single, in which the problem of selecting a suitable attribute is solved trivially. The definition of a hierarchy of classes is introduced. The property of monotonicity was clarified. The idea of monotonicity is that all the attributes that the class inherits either identified in its direct ancestors, or inherited them. Methods of conflict resolution used in programming languages CLOS and LOOPS were formally defined.

Key Words: Multiple objects inheritance, CLOS, LOOPS, linearization algorithms.

Статтю представив д.ф.-м.н., професор Анісімов А.В.

Introduction. In object-oriented programming languages that implement multiple inheritance, the situation when the same method or field is found in multiple parent classes can take place. In this case, there is the problem of choosing the "right" method or field from several possible. Since the respective algorithms perform equally for both fields and methods of classes, we will in further call fields and methods of class as attributes. If the class **X** contains an attribute **atr**, then we say that the attribute **atr** is defined in class **X** ($atr \in X$).

There are two basic approaches in the resolution of the conflict of names [1]. In the first case, the conflict is resolved trivially by explicit the indication of the class in which the desired attribute exists. This is, for example, in C++. In the second case, a special algorithm is used to select the "most appropriate" attribute. This approach is used in languages such as

CLOS, LOOPS, Python, Perl, Dylan and others. The basic idea is that the parent classes are linearly ordered as the list. Then the method takes the first class in the list in which the attribute is defined.

We can say that the multiple inheritance by means of such ordering, boils down to the single inheritance, in which the problem of name conflict is solved by the fact that each class defines its own scope of visibility, which overload the scope of visibility of its the parent classes.

The main advantage of the second approach is that the name conflict can be resolved dynamically, i.e. during the program execution. Considered technique is called classes' linearization. There are several algorithms to resolve a name conflict by means of linearization. We will consider algorithms used in programming languages CLOS and LOOPS.

Although the algorithms have been created in the 90-th years of the last century, their formal properties are not well understood. That does not allow to talk about creating safe programs in programming languages that use these algorithms for classes' linearization. In addition, when formal definition of the semantics of programs is given, different approaches are usually consider the semantics of basic programming constructions like branching, looping, sequencing, but the semantics of classes and objects are omitted. The obtained results form the basis for a formal model that describes the classes, inheritance and relationships between classes.

The purpose of this paper is to give the formal definitions of these linearization algorithms that make possible the second step – proofing the formal correctness of these algorithms. Other algorithms for linearization examined in [4, 5, 6].

Definitions. Let's introduce the needed definitions [1], [2]. Class X , which is directly inherited from class Y , will be called direct descendant of the class Y , and Y , respectively, the direct ancestor (parent) of X (see. Fig. 1).



Fig. 1. Direct inheritance: Class Y is the direct ancestor (parent) of the class X ; the class X is the direct descendant of the class Y

Such inheritance will be written in the form of a pair (X, Y) , i.e. the pair (a descendant, an ancestor). A class can have a few direct ancestors, for example, a pair $(X, (Y, Z))$ indicates that the class X inherits from classes Y and Z , wherein the set of ancestors is ordered. Ordered family of items will be called tuples. Graphically, described multiple inheritance is shown in Fig. 2.

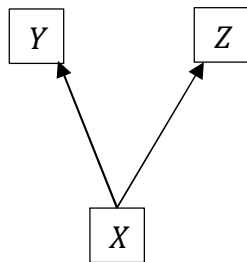


Fig. 2. Multiple inheritance: class X is a direct descendant of classes Y, Z , and direct ancestors are ordered from left to right

In turn classes Y and Z may be inherited from other classes, etc. Thus, we obtain a class hierarchy. It is said that two pairs of direct inheritances $(X_0, (X_1, \dots, X_n))$ and $(Y_0, (Y_1, \dots, Y_m))$ are linked if there exists $i = 1..n$, that $X_i = Y_0$. If $X_i = X_0$, the inheritance is called the self-linked. Graphically, the linked pairs of direct inheritances are shown in Fig. 3, self-linked inheritance is shown in Fig. 4.

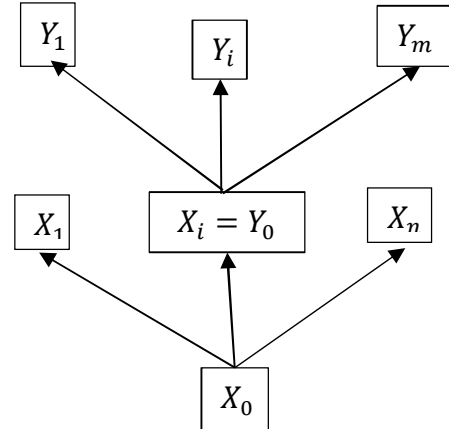


Fig. 3. Linked pairs of direct inheritances

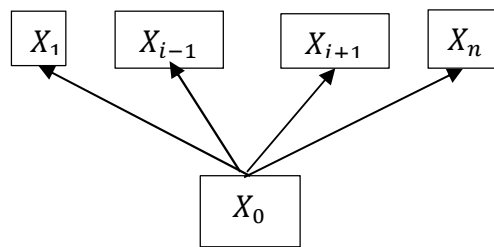


Fig. 4. Self-linked direct inheritance

Sequence of direct inheritances $r_1, r_2, \dots, r_k$ called a cycle when each i -th element is linked with the $(i+1)$ -th for $i = 1..k-1$ and the k -th element is linked with the first.

Let us denote by π_1^2 projection by the first component of the set of ordered pairs, and by π_2^2 – projection by the second component.

Let us denote by Δ the set of the classes, and by $\Delta_{<}$ the set of all classes of tuples Δ . Class hierarchy is called a binary relation $H \subseteq (\Delta \times \Delta_{<})$, such that:

- H is functionally, i.e. for each $X \in \pi_1^2(H)$ exists exactly one tuple $(X_1, \dots, X_n) \in \pi_2^2(H)$ that sets its direct ancestors;
- all the elements of the H are not self-linked;
- elements of the H does not form a cycles;
- elements from all tuples of projection $\pi_2^2(H)$ are not repeated.

The class hierarchy defines an acyclic direct simple graph whose vertices are the classes and the edges are defined as follows. For each direct inheritance $(X_0, (X_1, \dots, X_n))$ of the class hierarchy

the set of edges from X_0 will be defined as the set of pairs $\{(X_0, X_1), (X_0, X_2), \dots, (X_0, X_n)\}$. Union of the all of these sets will be the set of edges of the graph. Such graphs will be called the direct inheritance graphs. Each class hierarchy corresponds to a single graph of direct inheritance. But each direct inheritance graph corresponds, in general, several hierarchies of classes. They differ by the order of the elements in a tuple, since during the transition from a pair $(X_0, (X_1, \dots, X_n))$ to the set of pairs $\{(X_0, X_1), (X_0, X_2), \dots, (X_0, X_n)\}$ information about ordering of the classes $X_1, X_2, \dots, X_n$ lost.

The set of all edges of direct inheritance defines the anti-reflective and antisymmetric binary relation on classes. Its transitive closure is a strict partial order relation, denoted by $<_H$.

The relation $X <_H Y$ means that in the graph of direct inheritance there is a path from the class X to the class Y . In this case we say that the class Y is an ancestor of class X , and class X , respectively, a descendant of the class Y .

A linearization can be considered as a topological sorting (see. eg., [3], section 2.2.3) on a graph of direct inheritance or, another words, it specify linear order $<_\tau$ on the set of classes that would include a partial order $<_H$.

Such a topological sort is called linear extension of the graph of direct inheritance. Then we call linearization a unary function L , displaying a graph of direct inheritance of H in its linear extension $L(H)$. Thus, topological sorting is an example of a linearization.

The idea of linearization is, in fact, the reduction a multiple inheritance to a single inheritance, in which the problem of choosing the most suitable attribute from several ones is solved trivially.

For the class Y denote via H_Y the subgraph of direct inheritance graph of the hierarchy H , consisting of class Y and all of its ancestors (not only direct).

Under the method of conflict resolution we mean the mapping M , assigning to the class X , direct inheritance graph of the hierarchy H , and attribute atr the class Y from subgraph H_X , such that attribute atr is defined in this class: $X, H, atr \xrightarrow{M} Y$, attribute atr is defined in class Y . In particular, if the attribute atr is defined in the class X , then $M(X, H, atr) = X$.

The method of conflict resolution is not bound only to linearization of the direct inheritance graph and any other mapping satisfying definition can be used.

Adequate methods of conflict resolution must satisfy a number of natural properties, which essentially is not to produce results contrary to intuition.

The main feature is the so-called monotonicity. The idea is that all the attributes that a class inherits are defined in its direct ancestors, or inherited them. I.e. situation is impossible when a class inherits an attribute, but none of his direct ancestors; this attribute does not inherit or not defined. Monotonicity means that the class can be considered as a specialization (specification) of his direct ancestors, i.e. it has all the attributes that his direct ancestors have, plus the additional attributes that are unique to this class.

Methods of conflict resolution M is monotone if for any inheritance graph $H = (C, I)$, any class $X \in C$ and any attribute $atr \in X$ either $M(X, H, atr) = X$ or $M(X, H, atr) = Y, Y \neq X$, and there exists at least one direct ancestor Z of class X , such that $atr \in Z$ & $M(Z, H, atr) = Y$. In particular, it can be a situation that $Y = Z$.

We proceed to the definition of the monotonicity feature for linearization. Linearization L is monotone if for any inheritance graph $H = (C, I)$, and for any classes $X, Y \in C$ such that X is an ancestor of the Y , $L(H_X)$ included in the $L(H_Y)$.

Let's introduce the relation $pred_C$ of the precedence of classes on a hierarchy of classes H , where C is a set of classes. Two classes X_1 and X_2 are in relation of precedence if there is a direct inheritance $(Y_0, (Y_1, \dots, Y_n)) \in H$, that $X_1 = Y_i$ and $X_2 = Y_j$ and $i < j, i, j = 1..n$.

Relation of the precedence can be obtained as the union of all the relations of local linear orders, obtained by ordering direct ancestor classes.

The relation $pred_C$ must be acyclic for correct algorithms work. Then for each class you can specify an ordered set of his previous classes.

Linearization algorithms. Let's go directly to the linearization algorithm.

Algorithm of the linearization LOOPS.

For the hierarchy of the classes H build a direct inheritance graph $H = (C, I)$ and define relation of the precedence $pred_C$. Algorithm is defined in the triple $(C, I, pred_C)$, called the graph representation of the direct inheritance and denoted $R(H)$. It builds a linearization of the direct inheritance graph using the precedence relation to resolve the ambiguity. We denote this the linearization as $L(R(H))$.

For class X and representation graph of the inheritance $R(H_X)$ algorithm works as follows.

1) As an initial linearization it takes a sequence consisting of single element $L_X = [X]$.

2) We examine L_X from right to left and look for the first class X , which has at least one direct ancestor such that all its descendants already have been chosen, and it is not already chosen. If X has

few such direct ancestors, then take the first with respect to the relation $pred_C$. Add it to the sequence L_X .

3) Repeat step 2 until it is possible.

Algorithm CLOS.

1) As an initial linearization it takes a sequence consisting of a single element $L_X = [X]$.

2) We examine L_X from right to left and look for the first class, in which there is a direct ancestor such that all its descendants and all previous with respect to relation $pred_C$ classes have already been chosen, and it is not already selected. Add it to the sequence L_X .

Список використаних джерел

1. Ducournau R. Proposal for a Monotonic Multiple Inheritance Linearization / R. Ducournau, M. Habib, M. Huchard and M. L. Mugnier // OOPSLA'94 Proceedings of the ninth annual conference on Object-oriented programming systems, languages, and applications, 1994, P. 164–175.

2. Ducournau R. Monotonic conflict resolution mechanisms for inheritance / R. Ducournau, M. Habib, M. Huchard and M. L. Mugnier // OOPSLA'92 Proceeding conference proceedings on Object-oriented programming systems, languages, and applications, 1992, P. 16–24.

3. Knuth D.E. The art of computer programming. Vol. 1. Fundamental Algorithms / D.E. Knuth. – Addison-Wesley Publishing, Inc., 1998.

4. Simionato M. The Python 2.3 Method Resolution Order [Електронний ресурс] / M. Simionato – Режим доступу до ресурсу: <https://www.python.org/download/releases/2.3/mro/>

5. Guido V. R. Method Resolution Order [Електронний ресурс] / van Rossum Guido. – 2010. – Режим доступу до ресурсу: <http://python-history.blogspot.com/2010/06/method-resolution-order.html>.

6. Par Gaël Pegliasco Python Tutorial: Understanding Python MRO – Class search path [Електронний ресурс] / Pegliasco Par Gaël. – 2014. – <http://makina-corpus.com/blog/metier/2014/python-tutorial-understanding-python-mro-class-search-path>

3) Repeat step 2 until it is possible.

Conclusion. The paper clarified multiple inheritance classes (object-oriented): direct inheritance, linked couples, and cycles of direct inheritance, a class hierarchy, the graph of direct inheritance hierarchy, its transitive closure and linearization method of conflict resolution and its monotonicity. In these formal terms we specified linearization algorithms LOOPS and CLOS. The formal specification allows the formal proof of the correctness of these algorithms. It is the purpose of the subsequent work.

References

1. Ducournau, R., Habib, M., Huchard, M. and Mugnier, M.L. (1994) Proposal for a Monotonic Multiple Inheritance Linearization. *Proceedings of the ninth annual conference on Object-oriented programming systems, languages and applications*, OOPSLA'94, pp. 164–175, 1994.

2. Ducournau, R., Habib, M., Huchard, M. and Mugnier, M.L. (1992) Monotonic conflict resolution mechanisms for inheritance. *Proceeding conference proceedings on Object-oriented programming systems, languages, and applications*, OOPSLA'92, pp. 16–24. 1992.

3. Knuth, D.E. (1998) The art of computer programming. Vol. 1 Fundamental Algorithms Addison-Wesley Publishing, Inc., 1998.

4. Michele, Simionato The Python 2.3 Method Resolution Order” [Електронний ресурс]. – Режим доступу до ресурсу: <https://www.python.org/download/releases/2.3/mro/>

5. Guido, van Rossum (2010) Method Resolution Order [Електронний ресурс]. – Режим доступу до ресурсу: <http://python-history.blogspot.com/2010/06/method-resolution-order.html> June

6. Par Gaël, Pegliasco (2014) Python Tutorial: Understanding Python MRO – Class search path [Електронний ресурс]. – Режим доступу до ресурсу: <http://makina-corpus.com/blog/metier/2014/python-tutorial-understanding-python-mro-class-search-path>

УДК 004.658.2

Буй Д. Б., д.ф.-м.н., проф.,
Пузікова А. В., аспірантка.

Математична теорія нормалізації: нормальні форми 2-4 порядків

Київський національний університет імені
Тараса Шевченка, 03187, м. Київ, пр-т.
Глушкова 4д
e-mail: buy@unicyb.kiev.ua
e-mail: anna_inf@mail.ru

D. B. Buy, the doctor of physical and mathematical
sciences, the professor,
A. V. Puzikova, postgraduate.

Mathematical theory of normalization: 2-4 normal forms

Taras Shevchenko National University of Kyiv,
03187, Kyiv, Glushkova st., 4d,
e-mail: buy@unicyb.kiev.ua
e-mail: anna_inf@mail.ru

У статті викладено фрагмент математичної теорії нормалізації: наводяться строгі означення 2-4 нормальних форм (НФ) для табличних (реляційних) баз даних; визначаються потенційні ключі за умов, коли значення потужностей реляційної схеми і домена знаходяться у межах $|R| \leq 2$ або $|D| \leq 1$ (розглядаються умови, коли аксіоматика Армстронга не є повною); вказуються достатні умови для знаходження у 2-4 НФ для спеціальних випадків; встановлюються логічні зв'язки між сформульованими означеннями 2-4 НФ.

Ключові слова: табличні бази даних, нормальні форми.

This paper presents a fragment of normalization theory: rigorous definitions of Second Normal Form, Third Normal Form, Boyce-Codd Normal Form and Fourth Normal Form in table (relation) databases are given; finding out the meanings of candidate keys, provided that the values of cardinalities of the scheme R (finite set of attributes) and the universal domain D (the set, from which attributes take on values in interpretations) lie within $|R| \leq 2$ or $|D| \leq 1$ (considered conditions when Armstrong's axiomatic is not complete). Sufficient conditions which ensure that the table being in Second Normal Form, Third Normal Form, Boyce-Codd Normal Form and Fourth Normal Form for special cases are specified. Logical connections between the determined definitions of normal forms are established and it is shown that those definitions are satisfied the principle of "inclusion" higher normal forms in lower normal forms.

Key Words: table databases, normal forms.

Статтю представив д.ф.-м.н., проф. Анісімов А.В.

Різні підходи до понятійного апарату теорії нормалізації в реляційних базах даних призвели до різноманітності у визначеннях нормальних форм (НФ) в таких базах даних [1].

Метою роботи є побудова фрагменту математичної теорії нормалізації (стосовно 2-4 НФ для табличних баз даних), який би відповідав вимогам строгості для математичних теорій.

Всі неозначувані тут поняття та позначення використовуються в розумінні монографії [2]; зокрема, $t_\emptyset$ позначає порожню таблицю, яка не містить рядків і якій можна приписати довільну схему; t_ε позначає таблицю порожньої схеми $\emptyset$, яка містить тільки рядок ε також порожньої схеми: $t_\varepsilon \stackrel{\text{def}}{=} \{\varepsilon\}$.

Нехай t – таблиця, A, B, C – атрибути таблиці t , R – схема (довільна зафіксована скінченна множина атрибутів), X, Y, Z – підмножини схеми R , D – універсальний домен.

Суперключі, потенційні ключі, первинні та непервинні атрибути, повні та транзитивні ФЗ

Нагадаймо, що на таблиці t виконується, за означенням, функціональна залежність (ФЗ) $X \rightarrow Y$, якщо для двох довільних рядків s_1, s_2 таблиці t , які збігаються на множині атрибутів X , має місце їх рівність і на множині атрибутів Y [2]:

$$(X \rightarrow Y)(t) \stackrel{\text{def}}{=} \text{true} \Leftrightarrow \forall s_1, s_2 \in t (s_1|_X = s_2|_X \Rightarrow s_1|_Y = s_2|_Y).$$

Нехай F – множина ФЗ. Нагадаймо також, що таблиця t схеми R є моделлю множини ФЗ F , якщо кожна ФЗ $X \rightarrow Y \in F$ виконується на таблиці t :

$$t \text{ модель } F \stackrel{\text{def}}{\Leftrightarrow} \forall (X \rightarrow Y)(X \rightarrow Y \in F \Rightarrow (X \rightarrow Y)(t) = \text{true}) \quad [3].$$

ФЗ $X \rightarrow Y$ поставимо у відповідність множині ФЗ:

$$\Phi_{X \rightarrow Y} = \{X' \rightarrow Y \mid X' \subset X \wedge Y \not\subseteq X'\}.$$

Тобто, множина $\Phi_{X \rightarrow Y}$ містить усі нетривіальні ФЗ, такі, що їх права частина співпадає з множиною Y , а ліва частина є власною підмножиною множини X .

Очевидно, що $\Phi_{X \rightarrow Y} = \{X' \rightarrow Y \mid X' \in 2^X \setminus \{X\}\}$ за умови $X \subseteq Y$; $\Phi_{X \rightarrow Y} = \{X' \rightarrow Y \mid X' \in \{U \cup V \mid U \in 2^{X \setminus Y} \setminus \{X\}, V \in 2^Y \setminus \{Y\}\}\}$, за умови $Y \subseteq X$ або множини атрибутів X, Y непорівнювані відносно теоретико-множинного включення $\subseteq$. Вище 2^Y – булеан множини Y .

Означення 1. ФЗ $X \rightarrow Y$ називається *повною* (відносно множини ФЗ F), якщо вона виконується на довільній таблиці t , яка є моделлю множини F , а жодна ФЗ з множини $\Phi_{X \rightarrow Y}$ не виконується на таблиці t :

$$X \rightarrow Y \text{ – повна (відносно множини ФЗ } F) \stackrel{\text{def}}{\Leftrightarrow} \forall t(t \text{ – модель } F \Rightarrow (X \rightarrow Y)(t) = \text{true} \wedge \neg \forall \psi(\psi \in \Phi_{X \rightarrow Y} \Rightarrow \psi(t) = \text{false})). \quad \square$$

Враховуючи, що відношення синтаксичного і семантичного слідувань для аксіоматики ФЗ збігаються за умови $|D| \geq 2$ (див. критерій повноти аксіоматики Армстронга [4]), для вказаного випадку можна використовувати синтаксичну форму означення 1:

$$X \rightarrow Y \text{ – повна (відносно множини ФЗ } F) \stackrel{\text{def}}{\Leftrightarrow} X \rightarrow Y \in [F] \wedge \forall \psi(\psi \in \Phi_{X \rightarrow Y} \Rightarrow \psi \notin [F]),$$

де $[F]$ – синтаксичне замикання множини ФЗ F ¹.

Означення 2. Множина атрибутів $W \subseteq R$ називається *суперключем* (відносно множини ФЗ F), якщо ФЗ $W \rightarrow R$ виконується на довільній таблиці t , яка є моделлю множини F :

$$W \text{ – суперключ (відносно множини ФЗ } F) \stackrel{\text{def}}{\Leftrightarrow} \forall t(t \text{ – модель } F \Rightarrow (W \rightarrow R)(t) = \text{true}). \quad \square$$

Таким чином, наприклад, вся множина атрибутів схеми R є суперключем, оскільки ФЗ $R \rightarrow R$ є тривіальною і, значить, виконується на довільній таблиці.

Як і вище для випадку $|D| \geq 2$ можна використовувати синтаксичну форму означення 2:

$$W \text{ – суперключ (відносно множини ФЗ } F) \stackrel{\text{def}}{\Leftrightarrow} W \rightarrow R \in [F].$$

Означення 3. Множина атрибутів $K \subseteq R$ називається *потенційним ключем* (відносно множини ФЗ F), якщо K є суперключем і ФЗ $K \rightarrow R$ є повною. $\square$

Змістовно кажучи, потенційний ключ – це суперключ, ліву частину якого не можна зменшити, не втративши властивість бути суперключем.

Лема 1 (про існування потенційного ключа). Для довільної множини ФЗ F , заданої над схемою R , існує не менше одного потенційного ключа. $\square$

Доведення. Дійсно, якщо тривіальна ФЗ $R \rightarrow R$ є повною, то множина атрибутів R є шуканим потенційним ключем; інакше існує така власна підмножина $X^1 \subset R$, що на довільній моделі множини ФЗ F виконується ФЗ $X^1 \rightarrow R$. Якщо ФЗ $X^1 \rightarrow R$ є повною, то множина атрибутів X^1 є потенційним ключем, інакше існує така власна підмножина $X^2 \subset X^1 \subset R$, що $(X^2 \rightarrow R)(t) = \text{true}$, де t – довільна модель множини ФЗ F . Продовжуючи таким чином і враховуючи, що множина атрибутів R містить скінченну кількість атрибутів, на k -ому кроці (для певного $k \leq |R|$) знайдеться така множина атрибутів X^k , що $(X^k \rightarrow R)(t) = \text{true}$ і ФЗ $X^k \rightarrow R$ є повною. За означенням 3 множина атрибутів X^k і є шуканим потенційним ключем. $\square$

Задамо на множині $2^R \times 2^R$ (по суті, на множині всіх ФЗ, бо ФЗ є парою множин атрибутів) наступне бінарне відношення. Скажемо, що ФЗ $X' \rightarrow Y'$ і $X \rightarrow Y$ знаходяться у відношенні $\leq$, якщо $X' \subseteq X$ та $Y' = Y$:

$$X' \rightarrow Y' \leq X \rightarrow Y \stackrel{\text{def}}{\Leftrightarrow} X' \subseteq X \wedge Y' = Y.$$

Очевидно, що відношення $\leq$ на множині $2^R \times 2^R$ (тобто на множині всіх ФЗ) задає частковий порядок. Дійсно, відношення $\leq$ є:

¹ Див., наприклад, [3].

1) рефлексивним, оскільки
 $X \rightarrow Y \sqsubseteq X \rightarrow Y \Leftrightarrow X \subseteq X$;

2) антисиметричним, оскільки з того, що
 $X' \rightarrow Y \sqsubseteq X \rightarrow Y$ і $X \rightarrow Y \sqsubseteq X' \rightarrow Y$ відповідно
слідують включення $X' \subseteq X$ і $X \subseteq X'$, звідки
 $X' = X$;

3) транзитивним, оскільки з того, що
 $X'' \rightarrow Y \sqsubseteq X' \rightarrow Y$ і $X' \rightarrow Y \sqsubseteq X \rightarrow Y$ відповідно
слідують включення $X'' \subseteq X'$ і $X' \subseteq X$, звідки
 $X'' \subseteq X$, а отже, $X'' \rightarrow Y \sqsubseteq X \rightarrow Y$.

Розглянемо частково впорядковану множину
 $\langle \Phi_{R \rightarrow R} \cup \{R \rightarrow R\} \cap [F], \sqsubseteq \rangle$, де, як і раніше,
 $\Phi_{R \rightarrow R} = \{X \rightarrow R \mid X \subset R\}$. Тоді її кожний
мінімальний елемент вигляду $X \rightarrow R$ є повною
ФЗ і задає потенційний ключ X . Наявність
декількох мінімальних елементів визначає
декілька потенційних ключів. По суті це ще одне
визначення потенційного ключа.

Значення потенційних ключів при деяких
значеннях потужностей множин R та D вказані
у табл. 1. Дана таблиця містить три рядки та
чотири стовпчики. Комірки позначаються
парами, перша компонента – номер рядка, друга
– номер стовпчика.

однорядкові таблиці; оскільки ФЗ вигляду
 $\emptyset \rightarrow R$ виконується на порожній та
однорядковій таблицях (леми 1, 4 з [4]) з
очевидних включень $\emptyset \subset \{A\} \subseteq R$, де A –
довільний (або єдиний) атрибут схеми R , маємо,
що єдиним потенційним ключем є порожня
множина $\emptyset$; °

в) $|R|=1$, $|D| \geq 2$ (випадок (3,2)); в
інтерпретаціях не більше ніж одноатрибутні
таблиці; нехай для визначеності $R = \{A\}$.
Оскільки ФЗ $\emptyset \rightarrow \{A\}$ – не виконується на
таблиці з двох рядків, то таблиця має єдиний
потенційний ключ $\{A\}$ (ФЗ $\{A\} \rightarrow \{A\}$ є
тривіальною); °

г) $|R|=2$ і $|D| \geq 2$ (випадок (3,3)); нехай для
визначеності $R = \{A, B\}$ і $d_1, d_2 \in D$, де d_1, d_2 –
різні. Оскільки ФЗ вигляду $\emptyset \rightarrow R$ не
виконується на довільній дворядковій таблиці та
множини $\{A\}$, $\{A, B\}$, $(\{B\}, \{A, B\})$ не можуть
бути одночасно потенційними ключами з огляду
на мінімальність потенційних ключів, то
множина потенційних ключів може мати один
ключ: $\{A\}$ або $\{B\}$ або $\{A, B\}$, чи два: $\{A\}$ та
 $\{B\}$. Множина ключів залежить від множини ФЗ

Таблиця 1

Потенційні ключі при певних значеннях потужностей множин R та D

	1	2	3	4
$ R \backslash D $	$ R =0$	$ R =1$ $R = \{A\}$	$ R =2$ $R = \{A, B\}$	$ R >2$
1 $ D =0$	$\emptyset$	$\emptyset$	$\emptyset$	$\emptyset$
2 $ D =1$	$\emptyset$	$\emptyset$	$\emptyset$	$\emptyset$
3 $ D \geq 2$	$\emptyset$	Єдиний ключ: $\{A\}$	Єдиний ключ: $\{A\}$ або $\{B\}$ або $\{A, B\}$. Два ключі: $\{A\}$ і $\{B\}$	Ключі визначаються, виходячи з множини ФЗ F

Лема 2. Заповнення табл. 1 коректне. □

Доведення. Доведення проводиться
розглядом кожного з випадків.

а) $|R|=0$ (випадки (1,1), (2,1), (3,1), тобто
перший стовпчик); тоді, незалежно від
потужності домена, $t_\emptyset$, t_ε – всі можливі таблиці
в інтерпретації, бо атрибутів в даному випадку
немає. Єдиною можливою ФЗ є $\emptyset \rightarrow \emptyset$, яка є
тривіальною, звідси $\emptyset$ – потенційний ключ; °

б) $|R| \geq 1$, $|D| \leq 1$ (випадки (1,2), (1,3), (1,4),
(2,2), (2,3), (2,4), тобто комірки перших двох
рядків, окрім комірок першого стовпчика); в
інтерпретаціях розглядаються не більше ніж

F ; точніше кажучи: якщо $F = \emptyset$, то єдиний
потенційний ключ $\{A, B\}$; якщо $\{A\} \rightarrow \{A, B\} \in F$
($\{B\} \rightarrow \{A, B\} \in F$), то єдиний потенційний ключ
 $\{A\}$ (відповідно $\{B\}$); якщо
 $\{\{A\} \rightarrow \{A, B\}, \{B\} \rightarrow \{A, B\}\} \subseteq F$, то маємо два
потенційних ключа $\{A\}$ і $\{B\}$; °

д) $|R| \geq 2$ і $|D| \geq 2$ (випадок (3,4)); множина
потенційних ключів цілком визначається
множиною ФЗ F , тобто в цьому випадку нічого
не можна сказати про потенційні ключі. °□

Наслідок 1. Якщо $|R|=0$ або $|D|\leq 1$, то потенційним ключем буде тільки порожня множина $\emptyset$. $\square$

Доведення випливає з заповнення табл. 1. $\square$

Лема 3. Якщо $|R|>0$ і $|D|\geq 2$, то ФЗ виду $\{A\}\rightarrow Y$, де $A\in R$, $\emptyset\neq Y\subseteq R$ є повною за умови $\{A\}\rightarrow Y\in[F]$ (тобто виконання ФЗ $\{A\}\rightarrow Y$ на довільній моделі множини ФЗ F). $\square$

Для доведення достатньо врахувати той факт, що ФЗ вигляду $\emptyset\rightarrow Y$, де $Y\subseteq R$ і $Y\neq\emptyset$ не виконується на довільній дворядковій таблиці. $\square$

Означення 4. Атрибут $A\in R$ називається *транзитивно залежним* від множини атрибутів $X\subset R$ (відносно множини ФЗ F), якщо існує така множина атрибутів $Y\subset R$, що $A\notin X\cup Y$ і для довільної таблиці t , яка є моделлю множини ФЗ F , виконується $(X\rightarrow Y)(t)=true$, $(Y\rightarrow X)(t)=false$, $(Y\rightarrow\{A\})(t)=true^2$:

$$\exists Y(Y\subset R\wedge A\in R\setminus(X\cup Y)\wedge\forall t(t\text{ – модель}$$

$$F\Rightarrow(X\rightarrow Y)(t)=true\wedge$$

$$\wedge(Y\rightarrow X)(t)=false\wedge(Y\rightarrow\{A\})(t)=true). \square$$

Для випадку $|D|\geq 2$ запишемо синтаксичну форму означення 4:

атрибут A *транзитивно залежний* від множини атрибутів $X\subset R$ (відносно множини

$$\text{ФЗ } F) \stackrel{\text{def}}{\Leftrightarrow} \exists Y(Y\subset R\wedge A\in R\setminus(X\cup Y)\wedge \\ \wedge X\rightarrow Y\in[F]\wedge Y\rightarrow X\notin[F]\wedge Y\rightarrow\{A\}\in[F]).$$

Лема 4. Якщо K – потенційний ключ (відносно множини ФЗ F) і для деякого атрибута $A\in R$ ФЗ $K\rightarrow\{A\}$ не є повною, то атрибут A є транзитивно залежним від множини атрибутів K . $\square$

Доведення. Розглянемо довільну таблицю t , яка є моделлю множини ФЗ F . Нехай ФЗ $K\rightarrow\{A\}$ не є повною, оскільки ФЗ $K\rightarrow\{A\}$, очевидно, виконується на таблиці t (з огляду на те, що K – ключ), тоді існує така власна підмножина $K'\subset K$, $\{A\}\not\subset K'$ (тобто $A\notin K'$), що виконуються ФЗ $K'\rightarrow\{A\}$ на таблиці t ; але оскільки K – потенційний ключ, то ФЗ $K'\rightarrow K$ не виконується на таблиці t з огляду на мінімальність потенційного ключа K . Таким

чином, з виконання ФЗ $K\rightarrow K'$, $K'\rightarrow\{A\}$ і невиконання ФЗ $K'\rightarrow K$ на таблиці t (з огляду на її довільність) за означенням 4 атрибут A є транзитивно залежним від потенційного ключа K . $\square$

Означення 5. Атрибут $A\in R$ називається *первинним* (відносно множини ФЗ F), якщо A міститься в якому-небудь потенційному ключі³:

$$A\text{ – первинний} \stackrel{\text{def}}{\Leftrightarrow} A\in\bigcup_{K\in\Delta} K,$$

де $\Delta=\Delta(F)$ – сім'я всіх потенційних ключів (відносно множини ФЗ F). $\square$

Означення 6. Атрибут $A\in R$ називається *непервинним* (відносно множини ФЗ F), якщо A не міститься в жодному з потенційних ключів:

$$A\text{ – непервинний} \stackrel{\text{def}}{\Leftrightarrow} A\in R\setminus\bigcup_{K\in\Delta} K,$$

де Δ – як і раніше, сім'я всіх потенційних ключів. $\square$

Друга нормальна форма

Як і раніше фіксуємо схему R та множини ФЗ F .

Означення 7. Скажемо, що таблиця $t(R)$ знаходиться у *другій НФ* (2НФ), якщо, по-перше, таблиця $t(R)$ знаходиться у першій НФ (1НФ⁴), по-друге, таблиця t є моделлю множини ФЗ F та, по-третє, кожний непервинний атрибут характеризується повною ФЗ від кожного потенційного ключа, тобто:

$$\forall A\forall K(K\in\Delta\wedge A\in R\setminus\bigcup_{\bar{K}\in\Delta}\bar{K}\Rightarrow K\rightarrow\{A\}\text{ – повна ФЗ}). \square \quad (1)$$

З означення 7 випливає, що якщо множина непервинних атрибутів порожня, то умова (1) цього означення виконується автоматично (відповідна імплікація тривіально істинна, оскільки умова $A\in R\setminus\bigcup_{\bar{K}\in\Delta}\bar{K}$ хибна).

Твердження 1 (достатні умови для знаходження у 2НФ для спеціальних випадків). Таблиця t знаходиться у 2НФ, якщо $|R|\leq 2$ або $|D|\leq 1$. $\square$

Доведення. Використовуючи результати з табл. 1 розглянемо усі можливі випадки для умов

² Означення наведено відповідно до [5]. Дрібас у монографії [6] в означенні транзитивної ФЗ надає додаткову умову: $Y\not\subset X$. Тоді подальша лема 4 не є чинною, а в наведеному нижче означенні 3НФ потр*ібно вимагати, щоб таблиця задовольняла умовам 2НФ.

³ Означення 5-6 наводяться відповідно до [5-7].

⁴ Оскільки означення 1НФ не підлягає повній формалізації, пропонуємо використовувати один з його варіантів згідно [7], адаптований до термінології даної статті: «таблиця знаходиться у 1НФ тоді і тільки тоді, коли кожний її рядок містить тільки одне значення для кожного атрибуту».

$|R| \leq 2$ (перші три стовпчики) або $|D| \leq 1$ (перші два рядки):

а) $|R| = 0$ або $|D| \leq 1$ (комірки (1,1), (1,2), (1,3), (1,4), (2,1), (2,2), (2,3), (2,4), (3,1)); тобто два перших рядка та комірка (3,1); згідно з табл. 1 єдиним потенційним ключем є $\emptyset$, тобто $\Delta = \{\emptyset\}$, звідки всі атрибути схеми R є непервинними; оскільки в цьому випадку розглядаються не більше ніж однорядкові таблиці, то, очевидно, що довільна така таблиця буде моделлю множини $\Phi_3 F$; таким чином, залишається перевірити виконання умови (1), яка в даному випадку приймає вигляд:

$$\forall A(A \in R \Rightarrow \emptyset \rightarrow \{A\}) \text{ – повна } \Phi_3,$$

та, знову ж виконується очевидно; $\square$

б) $|D| \geq 2$ і $|R| = 1$ (комірка (3,2)); нехай для визначеності $R = \{A\}$ (як в табл. 1), тоді множина $\{A\}$ є потенційним ключем (див. табл. 1), а оскільки схема взагалі не містить непервинних атрибутів, то таблиця, яка є моделлю множини $\Phi_3 F$, знаходиться у 2НФ; $\square$

в) $|D| \geq 2$ і $|R| = 2$: (комірка (3,3)); нехай для визначеності $R = \{A, B\}$ (як в табл. 1). Нехай потенційним ключем є одноелементна множина, наприклад, $\{A\}$. Тоді $\Phi_3 \{A\} \rightarrow \{B\}$, де $\{B\}$ – єдиний непервинний атрибут, є повною згідно леми 3. Отже, таблиця t , яка є моделлю множини $\Phi_3 F$, знаходиться у 2НФ. $\square$

Нехай тепер множиною потенційних ключів є або $\Delta = \{\{A\}, \{B\}\}$ або $\Delta = \{\{A, B\}\}$; тоді схема взагалі не містить непервинних атрибутів, а отже, умова 2НФ (1) для таблиці t , яка є моделлю множини $\Phi_3 F$, виконується автоматично. $\square$

Змістовна інтерпретація твердження 1: не більше ніж двоатрибутні таблиці або не більше ніж однорядкові таблиці знаходяться у 2НФ (за умови, що вони є моделями множини $\Phi_3 F$).

В наступній лемі Δ – сім'я всіх потенційних ключів, як і раніше.

Лема 5 (достатні умови для знаходження у 2НФ). Таблиця t , яка є моделлю множини $\Phi_3 F$, знаходиться у 2НФ, якщо виконується одна з наступних умов:

1) всі потенційні ключі одноелементні;

2) $\bigcup_{K \in \Delta} K = R$, тобто сім'я потенційних ключів

є покриттям схеми R (еквівалентно: множина непервинних атрибутів порожня)⁵. $\square$

⁵ Формулювання умов наводиться згідно [8].

Доведення. Доведемо перше твердження. Оскільки при значеннях потужностей $|D| \leq 1$ або $|R| = 0$ потужність потенційного ключа нульова (наслідок 1), нам потрібно розглянути випадок, коли $|D| \geq 2$ і $|R| > 0$. Нехай всі потенційні ключі одноелементні. Згідно з лемою 3 усі непервинні атрибути повністю залежать від кожного потенційного ключа; отже, таблиця t знаходиться у 2НФ. $\square$

Доведемо друге твердження. Оскільки за умови 2) в формулюванні схема взагалі не містить непервинних атрибутів, то умова 2НФ для таблиці t не порушується. $\square$

Третя нормальна форма

Як і раніше фіксуємо схему R та множину $\Phi_3 F$.

Задамо на множині 2^R наступне бінарне відношення. Скажемо, що множини атрибутів $X, Y \subseteq R$ знаходяться у відношенні $\equiv$ (відносно множини $\Phi_3 F$), якщо $F \models X \rightarrow Y$ і $F \models Y \rightarrow X$.

Лема 6. Відношення $\equiv$ є відношенням еквівалентності. $\square$

Доведення. Дійсно, відношення $\equiv$ є:

1) рефлексивним, оскільки $X \equiv X \Leftrightarrow F \models X \rightarrow X$ ($\Phi_3 X \rightarrow X$ є тривіальною);

2) симетричним, оскільки з того, що $X \equiv Y$ випливає $Y \equiv X$;

3) транзитивним; покажемо, що виконується імплікація $X \equiv Y \wedge Y \equiv Z \Rightarrow X \equiv Z$. Розглянемо випадки:

а) $|D| \leq 1$ (перші два рядки табл. 1); оскільки на порожній та однорядковій таблицях виконуються довільні Φ_3 , то імплікація виконується очевидно;

б) $|R| = 0$ (перший стовпчик табл. 1); єдиною можливою Φ_3 є тривіальна $\Phi_3 \emptyset \rightarrow \emptyset$ отже, імплікація виконується тривіально;

в) $|D| \geq 2 \wedge |R| \geq 1$; враховуючи, що для даного випадку відношення синтаксичного слідування ($| - \rangle$) та семантичного слідування ($| = \rangle$) співпадають (див. критерій повноти аксіоматики Армстронга [4]), побудуємо доведення для $\Phi_3 X \rightarrow Z$:

1. $X \rightarrow Y$ (за припущенням $X \equiv Y$; отже, $F \models X \rightarrow Y$, звідси за критерієм повноти аксіоматики Армстронга маємо: $F | - X \rightarrow Y$);

2. $Y \rightarrow Z$ (за припущенням $Y \equiv Z$; отже, $F \models Y \rightarrow Z$, звідси за критерієм повноти аксіоматики Армстронга маємо: $F | - Y \rightarrow Z$);

3. $X \rightarrow Z$ (з 1 і 2 за правилом транзитивності).

Отже, $F \models X \rightarrow Z$, звідси випливає $F \models X \rightarrow Z$ (нагадаємо, що $|D| \geq 2$). Аналогічно будується доведення для $\Phi 3 \ Z \rightarrow X$, тобто, $F \models Z \rightarrow X$, отже $F \models Z \rightarrow X$. Звідси маємо: $X \equiv Z$. $\square$

Таким чином, відношення $\equiv$ розбиває множину 2^R (булеан схеми R) на класи еквівалентності.

Наслідок 2. За умов $|D| \leq 1$ або $|R| = 0$ усі елементи множини $2^R = \{X \mid X \subseteq R\}$ належать до одного класу еквівалентності. $\square$

Доведення випливає безпосередньо з доведення леми 6. $\square$

Лема 7. Множина суперключів належить до одного класу еквівалентності. $\square$

Доведення. Нехай множина атрибутів K є суперключем (відносно множини $\Phi 3 \ F$), тоді виконується $\Phi 3 \ K \rightarrow R$, тобто $F \models K \rightarrow R$. Враховуючи виконання тривіальної $\Phi 3 \ R \rightarrow K$ ($K \subseteq R$), тобто $F \models R \rightarrow K$, маємо $K \equiv R$. $\square$

Позначимо клас еквівалентності суперключів $[R]_{\equiv}$, де схема R є представником класу.

Лема 8. Виконується імплікація:

$$\forall A, \forall X (\forall Y (A \in R \setminus (X \cup Y) \wedge (Y \rightarrow \{A\})(t) = true \Rightarrow X \equiv Y) \Rightarrow$$

$$\Rightarrow A \text{ не є транзитивно залежним від } X). \square$$

Доведення проводиться безпосередньо. $\square$

Відомо, що наявність непервинних атрибутів, які є транзитивно залежними від потенційних ключів, приводить до виникнення аномалій (згідно визначення аномалій з [9]). З леми 8 випливає, що для вирішення вказаної проблеми достатньо вимагати, щоб ліві частини усіх нетривіальних $\Phi 3$ виду $X \rightarrow \{A\}$, де $\{A\}$ – непервинний атрибут, належали до класу $[R]_{\equiv}$ ⁶.

Означення 8. Скажемо, що таблиця $t(R)$ знаходиться у *третій НФ (ЗНФ)*, якщо, по-перше, таблиця $t(R)$ знаходиться у 1НФ, по-друге, таблиця t є моделлю множини $\Phi 3 \ F$ та, по-третє, для кожної нетривіальної $\Phi 3$ виду $X \rightarrow \{A\}$ такої, що $F \models X \rightarrow \{A\}$, де $\{A\}$ – непервинний атрибут, множина атрибутів X є суперключем, тобто:

$$\forall X, \forall A (X \subseteq R \wedge A \in R \setminus (\bigcup_{K \in \Delta} K \cup X) \wedge$$

$$\wedge (X \rightarrow \{A\})(t) = true \Rightarrow (X \rightarrow R)(t) = true). \square \quad (2)$$

З означення 8 випливає, що, якщо множина непервинних атрибутів порожня, то умова (2) цього означення виконується автоматично (відповідна імплікація тривіально істинна). Таким чином, доведена наступна лема.

Лема 9 (достатня умова для знаходження у ЗНФ). Таблиця t , яка є моделлю множини $\Phi 3 \ F$, знаходиться у ЗНФ якщо $\bigcup_{K \in \Delta} K = R$, тобто сім'я

потенційних ключів є покриттям схеми R . $\square$

Твердження 2 (достатні умови для знаходження у ЗНФ для спеціальних випадків). Таблиця знаходиться у ЗНФ якщо $|R| \leq 2$ або $|D| \leq 1$. $\square$

Доведення. Враховуючи результати з табл. 1 розглянемо усі можливі випадки для умов $|R| \leq 2$ (перші три стовпчики) або $|D| \leq 1$ (перші два рядки):

а) $|R| = 0$ (комірки (1,1), (2,1), (3,1), тобто перший стовпчик); оскільки порожня схема R не має непервинних атрибутів, то умови ЗНФ виконуються автоматично (лема 9); $\square$

б) $|R| \geq 1$ і $|D| \leq 1$ (комірки (1,2), (1,3), (1,4), (2,2), (2,3), (2,4); тобто перші два рядки); згідно з табл. 1 єдиним потенційним ключем є $\emptyset$, тобто $\Delta = \{\emptyset\}$, тоді будь-який елемент множини $2^R = \{X \mid X \subseteq R\}$ є суперключем; таким чином, умова (2) виконується для кожної відповідної $\Phi 3$ вигляду $X \rightarrow \{A\}$; оскільки в цьому випадку розглядаються не більше ніж однорядкові таблиці, то, очевидно, що довільна така таблиця буде моделлю множини $\Phi 3 \ F$ і знаходиться у ЗНФ; $\square$

с) $|R| = 1$, $|D| \geq 2$ (комірка (3,2)); нехай для визначеності $R = \{A\}$ (як в табл. 1), тоді множина $\{A\}$ є потенційним ключем (див. табл. 1), а оскільки схема не містить непервинних атрибутів, то за лемою 9 таблиця, яка є моделлю множини $\Phi 3 \ F$, знаходиться у ЗНФ; $\square$

д) $|R| = 2$, $|D| \geq 2$ (комірка (3,3)); нехай для визначеності $R = \{A, B\}$ (як в табл. 1). Нехай потенційним ключем є одноелементна множина, наприклад, $\{A\}$, тоді B – єдиний непервинний атрибут. Запишемо усі елементи множини $2^R = \{X \mid X \subseteq R\}$: $\{\emptyset, \{A\}, \{B\}, \{A, B\}\}$. Оскільки умова (2) виключає з розгляду тривіальні $\Phi 3$, то залишається перевірити лише імплікації:

$$1. (\emptyset \rightarrow \{B\})(t) = true \Rightarrow (\emptyset \rightarrow \{A, B\})(t) = true;$$

$$2. (\{A\} \rightarrow \{B\})(t) = true \Rightarrow (\{A\} \rightarrow \{A, B\})(t) = true.$$

⁶ Саме такий підхід здійснюється у класичних алгоритмах, описаних у [6].

В першій імплікації посилка є хибною, оскільки $\Phi_3 \emptyset \rightarrow \{B\}$ не виконуються на довільній дворядковій таблиці, отже, імплікація виконується.

Друга імплікація виконується, оскільки за припущенням множина $\{A\}$ є потенційним ключем. Отже, таблиця, яка є моделлю множини $\Phi_3 F$, знаходиться у ЗНФ. $\square$

Нехай тепер множиною потенційних ключів є або $\{\{A\}, \{B\}\}$ або $\{(A, B)\}$; тоді схема взагалі не містить непервинних атрибутів, а отже, за лемою 9 таблиця, яка є моделлю множини $\Phi_3 F$, знаходиться у ЗНФ. $\square$

Змістовна інтерпретація твердження 2: не більше ніж двоатрибутні таблиці або не більше ніж однорядкові таблиці знаходяться у ЗНФ (за умови, що вони є моделями множини $\Phi_3 F$).

Наслідок 3. $ЗНФ \Leftrightarrow 2НФ$ за умови $|R| \leq 2$ або $|D| \leq 1$. $\square$

Доведення випливає з тверджень 1 і 2. $\square$

Лема 10. Якщо таблиця t знаходиться у ЗНФ, то вона знаходиться у 2НФ: $ЗНФ \Rightarrow 2НФ$. $\square$

Доведення. З урахуванням наслідку 3 достатньо розглянути випадок, коли $|D| \geq 2$ і $|R| > 2$.

Нехай таблиця t , яка є моделлю множини $\Phi_3 F$, знаходиться у ЗНФ. Доведення проводиться від супротивного. Припустимо таблиця t не знаходиться у 2НФ, тоді для деякого непервинного атрибута A знайдеться такий потенційний ключ K , що $\Phi_3 K \rightarrow \{A\}$ не є повною, тобто, існує така підмножина $K' \subset K$, що виконується $\Phi_3 K' \rightarrow \{A\}$. Це порушує виконання умови (2), оскільки в силу мінімальності потенційного ключа множина $K' \subset K$ не є суперключем. Прийшли до протиріччя. $\square$

Нормальна форма Бойса-Кодда

Означення 9. Скажемо, що таблиця $t(R)$ знаходиться у нормальній формі Бойса-Кодда (НФБК), якщо, по-перше, таблиця $t(R)$ знаходиться у 1НФ, по-друге, таблиця t є моделлю множини $\Phi_3 F$ та, по-третє, для кожної нетривіальної Φ_3 виду $X \rightarrow \{A\}$ такої, що $F \models X \rightarrow \{A\}$, множина атрибутів X є суперключем:

$$\begin{aligned} & \forall X, \forall A \\ & (X \subset R \wedge A \in R \setminus X \wedge (X \rightarrow \{A\})(t) = \text{true} \Rightarrow \\ & \Rightarrow (X \rightarrow R)(t) = \text{true}). \square \end{aligned} \quad (3)$$

Твердження 3 (достатні умови для знаходження у НФБК для спеціальних випадків).

Таблиця знаходиться у НФБК якщо $|R| \leq 2$ або $|D| \leq 1$. $\square$

Доведення. Враховуючи результати з табл. 1 розглянемо усі можливі випадки для умов $|R| \leq 2$ (перші три стовпчики) або $|D| \leq 1$ (перші два рядки).

а) $|R| = 0$ або $|D| \leq 1$ (комірки (1,1), (2,1), (3,1), (1,2), (1,3), (1,4), (2,2), (2,3), (2,4); тобто перший стовпчик і перші два рядки); згідно з наслідком 2 усі елементи множини $2^R = \{X \mid X \subseteq R\}$ належать до одного класу еквівалентності – класу суперключів $[R]_{\equiv}$, де схема R є представником класу. Отже, умови НФБК виконуються автоматично і таблиця, яка є моделлю множини $\Phi_3 F$, знаходиться у НФБК; $\square$

б) $|R| = 1$, $|D| \geq 2$ (комірка (3,2)); нехай для визначеності $R = \{A\}$ (як в табл. 1), тоді множина $\{A\}$ є потенційним ключем (див. табл. 1). Імплікація з умови (3) для єдиної нетривіальної $\Phi_3 \emptyset \rightarrow \{A\}$ запишеться як:

$$(\emptyset \rightarrow \{A\})(t) = \text{true} \Rightarrow (\emptyset \rightarrow \{A\})(t) = \text{true},$$

і виконується тривіально. Отже, таблиця, яка є моделлю множини $\Phi_3 F$, знаходиться у НФБК. $\square$

с) $|R| = 2$, $|D| \geq 2$ (комірка (3,3)); нехай для визначеності $R = \{A, B\}$ (як в табл. 1). Можливі два випадки:

1) нехай множина потенційних ключів – одноелементна, тобто, $\Delta = \{\{A\}\}$ або $\Delta = \{\{B\}\}$, або $\Delta = \{\{A, B\}\}$. Враховуючи доведення з твердження 2 для даного випадку, залишається перевірити виконання умови (3) для нетривіальної $\Phi_3 \emptyset \rightarrow \{A\}$ ($\emptyset \rightarrow \{B\}$), де атрибут A (атрибут B) є первинним:

$$(\emptyset \rightarrow \{A\})(t) = \text{true} \Rightarrow (\emptyset \rightarrow \{A, B\})(t) = \text{true}.$$

Якщо посилка є хибною, то імплікація тривіально істинна. Якщо посилка є істинною, то всі рядки таблиці t мають однакове значення атрибута A . Оскільки $\{A\}$ – потенційний ключ, то у всіх рядків таблиці t однакові значення атрибута B . Отже, імплікація також істинна;

2) нехай тепер множина потенційних ключів двоелементна, тобто, $\Delta = \{\{A\}, \{B\}\}$; враховуючи попередній пункт доведення залишається перевірити виконання умови (3) для нетривіальної $\Phi_3 \{A\} \rightarrow \{B\}$ ($\{B\} \rightarrow \{A\}$):

$$(\{A\} \rightarrow \{B\})(t) = \text{true} \Rightarrow (\{A\} \rightarrow \{A, B\})(t) = \text{true}.$$

Імплікація виконується, оскільки за припущенням множина $\{A\}$ є потенційним

ключем. Отже, таблиця, яка є моделлю множини ФЗ F , знаходиться у НФБК. $\square$

Змістовна інтерпретація останнього твердження: не більше ніж двоатрибутні таблиці або не більше ніж однорядкові таблиці знаходяться у НФБК (за умови, що вони є моделями множини ФЗ F).

Наслідок 4. $НФБК \Leftrightarrow 3НФ$ за умови $|R| \leq 2$ або $|D| \leq 1$. $\square$

Доведення випливає з лем 11 і 13. $\square$

Наслідок 5. $НФБК \Leftrightarrow 3НФ \Leftrightarrow 2НФ$ за умови $|R| \leq 2$ або $|D| \leq 1$. $\square$

Доведення випливає з наслідків 3 і 4. $\square$

Лема 11. Якщо таблиця t знаходиться у НФБК, то вона знаходиться у 3НФ: $НФБК \Rightarrow 3НФ$. $\square$

Доведення. Нехай таблиця t , яка є моделлю множини ФЗ F , знаходиться у НФБК. З урахуванням включення $R \setminus (\bigcup_{K \in \Delta} K \cup X) \subseteq R \setminus X$ з виконання умови (3)

$$\forall X, \forall A (X \subset R \wedge A \in R \setminus X \wedge (X \rightarrow \{A\})(t) = true \Rightarrow (X \rightarrow R)(t) = true)$$

тривіально отримуємо виконання умови (2) для таблиці t :

$$\forall X, \forall A (X \subset R \wedge A \in R \setminus (\bigcup_{K \in \Delta} K \cup X) \wedge (X \rightarrow \{A\})(t) = true \Rightarrow (X \rightarrow R)(t) = true).$$

Отже, таблиця t , яка є моделлю множини ФЗ F , знаходиться у 3НФ. $\square$

Наслідок 6. Виконуються імплікації: $НФБК \Rightarrow 3НФ \Rightarrow 2НФ$. $\square$

Доведення випливає з лем 14 і 12. $\square$

Четверта нормальна форма

Нагадаймо, що на таблиці t схеми R виконується багатозначна залежність (БЗЗ) $X \twoheadrightarrow Y$, якщо для двох довільних рядків s_1, s_2 таблиці t , які збігаються на множині атрибутів X , існує рядок $s_3 \in t$, який дорівнює об'єднанню обмежень рядків s_1, s_2 на множині атрибутів $X \cup Y$ і $R \setminus (X \cup Y)$ відповідно:

$$(X \twoheadrightarrow Y)(t) = true \stackrel{def}{\Leftrightarrow} \forall s_1, s_2 \in t (s_1|_X = s_2|_X \Rightarrow \exists s_3 \in t (s_3|_X = s_1|_X \wedge s_3|_Y = s_2|_Y \wedge s_3|_{R \setminus (X \cup Y)} = s_1|_{R \setminus (X \cup Y)})) [2].$$

Нехай зафіксовані множини ФЗ F і БЗЗ G .

Нагадаймо, що таблиця $t(R)$ є моделлю множини $F \cup G$, якщо кожна залежність $\varphi \in F \cup G$ виконується на таблиці t :

$$t(R) \text{ модель } F \cup G \stackrel{def}{\Leftrightarrow}$$

$$\stackrel{def}{\Leftrightarrow} \forall \varphi (\varphi \in F \cup G \Rightarrow \varphi(t) = true) [10].$$

Означення 10. Скажемо, що таблиця $t(R)$ знаходиться у четвертій НФ (4НФ), якщо, по-перше, таблиця $t(R)$ знаходиться у 1НФ, по-друге, таблиця t є моделлю об'єднання множин ФЗ і БЗЗ $F \cup G$ та, по-третє, для кожної нетривіальної БЗЗ $X \twoheadrightarrow Y$, такої, що $F \cup G \models X \twoheadrightarrow Y$, множина атрибутів X є суперключем:

$$\forall X, \forall Y (X \subset R \wedge Y \neq R \setminus X \wedge X \wedge \wedge (X \twoheadrightarrow Y)(t) = true \Rightarrow (X \rightarrow R)(t) = true). \square (4)$$

Твердження 4 (достатні умови для знаходження у 4НФ для спеціальних випадків). Таблиця знаходиться у 4НФ якщо $|R| \leq 1$ або $|D| \leq 1$. $\square$

Доведення. Враховуючи результати з табл. 1 розглянемо усі можливі випадки для умов $|R| \leq 1$ (перші два стовпчики) або $|D| \leq 1$ (перші два рядки).

а) $|R| = 0$ або $|D| \leq 1$ (комірки (1,1), (2,1), (3,1), (1,2), (1,3), (1,4), (2,2), (2,3), (2,4)); тобто перший стовпчик і перші два рядки); згідно з табл. 1 єдиним потенційним ключем є $\emptyset$, тобто $\Delta = \{\emptyset\}$, звідки всі елементи множини $2^R = \{X \mid X \subseteq R\}$ є суперключами. Отже, умови 4НФ виконуються автоматично і таблиця, яка є моделлю множини $F \cup G$, знаходиться у 4НФ; $\square$

б) $|R| = 1$ і $|D| \geq 2$ (комірка (3,2)); нехай для визначеності $R = \{A\}$ (як в табл. 1). Можливими БЗЗ є $\emptyset \twoheadrightarrow \emptyset$, $\emptyset \twoheadrightarrow \{A\}$, $\{A\} \twoheadrightarrow \emptyset$, $\{A\} \twoheadrightarrow \{A\}$, кожна з яких є тривіальною. Отже, умови 4НФ виконуються автоматично (відповідна імплікація тривіально істинна) і таблиця, яка є моделлю множини $F \cup G$, знаходиться у 4НФ. $\square$

Зауважимо, що за умови $|R| = 2$ таблиця буде знаходитись у НФБК, але може не знаходитись в 4НФ, як показано в наступному прикладі.

Приклад 1. Нехай задані схема $R = \{A, B\}$ і множина БЗЗ $G = \{\emptyset \twoheadrightarrow \{A\}\}$. З вказаних умов випливає, що потенційним ключем є множина атрибутів $\{A, B\}$. Таблиця із схемою $\{A, B\}$, зображена на рис. 1, є моделлю заданої множини БЗЗ G . Очевидно, що таблиця знаходиться в НФБК, але, оскільки ліва частина БЗЗ $\emptyset \twoheadrightarrow \{A\}$

не є суперключем, таблиця не знаходиться в 4НФ.

A	B
1	1
1	2

Рисунок 1. Таблиця з прикладу 1.

Лема 12. Якщо таблиця $t(R)$ знаходиться у 4НФ, то вона знаходиться у НФБК: $4НФ \Rightarrow НФБК$. □

Доведення. З урахуванням тверджень 3 і 4 достатньо розглянути випадок, коли $|D| \geq 2$ і $|R| \geq 2$. Нехай таблиця t , яка є моделлю множини $F \cup G$, знаходиться у 4НФ і нехай $X \rightarrow \{A\}$ – довільна нетривіальна ФЗ (тобто, $A \in R \setminus X$) така, що $F \cup G \models X \rightarrow \{A\}$. Покажемо, що множина атрибутів X є суперключем. Розглянемо випадки:

а) $X \cup \{A\} \subset R$; оскільки, за припущенням, $F \cup G \models X \rightarrow \{A\}$, то за критерієм повноти аксіоматики БЗЗ і ФЗ [10] для випадку $|R| \geq 2$ маємо: $F \cup G \models \neg X \rightarrow \{A\}$; тоді за правилом продовження ФЗ до БЗЗ (див., наприклад, [2]) на таблиці t виконується БЗЗ $X \rightarrow \neg \{A\}$, яка є нетривіальною (нагадаємо, що $A \in R \setminus X$ і $X \cup \{A\} \subset R$). За припущенням таблиця t знаходиться у 4НФ, тобто виконується умова (4) і множина атрибутів X є суперключем. Отже, імплікація з умови (3)

$$\begin{aligned} & \forall X, \forall A \\ & (X \subset R \wedge A \in R \setminus X \wedge (X \rightarrow \{A\})(t) = \text{true} \Rightarrow \\ & \Rightarrow (X \rightarrow R)(t) = \text{true}) \end{aligned}$$

є істинною і таблиця t знаходиться у НФБК; □

б) $X \cup \{A\} = R$; аналогічно до попереднього пункту для випадку, що розглядається ($|R| \geq 2$), маємо: $F \cup G \models \neg X \rightarrow \{A\}$; тоді за правилом поповнення (див., наприклад, [3]) на таблиці t виконується ФЗ $X \rightarrow X \cup \{A\}$, тобто ФЗ $X \rightarrow R$, звідси множина атрибутів X є суперключем. Отже, таблиця t , яка є моделлю множини $F \cup G$, знаходиться у НФБК. □

Теорема 1. Виконуються імплікації: $4НФ \Rightarrow НФБК \Rightarrow 3НФ \Rightarrow 2НФ$. □

Доведення тривіально випливає з леми 12 і наслідку 6. □

Висновки

1. Таблиця $t(R)$, яка знаходиться у 1НФ і є моделлю множини ФЗ F , знаходиться у 2НФ, 3НФ і НФБК, якщо $|R| \leq 2$ або $|D| \leq 1$.

2. Сформульовані означення 2-4 НФ задовольняють принципу «включення» НФ вищих порядків в НФ нижчих порядків, тобто виконуються імплікації:

$$4НФ \Rightarrow НФБК \Rightarrow 3НФ \Rightarrow 2НФ.$$

3. Подальша робота полягає у побудові фрагменту теорії нормалізації стосовно НФ вищих порядків (зокрема, проективно-з'єднувальної НФ), який би відповідав вимогам строгості для математичних теорій.

Список використаних джерел

1. Буй Д. Б. Теорія нормалізації в реляційних базах даних (огляд) / Д. Б. Буй, А. В. Пузікова // Радіоелектронні і комп'ютерні системи. – 2014. – №5(69). – С. 45-49.
2. Редько В.Н. Реляційні бази даних: табличні алгебри та SQL-подібні мови / В.Н. Редько, Ю.Й. Брона, Д.Б. Буй, С.А. Поляков. – Київ: Видавничий дім “Академперіодика”, 2001. – 198 с.
3. Буй Д. Б. Повнота аксіоматики Армстронга / Д. Б. Буй, А. В. Пузікова // Вісник Київського національного університету ім. Т. Шевченка. Серія: фіз.-мат. науки. – 2011. – № 3 – С. 103-108.

References

1. BUI, D. B. and PUZIKOVA, A. V. (2014) Theory of Normalization in Relational Databases (A Survey). *Radioelektronni i Komp'yuterni Systemy*. 5(69). p. 45-49.
2. REDKO, V. N. et al. (2001) *Reliatsiini Bazy Danykh: Tablychni Alhebry Ta SQL-Podibni Movy*. Kyiv: Akademperiodyka.
3. BUI, D. B. and PUZIKOVA, A. V. (2011) Completeness of Armstrong's axiomatic. *Bulletin of Taras Shevchenko National University of Kyiv. Series Physics & Mathematics*. 3, 103-108.

4. Буй Д. Б. Критерій повноти аксіоматики Армстронга / Д. Б. Буй, А. В. Пузікова // Матеріали міжнародної конференції «Теоретичні та прикладні аспекти побудови програмних систем» – ТАAPSD'2011 (Україна, Ялта, 19-23 вересня 2011 року). – С. 30-34.
5. Мейер Д. Теория реляционных баз данных / Д. Мейер. – Москва: Мир, 1987. – 608 с.
6. Дрибас В. П. Реляционные модели баз данных / В. П. Дрибас. – Минск: Изд-во БГУ им. В. И. Ленина, 1982. – 192 с.
7. Дейт К. Дж. Введение в системы баз данных, 8-е издание.: Пер. с англ. – М.: Вильямс, 2005. – 1328 с.
8. Исаченко А. Н. Модели данных и СУБД / А. Н. Исаченко, С. П. Бондаренко. – Минск: БГУ, 2007. – 205 с.
9. Fagin R. A Normal Form for Relational Databases That Is Based on Domains and Keys / R. Fagin // Communications of the ACM. – 1981. – Vol. 6. – P. 387-415.
10. Буй Д. Б. Аксіоматика багатозначних залежностей табличних баз даних: повнота та її критерій / Д. Б. Буй, А. В. Пузікова // Матеріали міжнародної наукової конференції «Теоретичні та прикладні аспекти побудови програмних систем» – ТАAPSD'2014 (Україна, Київ, 15-17 грудня 2014 року). – С. 35-43.
4. BUI, D. B. and PUZIKOVA, A. V. (2013) Kryteriy Povnoty Aksiomatyky Armstronha. In *Theoretical and Applied Aspects of Program Systems Development Conference*, Monday 19th to Friday 23d September 2011. Yalta. pp. 22-23.
5. MAIER, D. (1983) *The theory of relational databases*. Computer Science Press.
6. DRIBAS, V. P. (1982) *Relyacionnye modeli bas dannyh*. Minsk: Izdatelskiy centr Belorusskiy universytet.
7. DATE, C. J. (2005) *An Introduction to Database Systems*. 8th Ed. Addison-Wesley.
8. YSACHENKO A. N. and BONDARENKO S. P. (2007) *Modeli danykh i SUBD*. Mins'k: BHU.
9. FAGIN, R. (1981) A Normal Form for Relational Databases That Is Based on Domains and Keys. *Communications of the ACM*. 6. p. 387-415.
10. BUI, D. B. and PUZIKOVA, A. V. (2014) Aksiomatyka bahatoznachnykh zalezhnostey tablychnykh baz danykh: povnota ta yiyi kryteriy. In *Theoretical and Applied Aspects of Program Systems Development Conference*, Monday 15th to Wednesday 17th December 2014. Kyiv. pp. 35-43.

Надійшла до редколегії 12.06.2015

УДК 681.3.062

Вознюк Т.Г., аспірант.

**Побудова класифікатора для
вирішення займенникової анафори на
основі тензорної моделі**

Київський національний університет імені
Тараса Шевченка, 83000, м. Київ, пр-т.
Глушкова 4д,
e-mail: taarraas@gmail.com

T.G. Vozniuk, Postgraduate Student.

**Building a classifier to solve pronominal
anaphora based on the Tensor model**

Taras Shevchenko National University of Kyiv,
83000, Kyiv, Glushkova st., 4d,
e-mail: taarraas@gmail.com

В роботі запропоновано підхід до вирішення займенникової анафори за допомогою побудови класифікатора, який на основі даних про синтаксичні та семантичні властивості слів анафори та антецедента робить висновок про їх сумісність чи не сумісність.

Ключові слова: метод опорних векторів, комп'ютерна лінгвістика, керуючий простір, проблема анафори.

The paper presents an approach for solving the pronominal anaphora. A short description of existing approaches to resolve the anaphora using classifiers and semantic knowledge was given. There was formulated twenty five heuristic syntactic criteria. Some of them take into the account the morphological properties of the anaphora and the antecedent. Other are based on the distance in words in general and distance in specific part of speech between the two. Criteria based on evaluation of the compatibility by the method of support vectors machine to the already recognized anaphora requires modification to training algorithm, because in this case training process can change training data. Eight criteria were obtained by modifying the estimation algorithm based on the Tensor Model of natural language texts. The basis of the syntactic component analysis with tensor model-based algorithm is control spaces of natural language sentences. Semantic analysis was done using the semantic ontology WordNet and the factorized tensor of typical control spaces. Proposed classifier builds hyperplane in thirty-three dimensional space between the points of matching and not matching pairs of anaphora. Problem of finding this hyperplane has been reduced to an optimization problem. Its dual problem can be solved by gradient descent.

Keywords: support vector machine, computational linguistics, control space, anaphora resolution.

Статтю представив д.ф.-м.н., проф. Анісімов А.В.

В наш час інформаційні технології все більше проникають в повсякденне життя людей. Більшість людей які користуються даними технологіями не є фахівцями в цій галузі, а є пересічними людьми різних професій. Тому інтерфейси програмних додатків стають все більш простими. Один з найбільш зрозумілих інтерфейсів для людини – це засоби спілкування природньою мовою. З одного боку це розробки, що дозволяють синтезувати звук з тексту, тобто читати текст в голос (також розвиваються продукти, що займаються протилежною задачею – розпізнавання звуків природньої мови, тобто «слухати» людей), з іншого боку необхідні засоби аналізу та синтезу природньомовних текстів. Останні дослідження пов'язані з

синтаксичним аналізом текстів показують високі результати. Тоді як семантичний аналіз текстів містить в собі багато не вирішених проблем. Одна з них – розв'язання анафори. Можна виділити наступні типи анафори: займенникова, іменникова, прислівна та нульова. Метою даної статті є покращення результатів вирішення займенникової анафори.

Серед робіт присвячених розрізненню займенникової анафори слід виділити алгоритм Міткова[1]. Представлена в його роботі модель узагальнює багато існуючих на той час евристик в одній зручній формі. А саме виділення деяких критеріїв оцінювання антецедентів по відношенню до анафори на основі синтаксичної та морфологічної інформації, надання кожному

критерію деякої ваги і сумування ваг критеріїв, яким задовільняє дана пара.

В роботі [2] представлений алгоритм, що дозволяє покращити роботу алгоритму Міткова за допомогою використання семантичної інформації.

В роботі [3] описаний критерій для алгоритму Міткова заснований на використанні семантичної інформації представленої тензорною моделлю.

В роботі [4] було побудовано класифікатор, який на основі синтаксичної інформації дає відповідь про сумісність чи несумісність пари антецедент-анафора.

Дана робота являє собою синтез підходів описаних в [3] та [4], тобто буде побудовано класифікатор, що враховує семантичну інформацію.

Критерії оцінювання на основі синтаксичної інформації

Серед критеріїв оцінювання можна виділити ті, що приймають бінарне, натуральне та дійсне значення.

Бінарні критерії кодуються стандартно: 0 – критерій не виконується, 1 – критерій виконується. Наведемо перелік використаних бінарних критеріїв:

1. Антецедент є підметом речення.
2. Антецедент є власною назвою.
- 3-11. Бінарні ознаки чоловічого, жіночого, середнього роду, множини анафори та антецеденту.

Розглянемо критерії, що представлені натуральними числами.

12. Кількість входжень антецедента в текст.
13. Кількість входжень антецедента в дане речення.
14. Кількість використань антецедента в анафоричних зв'язках.
15. Кількість власних назв між анафорою та антецедентом.
16. Кількість іменників між анафорою та антецедентом.
17. Кількість займенників між анафорою та антецедентом.
18. Кількість речень між анафорою та антецедентом.
19. Порядковий номер серед іменників антецедента в реченні.
20. Порядковий номер серед займенників анафори в реченні.

21. Кількість анафор, для яких було раніше визначена відповідність даному антецеденту на проміжку між антецедентом та анафорою.

22. Кількість анафор, для яких було раніше визначена не відповідність даному антецеденту на проміжку між антецедентом та анафорою.

Найскладнішими є критерії, значення яких представлене дійсним числом.

23. Кандидат входить в альфа-зв'язок від слова, що знаходиться в одній і тій же фактор-множині як і слова discuss, illustrate, identify тощо. Для перевірки критерію необхідно взяти значення з матриці альфа-бета зв'язків по парі слів (identify, антецедент). Також треба зробити запит до таблиці трійок за трійкою (антецедент, identify, -). Вибір іншого слова (illustrate чи discuss) не суттєво вплине на результати, оскільки всі вони будуть знаходитися в одній фактор-множині.

24. Середнє значення оцінки пари антецедент-анафора методом опорних векторів, для яких було раніше визначена відповідність на проміжку між антецедентом та анафорою.

25. Середнє значення оцінки пари антецедент-анафора методом опорних векторів, для яких було раніше визначена не відповідність на проміжку між антецедентом та анафорою.

Критерії 21, 22, 24 та 25 можуть бути введені в процес тренування тільки після початкового тренування на всіх інших критеріях, оскільки вони вимагають вже існуючого класифікатора. Крім того після кожного перетренування з врахуванням цих додаткових критеріїв вид класифікатора змінюється. Тому перерахунок результатів цих критеріїв необхідно включити в процес тренування.

Модифікація тензорного алгоритму оцінювання пари антецедент-анафора

Нехай дано слово чи словосполучення антецеденту w та речення v яке воно входить s . Також маємо анафоричний займенник w' та речення v' яке він входить s' .

Сформулюємо алгоритм підрахунку семантико-синтаксичних критеріїв для заданої пари w та w' .

- I. Побудувати керуючі простори cs та cs' для речень s та s' за допомогою алгоритму запропонованого в [5].
- II. Визначити двійку two та трійку $three$, якщо вони існують, такі що :
 - a) two in cs , $three$ in cs
 - б) w входить як компонент в two і $three$
- III. Визначити двійку two' та трійку $three'$, якщо вони існують, такі що :
 - a) two' in cs' , $three'$ in cs'
 - б) w' входить як компонент в two' і $three'$
- IV. Обчислити бінарний критерій:

$$\begin{cases} 1, & \text{якщо } two.type = two'.type \\ 0, & \text{інакше} \end{cases}$$
- V. Обрахувати натуральнозначні критерії на основі отриманої інформації та знаходження семантичної відстані між словами по онтології WordNet[6].
 1. $two.first$ та $two'.first$
 2. $two.second$ та $two'.second$
 3. $three.first$ та $three'.first$
 4. $three.second$ та $three'.second$
 5. $three.third$ та $three'.third$
- VI. Знайти дійснозначні критерії вживаності антецедента в контексті анафори за допомогою запитів до тензорної моделі за наступними векторами:
 1. Якщо $two'.first = w'$,
то $(w, two'.second)$
інакше $(two'.first, w)$
 2. Якщо $three'.first = w'$
то $(w, three'.second, three'.third)$
інакше якщо $three'.two = w'$
то $(three.first, w, three'.third)$
інакше $(three'.first, three'.second, w)$

Критерій IV є аналогом критерію синтаксичного паралелізму, оскільки він перевіряє, чи в однаковий тип зв'язку входить антецедент та анафора.

Критерії V.1-5 відповідають за синтаксично-семантичний паралелізм та перевіряють чи схожі за змістом слова знаходяться в однакових синтаксичних відношеннях відносно анафори та антецедента.

За допомогою критеріїв VI.1-2 за допомогою тензорної моделі мови визначається, чи можна вжити антецедент в реченні замість анафори. Якщо при заміні анафори антецедента в парі та трійці утворюються семантично не коректні фрази, то така пара повинна бути відкинута.

Таким чином ми сформували 8 критеріїв на основі аналізу контексту антецедента та анафори за допомогою тензорної моделі мови

Тренування класифікатора методу опорних векторів

Скористаємося навчанням з вчителем (supervised learning). Для даного типу навчання нам необхідний розмічена тренувальна вибірка. На основі тренувальної вибірки треба сформувати позитивні та негативні приклади для навчання. Для цього обрахуємо сформульовані вище ознаки всіх пар антецедентів та анафор. Кількість пар в негативних прикладах буде близькою до $p \cdot m$, де p – кількість анафоричних займенників в тренувальній вибірці, m – кількість кандидатів в антецеденти. Тому кількість негативних ознак буде значно більше ніж позитивних. Для пришвидшення навчання можна зменшити кількість негативних прикладів, викинувши з тренуючої вибірки пари негативні пари антецедент-анафора, відстань між якими більша ніж деяке порогове значення, вибране виходячи з ступеня зв'язності тренуючої вибірки. Маємо:

$$D = \{(x_i, y_i) \mid x_i \in \mathcal{R}^{33}, y_i \in \{-1, 1\}^n\}_{i=1}$$

де x_i - i -та точка тренувальної вибірки в просторі ознак,

$$y_i = \begin{cases} 1, & \text{якщо } \text{точка представляє сумісну пару} \\ -1, & \text{інакше} \end{cases}$$

,
 p – кількість пар в тренувальній вибірці.

В лінійному випадку методу опорних векторів поверхня має вигляд гіпер-площини. Цю гіперплощину можна задати наступним рівнянням:

$$w \cdot x - b = 0,$$

де $\cdot$ – скалярний добуток.

Сучасні алгоритми штучного інтелекту не дійшли до повного розуміння тексту. Сформульований вище простір ознак не є виключенням. Тому деякі точки в просторі можуть вказувати на сумісну або не сумісну в деякому контексті пару. Це означає, що побудована для класифікатора поверхня не зможе розділити простір на 2 частини, в кожній з яких будуть точки свого класу. Додатковою проблемою нероздільності класів може бути не правильно обраний вид поверхні, що їх розділює. В нашому випадку – це гіперплощина. Для нероздільних класів використовується модифікація алгоритму опорних векторів з м'яким зазором. В цьому випадку тренування полягає в вирішенні наступної оптимізаційної задачі:

$$\arg \min_{w,b,\xi} \left(\frac{1}{2} \|w\|^2 + C \sum_{i=1}^n \xi_i \right)$$

$$y_i \cdot (w \cdot x_i - b) \geq 1 - \xi_i, 1 \leq i \leq n$$

$$\xi_i \geq 0$$

де C – лінійний штрафний коефіцієнт

Така задача може бути вирішена будь-яким методом оптимізації.

Висновки

У роботі було представлено метод вирішення проблеми анафори за допомогою тренування класифікатора. Було сформовано 33-вимірний простір ознак. Серед них, 8 ознак використовують семантико-синтаксичну інформацію, представлену тензорною моделлю мови.

Список використаних джерел

1. Mitkov R. A New, Fully Automatic Version of Mitkov's Knowledge-Poor Pronoun Resolution Method // *CICLing*. — 2002. — PP. 168—186.
2. Марченко О.О. Semantic Modification of the Mitkov Algorithm for Anaphora Resolution // *Штучний інтелект*. — 2012 — № 3. — С. 106—110.
3. Вознюк Т.Г. Застосування керуючого простору синтаксичних структур природномовних текстів для вирішення проблеми анафори // *Вісник Київського національного університету імені Тараса Шевченка серія фізико-математичні науки*. — 2014. — № 2. — С. 100—103.
4. Алгоритм автоматизованого разрешення анафоры местоимений третьего лица на основе методов машинного обучения [Электронный ресурс] / Толпегин П.В., Ветров Д.П., Кропотов Д.А. // *Диалог* — 2006 — Режим доступа : <http://www.dialog-21.ru/digests/dialog2006/materials/html/Tolpegin.htm> (20.05.14). — Загл. с экрана
5. Вознюк Т.Г. Алгоритм побудови керуючого простору синтаксичних структур природномовних текстів // *Вісник Київського національного університету імені Тараса Шевченка, серія фізико-математичні науки*. — 2014. — № 1 — С. 122—127.
6. Марченко О.О. Методи оцінювання семантичної близькості-зв'язності слів природної мови // *Штучний інтелект*. — 2012 — №4 — С. 213—219.

References

1. MITKOV, R. (2002) Fully Automatic Version of Mitkov's Knowledge-Poor Pronoun Resolution Method, In *CICLing*, 2002. pp. 168-186.
2. MARCHENKO, O.O (2002) Semantic Modification of the Mitkov Algorithm for Anaphora Resolution, *Artificial intelligence* #3, pp. 106-110.
3. VOZNIUK, T.G.(2014) Application of the control space of syntactic structures for anaphora resolution, *Bulletin of Taras Shevchenko National University of Kyiv Series Physics & Mathematics*, #2, pp. 100-103
4. TOLPEHIN, P.V & VETROV, D.P & KROPOTOV D.A.(2006) Algorithm for automatic pronominal anaphora resolution based on machine learning, *Dialog*, [Online] Available from <http://www.dialog-21.ru/digests/dialog2006/materials/html/Tolpegin.htm>.
5. VOZNIUK, T.G.(2014) Algorithm for construction of the control space of syntactic structures, *Bulletin of Taras Shevchenko National University of Kyiv Series Physics & Mathematics*, #1, pp. 122-127.
6. MARCHENKO O.O.(2012) Methods for estimation of the semantic distance of natural language words, *Artificial intelligence*, №4, pp. 213- 219.

Надійшла до редколегії 18.06.14

Галкін О. А., к.ф.-м.н., м.н.с.

**Дослідження процедур розпізнавання на
основі розподільного розташування
багатовимірних даних та функцій
глибини**

Київський національний університет імені
Тараса Шевченка, 03680, м. Київ, пр-т.
Академіка Глушкова 4д,
e-mail: galkin.o.a@gmail.com

O. A. Galkin, candidate of physical and mathematical
sciences, junior researcher

**The research of recognition procedures based
on the distributive location of
multidimensional data and depth functions**

Taras Shevchenko National University of Kyiv,
03680, Kyiv, Glushkova st., 4d,
e-mail: galkin.o.a@gmail.com

Стаття присвячена дослідженню непараметричних методів класифікації та їх асимптотичних властивостей на основі використання функцій напівпросторової та регресійної глибини, що використовуються для побудови лінійних та нелінійних розділових поверхонь серед конкуруючих множин даних. Запропоновано повністю незалежний від розподілу підхід, де для мінімізації коефіцієнтів помилкової класифікації застосовується розподільне розташування багатовимірних даних. Реалізовано оптимізаційний алгоритм для вирішення проблеми наявності декількох локальних мінімумів.

Ключові слова: функція глибини, розділова гіперплощина, байєсівський класифікатор.

Linear and quadratic classifiers are optimal Bayesian classifiers under the multivariate normal distribution based on simple parametric surfaces, and provide effective low-field class resolution. However, these methods are not reliable enough and sometimes are quite sensitive to extreme values and reductions in sample. In addition, these methods may not lead to correct classification if the data correspond to a certain distribution with heavy tails. This usually occurs when the assumption of multivariate normal distribution of data is violated. Most non-parametric methods do not use the information about the distribution of data and forms of dividing surface, but such flexible methods of statistical analysis for recognition problems sometimes tend to overfitting of training data and their performance can decrease significantly in the absence of large data sets. The article is devoted to research nonparametric classification methods and their asymptotic properties through the use of half-space and regression depth functions that are used to build linear and nonlinear separating surface between competing sets of data. A fully distribution free approach is proposed where distributive location of multidimensional data is used to minimize misclassification coefficients. An optimization algorithm was implemented for solving the problem of the presence of several local minima.

Key Words: depth function, separating hyperplane, Bayesian classifier.

Статтю представив д.ф.-м.н., проф. Анісімов А.В.

Вступ

Лінійні та квадратичні класифікатори на основі простих параметричних поверхонь є оптимальними байєсівськими класифікаторами в рамках багатовимірного нормального розподілу та забезпечують ефективне низькорозмірне поле класової роздільності. Однак, вказані класифікатори не є достатньо надійними та є досить чутливими до екстремальних значень та викидів, що наявні в вибірці даних. Це є наслідком того, що дані процедури потребують

оцінок проекційного вектора Φ , постійної величини γ та симетричної матриці M , використовуючи моменти першого та другого порядку вибірки. Крім того, дані методи можуть призвести до некоректної класифікації, якщо дані відповідають певному розподілу з важкими хвостами. Як правило, це має місце, коли порушується припущення багатовимірного нормального розподілу даних. Більшість непараметричних методів не використовують інформацію щодо розподілу даних та форми

розділової поверхні, проте такі гнучкі методи статистичного аналізу для задач розпізнавання іноді мають тенденцію перенавчання на навчальних даних, а їх продуктивність може суттєво знизитись при відсутності великих множин даних.

У даній статті досліджуються деякі непараметричні методи класифікації на основі концепцій функцій напівпросторової та регресійної глибини, що використовуються для побудови лінійних та нелінійних розділових поверхонь серед конкуруючих множин даних. Досліджувані інструменти методів статистичного аналізу для задач розпізнавання є повністю незалежними від розподілу, оскільки ґрунтуються на функціях напівпросторової та регресійної глибини. Для мінімізації коефіцієнтів помилкової класифікації, дані класифікатори застосовують розподільне розташування багатовимірних даних. Крім того, вони є незалежними від будь-яких моделей для основних розподілів множин даних.

Виклад основного матеріалу

У багатовимірному просторі функція напівпросторової глибини визначає занурення певного елементу даних відносно багатовимірного розподілу, або заданих багатовимірних даних. Концепція регресійної глибини є поняттям глибини регресійної відповідності. Найпростішими формами розділових поверхонь є гіперплощини, що забезпечують лінійну роздільність між класами [1].

Далі у статті досліджуються процедури побудови вказаних глибинних інструментів лінійної класифікації на основі заданої навчальної вибірки з двома класами.

Напівпросторова глибина r -вимірних даних z відносно багатовимірного розподілу F знаходиться, як мінімальна ймовірність замкнутого напівпростору, що містить z , а саме:

$$Hnp.Гл.(z, F) = \inf_S P_F\{S\}, \quad (1)$$

де S є замкнутим напівпростором, а $z \in S$.

Емпіричний варіант даної функції глибини отриманий шляхом заміни F на емпіричну функцію розподілу F_m . Емпірична версія функції напівпросторової глибини, що є афінно-інваріантною, рівномірно збігається до функції глибини, коли F є неперервною функцією.

Розглянемо двокласову задачу з одномірними даними. Можна очікувати, що більшість різниць

$z_{1i} - z_{2j}$, де z_{1i} та z_{2j} належать до двох різних класів для $1 \leq i \leq m_1$ та $1 \leq j \leq m_2$, будуть мати однаковий знак у випадку, якщо класи є ефективно розділеними. Такий підхід може бути розвинутий до багатовимірних випадків, де можна отримати лінійну проекцію $\phi'z$, для якої більшість різниць $\phi'z_{1i} - \phi'z_{2j}$ будуть мати однаковий знак. Це має місце лише у випадку, коли два класи можуть бути ефективно розділеними за допомогою лінійної роздільної функції [2].

Ми будемо оцінювати ϕ , мінімізуючи

$$\Phi_m(\phi) = \frac{1}{m_1 m_2} \sum_{i=1}^{m_1} \sum_{j=1}^{m_2} \Lambda\{\phi'(z_{1i} - z_{2j}) > 0\}, \quad (2)$$

де $m = (m_1, m_2)$ є вектором розмірів вибірки для двох класів. Задача максимізації може бути обмежена на множині $\{\phi: \|\phi\|=1\}$, що є задачею максимізації на скінченній вибірці. Зауважимо, що оцінена лінійна проекція є ортогональною до гіперплощини, що визначає напівпросторову глибину координат відносно даних, утворених різницями $z_{1i} - z_{2j}$ в r -вимірному просторі. Дана узагальнена незміщена статистика $\Phi_m(\phi)$ є мірою лінійної роздільності між двома класами вздовж напрямку ϕ , а її максимальне значення серед різних можливих значень ϕ може розглядатися, як багатовимірна модель одномірної незміщеної статистики Манна-Уїтні.

Максимізуюча оцінка $\Phi_m(\phi)$, позначена як $\bar{\phi}_{nnp}$ може бути використана для побудови лінійного класифікатора $\bar{\phi}'_{nnp}z + \gamma = 0$ для деякої відповідно вибраної постійної величини γ . Слід зауважити, що правило класифікації, а отже і відповідні ймовірності помилкової класифікації залежать від вибору цієї постійної величини.

Після знаходження оцінки $\bar{\phi}_{nnp}$, $\bar{\gamma}_{nnp}$ може бути отримана шляхом мінімізації відносно γ середньої помилки помилкової класифікації $\Psi_{nnp}(\bar{\phi}, \gamma)$, що задається виразом

$$\Psi_{nnp}(\bar{\phi}, \gamma) = \frac{p_1}{m_1} \sum_{i=1}^{m_1} \Lambda\{\bar{\phi}'_i z_{1i} + \gamma < 0\} + \frac{p_2}{m_2} \sum_{i=1}^{m_2} \Lambda\{\bar{\phi}'_{nnp} z_{2i} + \gamma > 0\}, \quad (3)$$

де p_1 та p_2 є апіорними ймовірностями для двох класів.

Функція регресійної глибини визначає глибину занурення даних, що визначається за

допомогою вектора коефіцієнтів $\varphi_+ = (\varphi, \gamma) \in R^{r+1}$ в лінійній регресійній структурі. Маючи множину даних

$$\Delta_m = [\{z_i = (z_{i1}, z_{i2}, z_{ir}), y_i\}; i = 1, 2, \dots, m]$$

з m елементів даних, величина φ_+ буде незануренням до Δ_m тоді і тільки тоді, якщо існує така афінна гіперплощина S в z -просторі, що жоден z_i не належить до S , а остачі $y_i - \varphi'_+(z_i, 1)$ є всі додатні в одному відкритому напівпросторі (тобто, одна сторона S) в z -просторі та всі від'ємні в доповненні відкритого напівпростору (тобто, інша сторона S). Функція регресійної глибини занурення φ_+ визначається, як мінімальне число елементів даних, що повинні бути видалені для перетворення її в незанурення [3].

Такий підхід регресійної глибини був використаний для двокласових задач в контексті бінарної регресії для побудови лінійних класифікаторів [4]. Використовуючи мітки класу 0 або 1, як значення залежної змінної y та досліджуючи незанурення $\varphi_+ = (0, 0, \dots, 0, 0.7)$, величина φ_+ буде незануренням до Δ_m тоді і тільки тоді, якщо існує гіперплощина S в z -просторі, що повністю розділяє дані з двох класів. Як результат, регресійна глибина незанурення φ_+ може розглядатися, як мінімальне число помилково класифікованих даних, що може бути досягнуто за допомогою розділової гіперплощини S в z -просторі.

Оскільки в спеціалізованій літературі було досліджено лише задачу визначення розділової гіперплощини шляхом мінімізації загального числа помилково класифікованих даних [5], можна стверджувати, що лінійний класифікатор є емпірично оптимальним, коли два конкуруючих класи мають апіорні ймовірності, що пропорційні їх розмірам вибірки. На практиці необхідно коректно налаштувати ваги для різних даних та визначити виважену регресійну глибину занурення φ_+ , як мінімальне число ваг, що повинні бути видалені для перетворення її в незанурення. Тоді, виважена регресійна глибина в кінцевому підсумку стає середньою ймовірністю помилкової класифікації, що має наступний вигляд:

$$\Psi_m(\varphi, \gamma) = \frac{p_1}{m_1} \sum_{i=1}^{m_1} \Lambda\{\varphi' z_{1i} + \gamma < 0\} + \frac{p_2}{m_2} \sum_{i=1}^{m_2} \Lambda\{\varphi' z_{2i} + \gamma > 0\}. \quad (4)$$

Мінімізація $\Psi_m(\varphi, \gamma)$ відносно φ та γ дозволяє отримати оцінки $\bar{\varphi}_{рег}$ та $\bar{\gamma}_{рег}$, що визначають розділову гіперплощину для використання в задачах класифікації. Задача мінімізації може бути обмежена до $\{(\varphi, \gamma) : \|(\varphi, \gamma)\| = 1\}$. Також можна перевірити, що мінімізація $\Psi_m(\varphi, \gamma)$ виявляється оптимізаційною задачею на скінченній множині.

Крім того, було досліджено той факт, що оцінка максимальної ймовірності в задачі логістичної регресії існує лише тоді, коли є деякі збіги в коваріативному просторі між даними з двох класів, що відповідають значенням 0 та 1 залежної змінної. У випадках абсолютної роздільності не існує скінченної оцінки максимальної ймовірності для вектора регресійних коефіцієнтів φ_+ . Якщо дані з двох класів є повністю розділними, можна побачити, що $(\bar{\varphi}_{рег}, \bar{\gamma}_{рег})$ є мінімізуючою оцінкою $\Psi_m(\varphi, \gamma)$ тоді і тільки тоді, коли $\bar{\varphi}_{рег}$ максимізує $\Phi_m(\varphi)$, а тому $\bar{\varphi}_{рег}$ є також $\bar{\varphi}_{нпр}$.

У випадку, коли межі класу є досить складними, лінійні класифікатори можуть бути недоцільними для використання. В такому разі, для розрізнення між конкуруючими класами необхідна залежність від нелінійних розділових поверхонь. Для побудови таких поверхонь, можна спроектувати дані $z_i = (z_{i1}, z_{i2}, \dots, z_{id})$ в простір більш високої розмірності для отримання нового вектора вимірюваних величин $t_i = (\omega_1(z_i), \omega_2(z_i), \dots, \omega_r(z_i))$ та виконати лінійну класифікацію в l -вимірному просторі.

Якщо ми проектуємо дані в простір квадратичних функцій, це може розглядатися, як лінійна класифікація з $l = r + \binom{r}{2}$ вимірюваними величинами, які в кінцевому підсумку призводять до квадратичного розділення в початковому r -вимірному просторі. Величини

$$\Phi_m(\varphi) = \frac{1}{m_1 m_2} \sum_{i=1}^{m_1} \sum_{j=1}^{m_2} \Lambda\{\varphi'(t_{1i} - t_{2j}) > 0\} \quad (5)$$

та

$$\Psi_m(\varphi, \gamma) = \frac{p_1}{m_1} \sum_{i=1}^{m_1} \Lambda\{\varphi' t_{1i} + \gamma < 0\} + \frac{p_2}{m_2} \sum_{i=1}^{m_2} \Lambda\{\varphi' t_{2i} + \gamma > 0\} \quad (6)$$

можуть бути оптимізовані для отримання відповідних оцінок φ та γ , які

використовуються для формування розділової поверхні в двокласовій задачі.

Традиційні лінійні та квадратичні методи статистичного аналізу для задач розпізнавання головним чином мотивовані багатовимірними нормальними розподілами. В двокласовій задачі лінійна розділова функція на основі моменту тісно пов'язана з відстанню Махаланобіса, що є досить чутливою до можливих викидів, присутніх в даних. З іншої сторони, незалежні від розподілу глибинні класифікатори є досить стійкими по відношенню до викидів [6].

Теорема 1. Припустимо, що $h_i(z) = c(z - \varepsilon_i)$, де $i = (1, 2)$ для деяких параметрів локалізації ε_i та загальної еліптично-симетричної функції щільності c , що задовольняє умову $c(kz) \geq c(z)$ для кожного z та $0 < k < 1$. У даному випадку, функції щільності h_1 та h_2 двох конкуруючих класів є еліптично-симетричними із загальною матрицею розсіювання Ξ .

Припускаючи, що апіорні ймовірності двох конкуруючих класів даних є рівними, ми можемо стверджувати, що:

(а) $\varphi = \Xi^{-1}(\varepsilon_1 - \varepsilon_2)$ є максимізуючою оцінкою $\Phi(\varphi)$, а також мінімальною оцінкою $\Psi(\varphi, \gamma)$ при належному виборі значення γ ;

(б) існує оптимальний лінійний байєсівський класифікатор.

Доведення. (а) Характеристичними функціями розподілу є функції наступного вигляду:

$$O_{h_1}(\kappa) = e^{i\kappa'\varepsilon_1} \alpha(\kappa'\Xi\kappa) \quad (7)$$

та

$$O_{h_2}(\kappa) = e^{i\kappa'\varepsilon_2} \alpha(\kappa'\Xi\kappa) \quad (8)$$

оскільки розподіли мають загальну матрицю розсіювання Ξ та спільну еліптично-симетричну форму з параметрами локалізації ε_1 та ε_2 , а α є деякою загальною скалярною функцією.

Визначаючи

$$X = \varphi' \{ (Z_1 - Z_2) - (\varepsilon_1 - \varepsilon_2) \} / (\varphi'\Xi\varphi)^{1/2}, \quad (9)$$

де $Z_1 \approx h_1$ та $Z_2 \approx h_2$, можна побачити, що характеристична функція X задається, як $O_X(\kappa) = \{ \alpha(\kappa^2) \}^2$. Оскільки розподіл X є симетричним відносно нуля та незалежний від таких параметрів множини, як ε та Ξ , $P\{\varphi'(Z_1 - Z_2) > 0\}$ може бути виражена, як

$$P\{\varphi'(Z_1 - Z_2) > 0\} = F_X \left(\left[\frac{\{\varphi'(\varepsilon_1 - \varepsilon_2)\}^2}{\varphi'\Xi\varphi} \right]^{1/2} \right) \quad (9)$$

де F_X є кумулятивною функцією розподілу X . В результаті, $P\{\varphi'(Z_1 - Z_2) > 0\}$ матиме максимальне значення для деякого φ , якщо дане φ максимізує $\{\varphi'(\varepsilon_1 - \varepsilon_2)\}^2 / \varphi'\Xi\varphi$. Це означає, що $\varphi = \Xi^{-1}(\varepsilon_1 - \varepsilon_2)$ є максимізуючою оцінкою $\Phi(\varphi)$.

(б) Функції щільності h_1 та h_2 можуть бути виражені, як

$$h_1(z) = V_r |\Xi|^{-1/2} \varpi \{ (z - \varepsilon_1)' \Xi^{-1} (z - \varepsilon_1) \} \quad (10)$$

та

$$h_2(z) = V_r |\Xi|^{-1/2} \varpi \{ (z - \varepsilon_2)' \Xi^{-1} (z - \varepsilon_2) \}, \quad (11)$$

що є результатом еліптичної симетрії з локалізаційним зсувом, де ε є монотонно спадною функцією на $[0, \infty)$, а V_r - постійною величиною, що залежить від розмірності r . Оптимальне байєсівське правило класифікує дані до пергошо класу тоді і тільки тоді, якщо

$$\begin{aligned} h_1(z) \geq h_2(z) &\Leftrightarrow (z - \varepsilon_1)' \Xi^{-1} (z - \varepsilon_1) \leq \\ &\leq (z - \varepsilon_2)' \Xi^{-1} (z - \varepsilon_2) \Leftrightarrow \\ &\Leftrightarrow (\varepsilon_1 - \varepsilon_2)' \Xi^{-1} z \geq \frac{1}{2} [\varepsilon_1' \Xi^{-1} \varepsilon_1 - \varepsilon_2' \Xi^{-1} \varepsilon_2]. \end{aligned} \quad (12)$$

Даний результат має місце у випадку рівних апіорних ймовірностей. Отже, оптимальний лінійний класифікатор є байєсівським класифікатором, а $\varphi = \Xi^{-1}(\varepsilon_1 - \varepsilon_2)$ - мінімізуючою оцінкою $\Psi(\varphi, \gamma)$ при належному виборі значення γ .

Теорема 2. Нехай щільності h_1 та h_2 належать до класу з еліптично-симетричним багатовимірним нормальним розподілом, а також вони мають однакову форму, за винятком, можливо, їх параметрів локалізації та розкиду. Тоді, середній коефіцієнт помилкової класифікації квадратичного класифікатора, побудованого з використанням регресійної глибини збігається до оптимальної байєсівської помилки, а квадратичний класифікатор збігається майже напевно до оптимального байєсівського класифікатора, коли розмір навчальної вибірки прямує до нескінченності.

Доведення. У випадку, коли два конкуруючих розподіли множин даних є багатовимірними нормальними розподілами з параметрами локалізації та розсіювання (ε_1, Ξ_1) та (ε_2, Ξ_2) ,

$$\begin{aligned}
 p_1 h_1(z) &> p_2 h_2(z) \Leftrightarrow \\
 \Leftrightarrow p_1 |\Xi_1|^{-1/2} e^{-\frac{1}{2}(z-\varepsilon_1)' \Xi_1^{-1} (z-\varepsilon_1)} &> \\
 > p_2 |\Xi_2|^{-1/2} e^{-\frac{1}{2}(z-\varepsilon_2)' \Xi_2^{-1} (z-\varepsilon_2)} &\Leftrightarrow \\
 \Leftrightarrow (z-\varepsilon_2)' \Xi_2^{-1} (z-\varepsilon_2) - (z-\varepsilon_1)' \Xi_1^{-1} (z-\varepsilon_1) &> V,
 \end{aligned} \quad (13)$$

де $V = 2 \log \left(\frac{p_2 |\Xi_1|^{1/2}}{p_1 |\Xi_2|^{1/2}} \right)$.

Оскільки достатньо показати, що оптимальний квадратичний класифікатор є унікальним байєсівським класифікатором за заданих умов, ми отримали, що оптимальне байєсівське правило є унікальним та квадратичним [7].

Ймовірнісна функція щільності $h(z)$ r -вимірного еліптично-симетричного багатомірного нормального розподілу задається, як

$$h(z) = V_r |\Xi|^{-1/2} \{1 + \mu^{-1}(z - \varepsilon)' \Xi^{-1} (z - \varepsilon)\}^{-U}, \quad (14)$$

де $\mu > 0$, $U > r/2$, $V_r = (p\mu)^{-r/2} \Gamma(U) / \Gamma(U - r/2)$, а ε та Ξ є параметрами локалізації та розсіювання, відповідно. Далі проведемо дослідження двох однакової форми еліптично-симетричних багатомірних нормальних розподілів за винятком їх параметрів локалізації та розсіювання.

Припустимо, що ε_i є параметром локалізації, а Ξ_i - матрицею розсіювання для i -ї ($i=1,2$) множини даних, а p_i є її апіорною ймовірністю. Тоді, має місце наступна нерівність:

$$\begin{aligned}
 p_1 h_1(z) &> p_2 h_2(z) \Leftrightarrow \\
 \Leftrightarrow p_1 |\Xi_1|^{-1/2} \{1 + \mu^{-1}(z - \varepsilon_1)' \Xi_1^{-1} (z - \varepsilon_1)\}^{-U} &> \\
 > p_2 |\Xi_2|^{-1/2} \{1 + \mu^{-1}(z - \varepsilon_2)' \Xi_2^{-1} (z - \varepsilon_2)\}^{-U} & \\
 \Leftrightarrow \left\{ \frac{1 + \mu^{-1}(z - \varepsilon_1)' \Xi_1^{-1} (z - \varepsilon_1)}{1 + \mu^{-1}(z - \varepsilon_2)' \Xi_2^{-1} (z - \varepsilon_2)} \right\}^{-U} &> n \text{ для } n = \frac{p_2 |\Xi_2|^{-1/2}}{p_1 |\Xi_1|^{-1/2}} \quad (15) \\
 \Leftrightarrow \left\{ \frac{\mu + (z - \varepsilon_1)' \Xi_1^{-1} (z - \varepsilon_1)}{\mu + (z - \varepsilon_2)' \Xi_2^{-1} (z - \varepsilon_2)} \right\} < n_0 = n^{-1/U} & \\
 \Leftrightarrow (z - \varepsilon_1)' \Xi_1^{-1} (z - \varepsilon_1) - n_0 (z - \varepsilon_2)' \Xi_2^{-1} (z - \varepsilon_2) - (n_0 - 1)\mu < 0. &
 \end{aligned}$$

В результаті, оптимальне байєсівське правило має властивість унікальності та є квадратичним класифікатором, оскільки ліва частина цієї нерівності є квадратичною функцією від z .

Зауважимо, що ймовірнісна функція щільності r -вимірного еліптично-симетричного

багатомірного нормального розподілу задається, як

$$\begin{aligned}
 h(z) &= (p\mu)^{-r/2} \Gamma(U) / \Gamma(U - r/2) |\Xi|^{-1/2} \times \\
 &\times \{1 + \mu^{-1}(z - \varepsilon)' \Xi^{-1} (z - \varepsilon)\}^{-U}, \quad (16)
 \end{aligned}$$

де ε та Ξ є параметрами локалізації та розкиду, $\mu > 0$, а $U > r/2$. Коли $U = (\mu + r)/2$, а μ є цілим числом, відповідний розподіл є також багатомірним t -розподілом з функцією щільності μ . Як частковий випадок, має місце багатомірний розподіл Коші, коли $\mu = 1$.

Результати практичних експериментів

Максимізація $\Phi_m(\varphi)$ може бути зведена до φ

з $\|\varphi\| = 1$, а мінімізація $\Psi(\varphi, \gamma)$ - до (φ, γ) з $\|(\varphi, \gamma)\| = 1$. Враховуючи те, що порядок обчислювальної складності швидко зростає з розмірністю l , точна оптимізація $\Phi_m(\varphi)$ та $\Psi_m(\varphi, \gamma)$ не може бути застосованою за задач великої розмірності, тому можна виконувати лише приблизну оптимізацію.

В практичних експериментах було застосовано підхід, де індикаторні функції, що фігурують у виразах для Φ_m та Ψ_m апроксимуються за допомогою відповідно визначених функцій згладжування. Використовуючи вказану апроксимацію, було використано похідні для визначення напрямку найшвидшого підйому, або спуску цільової функції, яку необхідно оптимізувати. Логістична функція $1/(1 + e^{-\kappa z})$ з великим додатнім κ є апроксимацією для індикаторної функції $\Lambda(z > 0)$.

Зрозуміло, що недостатньо велике значення κ призведе до неточної апроксимації. Однак, дуже велике значення κ призведе до достатньо точної апроксимації, однак числова оптимізація на основі найшвидшого підйому, або спуску буде чисельно нестійкою. Встановлено, що високий коефіцієнт чисельної стійкості в оптимізаційному алгоритмі може бути отриманий навіть для дуже великого значення t . Це стало можливим при умові, коли всі вимірювані величини були нормалізованими до процесу апроксимації. На основі практичних експериментів було встановлено, що якщо використовується значення $10 \leq \kappa \leq 15$ після процесу нормалізації вимірюваних величин, середні коефіцієнти помилкової класифікації для отриманих алгоритмів залишаються майже однаковими та достатньо малими. В результаті, були виведені

найбільш прийнятні значення, отримані в цьому діапазоні.

В двовимірному випадку для лінійних методів статистичного аналізу для задач розпізнавання, де точне визначення $\Phi_m(\varphi)$ та $\Psi_m(\varphi, \gamma)$ є досить простим, було порівняно ефективність точних та наближених глибинних методів класифікації, а також встановлено, що вони досягають практично однакових середніх коефіцієнтів помилкової класифікації.

Починаючи з різних випадкових початкових даних, наближені реалізації оптимізаційних алгоритмів виконувалися декілька разів для вирішення проблеми можливої наявності декількох локальних мінімумів.

Стосовно застосування класифікатора на основі напівпросторової глибини, необхідно було оцінити γ з навчальної вибірки після оцінки φ . Це було реалізовано в результаті перерахування порядкових статистик спроектованих даних $\bar{\varphi}'_{np} t_{1i}$ та $\bar{\varphi}'_{np} t_{2j}$ ($1 \leq i \leq m_1$, $1 \leq j \leq m_2$) уздовж оціненого напрямку $\bar{\varphi}_{np}$. Було зроблено висновок, що обчислювальна складність при отриманні оцінки $\bar{\gamma}_{np}$ не збільшується з розмірністю l у результаті використання лінійних проекцій.

Список використаних джерел

1. Mizera I. On depth and deep points: a calculus / I. Mizera // The Annals of Statistics. – 2002. – 30. – P. 1681-1736.
2. Mosler K. Multivariate Dispersions, Central Regions and Depth / K. Mosler. – New York: Springer-Verlag, 2002.
3. Serfling R. A depth function and a scale curve based on spatial quantiles / R. Serfling. – Statistics and Data Analysis based on L_1 -Norm and Related Methods. – 2002. – Birkhaeuser, Boston. – P. 25-38.
4. Vardi Y. The multivariate on L_1 -median and associated data depth / Y. Vardi, C.H. Zhang // Proceedings of the National Academy of Sciences (USA). – 2000. – 97. – P. 1423-1426.
5. Zuo Y. Structural properties and convergence results for contours of sample statistical depth functions / Y. Zuo, R. Serfling // The Annals of Statistics. – 2000. – 28. – P. 483-499.
6. Godtliebsen F. Significance in scale space for bivariate density estimation / F. Godtliebsen, J.S. Marron, P. Chaudhuri // Journal of Computational and Graphical Statistics. – 2002. – 11. – P. 1-22.
7. Holmes C.C. A probabilistic nearest neighbor method for statistical pattern recognition / C.C. Holmes, N.M. Adams // Journal of the Royal Statistical Society. – 2002. – 64. – P. 295-306.

References

1. MIZERA, I. (2002) On depth and deep points: a calculus. *The Annals of Statistics*. 30. p. 1681-1736.
2. MOSLER, K. (2002) *Multivariate Dispersions, Central Regions and Depth*. New York: Springer-Verlag.
3. SERFLING, R. (2002) *A depth function and a scale curve based on spatial quantiles*. Statistics and Data Analysis based on L_1 -Norm and Related Methods. Birkhaeuser, Boston. p. 25-38.
4. VARDI, Y. & ZHANG, C.H. (2000) The multivariate on L_1 -median and associated data depth. *Proceedings of the National Academy of Sciences (USA)*. 97. p.1423-1426.
5. ZUO, Y., SERFLING R. (2000) Structural properties and convergence results for contours of sample statistical depth functions. *The Annals of Statistics*. 28. p. 483-499.
6. GODTLIEBSEN, F., MARRON J.S., CHAUDHURI P. (2002) Significance in scale space for bivariate density estimation. *Journal of Computational and Graphical Statistics*. 11. p. 1-22.
7. HOLMES, C.C., ADAMS N.M. (2002) Significance in scale space for bivariate density estimation. *Journal of the Royal Statistical Society*. 64. p. 295-306.

Надійшла до редколегії 17.04.15

УДК 517.929.4

Гаркуша Н.І., к. е. н., пров.інж.

Отримання оцінки області стійкості положення рівноваги моделі популяції з післядією

Київський національний університет імені
Тараса Шевченка, 03680, м. Київ, пр.-т.
Глушкова, 4д,
e-mail: ngarkusha@gmail.com

N. I. Garkusha, PhD, lead engineer

Obtaining the estimation of the stability area of equilibrium in the model of population with aftereffect

Taras Shevchenko National University of Kyiv,
03680, Kyiv, Glushkova av., 4d,
e-mail: ngarkusha@gmail.com

У статті досліджуються положення рівноваги моделі популяції з запізненням. Розглядається система диференціальних рівнянь із запізненням. Дослідження динаміки проводиться з використанням лінеаризації системи в околиці ненульовий особливої точки. Показано, що особлива точка асимптотично стійка.

Ключові слова: математична модель, динамічна система, диференціальні рівняння, запізнювання.

The article investigates the equilibrium population models with delay. The system of differential equations with delay is investigated. Nonlinear right part is described by the variables that determine the dependence of the density of the populations of the victim and predator, respectively, rate of breeding populations of victims in the absence of predators, the specific rate of consumption of predator population offerings at unit density, natural mortality predator ratio processing consumed predator biomass victim of its own biomass and others. Two special points system are identified. The first point corresponds to the absence of populations. The second special point corresponds to sustainable regime. Research of dynamics is conducted using linearization system in a neighborhood of nonzero singular point. The characteristic equation are constructed. Showing that particular point asymptotically stable. Guaranteed region of asymptotic stability assessment equilibrium is computed using the second method of Lyapunov. Region asymptotic stability of equilibrium nonlinear system is calculated using a quadratic Lyapunov function constructed for the linear system. In system the delayed derivative is measured using the provided B.S. Razumyhn.

Key words: mathematical model, dynamic system, differential equations, delay.

Статтю представив д.ф.-м.н, проф. Хусаїнов Д. Я.

Ця стаття є продовженням досліджень динаміки поведінки моделі популяції з запізненням, розпочата в [5,6]. Дослідження з розробки математичних моделей динаміки популяцій проводилися в роботах Ферхюльста П., Лотки А. [1], Базикіна А.Д. [3], Вольтера В. [4], Сміта Дж. [5], Колмогорова А.Н. [10]. Було складено ряд математичних моделей взаємодіючих популяцій типу «хижак-жертва». Вони мали вигляд систем диференціальних рівнянь з нелінійною правою частиною зі змінними, визначальними залежністю щільності популяцій жертви і хижака, відповідно, швидкість розмноження популяції жертви у відсутності хижака, питому швидкість споживання популяцією хижака жертви при одиничній щільності, природну

смертність хижака, коефіцієнт переробки спожитої хижаком біомаси жертви у власну біомасу та ін.

Одна з моделей описувалася системою двох диференціальних рівнянь без запізнювання

$$\begin{aligned}\dot{x}(t) &= \left[\frac{ax(t)}{N + x(t)} - by(t) \right] x(t), \\ \dot{y}(t) &= -[c - dx(t)]y(t).\end{aligned}\quad (1)$$

Тут a – швидкість розмноження популяції жертви у відсутності хижака, b – питома швидкість споживання популяції хижака жертви при одиничній щільності, c – природна смертність хижака, d/b – коефіцієнт

переробки спожитої хижакіом біомаси жертви у власну біомасу, N – максимальну кількість популяції жертви.

У роботах [5,6] було запропоновано розглянути вплив запізнювання, обумовлений часом дозрівання особин. Диференціальні рівняння (1) в цьому випадку мали вигляд

$$\begin{aligned}\dot{x}(t) &= \left[\frac{ax(t-\tau)}{N+x(t-\tau)} - by(t-\tau) \right] x(t), \\ \dot{y}(t) &= -[c - dx(t-\tau)]y(t).\end{aligned}\quad (2)$$

Система рівнянь (2) має ті ж положення рівноваги, що і система без запізнювання (1). Дослідження динаміки проводилося з використанням лінійного наближення в околиці кожної з особливих точок. Особливих точок дві. Ними є

$$O(0,0), O_1(x_1, y_1), x_1 = \frac{c}{d}, y_1 = \frac{ac}{b(Nd+c)}.$$

Стационарні точки визначаються із системи

$$\left[\frac{ax(t)}{N+x(t)} - by(t) \right] x(t) = 0, -[c - dx(t)]y(t) = 0.$$

Ними будуть

$$O(0,0), M(x_1, y_1), x_1 = \frac{c}{d}, y_1 = \frac{ac}{b(Nd+c)}.$$

Перша особлива точка $O(0,0)$ не представляє інтересу. Вона відповідає відсутності популяції. Розглянемо другу особливу точку. Система лінійного наближення в околі цієї особливої точки має вигляд

$$\begin{aligned}\dot{z}(t) &= Az(t) + Bz(t-\tau), \\ A &= \begin{bmatrix} \frac{ax_1}{N+x_1} - by_1 & 0 \\ 0 & -c + dx_1 \end{bmatrix}, \\ B &= \begin{bmatrix} \frac{aNx_1}{(N+x_1)^2} & -bx_1 \\ dy_1 & 0 \end{bmatrix}, z(t) = \begin{pmatrix} x(t) \\ y(t) \end{pmatrix}, \\ z(t-\tau) &= \begin{pmatrix} x(t-\tau) \\ y(t-\tau) \end{pmatrix}.\end{aligned}\quad (3)$$

Підставивши значення $O_1(x_1, y_1)$, отримаємо

$$A = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}, B = \begin{bmatrix} b_{11} & b_{12} \\ b_{21} & 0 \end{bmatrix}, b_{11} = -\frac{acdN}{(Nd+c)^2},$$

$$b_{12} = -c, b_{21} = \frac{acd}{b(Nd+c)}.\quad (4)$$

Таким чином, система лінійного наближення в околиці особливої точки $O_1(x_1, y_1)$ є «системою з чистим запізненням». Її характеристичне рівняння має вигляд

$$\lambda^2 + \frac{ac}{dK} e^{-\lambda\tau} \lambda + \frac{ac(dK-c)}{dK} e^{-2\lambda\tau} = 0.$$

Умовами асимптотичної стійкості будуть [5]

$$-\frac{\pi}{2\tau} < s_1 < 0, -\frac{\pi}{2\tau} < s_2 < 0,\quad (5)$$

де

$$s_{1,2} = -\frac{ac}{2dK} \pm \sqrt{\left(\frac{ac}{2dK} + 1\right)^2 - (a+1)}$$

Оскільки вихідна система нелінійна, то асимптотична стійкість положення рівноваги є локальною, тобто в деякій околиці $U_\delta(O_1)$ точки $O_1(x_1, y_1)$.

Гарантовану оцінку області асимптотичної стійкості положення рівноваги $O_1(x_1, y_1)$ можна обчислити, використовуючи другий метод Ляпунова. Розглянемо функцію Ляпунова квадратичного виду

$$\begin{aligned}V(z_1(t)) &= z_1^T(t) H z_1(t), z_1(t) = \begin{pmatrix} x_1(t) \\ y_1(t) \end{pmatrix}, \\ H &= \begin{pmatrix} h_{11} & h_{12} \\ h_{21} & h_{22} \end{pmatrix}\end{aligned}\quad (5)$$

з позитивно визначеною матрицею H .

Функцію Ляпунова для нелінійної системи побудувати, в загальному випадку, важко. Як правило, будують функцію Ляпунова для лінеаризованої системи, а її повну похідну обчислюють в силу нелінійної системи. Для лінійних систем з запізненням необхідною умовою стійкості є стійкість при відсутності запізнювання, тобто системи

$$\dot{z}_1(t) = (A+B)z(t), A+B = \begin{bmatrix} b_{11} & b_{12} \\ b_{21} & 0 \end{bmatrix}.\quad (6)$$

Оскільки характеристичне рівняння системи без запізнювання має вигляд

$$\det(A + B - \lambda E) = \begin{vmatrix} b_{11} - \lambda & b_{12} \\ b_{21} & -\lambda \end{vmatrix},$$

$$= \lambda^2 - b_{11}\lambda - b_{12}b_{21} = 0$$

то, згідно з критерієм Гурвіца, для двовимірних систем необхідною і достатньою умовою стійкості є позитивність коефіцієнтів. Оскільки

$$-b_{11} = \frac{acdN}{(Nd + c)^2} > 0,$$

$$-b_{12}b_{21} = \frac{ac^2d}{b(Nd + c)} > 0,$$

то система без запізнювання (6) буде асимптотично стійкою. Розглянемо побудову функції Ляпунова (5) для лінеаризованої системи (6). Матричне рівняння Ляпунова має вигляд

$$\begin{bmatrix} b_{11} & b_{21} \\ b_{12} & 0 \end{bmatrix} \begin{bmatrix} h_{11} & h_{12} \\ h_{12} & h_{22} \end{bmatrix} + \begin{bmatrix} h_{11} & h_{12} \\ h_{12} & h_{22} \end{bmatrix} \begin{bmatrix} b_{11} & b_{12} \\ b_{21} & 0 \end{bmatrix}$$

$$= - \begin{bmatrix} c_{11} & c_{12} \\ c_{12} & c_{22} \end{bmatrix}$$

де $c_{11} > 0$, $c_{11}c_{12} - c_{12}^2 > 0$ - довільні числа.

Поклавши $c_{11} = c_{22} = 2c$, $c_{12} = 0$, отримаємо систему лінійних алгебраїчних рівнянь

$$b_{11}h_{11} + b_{12}h_{12} = -c, \quad b_{21}h_{11} + b_{11}h_{12} + b_{12}h_{22} = 0,$$

$$b_{21}h_{12} = -c.$$

Її рішенням буде

$$h_{11} = c \frac{b_{21} - b_{12}}{b_{11}b_{21}}, \quad h_{12} = -c \frac{1}{b_{21}},$$

$$h_{22} = c \frac{b_{11}^2 + b_{21}^2 - b_{12}b_{21}}{b_{11}b_{12}b_{21}}. \quad (7)$$

Відповідно, екстремальними власними числами отриманої симетричної, позитивно визначеною матриці H , будуть

$$\lambda_{\max}(H) = \frac{1}{2} \left\{ (h_{11} + h_{22}) + \sqrt{(h_{11} - h_{22})^2 + 4h_{12}^2} \right\},$$

$$\lambda_{\min}(H) = \frac{1}{2} \left\{ (h_{11} + h_{22}) - \sqrt{(h_{11} - h_{22})^2 + 4h_{12}^2} \right\}.$$

І для функції Ляпунова квадратичного виду (5) з коефіцієнтами h_{11} , h_{12} , h_{22} ,

визначеними в (8), повна похідна, в силу системи (7), дорівнює

$$\frac{d}{dt} V(z_1(t)) = -c[x_1^2(t) + y_1^2(t)] < 0,$$

а нульове положення рівноваги системи лінійного наближення (6) глобально асимптотично стійке. Як впливає з теореми про стійкість за лінійним наближенням, буде асимптотично стійким (але локально) і нульове положення вихідної нелінійної «обуреної системи». Обчислимо оцінку області асимптотичної стійкості. Запишемо систему, «обурену» до системи (2) в точці $O_1(x_1, y_1)$.

Відомо, що систему загального вигляду

$$\dot{z}_1(t) = F(z_1(t), z_1(t - \tau))$$

можна переписати, виділивши лінійну частину, наступним чином

$$\dot{z}_1(t) = \frac{\partial}{\partial z_1(t)} F(z_1(t), z_1(t - \tau))|_{(z_1, z_1)} z_1(t) +$$

$$+ \frac{\partial}{\partial z_1(t - \tau)} F(z_1(t), z_1(t - \tau))|_{(z_1, z_1)} z_1(t - \tau) +$$

$$+ F(z_1(t), z_1(t - \tau)) -$$

$$\left[\frac{\partial}{\partial z_1(t)} F(z_1(t), z_1(t - \tau))|_{(z_1, z_2)} z_1(t) + \right.$$

$$\left. + \frac{\partial}{\partial z_1(t - \tau)} F(z_1(t), z_1(t - \tau))|_{(z_1, z_1)} z_1(t - \tau) \right]$$

Або в вигляді

$$\dot{z}_1(t) = (A + B)z_1(t) + B[z_1(t - \tau) - z_1(t)] +$$

$$\{F(z_1(t), z_1(t - \tau)) - Az_1(t) - Bz_1(t - \tau)\} \quad (8)$$

$$A = \frac{\partial}{\partial z_1(t)} F(z_1(t), z_1(t - \tau))|_{(z_1, z_1)} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix},$$

$$B = \frac{\partial}{\partial z_1(t - \tau)} F(z_1(t), z_1(t - \tau))|_{(z_1, z_1)} = \begin{bmatrix} b_{11} & b_{12} \\ b_{21} & 0 \end{bmatrix}.$$

Повна похідна Ляпунова (5) в силу нелінійної системи з запізненням (8) в зміщеній точці $O_1(x_1, y_1)$ має вигляд

$$\frac{d}{dt} V(z_1(t)) = \dot{z}_1^T(t) H z_1(t) + z_1^T(t) H \dot{z}_1(t) =$$

$$z_1^T(t) (A + B)^T H z_1(t) + [z_1(t - \tau) - z_1(t)]^T B^T H z_1(t) +$$

$$+ \{F(z_1(t), z_1(t - \tau)) - Az_1(t) - Bz_1(t - \tau)\}^T H z_1(t) +$$

$$+ z_1^T(t)H(A+B)z_1(t) + z_1^T(t)HB[z_1(t-\tau) - z_1(t)] + \\ + z_1^T(t)H\{F(z_1(t), z_1(t-\tau)) - Az_1(t) - Bz_1(t-\tau)\}^T.$$

Використовуючи матричне рівняння Ляпунова, отримаємо

$$\frac{d}{dt}V(z_1(t)) \leq -z_1^T(t)Cz_1(t) + \\ 2|HB||z_1(t)||z_1(t) + z_1(t-\tau)| + \\ + 2|H||z_1(t)||F(z_1(t), z_1(t-\tau)) - Az_1(t) - Bz_1(t-\tau)|$$

Припустимо, що виконуються умови Б.С. Разуміхіна. Вони мають вигляд

$$\lambda_{\min}(H)||z_1(s)||^2 \leq V(z_1(s)) < V(z_1(t)), \quad s < t,$$

$$\leq \lambda_{\max}(H)||z_1(t)||^2$$

або

$$|z_1(t-\tau)| < \sqrt{\varphi(H)}|z_1(t)|, \\ \varphi(H) = \lambda_{\max}(H)/\lambda_{\min}(H).$$

Тоді для повної похідної функції Ляпунова в силу «обуреної системи» (8) матиме місце наступна нерівність

$$\frac{d}{dt}V(z_1(t)) \leq -\{\lambda_{\min}(C) - 2|HB|(1 + \sqrt{\varphi(H)})\}|z_1(t)|^2 + \\ + 2|H||z_1(t)||F(z_1(t), z_1(t-\tau)) - Az_1(t) - Bz_1(t-\tau)|$$

Оцінимо останній доданок, що визначає оцінку області стійкості.

$$R[z_1(t), z_1(t-\tau)] = \\ 2|H||z_1(t)||F(z_1(t), z_1(t-\tau)) - Az_1(t) - Bz_1(t-\tau)| =$$

Векторна функція $F(z_1(t), z_1(t-\tau))$ має вигляд

$$F(z_1(t), z_1(t-\tau)) = \\ = \begin{pmatrix} F_1(x_1(t), y_1(t), x_1(t-\tau), y_1(t-\tau)) \\ F_2(x_1(t), y_1(t), x_1(t-\tau), y_1(t-\tau)) \end{pmatrix}$$

де

$$F_1(x_1(t), x_1(t-\tau), y_1(t), y_1(t-\tau)) = \\ = \left[\frac{a(x_1(t-\tau) + x_1)}{N + (x_1(t-\tau) + x_1)} - b(y_1(t-\tau) + y_1) \right] \\ (x_1(t) + x_1) -$$

$$- \left[\frac{ax_1}{N + x_1} - by_1 \right] x_1,$$

$$F_2(x_1(t), x_1(t-\tau), y_1(t), y_1(t-\tau)) = \\ = -[c - d(x_1(t-\tau) + x_1)](y_1(t) + y_1) + \\ [c - d_1x_1]y_1,$$

$$Az_1(t) + Bz_1(t-\tau) = \begin{pmatrix} b_{11}x_1(t-\tau) + b_{12}y_1(t-\tau) \\ b_{21}x_1(t-\tau) \end{pmatrix}.$$

Розглянемо перший елемент $R_1[z_1(t), z_1(t-\tau)]$ «вектора збурень»

$$R[z_1(t), z_1(t-\tau)] = \begin{Bmatrix} R_1[z_1(t), z_1(t-\tau)] \\ R_2[z_1(t), z_1(t-\tau)] \end{Bmatrix}^T = \\ = F_1(z_1(t), z_1(t-\tau)) - Az_1(t) - Bz_1(t-\tau).$$

Враховуючи, що

$$x_1 = \frac{c}{d}, \quad y_1 = \frac{ac}{b(Nd + c)}, \quad b_{11} = -\frac{acdN}{(Nd + c)^2},$$

$$b_{12} = -c, \quad b_{21} = \frac{acd}{b(Nd + c)},$$

отримуємо

$$R_1[z_1(t), z_1(t-\tau)] = \\ = \left[\frac{a(x_1(t-\tau) + x_1)}{N + (x_1(t-\tau) + x_1)} - b(y_1(t-\tau) + y_1) \right] \\ (x_1(t) + x_1) - \\ \left[\frac{ax_1}{N + x_1} - by_1 \right] x_1 - \\ - b_{11}x_1(t-\tau) - b_{12}y_1(t-\tau). \quad (9)$$

Розглянемо другий елемент

$$R_2[z_1(t), z_1(t-\tau)].$$

Отримуємо

$$R_2[z_1(t), z_1(t-\tau)] = \\ -[c - d(x_1(t-\tau) + x_1)](y_1(t) + y_1) + \\ [c - d_1x_1]y_1 - b_{21}x_1(t-\tau) = \\ = -cy_1(t) + dx_1(t-\tau)y_1(t) + \\ + dx_1y_1(t) - cy_1 + dx_1(t-\tau)y_1 + \\ + dx_1y_1 + cy_1 - dx_1y_1 - b_{21}x_1(t-\tau) = \\ = dx_1(t-\tau)y_1(t). \quad (10)$$

Таким чином,

$$\frac{d}{dt}V(z_1(t)) \leq -\{\lambda_{\min}(C) - 2|HB|(1 + \sqrt{\varphi(H)})\}|z_1(t)|^2 + \\ 2|H||z_1(t)||R[z_1(t), z_1(t-\tau)]|.$$

І для отримання оцінки «гарантованої області асимптотичної стійкості» отримуємо нерівність

$$U_R = \{z : |z| \leq r\},$$

$$r = \max_{t \geq 0} \left\{ z : |z| \leq \frac{\lambda_{\min}(C) - 2|HB|(1 + \sqrt{\varphi(H)})}{2|H|R[z_1(t), z_1(t - \tau)]} \right\}$$

$$\text{при умові } |z(t - \tau)| < \sqrt{\varphi(H)}|z(t)|.$$

Список використаних джерел

1. Lotka A.J. Elements of physical biology / A.J. Lotka. - Baltimore: Williams and Wilkins, 1925. - 460 с.
2. Malthus T.R. An Essay on the Principle of Population / T.R. Maltus.- London: Johnson, 1798. - 134 с.
3. Базыкин А.Д. Математическая биофизика взаимодействующих популяций / А.Д. Базыкин. - Москва: Наука, 1985. - 181 с.
4. Вольтерра В. Математическая теория борьбы за существование / В. Вольтерра. - Москва: Наука, 1976. - 286 с.
5. Смит Дж. Модели в экологии / Дж. Смит. - Москва: Мир, 1976. - 182 с.
6. Эльсгольц Л.Э. Введение в теорию дифференциальных уравнений с отклоняющимся аргументом / Л.Э.Эльсгольц, С.Б. Норкин. - Москва: Наука, 1970. - 240 с.
7. Гаркуша Н.І. Про одну математичну модель динаміки взаємодіючих популяцій / Н.І. Гаркуша // Вісник Київського національного університету імені Тараса Шевченка. Серія фізико-математичні науки. - 2014. - №4. - С. 135-138.
8. Гаркуша Н.І. Динаміка однієї екологічної моделі «хижак-жертва» без врахування вікової структури / Н.І. Гаркуша // Вісник Київського національного університету імені Тараса Шевченка. Серія кібернетика. - 2014. - №14. - С. 22-24.
9. Verhulst P.F. Natice sur la loi que la population suit dans son accroissement / P.F. Verhulst // Corr. Math. Et. Phys. - 1838. - V.10. - P. 113-121.
10. Колмогоров А.Н. Качественное изучение математических моделей динамики популяций / А.Н. Колмогоров // Проблемы кибернетики. - 1972. - №25. - С. 100-106.

References

1. LOTKA, A. (1925) *Elements of Physical Biology*. Baltimore: Williams and Wilkins.
2. MALTHUS, T. (1798) *An Essay on the Principle of Population as it Affects the Future Improvement of Society, with Remarks on the Speculations of Mr. Godwin, M. Condorcet, and Other Writers*. London: Johnson.
3. BAZYKIN, A. (1985) *Mathematical Biophysics of Interacting Populations*. Moskva: Nauka.
4. VOLTERRA V. (1976) *The Mathematical Theory of the Struggle for Existence*. Moskva: Nauka.
5. SMITH, J. (1974) *Models in Ecology*. Cambridge: CambridgeUniversity Press.
6. ELSGOLTZ, L., NORKIN, S. (1970) *Introduction to the Theory of Differential Equations with Deviating Argument*. Moskva: Nauka.
7. GARKUSHA, N. (2014) About One Mathematical Model of the Dynamics of Interacting Population. In *Bulletin of Taras Shevchenko National University of Kyiv. Series Physcs&Mathematics*. №4.- p.135-138.
8. GARKUSHA, N. (2014) The Dynamics of the Ecological Model "Predator-Prey" without Age Structure. In *Bulletin of Taras Shevchenko National University of Kyiv Series Cybernetics*, №14. - p. 22-24.
9. VERHULST, P. (1838) Natice sur la loi que la population suit dans son accroissement. In *Corr. math. et phys.*, v. 10. - p.113-121.
10. KOLMOGOROV, A. (1972) A Qualitative Study of Mathematical Models of Population Dynamics. - In *Problems of Cybernetics*. №25. - p. 100-106.

Надійшло до редколегії 10.05.2015 р.

УДК 004.75

O. V. Hordiichuk, postgraduate

Location oriented peer-to-peer networks

Taras Shevchenko National University of Kyiv,
83000, Kyiv, Glushkova st., 4d,

e-mail: oleg.gordiychuck@gmail.com

Гордійчук О. В., аспірант

Однорангові мережі з урахуванням місцезнаходження

Київський національний університет імені Тараса
Шевченка, 83000, м. Київ, пр-т. Глушкова 4д,

e-mail: oleg.gordiychuck@gmail.com

Додатки побудовані на основі однорангових мереж допомагають розв'язати проблему доставки інформації багатьом абонентам, шляхом розподілення мережних навантажень серед усіх учасників. Проте із ростом мережі, також зростає і затримка доставки даних, яка виникає внаслідок неоптимального з географічної точки зору розташування вузлів в топології. В цій статті запропонований новий алгоритм пошуку найближчих сусідів, а також топологія, яка дозволяє робити подібні запити, в основі яких лежить принцип пошуку найближчих в мережевому розумінні вузлів за допомогою розподіленої системи мережеских координат. Даний підхід дозволяє більш ефективно використовувати мережеві ресурси, звести до мінімуму надлишкову затримку, яка виникає при передачі даних, а також знизити навантаження мереж інтернет-провайдерів. В свою чергу використання алгоритму побудови топології на основі 7-вимірного простору надає змогу вузлам знайти найближчих сусідів за кількістю запитів, яка логарифмічно залежить від кількості вузлів в мережі, що робить систему швидшою, порівняно з існуючими розв'язками.

Ключові слова: однорангові мережі, пошук найближчих сусідів, мережеві координати.

Applications that use peer-to-peer networks help to solve a data delivery problem to many endpoints by distributing network load among all participants. However with network growing also a delivery data delay increased, which occurs due to unoptimal location in geographical meaning of nodes in a topology. In this paper proposed a new algorithm of the nearest neighbor search as well as a topology that enables these queries that uses a principle of searching the nearest nodes in the network by using a distributed coordinate system. This approach helps to use network resource more efficiently, decrease to a minimum unwanted delay that appears during data dissemination, and at the same time decrease a network load for Internet providers. Also usage of this algorithm, which is based on the 7-dimensional space, makes possible to find nearest neighbors with a count of queries that is equal to a logarithm of the total amount of nodes in the network and makes this system faster compared to existing solutions.

Key Words: peer-to-peer networks, nearest neighbors search, network coordinates.

Статтю представив д.т.н., проф. Погорілий С.Д.

Introduction

Peer-to-peer network is a promising technology for solving data dissemination problems. Its distributed feature makes possible to unload a network bandwidth of a central server. While for file transmission there are a lot of efficient existing solutions like BitTorrent, for video streaming and other latency dependant data this approach is still too far away from meeting quality of experience (QoE) criteria. Main reasons are unpredictable behaviour of users that join and leave the network and a network

delay that appears during data transmission. Moreover this approach is not scalable in general, because the bigger size of the network, the bigger quality impact for the end users. With a network growth participants of the network also notice sudden interruptions and a jitter until new data will be available. Another possible side effect is a time lag of the executed query in the network like accessing key-value resource or searching another nodes. In case of real-time peer-to-peer streaming event like football or basketball match, users in adjacent rooms will observe a synchronization

problem. While in one room viewers will see a score, in another room they will wait up to several minutes until this event occur. These reasons make peer-to-peer networks for real-time applications rather proof of concepts than production ready solutions.

In this paper proposed a new distributed location service that has a goal for the decreasing end-user delays, which in a result improves quality of experience. The key idea is to make things completely distributed and thus scalable to networks of any size. At the same time proposed schema significantly decreases query latency by avoiding inefficient transmitting between peers that are located far away from each other. It is mainly based on an approach of finding nearest neighbours in network meaning that is achieved by using a special topology structure, which enables this query for every node.

Related work

The task of finding nearest neighbours in the network is a known problem in peer-to-peer networks. A solution of this task is hard due to the fact that ISP networks are connected in a different manner with each other and even in case when users located in the same network they can have big value of RTT (round-trip time) compared to peers located in other networks. This situation is typical for cellular networks (2G-4G) as peers located in such network will be farther between each other than with another peer that uses a cable connection. It is also complicated due to a huge amount of possible distance pairs between peers and complexity of its measurement. However the last problem is partially solved by using King [1] tool that helps to evaluate distance between two hosts by using recursive DNS queries that return RTT values. This tool allowed measuring network distances between two hosts without direct querying them. However for answering the nearest neighbours query it still requires an efficient data structure and $n(n-1)$ "king" measurements, where n – is a number of nodes in the network. It means that using this approach for the problem solution requires a central server and a query complexity grows with every new node entrance and departure.

These problems are solved in Meridian [2]. It is a completely distributed system that helps peers to find nearest neighbors in a logarithm of network size amount of queries. Meridian has its own network topology that organizes neighbors into groups at different distances that are exponentially increased. Every group contains peers of a maximum possible

observed square calculated every time new peer tries to join it. In the proposed paper we also use an idea of building topology, where every peer has a small set of neighbors at different distances that helps newcomers to find their nearest neighbors. But instead of requesting every subset of neighbors for retrieving distance information, here we use network coordinates that make possible to avoid additional operations and increase speed of the query.

Another approach is to use virtual network coordinate systems. A key idea here is assuming that all peers live in Euclidean space and in result every node has its own coordinates in such space. There exist two different approaches: land-marking and distributed network coordinates. In the land-marking case every participant of the network measures distances to predefined nodes that are located as far as possible between each other. Every land-marked node represents separate axis, thus naturally the more predefined nodes the more accurate results could be achieved. This idea is used in GNP [3], Lighthouse [4] and PCoord [5] systems. Mainly after landmarks measurements accomplished every peer applies some optimization method like Simplex Downhill that minimizes possible errors. However these systems require lot of landmarks over the world for the really accurate results, also this solution requires centralized server and thus is not scalable. That is why the distributed network coordinates approach was proposed. Here each node acts as a landmark and is used for coordinates calculations. Perhaps the most known system that implements this idea is Vivaldi [6]. This system consists of a 2-d coordinate system that is enriched by a height parameter that represents a jitter or error that occurs in the network space as the triangle inequality is often violated. The Vivaldi algorithm continuously updates coordinates of every node, however they use only 2-d coordinates system due to inability to decrease network overhead for greater dimensions. The approach proposed in this paper uses 7-d coordinate system that achieves at least same accuracy, but it doesn't require computing the height parameter and has constant network overhead. Also in the Vivaldi the nearest neighbor query is not implemented. Another existing implementation is PIC (Practical Internet Coordinates) [7]. Here every node is used for the network coordinates estimation like in the Vivaldi. But comparing to the previous approach it has an ability to request the system with the nearest neighbor query by using an active node discovery protocol. It combines 3 strategies: random nodes, closest nodes and a hybrid of both for minimizing errors that appears in the Simplex

Downhill method. This idea improves overall accuracy and makes system reliable for both – near and far peers. In the proposed paper we also provide an ability to answer the nearest neighbor query in a completely decentralized manner and for achieving better accuracy we use sets of neighbors at different distances that represent a better sampling than in PIC.

An algorithm for building network coordinates

We assume that a network space could be represented as an n -dimensional Euclidian space, where distance function $r(n_1, n_2) = ||n_1 - n_2||$ represents RTT (round-trip time) between a pair of n_1 and n_2 nodes in the swarm. While it's known [6] that a real network space is not the Euclidian one due to a triangle inequality violations, it's still possible to accurately predict network distances with a high probability (evaluation of this method is described in the section below). The network coordinate of node n denoted as a tuple $(x_1, x_2, \dots, x_p)$, where every x_i represents a value of a dimension an p is a cardinality of x . Initially, when the peer joins the network, it assigns to itself a coordinate where all dimensions are equal to 0: $(0, 0, \dots, 0)$ and starts communicating with other nodes. Periodically (every k seconds) every peer requests the neighborhood current coordinates and updates information about RTT using response time. It triggers updating of its own coordinates and rebuilding local searching topology. As only coordinates values need to be sent, a network overhead of this algorithm is constant and configurable via k and p parameters. In this case for a valid operating of the algorithm every peer needs to additionally allocate b bytes for the upload and download bandwidth that may be evaluated as following equation:

$$b = c p b \quad (1)$$

Here b is an amount of bytes required to encode one dimension and c is a cardinality of the neighborhood set. Naturally the more neighbors every peer has and the greater cardinality of coordinates is, the more accurate results may be obtained. However, increasing of these values also increase the network overhead. That is why it's important to choose a good tradeoff. In this paper we use $c=10$, $b=2$, $p=7$ and $k=5$ that means it's only need 140 bytes of the bandwidth for every 5 seconds to accomplish next stage of the network coordinates updating. These values provide reasonable accuracy

and overhead (comparison with other parameters described in the next section).

Whenever peer receives new information about coordinates, it's possible to evaluate current error e that is defined as following equation:

$$e(n_i, n_j) = |r(n_i, n_j) - d(n_i, n_j)|, \quad n_i, n_j \in N, i \neq j. \quad (2)$$

Where N – is a set of all nodes, r function that returns the round-trip time between two nodes and d is a distance function on the defined Euclidean space that calculates distance by using network coordinates. Ideal case when all error values are equal to 0. However, as it was mentioned before, generally it's known to be an impossible task due to the triangle inequality violations in the Internet network. So in this case network coordinate will always contain error information and therefore a key idea of the algorithm is updating all coordinates by minimizing possible errors. It could be represented as a task of minimizing a square sum of observable errors for every node n_i :

$$\min \sum_{n_j \in N_i} e^2(n_i, n_j), \quad N_i \subseteq N. \quad (3)$$

Where N_i is a set that represents the neighborhood of a peer n_i . So the defined task is a common non-linear least squares problem. It could be solved by using Levenberg-Marquardt algorithm (LMA) [7]. In the input the current network coordinates of every neighbor, their values of RTT and the current node's coordinate are provided. And the algorithm will modify values of the current peer's coordinates to satisfy condition described in (3). In a result network coordinate of the node updated and available for requesting by other nodes and possible errors are minimized. Next time, when neighbors of the peer will request new coordinates they also update their ones, which will produce more accurate result as errors are minimized. It means that this algorithm is iterative and at every step accuracy is improved. Also in a case when peer churn appeared or new participants entered the network coordinate topology will be rebuilt in a couple of updates that makes it robust to dynamic changes in the network. These approach enables possibility to use network coordinates for answering the nearest neighbor query for both: server based and distributed implementations. While both of implementations have their advantages and disadvantages, we

recommend to use network coordinates obtained from the described algorithm and use them in the Meridian topology. In this case it's possible to avoid direct measurements for the nearest neighbor query and make things completely in a distributed manner.

Evaluation of the algorithm

For evaluating the algorithm that described in the previous section we used a statistical table data for 168 cities that contains ping measurements for every pair in the table (28056 records totally) [9]. Cities are located on all continents of the Earth and mostly represent capitals and big cities. It's not hard to notice that main parameters of the algorithm, which will affect accuracy are number of dimensions and a size of the neighborhood. The more dimensions and the greater size are, the better accuracy and higher performance penalty. While the neighborhood size may vary in different peer-to-peer applications, we assume that in most cases the minimum value equal to 10 neighbors. This value doesn't mean that it must be used in implementations of the proposed method, rather it helps to compare how number of dimensions will impact in a typical case. Every neighbor in every test has chosen randomly, all tests repeated for 10 times and in the end obtained values are averaged. Such approach shows results in different environments and thus eliminates possibility of "well-tuned" effect. Below represented a column chart that compares average error of every dimension type after 20 cycles of the updating coordinates:

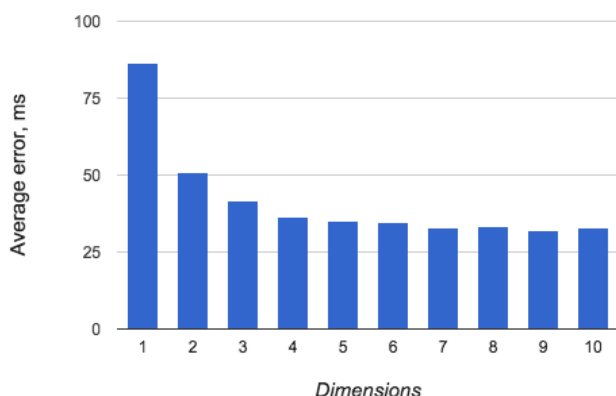


Figure 1. Average error compared to dimensions.

As it's shown in the figure above starting from the 7th dimension overall accuracy doesn't significantly change. On the average it's equal to 33.07 ms for and in the case with 10 dimensions it's equal to 32.86 ms. Therefore it's recommended to use p parameter (which represents amount of dimensions) equal to 7. As in this case it's possible

to achieve reasonable accuracy without overloading CPU for applying the Levenberg-Marquardt algorithm. Separately for the case with 7 dimensions presented charts, which show how many (in percentage) nodes are covered with relative and absolute errors:

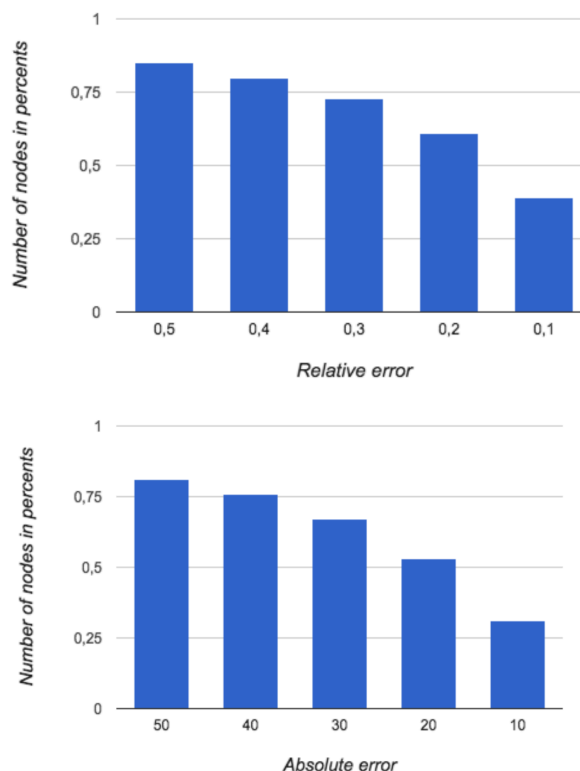


Figure 2. Relative and absolute errors compared to number of nodes.

Moreover these results may be improved if apply following heuristic. It's known that some of the nodes are affected by network jitter that may appear in mobile carrier networks, DSL or due to the network congestion. If assume that these nodes don't have any neighbors closer than 100 ms, than it's possible to create a second layer of network coordinates, where such nodes will be removed. In this case they still can use the first layer, while others use both layers for more precise calculations. In this case the second layer of our dataset contains 130 cities from 168. Applying the proposed algorithm results to an average relative error being equal to 17.12 milliseconds and nearly in 50% of cases peers have relative error less than 10 milliseconds.

Conclusion

In this paper proposed a new algorithm for building network coordinates. It's based on the idea of minimizing visible errors in a 7 dimensional space on the each side of the swarm. It requires a minimum of additional bandwidth allocation that will not be

noticed in most peer-to-peer applications. And its algorithm is simple in implementation and compared to analogues doesn't require to transmit or evaluate additional information except network coordinates. Also the proposed approach executed in a completely distributed manner and may be used in a existing engines for searching nearest neighbors. At the same time it demonstrates very accurate results: mean relative error is equal to 33.07 milliseconds on the proposed dataset. And in a case of using a second additional layer with the reducing heuristic the mean relative error is almost halved and equal to 17.12 milliseconds. Achieved results enable usage of the algorithm in all peer-to-peer locations, specially video-streaming ones, that improves quality of experience of such systems.

Список використаних джерел

1. Gummadi K. P. King: Estimating latency between arbitrary internet end hosts / Gummadi K.P., Saroiu S., Gribble S. D. Proceedings of the 2nd ACM SIGCOMM Workshop on Internet measurement. – ACM, 2002. – P. 5-18.
2. Wong B. Meridian: A lightweight network location service without virtual coordinates / Wong B., Slivkins A., Sirer E. G. - ACM SIGCOMM Computer Communication Review. – ACM, 2005. – Vol. 35. – No. 4. – P. 85-96.
3. Ng T. S. E. Predicting Internet network distance with coordinates-based approaches / Ng T. S. E., Zhang H. - INFOCOM 2002. Twenty-First Annual Joint Conference of the IEEE Computer and Communications Societies. Proceedings. IEEE. – IEEE, 2002. – Vol. 1. – P. 170-179.
4. Pias M. Lighthouses for scalable distributed location / Pias M. - Peer-to-Peer Systems II. – Springer Berlin Heidelberg, 2003. – P. 278-291.
5. Lehman L. Pcoord: Network position estimation using peer-to-peer measurements / Lehman L., Lerman S. - Network Computing and Applications, 2004. (NCA 2004). Proceedings. Third IEEE International Symposium on. – IEEE, 2004. – P. 15-24.
6. Dabek F. Vivaldi: A decentralized network coordinate system / Dabek F. et al. ACM SIGCOMM Computer Communication Review. – ACM, 2004. – Vol. 34. – No. 4. – P. 15-26.
7. Cox R. Practical, distributed network coordinates / Cox R. et al. ACM SIGCOMM Computer Communication Review. – 2004. – Vol. 34. – No. 1. – P. 113-118.
8. Moré J. J. The Levenberg-Marquardt algorithm: implementation and theory / More J. J.

Numerical analysis. – Springer Berlin Heidelberg, 1978. – P. 105-116.

9. Wondernetwork statistical <https://wondernetwork.com/pings/>

References

1. GUMMADI K. P., SAROIU S., GRIBBLE S. D. (2002) King: Estimating latency between arbitrary internet end hosts. *Proceedings of the 2nd ACM SIGCOMM Workshop on Internet measurement*. p. 5-18.
2. WONG B. SLIVKINGS A., SIRER E. G. (2005) Meridian: A lightweight network location service without virtual coordinates. *Computer Communication Review* 35 (4). p. 85-96.
3. NG T. S. E., ZHANG H. (2002). Predicting Internet network distance with coordinates-based approaches. *Twenty-First Annual Joint Conference of the IEEE Computer and Communications Societies I*. p. 170-179.
4. PIAS M. (2003). Lighthouses for scalable distributed location. *Peer-to-Peer Systems II. – Springer Berlin Heidelberg*. p. 278-291.
5. LEHMAN L., LEHMAN S. (2004) Pcoord: Network position estimation using peer-to-peer measurements. *Network Computing and Applications. Proceedings. Third IEEE International Symposium on. – IEEE, 2004. – p. 15-24.*
6. DABEK F. (2004) Vivaldi: A decentralized network coordinate system. *ACM SIGCOMM Computer Communication Review* 34 (4). p. 15-26.
7. COX R. (2004) Practical, distributed network coordinates. *ACM SIGCOMM Computer Communication Review* 34 (1). p. 113-118.
8. MORE J. J. (1978) The Levenberg-Marquardt algorithm: implementation and theory. *Numerical analysis. – Springer Berlin Heidelberg, 1978. – p. 105-116.*
9. WONDERNETWORK STATISTICAL <https://wondernetwork.com/pings/>

Надійшла до редколегії 28.05.15

УДК 51-77+519.854

Донченко В.С., д.ф.-м.н., проф.,
Тарасова О.В., аспірант

V. S. Donchenko, Doctor of Sciences (Physics &
Mathematics), Full Professor,
O. V. Tarasova, Postgraduate Student.

Матрична множинна регресія

Matrix multiple regression

Київський національний університет імені
Тараса Шевченка, 83000, м. Київ, пр-т.
Глушкова 4д,
e-mail: voldon@univ.kiev.ua
e-mail: olga_ta@bigmir.net

Taras Shevchenko National University of Kyiv,
83000, Kyiv, Glushkova st., 4d,
e-mail: voldon@univ.kiev.ua
e-mail: olga_ta@bigmir.net

Запропоновано до розгляду клас матричних функцій набору матричних аргументів: матричну множинну регресію. У рамках розвитку концепції кортежних операторів розроблено математичний апарат сингулярного (SVD) розкладу та техніки псевдообернення (ПдО) за Муром-Пенроузом, що дозволяє поставити та конструктивно розв'язати задачу оцінки методом найменших квадратів (МНК) для множинної матричної регресії.

Ключові слова: матрична множинна регресія, псевдоінверсія, медіабізнес.

The class of matrix functions with set of matrix arguments (matrix multiple regression) is offered for consideration. The mathematical apparatus of singular (SVD) decomposition and of pseudoinverse technique (PDO) by Moore-Penrose is developed in this paper. The development of mathematical apparatus is made within the concept of cortege operators. The materials of this paper allow to formalize and to solve constructively the estimation problem by the method of least squares (OLS) for multiple matrix regression. OLS-estimation algorithm is proposed and realized for with unknown parameters vector of the matrix functions class in solving forecast. The example of forecasting with the use of the matrix multiple regression is shown for TV Viewing Ukrainian data. As shown in the example, matrix multiple regression can be used for mediabusiness parameters forecasting problems with acceptable business planning processes accuracy. Moreover the forecast indicators accuracy for commercial time slots higher than for noncommercial; that fact is especially important for commercial media planning and planning the TV-media business processes.

Key Words: multiple regression matrix, pseudoinversion, mediabusiness.

Статтю представив д.т.н., проф.. Гаращенко Ф.Г.

Задача групування інформації [1]: задача відновлення залежності і задача класифікації, кластеризації і розпізнавання образів, - це задача принципової важливості у прикладних дослідженнях. Важливим класом математичних структур ([2]) є евклідові простори числових векторів. Їх використання визначається багатством структурних зв'язків евклідових просторів в цілому та розвинутою технікою оперування у цих просторах, в тому числі сингулярний (SVD) розклад в псевдообернення (ПдО) у цих просторах [3-5]. Відмітимо значну роль М.Ф. Кириченко що розвиває ПдО-техніку [6]. Тим не менше, використання евклідових просторів числових векторів стає обмежуючим у розв'язанні прикладних задач, якими є, зокрема, задачі розпізнавання мови (speech recognition) та

обробки зображень. У таких задачах природний представник досліджуваних об'єктів – це матриці: зокрема, спектрограми і матриці зображень. У роботі [7] запропоновано концепцію кортежних операторів, що дозволяє перенести SVD – і ПдО-техніки на евклідові простори матриць фіксованої розмірності; згадані техніки використано для побудови відстані відповідності у задачах кластеризації. У роботах [8-9] запропоновано основи векторного і матричного методів найменших квадратів з прикладами прикладного застосування для макро- та мікроекономічних показників.

Теоретичні основи матричної множинної регресії

1. **Теорема В1.** (Базова теорема про SVD-розклад [7].) Нехай A - лінійний оператор між

двома абстрактними евклідовими просторами $A: E_1 \rightarrow E_2$, $A^*: E_2 \rightarrow E_1$ - спряжений до нього, $r = \text{rank} A = \dim \mathfrak{R}(A)$. Тоді:

- 1) загальна кількість ненульових сингулярностей (пари власний вектор – власне число) двох операторів AA^* , A^*A дорівнює r ;
- 2) набір власних чисел $\lambda_1^2 \geq \dots \geq \lambda_r^2 > 0$ згаданих операторів є спільним;
- 3) власні вектори відповідних наборів сингулярностей $(u_i, \lambda_i^2), (v_i, \lambda_i^2), i = \overline{1, r}$ пов'язані співвідношенням:

$$u_i = \frac{Av_i}{\lambda_i}, v_i = \frac{A^*u_i}{\lambda_i}, i = \overline{1, r};$$

- 4) оператор може бути представлений у вигляді

$$Ax = \sum_{i=1}^r \lambda_i u_i (v_i, x), x \in E_1$$

2. **Визначення.** Псевдооберненим до оператора $A: E_1 \rightarrow E_2$ - позначатимемо A^+ - називається оператор $A^+: E_2 \rightarrow E_1$, що визначається співвідношенням $A^+y = \sum_{i=1}^r \lambda_i v_i (u_i, y), y \in E_2$.

3. **Теорема В3.** Значення A^+y є мінімальним за нормою рішенням оптимізаційної задачі мінімізації для $x \in E_1$ функціоналу $\|Ax - y\|^2$:

$$A^+y = \underset{\substack{z \in \text{Argmin} \|Ax - y\|^2 \\ x \in E_1}}{\arg \min} \|z\|^2.$$

4. **Теорема В4.** (Дослідження СЛАР.) У прийнятих вище позначеннях, необхідною і достатньою умовою розв'язуваності відносно системи лінійних алгебраїчних рівнянь (СЛАР) $Ax = y, y \in E_2$ є умова $(y, (I_{E_2} - AA^+)y) = 0$, а множина всіх розв'язків має такий вигляд:

$$A^+y + (I_{E_1} - A^+A)E_1 \quad (1)$$

У разі нерозв'язуваності СЛАР: $(y, (I_{E_2} - AA^+)y) > 0$, - множина, що визначається співвідношенням (1), описує множину $\text{Argmin} \|Ax - y\|^2$, тобто множина всіх $x \in E_1$

розв'язків оптимізаційної задачі найкращого квадратичного наближення правої частини СЛАР значеннями лівої частини. Усі елементи множини (1) у такому разі називаються псевдорозв'язками.

5. Множинну матричну регресію з матричними значеннями визначимо у детермінованій постановці: як матричну функцію

Y матричного кортежа-рядка $\alpha_{1,K} = (X_1, \dots, X_K)$,

$X_k \in R^{m \times n}$, $k = \overline{1, K}$ такого виду

$Y(\alpha_{1,K}) = \beta_1 X_1 + \dots + \beta_K X_K, \beta_k \in R^1, k = \overline{0, K}$, що

визначаються на основі спостережень $(\alpha_{1,K}^{(i)}, Y_i), i = \overline{1, N}$, де в свою чергу

$Y_i = Y(\alpha_{1,K}^{(i)}), \alpha_{1,K}^{(i)} = (X_1^{(i)}, \dots, X_K^{(i)}), i = \overline{1, N}$.

6. Відображення $\wp_{\alpha_{1,K}}: R^K \rightarrow R^{m \times n}$, що визначається таким співвідношенням: $Y(\alpha_{1,K}) = \beta_1 X_1 + \dots + \beta_K X_K \equiv \wp_{\alpha_{1,K}} \beta$, де кортежним оператором називається $\beta^T = (\beta_1, \dots, \beta_K), \beta \in R^K$, у відповідності з [7].

7. Будемо позначати так само:

- $\mathcal{X}$ - матричний стовпчиковий кортеж, отриманий зі значень спостережень $Y_i, i = \overline{1, N}$:

$$\mathcal{X}_Y = \begin{pmatrix} Y_1 \\ \dots \\ Y_N \end{pmatrix}, i = \overline{1, N};$$

- $\mathcal{X}_{N,1}^{(j)}, j = \overline{1, K}$ - стовпчиковий кортеж довжини N , отриманий із j -тих компонент рядкових кортежів $\mathcal{X}_{1,K}^{(i)}, i = \overline{1, K}$;

- $\alpha_{N,K}$ - строковий кортеж довжини K , отриманий зі стовпчикових кортежів $\mathcal{X}_{N,1}^{(j)}, j = \overline{1, K}$: $\alpha_{N,K} = (\mathcal{X}_{N,1}^{(1)}, \dots, \mathcal{X}_{N,1}^{(K)})$;

- $\wp_{\alpha_{N,K}}: R^K \rightarrow R^{N, m \times n}$ - матричний кортежний оператор з евклідового простору R^K в евклідовому просторі стовпчикових кортежів довжини N з матриць розмірності $m \times n$ і покомпонентним скалярним добутком для

$$\mathcal{X}_1 = \begin{pmatrix} Y_1^{(1)} \\ \dots \\ Y_N^{(1)} \end{pmatrix} \in R^{N, m \times n}, \mathcal{X}_2 = \begin{pmatrix} Y_1^{(2)} \\ \dots \\ Y_N^{(2)} \end{pmatrix} \in R^{N, m \times n} \quad \text{у}$$

$$\text{такому вигляді } (\mathcal{X}_1, \mathcal{X}_2)_{N, m \times n} = \sum_{i=1}^N (Y_i^{(1)}, Y_i^{(2)})_{tr}.$$

8. Оцінка $\hat{\beta} \in R^K$ методу найменших квадратів невідомого параметру матричної множинної регресії з матричними значеннями природнім чином визначається як розв'язок оптимізаційної задачі мінімізації функціоналу

$\sum_{i=1}^N \|Y_i - \varphi_{\alpha_{1,K}^{(j)}} \beta\|_{m \times n}^2$ методу найменших квадратів:

$$\hat{\beta} = \underset{\beta \in R^K}{\operatorname{argmin}} \sum_{i=1}^N \|Y_i - \varphi_{\alpha_{1,K}^{(j)}} \beta\|_{m \times n}^2 \quad (2)$$

9. Стандартним чином МНК-оцінку $\hat{Y}(\alpha_{1,K}), \alpha_{1,K} = (X_1, \dots, X_K), X_k \in R^{m \times n}, k = \overline{1, K}$ матричної множинної регресії будемо визначати співвідношенням

$$\hat{Y}(\alpha_{1,K}) = \varphi_{\alpha_{1,K}} \hat{\beta} = \sum_{k=1}^K \hat{\beta}_k X_k, X_k \in R^{m \times n}, k = \overline{1, K}.$$

10. Очевидно, в силу покомпонентного визначення скалярного добутку матричних кортежів, все одно стовпчикових чи рядкових кортежів, функціонал методу найменших квадратів

$$A(\beta) = \|Y - \varphi_{\alpha_{1,K}} \beta\| = \|\chi - \beta \chi\|, \chi = \begin{pmatrix} Y \\ \dots \\ Y \end{pmatrix}$$

матричної регресії може бути записано за стовпчиковими матричними кортежами.

11. **Лема.** Функціонал методу найменших квадратів матричної регресії $A(\beta)$ може бути представлено у вигляді:

$$A(\beta) = \sum_{i=1}^N \|Y_i - \varphi_{\alpha_{1,K}^{(j)}} \beta\|_{m \times n}^2 = \|\chi_Y - \sum_{j=1}^K \beta_j \chi_{N,1}^{(j)}\|_{N, m \times n}^2 = \|\chi_Y - \varphi_{\alpha_{N,K}} \beta\|^2$$

12. **Теорема 1.** Розв'язком оптимізаційної задачі (2) є $\hat{\beta}$, що визначається співвідношенням:

$$\hat{\beta} = \varphi_{\alpha_{N,K}}^+ \chi_Y \quad (3)$$

Доведення Теорема 1 витікає із загального варіанту SVD-розкладу, визначення псевдоінверсії та дослідження СЛАР для лінійних операторів між абстрактними евклідовими просторами (Теорема ВЗ).

13. **Теорема 2.** Оператор $\varphi_{\alpha_{1,K}}^* \varphi_{\alpha_{1,K}} : R^K \rightarrow R^K$ задається стандартним чином матрицею F , що є матрицею Грама Γ для набору векторів $\chi_{N,1}^{(j)}, j = \overline{1, K}$:

$$\Gamma = \left((\chi_{N,1}^{(i)}, \chi_{N,1}^{(j)})_{N, m \times n} \right)_{i,j=\overline{1, K}}.$$

Доведення Теорема 2 проводиться так само, як і у [7]: спочатку обчисленням спряженого оператора

$\varphi_{\alpha_{N,K}}^*$ до оператора $\varphi_{\alpha_{1,K}}$, що визначається співвідношенням:

$$\varphi_{\alpha_{N,K}}^* \chi = \begin{pmatrix} (\chi, \chi_{N,1}^{(1)})_{N, m \times n} \\ \dots \\ (\chi, \chi_{N,1}^{(K)})_{N, m \times n} \end{pmatrix}, \chi \in R^{N, m \times n}. \quad (3)$$

Підстановка у значення $\chi = \varphi_{\alpha_{N,K}} \beta = \sum_{j=1}^K \beta_j \chi_{N,1}^{(j)}$ завершує доведення.

Позначимо $(v, \lambda), i = \overline{1, r}$ набір ненульових сингулярностей – пар «власні вектори – власні числа», - оператору, що задається матрицею F : $r = \operatorname{rank} F, \lambda_1^2 \geq \dots \geq \lambda_r^2 > 0$. Нагадаємо, мова йде про власні вектори і власні числа матриці.

14. **Теорема.** Задача SVD-розкладу кортежного оператора $\varphi_{\alpha_{N,K}}$ зводиться до стандартної задачі на власні числа для матриць: при $(v_i, \lambda_i^2), i = \overline{1, r}$ - ненульових сингулярностей матриці F набір «спряжених сингулярностей»: $(\chi_{U,i}, \lambda_i^2), i = \overline{1, r}$, при збереженні набору власних чисел, визначається співвідношенням

$\chi_{U,i} = \frac{\varphi_{\alpha_{N,K}} v_i}{\lambda_i}, i = \overline{1, r}$, а SVD-розклад кортежних операторів $\varphi_{\alpha_{N,K}}, \varphi_{\alpha_{N,K}}^+$ представляються співвідношеннями, відповідно:

$$\varphi_{\alpha_{N,K}} = \sum_{k=1}^r \lambda_k \chi_{U,i} v_k^T : R^K \rightarrow R^{m \times n},$$

$$\varphi_{\alpha_{N,K}}^+ \chi = \sum_{k=1}^r \lambda_k^{-1} v_k (\chi_{U,i}, \chi)_{N, m \times n} : R^{N, m \times n} \rightarrow (R^K)^*,$$

де $\chi \in R^{N, m \times n}$

15. **Теорема 3.** Оцінка МНК вектору невідомих параметрів визначається співвідношенням

$$\hat{\beta} = \varphi_{\alpha_{N,K}}^+ \chi_Y = \sum_{k=1}^r \lambda_k^{-1} v_k (\chi_{U,i}, \chi_Y)_{N, m \times n}.$$

Доведення Теорема 3 отримується із прямого застосування Теорема 1.

16. **Алгоритм обчислення $\hat{\beta}$:**

1. Побудова стовпчикових «компонентних» кортежів $\chi_{N,1}^{(j)}, j = \overline{1, K}$.
2. Побудова матриці Грама Γ стовпчикових «компонентних» кортежів.

3. Обчислення ненульових сингулярностей $(v_i, \lambda_i^2), i = \overline{1, r}$ матриці Грама Γ стовпчикових «компонентних» кортежів.

4. Обчислення власних кортежів

$$\chi_{Uj} : \chi_{Uj} = \frac{\vartheta_{\alpha N, K} v_i}{\lambda_i}, i = \overline{1, r} \quad \text{оператора}$$

$$\vartheta_{\alpha N, K} \vartheta_{\alpha N, K}^*.$$

5. Обчислення скалярних добутків $(\chi_{Uj}, \chi_Y)_{N, m \times n}, i = \overline{1, r}$.

6. Обчислення $\hat{\beta}$ за формулою

$$\hat{\beta} = \sum_{k=1}^r \lambda_k^{-1} v_k (\chi_{Uj}, \chi_Y)_{N, m \times n}.$$

17. Варіант обчислювальної формули розв'язку через задачу на власні значення для матриці Грама:

$$\begin{aligned} \hat{\beta} &= \sum_{k=1}^r \lambda_k^{-1} v_k (\lambda_k^{-1} \vartheta_{\alpha} v_k, \chi_Y)_{N, m \times n} = \\ &= \sum_{k=1}^r \lambda_k^{-2} v_k (v_k, \vartheta_{\alpha}^* \chi_Y) \end{aligned}$$

Далі залишається врахувати, що

$$\vartheta_{\alpha N, K}^* \chi = \begin{pmatrix} (\chi, \chi_{N,1}^{(1)})_{N, m \times n} \\ \dots \\ (\chi, \chi_{N,1}^{(K)})_{N, m \times n} \end{pmatrix}, \chi \in R^{N, m \times n}.$$

Приклад використання матричної множинної регресії для прогнозування показників телеперегляду

Частка телеканалу – показник, що характеризує частку телевізійної аудиторії, яка дивилася певний телеканал у певному часовому інтервалі відносно загальної телевізійної аудиторії у той самий часовий період [10]. Розв'язується задача прогнозування річної частки шести телеканалів для годинних інтервалів у межах усередненого тижня 2013 року з дискретизацією за днями тижня (понеділок, вівторок, ..., неділя). Використано дані 2006-2013 років з погодинною дискретизацією за днями тижня для усередненого тижня; дані 2013 року використовуються для аналізу модельної точності прогнозування. Повна таблиця даних не наводиться з причини її великого розміру. Задача розв'язується з використанням вище описаного алгоритму обчислення оцінки вектора невідомих параметрів, що застосовується для річних часток

каналів, дискретизованих до матриць розмірності 24x7 (години х дні тижня).

1. Побудова стовпчикових «компонентних» кортежів $\chi_{N,1}^{(j)}, j = \overline{1, K}$. Формуються на основі річних даних. Матричне рівняння для кожного з каналів:

$$\begin{aligned} Y(i) &= \beta_1 \chi_{6,1}^{(2006)}(i) + \beta_2 \chi_{6,1}^{(2007)}(i) + \dots + \\ &+ \beta_6 \chi_{6,1}^{(2011)}(i) \equiv Y^{2012}(i), i = \overline{1, 6} \end{aligned} \quad (4)$$

Для знаходження $\hat{Y}^{2013}(i), i = \overline{1, 6}$ із (4) шукаємо оцінку вектора параметрів $\hat{\beta}$.

2. Побудова матриці Грама Γ стовпчикових «компонентних» кортежів – у нашому прикладі матриця покомпонентних скалярних добутків розмірності 6x6.

3. Обчислення ненульових сингулярностей $(v_i, \lambda_i^2), i = \overline{1, r}$ матриці Грама Γ стовпчикових «компонентних» кортежів.

$$\begin{aligned} v_1 &= (0.5, -0.3, -0.6, -0.3, 0.1, 0.5), \lambda_1 = 71.7 \\ v_2 &= (0.5, 0.4, 0.4, 0.4, 0.4, 0.4), \lambda_2 = 759.9 \\ v_3 &= (-0.6, -0.4, 0.0, 0.3, 0.5, 0.4), \lambda_3 = 141.5 \\ v_4 &= (0.0, -0.1, -0.1, 0.6, -0.7, 0.3), \lambda_4 = 31.6 \\ v_5 &= (-0.1, 0.5, -0.7, 0.4, 0.3, -0.3), \lambda_5 = 40.5 \\ v_6 &= (0.4, -0.6, 0.0, 0.4, 0.2, -0.5), \lambda_6 = 50.7 \end{aligned}$$

$$4. \quad \text{Обчислення} \quad \chi_{Uj} : \chi_{Uj} = \frac{\vartheta_{\alpha N, K} v_i}{\lambda_i}, i = \overline{1, r}.$$

Отримуємо шість власних кортежів-матриць розмірності 144x7, які не наводимо з причини їх великого розміру.

5. Обчислення скалярних добутків $sk_i = (\chi_{Uj}, \chi_Y)_{N, m \times n}, i = \overline{1, r}$.

$$\begin{aligned} sk_1 &= 26.39; sk_2 = 282.31; sk_3 = 53.30 \\ sk_4 &= 9.99; sk_5 = -6.52; sk_6 = -29.59 \end{aligned}$$

6. Обчислення $\hat{\beta}$, що дозволяє обчислити прогноз показників частки шести каналів для 2013 року з погодинною дискретизацією для днів тижня $\hat{\beta} = (-0.11, 0.16, 0.01, 0.05, -0.05, 0.92)$.

Точність прогнозу розраховуємо за критерієм $APE = \left| \frac{y - \hat{y}}{y} \right|$. Середнє значення критерію для кожного з каналів наведено у Таблиці 2 для доби

і для часового інтервалу 18:00-24:00 – найбільш комерційно значимого інтервалу. Найбільше значення критерію APE досягається у нічних часових інтервалах, що не є суттєвим для комерційного планування. Як бачимо, точність у комерційних часових інтервалах більш висока, що особливо важливо для комерційного планування бізнес-процесів.

Таблиця 1

Значення критерію APE для частки шести каналів 2013 року

Канал	сутки	18:00-24:00
Канал 1	15,5%	17,1%
Канал 2	19,3%	16,0%
Канал 3	10,1%	8,3%
Канал 4	18,3%	13,4%
Канал 5	12,6%	12,4%
Канал 6	13,0%	10,0%

Порівняємо середні погодинні частки прогнози, планові і фактичні на прикладі окремих каналів. Канал 3: усереднені частки вихідного дня наведено на Рис 1, буднів – на Рис 2. У середньому значенні критерію APE для планових показників Каналу 3 складає 11,5%, а для прогнозного з використанням матричної множинної регресії і апарату псевдоінверсії 6,9%. Така різниця є зрозумілою, якщо канал планує ріст. Проте якщо ріст не планується, то план є оптимістичним, що і підтверджує прогноз. На цьому прикладі показано, що використання прогнозних показників може допомогти перевірити реалістичність планових показників.

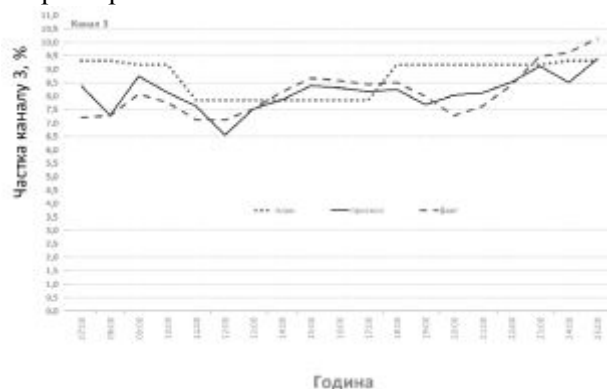


Рис. 1 Частка каналу 3 у вихідні дні з погодинною дискретизацією

Для каналу 5: на Рис. 3-4 наведено результати прогнозування. Середнє значення критерію APE для планових показників 10,7%, для прогнозних - 7,2%. Таким чином, прогнозні показники більш точні по відношенню до факту 2013 року. У той самий час, планові частки Каналу 5 більш реалістичні, ніж планові показники Каналу 3. Таким чином, матрична множинна регресія може

бути використана для задач прогнозування показників медіабізнесу з прийнятною для планування бізнес-процесів точністю.

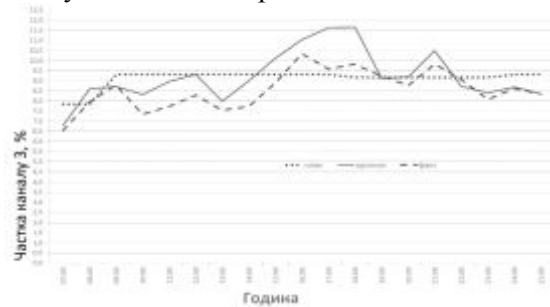


Рис. 2 Частка каналу 3 у будні дні з погодинною дискретизацією

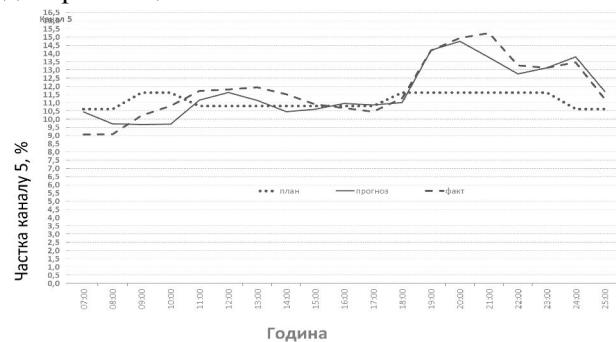


Рис. 3 Частка каналу 5, вихідні дні

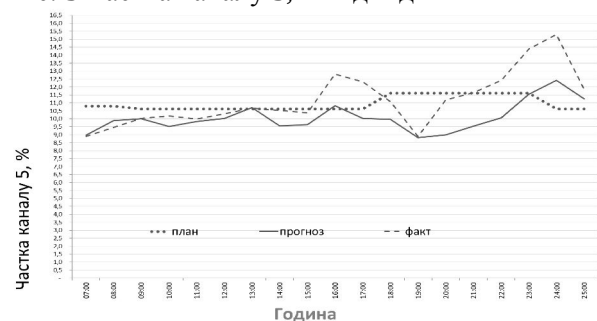


Рис. 4 Частка каналу 5, будні дні

Висновки

Запропоновано до розгляду клас матричних функцій набору матричних аргументів: матричну множинну регресію. У рамках розвитку концепції кортежних операторів розроблено математичний апарат сингулярного (SVD) розкладу і техніку псевдообернення (ПдО) за Муром-Пенроузом, що дозволяє поставити і конструктивно розв'язати задачу оцінки методом найменших квадратів (МНК) для множинної матричної регресії. Запропоновано і реалізовано алгоритм МНК-оцінювання для вектора невідомих параметрів згаданого класу матричних функцій у розв'язуванні задач прогнозу.

Список використаних джерел

1. Donchenko V. Recurrent procedure in solving the Grouping Information Problem in Applied Mathematics / V. Donchenko, Ju. Krak, Yu. Krivonos // *ITHEA: International Journal Information Theories and Applications*. – 2012. – Vol.1, No 1. – P. 62 -77.
2. Донченко В. Евклидовы пространства числовых векторов и матриц: конструктивные методы описания базовых структур и их использование / В. Донченко // *International Journal “Information technologies & Knowledge”*. – 2011. – Vol. 5, No 3. – P. 203-216.
3. Moore E.H. On the reciprocal of the general algebraic matrix. / E.H. Moore // *Bulletin of the American Mathematical Society*. – 1920. – No 26, – P. 394-395.
4. Penrose R. A generalized inverse for matrices / R. Penrose // *Proceedings of the Cambridge Philosophical Society* – 1955. – 51. – P. 406-413.
5. Albert A.E. Regression and the Moore-Penrose pseudoinverse / A.E. Albert. – Academic Press, 1972, – 180 p.
6. Kirichenko N.F. Analytical Representation of Perturbation of Pseudoinverse Matrices / N.F. Kirichenko // *Cybernetics and Systems Analysis*. – 1997. – Vol.33, No 2. – P. 230-239.
7. Donchenko V. “Feature Vectors” in Grouping Information Problem in Applied Mathematics: Vectors and Matrixes / V. Donchenko, T. Zinko, F. Skotarenko // *Problems of Computer Intellectualization: international conference, Institute of Cybernetics NASU, ITHEA*. – Kyiv, Ukraine, - Sofia, Bulgaria. – 2012. – P. 111-124.
8. Donchenko V. Vectors and matrixes least square method: foundation and application examples / V. Donchenko, I. Nazaraga, O. Tarasova // *International Journal Information Theories and Applications*. – 2013. – Vol.20, No 4. – P.311-322.
9. Donchenko V. Matrixes least squares method and examples of its application / V. Donchenko, I. Nazaraga, O. Tarasova // *International Journal Information Technologies & Knowledge*. – 2013. – Vol.7, No 4. – P.325–336.
10. Назаров М. Зарубежные рынки телевизионной рекламы: сравнительное

исследование / М. Назаров – Москва: ООО «НИПКЦ Восход–А», 2011. – 364 с.

References

1. DONCHENKO, V., KRAK, JU., KRIVONOS, YU. (2012) Recurrent procedure in solving the Grouping Information Problem in Applied Mathematics *ITHEA: International Journal Information Theories and Applications*. 1(1). p. 62 -77.
 2. DONCHENKO, V. (2011) Evklidovy prostranstva chislovyh vektorov I matrix: konstruktivniie metody opisaniia bazovyh struktur i ih *International Journal “Information technologies & Knowledge”*. 5(3). p. 203-216.
 3. MOORE, E.H. (1920) On the reciprocal of the general algebraic matrix *Bulletin of the American Mathematical Society*. 26. p. 394-395.
 4. PENROSE, R. (1955) A generalized inverse for matrices *Proceedings of the Cambridge Philosophical Society*. 51. p. 406-413.
 5. ALBERT, A.E. (1972) *Regression and the Moore-Penrose pseudoinverse*. Academic Press.
 6. KIRICHENKO, N.F. (1997) Analytical Representation of Perturbation of Pseudoinverse Matrices *Cybernetics and Systems Analysis*. 33(2). p. 230-239.
 7. DONCHENKO, V., ZINKO, T., SKOTARENKO, F. (2012) “Feature Vectors” in Grouping Information Problem in Applied Mathematics: Vectors and Matrixes. In *Problems of Computer Intellectualization Conference, ITHEA*, Kyiv, Ukraine, - Sofia, Bulgaria: Institute of Cybernetics NASU. pp. 111-124.
 8. DONCHENKO, V., NAZARAGA, I., TARASOVA, O. (2013) Vectors and matrixes least square method: foundation and application examples *International Journal Information Theories and Applications*. 20(4) p. 311-322.
 9. DONCHENKO, V., NAZARAGA, I., TARASOVA, O. (2013) Matrixes least squares method and examples of its application *International Journal Information Technologies & Knowledge*. 7(4) p. 325–336.
 10. NAZAROV, M. (2011) *Zarubezhniie rynki televizionnoj reklamy: sravnitelnoe isledovani*. Moskva: ООО NIKPTs Voshod-A.
- Надійшла до редколегії 24.03.2015

УДК 519.72

І.О. Завадський, к. ф.-м. н., доцент

Igor O. Zavadskyi, Associate Professor, Ph. D.

Префіксні (Δ, k) -коди

Prefix (Δ, k) -codes

Київський національний університет імені
Тараса Шевченка, 03680, м. Київ, пр-т.
Глушкова 4д, e-mail: zava@ukr.net

Taras Shevchenko National University of Kyiv,
03680, Kyiv, Glushkova av., 4d

Описано новий клас універсальних префіксних кодів, відмінною рисою яких є можливість використання кількох роздільників кодових слів в одному коді. Доведено такі властивості кодів описаного класу, як повнота та універсальність.

Ключові слова: префіксні коди, (Δ, k) -коди, повнота коду, універсальні коди.

The new family of universal prefix codes is introduced. The distinctive feature of this family is the use of multiple codeword delimiters in the same code. Such properties of representatives of this code family as completeness and universality are proved. The delimiters are the bit groups of the form $01\dots 10$ containing m_1 or... or m_t ones and the code with such delimiters is denoted by $C_{m_1, \dots, m_t}$. The bit sequence of the form $1\dots 10$ containing m_1 or... or m_t ones is called delimiting sequence for the code $C_{m_1, \dots, m_t}$. This code contains all codewords that consist of only some delimiting sequence and also all codewords satisfying the following conditions: (i) codeword doesn't start with the delimiting sequence; (ii) codeword ends with the bit sequence $0D$ and (iii) codeword doesn't contain the bit sequence $0D$ anywhere except the end, where D is the delimiting sequence. The name " (Δ, k) -codes" is due to representation of codewords as sequences of (Δ, k) -groups of bits. Each group is constructed as a concatenation of codes of Δ - and k -values. The value of Δ is represented by one bit in fact, while the code of k -value has the form $1\dots 10$ containing zero or more ones.

Key words: compression code, $(2, 3)$ -representation, text compression, prefix code.

Статтю представив д.ф.-м.н., проф. Анісімов А.В.

Вступ

У роботі [1] було розглянуто префіксний код C_1 , що базується на поданні цілих чисел у двобазисній системі з основним базисом 2 і додатковим базисом 3, яке було введено у [2]. Цей код можна досить ефективно застосовувати для стиснення інформації, передусім природномовних текстів. Оскільки код C_1 – префіксний, то він також і однозначно декодований. Однак цей код не є повним. Це означає, що множину кодових слів можна розширити так, що властивість однозначної декодованості не втратиться. Таке розширення приводить до коду C_2 , описаного в [1].

Код C_2 з [1] можна описати як код із роздільником 0110. У цій статті буде розглянуто цілу родину подібних стискальних кодів, кожен із яких має кілька роздільників вигляду $01\dots 10$, що містять m_1 або m_2 або... або m_t одиниць. Їх ми називатимемо (Δ, k) -кодами, а представника такого сімейства кодів позначатимемо як $C_{m_1, \dots, m_t}$. Наприклад, $C_{2,3}$ – це код із роздільниками 0110 та 01110.

Означення (Δ, k) -кодів

Означення 1. Код $C_{m_1, \dots, m_t}$ містить усі слова вигляду $1\dots 10$ із m_1 або... або m_t одиницями, а також усі слова, що задовольняють таким умовам:

- слово не починається з послідовності $1\dots 10$, що містить m_1 або m_2 або... або m_t одиниць;
- слово закінчується послідовністю $01\dots 10$, що містить m_1 або m_2 або... або m_t одиниць;
- слово не містить послідовності $01\dots 10$ із m_1 або m_2 або... або m_t одиницями ніде, крім кінця.

Тут і надалі ми припускаємо, що $0 < m_1 < \dots < m_t$.

Назва « (Δ, k) -коди» пов'язана з тим, що слова цих кодів можуть бути подані у вигляді послідовностей так званих (Δ, k) -груп бітів, що означаються в такий спосіб.

Означення 2. (Δ, k) -група утворюється з'єднанням двійкових кодів значень Δ і k . Для коду $C_{m_1, \dots, m_t}$ величина Δ може набувати $2^d + 1$ значень, де d – деякий параметр, $0 \leq d < m_1$. Значення Δ подаються рядком “0” та всіма двійковими рядками довжини $d + 1$ із першим символом “1”. Параметр k – це будь-яке додатне ціле число. Його значення подається у вигляді послідовності $1 \dots 10$, що містить $k-1$ одиниць.

Зв'язок між (Δ, k) -групами та кодами $C_{m_1, \dots, m_t}$ встановлює така теорема.

Теорема 1. Будь-яке слово з будь-якого коду $C_{m_1, \dots, m_t}$ може бути інтерпретоване як послідовність (Δ, k) -груп бітів.

Доведення. Якщо слово має вигляд $1 \dots 10$, то, згідно з означенням 1, його довжина є не меншою за $m_1 + 1$. Але, згідно з означенням 2, довжина коду Δ не перевищує m_1 і код Δ не включає в себе всього слова вигляду $1 \dots 10$, а отже, решта цього слова може бути інтерпретована як код k .

Розглядаючи інші кодові слова, зауважимо, що початок будь-якої двійкової послідовності може бути інтерпретовано і як код деякого значення Δ , і як код деякого значення k . Вилучимо з початку кодового слова код Δ , потім – код k і отримаємо коротше кодове слово, до якого застосуємо ту саму процедуру, аж поки не поділимо все слово на (Δ, k) -групи бітів або не отримаємо залишок, у який не вкладається ціла (Δ, k) -група. Подивимось, як згідно з цією процедурою оброблятиметься закінчення кодового слова, що має вигляд $01 \dots 10$ і містить не менше m_1 одиниць. Перший біт “0” цього закінчення може бути або кінцем деякого коду k , або належати коду Δ . У першому випадку на останній ітерації отримуємо залишок $1 \dots 10$ із не менш ніж m_1 одиницями, що, як показано вище, є повною (Δ, k) -групою. У другому випадку зазначимо, що код Δ містить не більш ніж m_1 бітів і після його вилучення отримаємо послідовність вигляду $1 \dots 10$, яка представляє певне значення k . Таким чином, ситуація, коли на останній ітерації отримується залишок, у який не вкладається ціла (Δ, k) -група, неможлива, і будь-яке слово з будь-якого коду $C_{m_1, \dots, m_t}$ складається з цілої кількості (Δ, k) -груп бітів. ■

Зауважимо, що теорема 1 справедлива для будь-яких значень параметра d , що задовольняють нерівності $0 \leq d < m_1$. Тому

надалі для спрощення вважатимемо, що $d=0$, тобто код величини Δ складається з одного біта.

Алгоритми кодування та декодування

На вхід кодувального алгоритму надходить натуральне число x , двійковий вигляд якого – $x_n x_{n-1} \dots x_0$, а на виході отримуємо слово з коду $C_{m_1, \dots, m_t}$. Нехай $J = \{j_1, j_2, \dots\}$ – така нескінченна зростаюча послідовність натуральних чисел, що j_1 – найменше число, яке не належить множині $\{m_1, \dots, m_t\}$, а j_i – таке найменше число, що $j_i > j_{i-1}$ і $j_i \notin \{m_1, \dots, m_t\}$. Стандартний алгоритм кодування є таким:

1) $x \leftarrow x - 2^n$, тобто видаляємо найстарший біт числа x , який завжди дорівнює 1.

2) Якщо $x=0$, до рядка $x_{n-1} \dots x_0$, який містить лише нулі або порожній, дописуємо справа послідовність $1 \dots 10$ з m_1 одиницями. Шукане кодове слово має вигляд $x_{n-1} \dots x_0 1 \dots 10$. Кінець алгоритму.

3) Якщо двійкове подання x набуває вигляду рядка $1 \dots 10$ або $0 \dots 01 \dots 10$ із m_2 або... або m_t одиницями, це і є шукане кодове слово. Кінець алгоритму.

4) Розглядаємо кожну бітову послідовність вигляду $01 \dots 10$, що не є розташованою в кінці слова послідовністю із m_2 або ... або m_t одиницями. Якщо в цій послідовності d одиниць, замінюємо її подібною послідовністю $01 \dots 10$, що містить j_d одиниць.

5) Якщо слово починається з послідовності $1 \dots 10$, закінчується послідовністю $01 \dots 1$, або цілком складається з послідовності вигляду $1 \dots 1$, замінюємо кількість одиниць у цих послідовностях із d на j_d .

6) Якщо слово закінчується послідовністю $01 \dots 10$, яка містить m_2 або ... або m_t одиниць, це шукане слово. Кінець алгоритму.

7) Допишуємо до слова справа рядок $01 \dots 10$, що містить m_1 одиниць.

Згідно з цим алгоритмом, якщо $x - 2^n \neq 0$, роздільник $01 \dots 10$ із m_1 одиницями дописується штучно і тому має бути видалений під час декодування, а роздільники вигляду $01 \dots 10$ із m_2 або ... або m_t одиницями є інформативними частинами кодових слів і тому під час декодування повинні оброблятися. Якщо ж $x - 2^n = 0$, останні $m_1 + 1$ бітів вигляду $1 \dots 10$ мають бути видалені.

Далі описано стандартний алгоритм декодування, на вхід якого подається кодове слово, а на виході отримуємо число.

1) Якщо кодове число має вигляд $0\dots 01\dots 10$ або $1\dots 10$ і містить m_1 одиниць, видаляємо останні m_1+1 бітів і переходимо на крок 5.

2) Якщо число закінчується послідовністю $01\dots 10$, яка містить m_1 одиниць, видаляємо останні m_1+2 бітів.

3) Замінюємо кожну послідовність вигляду $01\dots 10$ із $j_d \in J$ одиницями, кінець якої не збігається з кінцем кодового слова, послідовністю подібного вигляду $01\dots 10$, що містить d одиниць.

4) Якщо слово починається з рядка $1\dots 10$ із $j_d \in J$ одиницями, замінюємо його подібним рядком $1\dots 10$, що містить d одиниць.

5) Допишуємо символ "1" до початку числа.

Повнота (Δ, k) -кодів

Із самого факту існування наведеної вище декодувальної процедури випливає, що (Δ, k) -коди є роздільними, однак це можна також довести, перевіривши виконання нерівності Крафта-Макмілана: $\sum_{c \in C} 2^{-|c|} \leq 1$, де C – певний код. Крім того, усі (Δ, k) -коди мають ще одну важливу властивість: вони повні. Це означає, що до будь-якого коду з цього сімейства неможливо додати жодного слова так, щоб властивість роздільності не втратилася. Усі ці властивості доведено в теоремі 2.

Теорема 2. Будь-який (Δ, k) -код $C_{m_1, \dots, m_t}$ є роздільним і повним.

Доведення. Необхідною та достатньою умовою роздільності та повноти деякого коду C є виконання нерівності Крафта-Макмілана як рівності: $\sum_{c \in C} 2^{-|c|} = 1$. Якщо f_n – кількість кодових слів довжини n , то цю рівність можна переписати у вигляді

$$\sum_{n=1}^{\infty} 2^{-n} f_n = 1 \quad (1)$$

Розглянемо (Δ, k) -код $C_{m_1, \dots, m_t}$. Для будь-якого $n \geq 2$ існує дві (Δ, k) -групи довжини n : $1\dots 10$ і $01\dots 10$, що містять $n-1$ та $n-2$ одиниць відповідно. Серед усіх (Δ, k) -груп ті, що містять m_1, m_2 або $\dots$ або m_t одиниць, є термінальними, тобто такими, що можуть трапитися лише в кінці кодового слова. Отже, для коду $C_{m_1, \dots, m_t}$ існує $2t$ термінальних груп, довжини яких – $m_1+1, m_1+2, \dots, m_t+1, m_t+2$. Позначимо кількість термінальних груп довжини n через T_n . Очевидно, T_n дорівнює кількості входжень числа n в набір $\{m_1+1, m_1+2, \dots, m_t+1, m_t+2\}$ (ця кількість становить 0, 1 або 2), а кількість нетермінальних груп довжини n дорівнює $2-T_n$.

Розглянемо кодові слова довжини n , що містять принаймні дві (Δ, k) -групи. Кожне таке слово можна отримати додаванням зліва його першої нетермінальної (Δ, k) -групи до коротшого кодового слова. З іншого боку, додавання до будь-якого кодового слова зліва довільної нетермінальної групи утворює довше кодове слово. Якщо ж кодове слово містить тільки одну групу, то ця група є термінальною. Отже, врахувавши, що довжина найкоротшої (Δ, k) -групи становить 2, отримуємо таку рекурентну формулу для обчислення кількості кодових слів довжини n :

$$\begin{aligned} f_n &= T_n + \sum_{k=0}^{n-2} (2 - T_k) f_k = \\ &= T_n + 2(f_{n-2} + f_{n-3} + \dots) - \\ &f_{n-(m_1+1)} - \dots - f_{n-(m_t+1)} - \\ &f_{n-(m_1+2)} - \dots - f_{n-(m_t+2)} \end{aligned} \quad (2)$$

Застосуємо цю формулу для обчислення f_{n-1} :

$$\begin{aligned} f_{n-1} &= T_{n-1} + \sum_{k=0}^{n-3} (2 - T_k) f_k = \\ &= T_{n-1} + 2(f_{n-3} + f_{n-4} + \dots) - \\ &f_{n-(m_1+2)} - \dots - f_{n-(m_t+2)} - \\ &f_{n-(m_1+3)} - \dots - f_{n-(m_t+3)} \end{aligned} \quad (3)$$

Знайдемо праву частину (3) в (2) і замінимо її на f_{n-1} :

$$\begin{aligned} f_n &= T_n - T_{n-1} + 2f_{n-2} + f_{n-1} - \\ &f_{n-m_1-1} - \dots - f_{n-m_t-1} + \\ &f_{n-m_1-3} + \dots + f_{n-m_t-3} \end{aligned} \quad (4)$$

Позначивши ліву частину (1) через s та врахувавши, що $f_0 = f_{-1} = \dots = 0$, для будь-якого $p > 0$ матимемо такі рівності: $\sum_{n=1}^{\infty} 2^{-n} f_{n-p} = 2^{-p} \sum_{n=1}^{\infty} 2^{-(n-p)} f_{n-p} = s 2^{-p}$. Враховуючи їх та підставляючи вираз (4) в (1), отримаємо:

$$\begin{aligned} s &= \sum_{n=1}^{\infty} 2^{-n} f_n = \sum_{n=1}^{\infty} 2^{-n} (T_n - T_{n-1} + \\ &f_{n-1} + 2f_{n-2} - f_{n-(m_1+1)} - \dots - f_{n-(m_t+1)} + \\ &f_{n-(m_1+3)} + \dots + f_{n-(m_t+3)}) = \\ &\sum_{n=1}^{\infty} 2^{-n} T_n - \frac{1}{2} \sum_{n=1}^{\infty} 2^{-(n-1)} T_{n-1} + \\ &s \left(\frac{1}{2} + \frac{1}{2} - 2^{-m_1-1} - \dots - 2^{-m_t-1} + \right. \\ &\left. 2^{-m_1-3} + \dots + 2^{-m_t-3} \right) \end{aligned} \quad (5)$$

Взявши до уваги, що $2^{-m_i-3} - 2^{-m_i-1} = -3 \cdot 2^{-m_i-3}$ для будь-якого i , $\sum_{n=1}^{\infty} 2^{-n} T_n = \sum_{n=1}^{\infty} 2^{-(n-1)} T_{n-1}$ і скоротивши s в обох частинах (5), отримаємо таку формулу:

$$3s \sum_{i=1}^t 2^{-m_i-3} = \frac{1}{2} \sum_{n=1}^{\infty} 2^{-n} T_n \quad (6)$$

Оскільки довжини термінальних (Δ, k) -груп становлять $m_1 + 1, m_1 + 2, \dots, m_t + 1, m_t + 2$,

виконується рівність $\sum_{n=1}^{\infty} 2^{-n} T_n = \sum_{i=1}^t (2^{-m_i-1} + 2^{-m_i-2}) = \frac{3}{4} \sum_{i=1}^t 2^{-m_i}$. Тому нерівність (6) набуває вигляду $\frac{3}{8} s \sum_{i=1}^t 2^{-m_i} = \frac{3}{8} \sum_{i=1}^t 2^{-m_i}$, звідки випливає $s = 1$. ■

Універсальність (Δ, k) -кодів

Набір кодових слів із довжинами l_i ($l_1 \leq l_2 \leq \dots$) називається універсальним у розумінні Еліаса [3], якщо існує така константа K , що для будь-якого скінченного розподілу ймовірностей $P = (p_1, \dots, p_n)$, де $p_1 \geq p_2 \geq \dots$ виконується нерівність

$$\sum_{i=1}^n l_i p_i \leq K \cdot \max(1, E(P)), \quad (7)$$

де $E(P) = -\sum_{i=1}^n p_i \log_2 p_i$ – ентропія розподілу P , а K – незалежна від P константа.

Теорема 3. Будь-який (Δ, k) -код $C_{m_1, \dots, m_t}$ є універсальним.

Доведення. Оцінимо величину s_n – кількість кодових слів, довжина яких не перевищує n . Для цього скористаємося рівністю (4):

$$s_n = \sum_{i=1}^n f_i = \sum_{i=1}^n (T_i - T_{i-1}) + 2 \sum_{i=1}^n f_{i-2} + f_{n-1} + \sum_{i=1}^{n-1} f_{i-1} - f_{n-m_1-1} - \dots - f_{n-m_t-1} - f_{n-m_1-2} - \dots - f_{n-m_t-2} - \sum_{i=1}^{n-2} (f_{i-m_1-1} + \dots + f_{i-m_t-1}) + \sum_{i=1}^n (f_{i-m_1-3} + \dots + f_{i-m_t-3}) \quad (8)$$

Зауважимо, що $\sum_{i=1}^n (T_i - T_{i-1}) = T_n - T_0 = T_n$, оскільки $T_0 = 0$; $2 \sum_{i=1}^n f_{i-2} + \sum_{i=1}^{n-1} f_{i-1} = 3 \sum_{i=1}^{n-2} f_i + 3f_0 + 2f_{-1} = 3 \sum_{i=1}^{n-2} f_i$, оскільки $f_0 = f_{-1} = 0$; $-\sum_{i=1}^{n-2} (f_{i-m_1-1} + \dots + f_{i-m_t-1}) + \sum_{i=1}^n (f_{i-m_1-3} + \dots + f_{i-m_t-3}) = f_{-m_1-1} + \dots + f_{-m_t-1} + f_{-m_1-2} + \dots + f_{-m_t-2} = 0$, оскільки $f_i = 0$, якщо $i < 0$. Таким чином, рівність (8) можна перетворити так:

$$s_n = T_n + 3 \sum_{i=1}^{n-2} f_i - f_{n-m_1-1} - \dots - f_{n-m_t-1} - f_{n-m_1-2} - \dots - f_{n-m_t-2} > 3s_{n-m_t-2}.$$

Безпосередньо з означення 1 випливає, що $s_i > 0$, якщо $i \geq m_1$. Тому як тільки

$$n - m_t - 2 \geq m_1, \quad (9)$$

виконується нерівність $s_n > \left(\sqrt[m_t+2]{3} \right)^{n-m_1}$. Позначивши $\sqrt[m_t+2]{3}$ через α , а $1/\alpha^{m_1}$ – через β , отримаємо нерівність $s_n > \beta \alpha^n$, де $\alpha > 1$ і $\beta > 0$ – деякі константи, що залежать від m_1 і m_t . Таким чином, якщо l_i – це довжина i -го кодового слова у впорядкованому за зростанням довжин наборі кодових слів, то $i > s_{l_i-1} > \beta \alpha^{l_i-1}$. Звідси отримуємо

$$l_i < \log_{\alpha}(i/\beta) + 1 = \log_{\alpha} 2 \cdot \log_2 i - \log_{\alpha} \beta + 1 \quad (10)$$

Ця нерівність виконується, якщо $l_i \geq m_1 + m_t + 2$ (умова (9)). Тоді $i \geq 2$, а отже, $\log_2 i \geq 1$ і нерівність (10) можна переписати у вигляді $l_i < (\log_{\alpha} 2 - \log_{\alpha} \beta + 1) \log_2 i = (\log_{\alpha} (2/\beta) + 1) \log_2 i$. Але оскільки величини p_i розташовано у незростаючому порядку, а їх сума дорівнює 1, то $p_i \leq 1/i$, а отже, $\log_2 i \leq -\log_2 p_i$; $l_i p_i < -(\log_{\alpha} (2/\beta) + 1) p_i \log_2 p_i$, а отже, константою K в нерівності (7) може бути величина $K_0 = \log_{\alpha} (2/\beta) + 1 = \log_{\alpha} 2 \alpha^{m_1} + 1$. Враховуючи, що $\alpha = \sqrt[m_t+2]{3}$, маємо $K_0 = \log_3 2 (m_t + 2) + m_1 + 1$. Якщо ж умова (9) не виконується, то ентропія $E(P)$ з нерівності (7) обмежена величиною $\log_2(m_1 + m_t + 2)$, а тому нерівність (7) виконується для певної величини $K = K_1$, що залежить від параметрів коду m_1 і m_t і не залежить від розподілу P . Таким чином, константою K з нерівності (7) може бути більша з величин K_0, K_1 . ■

Список літератури

1. Завадський І.О. Префіксний стискальний код на основі нижнього (2,3)-подання чисел / І.О. Завадський // Вісник Київського національного університету ім. Т. Шевченка. Серія: фіз.-мат. науки. – 2015. – №1. – с. 124-129.
2. Анисимов А.В. Помехоустойчивое префиксное кодирование на основе нижнего (2,3)-представления чисел / А.В. Анисимов, И.А. Завадский // Кибернетика и системный анализ. – 2014. – № 2. – С. 3–14.
3. Elias P. Universal codeword sets and representation of the integers / P.Elias // IEEE Trans. Inform. Theory – 1975. – v. 12, p. 194–203.

References

1. ZAVADSKYI, I.O. (2015) Prefix Compressing Code by Means of the 2,3-Representation of Numbers, *Bulletin of Taras Shevchenko National University of Kyiv. Series Physics & Mathematics*. No 3. pp. 124-129.
2. ANISIMOV, A.V., ZAVADSKYI, I.O. (2014) Pomekhoustojchivoe prefiksnoe kodirovanie na osnove nizhnego (2,3)-predstavlenija chisel. *Kibernetika i systemnyj analiz*. No 2, pp. 3–14.
3. ELIAS P. (1975) Universal codeword sets and representation of the integers. *IEEE Trans. Inform. Theory*. vol. 12, pp. 194–203.

Надійшла до редколегії 10.03.2015

УДК 530.145

Завадський І.О.¹, к.ф.-м.н., доц.,
Петльована М.В.², студентка.

Квантовий вентиль Тоффолі, його реалізація та застосування

Київський національний університет імені
Тараса Шевченка, 83000, м. Київ, Глушкова 4д

¹e-mail: ihorza@gmail.com

²e-mail: mariapetlovana@gmail.com

I.O. Zavadskyi¹, k. ph.-mat. sc., docent,
M.V. Petlovana², student.

The quantum Toffoli gate, its implementation and application

Taras Shevchenko National University of Kyiv,
83000, Kyiv, Glushkova st., 4d

¹e-mail: ihorza@gmail.com

²e-mail: mariapetlovana@gmail.com

В статті розглядається поняття квантового вентиля Тоффолі та запропоновано його реалізацію через 1- та 2-кубітні вентиля.

Ключові слова: квантовий вентиль, квантова схема, вентиль Тоффолі.

The concepts of quantum circuit, quantum gate including Toffoli gate are considered in the article. The implementations of quantum gates FANOUT, NOT (negation) and AND (conjunction) with only one Toffoli gate are presented. Thus one of the effective ways to solve the problem of modeling classical computation on a quantum computer is proposed. Toffoli gate is scrutinized as an element of a set of universal quantum gates to which any operation possible on a quantum computer can be reduced. The importance of the reversibility of the Toffoli gate is explained. The concept of the complexity of quantum computation is defined. The implementation of the Toffoli gate by circuit containing only 1- and 2-qubit gates is presented and compared with other known implementation. The proposed circuit requires 5 2-qubit gates and 2 1-qubit gates, while known implementation requires 6 2-qubit controlled-NOT gates and 10 1-qubit operations. Thus the proposed circuit significantly reduces the number of elementary gates needed for the implementation of the Toffoli gate.

Key Words: quantum gate, quantum circuit, Toffoli gate.

Статтю представив д.ф.-м.н., проф. Анісімов А.В.

1. Вступ

У схематичній моделі класичного комп'ютера будь-яке обчислення розглядається як дія схеми, побудованої з різних типів булевих логічних елементів (вентилів), що обробляють деякі бінарні вхідні дані. Кожний логічний елемент має входи та виходи й перетворює вхідні біти в один або кілька вихідних за певними формулами. Виходи одних вентилів ототожнюються із входами інших. Подібним чином можна змоделювати й обчислення на квантовому комп'ютері. Головна відмінність полягає в тому, що в той час як класичні логічні вентиля оперують класичними значеннями біта – 0 або 1, квантові вентиля оперують так званими кубітами, для яких 0 та 1 – це лише базисні стани. Загалом стан квантової частки – це суперпозиція базисних станів із певними комплексними коефіцієнтами, що називаються амплітудами. Стан системи з n квантових часток описується в загальному випадку вектором із 2^n амплітуд, а

отже, квантові обчислення значно різноманітніші, ніж класичні.

З іншого боку, квантові вентиля мають обмеження, не властиві класичним. Насамперед це оборотність. Кожному квантовому вентилю відповідає деяке унітарне перетворення, а такі перетворення завжди оборотні. Для класичних обчислень, навпаки, властива необоротність. Наприклад, логічний вентиль AND з двома входами й одним виходом реалізує необоротне обчислення, оскільки якщо вихід дорівнює 0, то відновити за ним значення входів неможливо. Тому задача реалізації довільного класичного обчислення на квантовому комп'ютері є нетривіальною. Ця задача має суттєве практичне значення, оскільки, яким би не був квантовий алгоритм, його невід'ємною складовою є виконання того чи іншого «класичного» обчислення. Наприклад, в квантовому алгоритмі факторизації великих чисел це піднесення до степеня. Один із ефективних способів

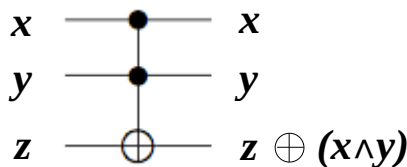
розв'язання задачі моделювання класичних обчислень на квантовому комп'ютері запропоновано в цій статті.

2. Реалізації базових класичних вентилів за допомогою вентиля Тоффолі

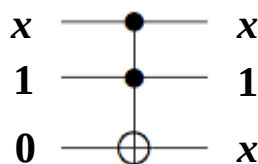
Відомо, що будь-який складний логічний вираз можна звести до однієї з двох стандартних форм: диз'юнктивної або кон'юнктивної. Крім того, диз'юнкція виражається через кон'юнкцію та заперечення. Тому якщо ми реалізуємо вентиля NOT та AND, то зможемо реалізувати будь-яку класичну схему. Крім того, в реальних обчисленнях використовується операція присвоєння, тобто вентиль копіювання значення, який позначимо як FANOUT.

Виявляється, що усі вищеперераховані операції можуть бути реалізовані за допомогою оборотного вентиля Тоффолі (рис. 1). Вентиль Тоффолі також називають Controlled Controlled NOT (CCNOT), оскільки його дія полягає в тому, що третій біт змінює своє значення на протилежне тоді й тільки тоді, коли перші два біти дорівнюють 1. Інакше кажучи, значення перших двох бітів контролює значення третього біта.

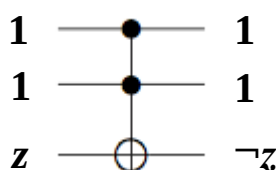
Вентиль Тоффолі



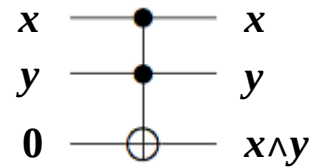
FANOUT



NOT (заперечення)



AND (кон'юнкція)



Оборотність

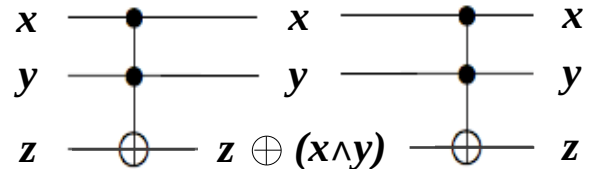


Рис. 1. Властивості вентиля Тоффолі

Унітарна матриця, що відповідає вентилю Тоффолі, має такий вигляд:

$$T = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \end{bmatrix}$$

3. Реалізація вентиля Тоффолі через 1- та 2-кубітні вентиля

Можливість реалізувати вентиль Тоффолі за допомогою 1- та 2-кубітних вентилів є дуже важливою, оскільки фізичні властивості нині відомих пристроїв такі, що суперпозиція перетворень, які обробляють 3 і більше кубітів, веде до експоненційного зростання величини похибок.

На даний момент існує декілька відомих реалізацій вентиля Тоффолі, кожна з яких використовує різні 1- та 2-кубітні перетворення, має свої переваги та недоліки, а також різну складність схеми. Одна з таких реалізацій представлена у джерелі [4] та представлена на рисунку 1.

За допомогою розкладання N -кубітного квантового обчислення на послідовність 1- та 2-кубітних квантових вентилів, можна охарактеризувати складність квантових обчислень як глибину мінімальної квантової схеми, що реалізує його.

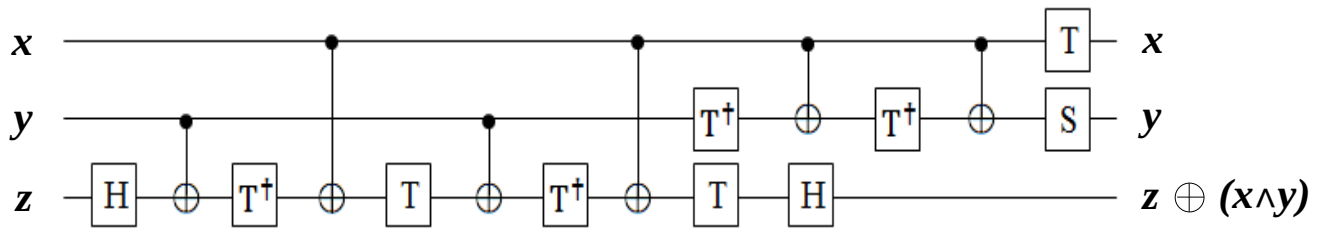


Рис. 2. Реалізація вентиля Тоффолі через 1- та 2-кубітні вентиля H, CNOT, S, T, T†

Реалізація, наведена на рисунку 2, складається з вентилів Адамара, CNOT, вентиля фази S та вентиля $\pi/8$ (в схемі символом $T^\dagger$ позначено транспоноване комплексно спряжене перетворення до вентиля $\pi/8$, або T). Загалом її ємнісна складність становить 16 вентилів.

$$H = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} \quad CNOT = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{bmatrix}$$

$$S = \begin{bmatrix} 1 & 0 \\ 0 & i \end{bmatrix} \quad T = \begin{bmatrix} 1 & 0 \\ 0 & e^{i\pi/4} \end{bmatrix}$$

Перевага даної реалізації полягає в тому, що серед 2-кубітних вентилів вона використовує лише базове перетворення CNOT, яке легко

реалізувати на квантовому комп'ютері. Проте недоліком є те, що використовується багато 1-кубітних перетворень, які збільшують складність і час спрацювання схеми.

У цій статті пропонується квантова схема (рис. 3), що реалізує вентиль Тоффолі і містить лише 7 1- і 2-кубітових вентилів, серед яких відомі вентиля Адамара, CNOT, а також два спеціальних введених нами вентиля N та V:

$$V = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & i \end{bmatrix} \quad N = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & -i \end{bmatrix}$$

Недоліком запропонованої схеми є те, що питання фізичної реалізації перетворень V та N залишається відкритим.

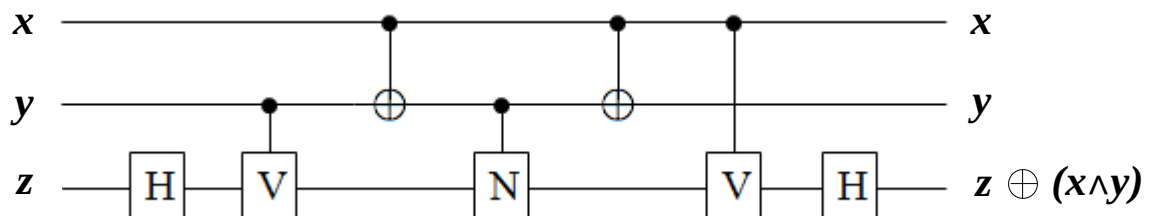


Рис. 3. Реалізація вентиля Тоффолі через 1- та 2-кубітні вентиля H, CNOT, N, V

Покажемо тепер, що наведена вище квантова схема дійсно реалізує перетворення Тоффолі.

Для цього розглянемо перетворення всіх можливих 8 базисних станів трьох кубітів:

$$\begin{aligned} |000\rangle &\xrightarrow{H} |00\rangle \frac{1}{\sqrt{2}} (|0\rangle + |1\rangle) \xrightarrow{V} |00\rangle \frac{1}{\sqrt{2}} (|0\rangle + |1\rangle) \xrightarrow{CNOT} |00\rangle \frac{1}{\sqrt{2}} (|0\rangle + |1\rangle) \xrightarrow{N} |00\rangle \frac{1}{\sqrt{2}} (|0\rangle + |1\rangle) \\ &\xrightarrow{CNOT} |00\rangle \frac{1}{\sqrt{2}} (|0\rangle + |1\rangle) \xrightarrow{V} |00\rangle \frac{1}{\sqrt{2}} (|0\rangle + |1\rangle) \xrightarrow{H} |000\rangle \\ |001\rangle &\xrightarrow{H} |00\rangle \frac{1}{\sqrt{2}} (|0\rangle - |1\rangle) \xrightarrow{V} |00\rangle \frac{1}{\sqrt{2}} (|0\rangle - |1\rangle) \xrightarrow{CNOT} |00\rangle \frac{1}{\sqrt{2}} (|0\rangle - |1\rangle) \xrightarrow{N} |00\rangle \frac{1}{\sqrt{2}} (|0\rangle - |1\rangle) \\ &\xrightarrow{CNOT} |00\rangle \frac{1}{\sqrt{2}} (|0\rangle - |1\rangle) \xrightarrow{V} |00\rangle \frac{1}{\sqrt{2}} (|0\rangle - |1\rangle) \xrightarrow{H} |001\rangle \end{aligned}$$

$$\begin{aligned}
 |010\rangle &\xrightarrow{H} |01\rangle \frac{1}{\sqrt{2}}(|0\rangle + |1\rangle) \xrightarrow{V} |01\rangle \frac{1}{\sqrt{2}}(|0\rangle + i|1\rangle) \xrightarrow{CNOT} |01\rangle \frac{1}{\sqrt{2}}(|0\rangle + i|1\rangle) \xrightarrow{N} |01\rangle \frac{1}{\sqrt{2}}(|0\rangle + |1\rangle) \\
 &\xrightarrow{CNOT} |01\rangle \frac{1}{\sqrt{2}}(|0\rangle + |1\rangle) \xrightarrow{V} |01\rangle \frac{1}{\sqrt{2}}(|0\rangle + |1\rangle) \xrightarrow{H} |010\rangle \\
 |011\rangle &\xrightarrow{H} |01\rangle \frac{1}{\sqrt{2}}(|0\rangle - |1\rangle) \xrightarrow{V} |01\rangle \frac{1}{\sqrt{2}}(|0\rangle - i|1\rangle) \xrightarrow{CNOT} |01\rangle \frac{1}{\sqrt{2}}(|0\rangle - i|1\rangle) \xrightarrow{N} |01\rangle \frac{1}{\sqrt{2}}(|0\rangle - |1\rangle) \\
 &\xrightarrow{CNOT} |01\rangle \frac{1}{\sqrt{2}}(|0\rangle - |1\rangle) \xrightarrow{V} |01\rangle \frac{1}{\sqrt{2}}(|0\rangle - |1\rangle) \xrightarrow{H} |011\rangle \\
 |100\rangle &\xrightarrow{H} |10\rangle \frac{1}{\sqrt{2}}(|0\rangle + |1\rangle) \xrightarrow{V} |10\rangle \frac{1}{\sqrt{2}}(|0\rangle + |1\rangle) \xrightarrow{CNOT} |11\rangle \frac{1}{\sqrt{2}}(|0\rangle + |1\rangle) \xrightarrow{N} |11\rangle \frac{1}{\sqrt{2}}(|0\rangle - i|1\rangle) \\
 &\xrightarrow{CNOT} |10\rangle \frac{1}{\sqrt{2}}(|0\rangle - i|1\rangle) \xrightarrow{V} |10\rangle \frac{1}{\sqrt{2}}(|0\rangle + |1\rangle) \xrightarrow{H} |100\rangle \\
 |101\rangle &\xrightarrow{H} |10\rangle \frac{1}{\sqrt{2}}(|0\rangle - |1\rangle) \xrightarrow{V} |10\rangle \frac{1}{\sqrt{2}}(|0\rangle - |1\rangle) \xrightarrow{CNOT} |11\rangle \frac{1}{\sqrt{2}}(|0\rangle - |1\rangle) \xrightarrow{N} |11\rangle \frac{1}{\sqrt{2}}(|0\rangle + i|1\rangle) \\
 &\xrightarrow{CNOT} |10\rangle \frac{1}{\sqrt{2}}(|0\rangle + i|1\rangle) \xrightarrow{V} |10\rangle \frac{1}{\sqrt{2}}(|0\rangle - |1\rangle) \xrightarrow{H} |101\rangle \\
 |110\rangle &\xrightarrow{H} |11\rangle \frac{1}{\sqrt{2}}(|0\rangle + |1\rangle) \xrightarrow{V} |11\rangle \frac{1}{\sqrt{2}}(|0\rangle + i|1\rangle) \xrightarrow{CNOT} |10\rangle \frac{1}{\sqrt{2}}(|0\rangle + i|1\rangle) \xrightarrow{N} |10\rangle \frac{1}{\sqrt{2}}(|0\rangle + i|1\rangle) \\
 &\xrightarrow{CNOT} |11\rangle \frac{1}{\sqrt{2}}(|0\rangle + i|1\rangle) \xrightarrow{V} |11\rangle \frac{1}{\sqrt{2}}(|0\rangle - |1\rangle) \xrightarrow{H} |110\rangle \\
 |111\rangle &\xrightarrow{H} |11\rangle \frac{1}{\sqrt{2}}(|0\rangle - |1\rangle) \xrightarrow{V} |11\rangle \frac{1}{\sqrt{2}}(|0\rangle - i|1\rangle) \xrightarrow{CNOT} |10\rangle \frac{1}{\sqrt{2}}(|0\rangle - i|1\rangle) \xrightarrow{N} |10\rangle \frac{1}{\sqrt{2}}(|0\rangle - i|1\rangle) \\
 &\xrightarrow{CNOT} |11\rangle \frac{1}{\sqrt{2}}(|0\rangle - i|1\rangle) \xrightarrow{V} |11\rangle \frac{1}{\sqrt{2}}(|0\rangle + |1\rangle) \xrightarrow{H} |110\rangle
 \end{aligned}$$

4. Висновок

Запропоновано реалізацію вентиля Тоффолі через 1- та 2-кубітні вентиля, при цьому ємнісна складність квантової схеми становить 7, що суттєво менше за ємнісну складність відомої реалізації перетворення Тоффолі, наведеної в [4].

Список використаних джерел

1. Williams C.P. Explorations in Quantum Computing / C.P. Williams. – Ed. 2-nd. – London: Springer Science & Business Media, 2011. – 717p.
2. Muthukrishnan A. Classical and Quantum Logic Gates [Online] / A. Muthukrishnan. – 1999. – 22p. – Available from: <http://www.optics.rochester.edu/~stroud/presentations/muthukrishnan991/LogicGates.pdf3>.
3. Metodi T.S. Quantum Computing for Computer Architects / T.S. Metodi, A.I. Faruque, F.T. Chong. – Ed. 2-nd. – Morgan & Claypool Publishers, 2011. – 202p.

4. Nielsen M.A. Quantum Computation and Quantum Information / M.A. Nielsen, I.L. Chuang. – Ed. 10-th. – New York: Cambridge University Press, 2011. – 702p.

References

1. WILLIAMS, C.P. (2011) *Explorations in Quantum Computing*. 2nd Ed. London: Springer Science & Business Media.
2. MUTHUKRISHNAN, A. (1999) *Classical and Quantum Logic Gates*. [Online] Available from: <http://www.optics.rochester.edu/~stroud/presentations/muthukrishnan991/LogicGates.pdf3>.
3. METODI, T.S., FARUQUE, A.I., CHONG, F.T. (2011) *Quantum Computing for Computer Architects*. 2nd Ed. Morgan & Claypool Publishers.
4. NIELSEN, M.A. and CHUANG, I.L. (2011) *Quantum Computation and Quantum Information*. 10th Ed. New York: Cambridge University Press.

Надійшла до редколегії 30.04.2015

УДК 004.7

Катеринич Л.О., к. ф.-м. .н., доц.,
Калашнікова А.А , студ.,
Марголін О.Г, студ.

Система прийняття рішень для ідентифікації користувачів

Київський національний університет імені
Тараса Шевченка, 83000, м. Київ, пр-т.
Глушкова 4д,
e-mail: katerinich@rambler.ru
e-mail: alisakalash@yandex.ru
e-mail: ag.margolin@gmail.com

L.O. Katerinich, PhD,
A. A. Kalashnikova, stud.,
A.G. Margolin, stud.

Information analysis system for user authentication

Taras Shevchenko National University of Kyiv,
83000, Kyiv, Glushkova st., 4d,
e-mail: katerinich@rambler.ru
e-mail: alisakalash@yandex.ru
e-mail: ag.margolin@gmail.com

Стаття присвячена дослідженню методів та засобів реалізації ідентифікації та аутентифікації суб'єктів автоматизованої системи . Для вирішення цієї проблеми запропонована реалізація системи прийняття рішень. Однією з найбільших переваг системи – є можливість її масштабування. Система може бути використана на практиці, наприклад для виявлення декількох сторінок однієї людини в соціальних мережах, форумах та порталах новин.

Ключові слова: ідентифікація користувача, інформаційне середовище, інформаційні потоки

This article is devoted to research of methods and tools for identifying subjects of an automated system. This problem remains unresolved, because there is no single method for identifying a person who is using the computer at the moment. In conditions of modernity, while information warfare and a variety of cybercrime, there is a great relevant of this subject researching and development of new methods of identification and identify trends. A multifactor decision-making system is proposed to solve this problem. System can answer the question whether the owner of two accounts is the same person with a certain percentage of accuracy. One of the main advantages of the system - is its ability to be expanded. Factors that are analyzed are fields that the user enters about himself during the registration and fields (hidden from the user) that the system gets from users browser, computer, software, any behavioral factors of the user etc. The system can be used in practice, for example to identify duplicate user pages in social networks, forums or news portals.

Key words: user identification, information environment, decision-making system, information flows

Статтю представив д. т. н., проф. Заславський В.А.

Вступ

Одна з основних проблем, яка постає перед розробниками програмних додатків, - це ідентифікація користувача. Реєстрація користувача полегшує збір статистики і забезпечує цілісність даних користувача, що дає можливість власникам додатків вигідно для себе

і для користувача використовувати цю інформацію. Ідентифікацію та аутентифікацію можна вважати основою програмно-технічних засобів безпеки, оскільки інші сервіси розраховані на обслуговування іменованих суб'єктів [1]. Ідентифікація та аутентифікація - це «перша лінія оборони» інформаційного простору організації.

Ідентифікація дозволяє суб'єкту (користувачеві, процесу, що діє від імені користувача чи іншого апаратно-програмному компоненту) назвати себе (повідомити своє ім'я). За допомогою аутентифікації друга сторона переконується, що суб'єкт дійсно той, за кого він себе видає. Як синонім слова "аутентифікація" іноді використовують словосполучення "перевірка справжності" [2].

Проблема ідентифікації та аутентифікації користувачів залишається відкритою, адже поки що не існує єдиного методу для виявлення суб'єкта, що користується системою у поточний момент часу [3]. Звідси випливає актуальність в дослідженні даної теми та розробці нових методів ідентифікації.

Позитивним результатом аутентифікації є авторизація користувача, тобто надання йому прав доступу до ресурсів. Для доступу до різних систем можуть застосовуватися різні методи аутентифікації.

В залежності від міри довірчих стосунків, структури, особливостей мережі і віддаленості об'єкта перевірка даних наданих до системи може бути односторонньою або взаємною. Розрізняють однофакторну і строгу (двофакторну) аутентифікації. В однофакторних системах, найпоширенішими в цей час є паролі системи аутентифікації [2].

Для виявлення дублікатів акаунтів користувача запропонована багатфакторна система аналізу інформації та прийняття рішень.

Основні принципи роботи багатфакторної СПР

Основна ідея системи полягає в тому, щоб прийняти рішення щодо належності двох акаунтів одному користувачу.

Система має багатфакторну структуру (поєднання одразу цілої низки методів та джерел оброблюваних даних) і можливість масштабування. Факторами, що аналізуються, є як поля, що вводить користувач про себе при реєстрації, так і поля, що виявляються приховано від користувача з даних його браузера, комп'ютера, програмного забезпечення, поведінкових факторів самого користувача тощо.

Система отримує n вхідних параметрів – поля, що вводить користувач при реєстрації, та параметри, що визначаються системою самостійно. Кожен з параметрів аналізується по всіх користувачах у системі та видається результат.

Система є напівавтоматичною, тобто є можливість автоматичної та адміністративної перевірки користувачів. Автоматична перевірка видає приблизний список підозрюваних на схожість користувачів. Далі пропонується перевірити цих користувачів за допомогою адміністративної перевірки. Адміністративна перевірка наводить таблицю порівняння двох користувачів за усіма факторами, які є у системі, та імовірність того, що два акаунти належать одній людині.

Останнє рішення, чи блокувати користувача, завжди залишається за людиною. Тобто система лише може зробити припущення та натякнути людині (наприклад адміністратору форуму), щоб він звернув увагу на підозрювані 2 акаунти.

Однією з найважливіших переваг системи – є можливість її масштабування. Структура системи передбачає можливість додавання довільної кількості факторів перевірки суб'єкта системи.

Кожному параметру присвоюється коефіцієнт впливу. Зрозуміло, що збіг деяких параметрів є більш важливим фактором для прийняття рішень ніж інший. На даний момент кожен з факторів має ваговий коефіцієнт, що дорівнює 1. Для того, щоб визначити вагу для кожного з факторів, які використовуються в системі, треба побудувати статистику для великої кількості користувачів на реальному проекті.

Демонстраційна версія системи написана мовою PHP [5], за допомогою фреймворку CakePhp [6] та з використання JavaScript. Для збереження даних використовується MySql база даних [7]. Для демонстрації запропонована панель адміністратора, в якій можна додати користувачів та зробити перевірку акаунтів цих користувачів на схожість.

Автоматична перевірка

Автоматична перевірка надає інформацію про те, що один користувач схожий на іншого на деякий процент. Далі адміністратору пропонується перейти до адміністративної перевірки.

Алгоритм автоматичної перевірки користувачів.

Автоматична перевірка працює за алгоритмом ART1 [8]. Обирається перший користувач і робиться запит до бази даних на ті поля, що потрібні при перевірці. За цими полями будується вектор-ознак E .

Далі виконується перевірка на схожість за нерівністю (1).

$$\frac{\|E\|}{(\beta + d)} < \frac{\|P_i \cap E\|}{(\beta + \|P_i\|)}, \quad (1)$$

де P_i - вектор-прототип, β - параметр, який обирає прототипи, в яких більше значень 1, за умови, що усі значення 1 в векторі-прототипі також наявні у векторі-ознак, що тестується, d - розмір вектора (його довжина). $\|u\|$ - значимість вектора u .

На першому етапі відсутній вектор-прототип. Тобто першим вектором-прототипом є вектор-ознак (2)

$$P_0 = E \quad (2)$$

Якщо тест на схожість пройшов успішно, виконується порівняння вектора ознак і вектор-прототипу з параметром уважності (нерівність (3)).

$$\frac{\|P_i \cap E\|}{\|E\|} < \rho \quad (3)$$

Завданням даного параметра є визначення розміру класу. Якщо значення параметра велике, утворюються більші класи (кластер з великою кількістю даних). При зменшенні значення створюються кластери з меншою кількістю даних. Якщо параметр уважності заданий досить низьким ($<0,1$), для допуску вектори ознак повинні збігатися з вектором-прототипом. Параметр для даної задачі був обраний експериментально: $\rho = 0.7$.

Якщо пройдений тест на уважність, поточний вектор ознак додається в поточний вектор-прототип (рівняння (4)).

$$P_i = P_i \cap E \quad (4)$$

В результаті кожен користувач належить певному кластеру за схожістю векторів ознак. Система надає ці результати адміністратору та пропонує виконати адміністративну перевірку.

Адміністративна перевірка

Адміністративна перевірка надає можливість адміністратору обрати два акаунти на перевірку і подивитися результати перевірки самостійно. Адміністративна перевірка представляє собою таблицю, яка складається з таких стовпчиків: номер критерію, критерій (поле, за яким іде

перевірка), перший користувач, другий користувач (тобто значення полів з попереднього стовпчика у кожного акаунта) та результат перевірки (+/-).

Алгоритм перевірки схожий на автоматичний, але тут перевіряється тільки 2 користувачі, значення яких передаються POST-запитом з форми вибору користувачів для перевірки.

Алгоритм ручної перевірки користувачів.

1. Обираються два довільні користувачі з бази даних.
2. Робиться запит до бази даних на ті поля, що нам потрібні при перевірці першого користувача.
3. Виконується перевірка по кожному полю (фактору), що надає нам інформацію про користувача, з іншими аналогічними полями у іншого користувача.
4. Якщо дані збігаються, то виводимо знак «+», в іншому випадку – «-».
5. Обчислюється і виводиться відповідь – ймовірність того, що акаунти належать одній людині.

Ймовірність система обчислює наступним чином. Відомо, що сума ваг усіх параметрів дорівнює S . Якщо при перевірці за параметром a значення в обох користувачів збігаються, то до лічильника додається вага цього параметра. В результаті, після перевірки за всіма параметрами, отримуємо значення S_1 . Ймовірність P обчислюється за формулою (5):

$$P = \frac{S_1 \cdot 100\%}{S} \quad (5)$$

Параметри перевірки

Ручна перевірка виконується за полями, які вводить користувач, а також за параметрами, які система дізнається з браузера користувача самостійно. На базі результатів порівняння параметрів можна вирішити з певною точністю, маємо ми справу з однією людиною чи з різними.

За допомогою JavaScript [9] можна отримати повну інформацію про браузер користувача та його операційну систему. Авжеж, ця інформація не буде його ідентифікувати на 100%, але як компонент так званого fingerprint, стане дуже корисною.

Параметри, значення яких визначає система:

- операційна система;
- браузер, що наразі використовується;
- IP адреса;

- ширина та висота дисплею, глибина кольору;
- платформа браузера;
- дата та часовий пояс;
- плагіни, що встановлені у браузері;
- системні шрифти.

З параметрів, що вводить користувач аналізує наступні: пароль, пошту, державу, адресу тощо.

Порівняння за цими параметрами дає змогу з певною точністю визначити, чи одна людина є власником двох різних акаунтів. При більшій кількості користувачів є можливість побудувати статистику частоти збіжності кожного з параметрів і присвоїти їм вагові коефіцієнти. Якщо параметр a зустрічається в n

відсотках випадків, то йому присвоюється коефіцієнт k (формула (6))

$$k = 1 - \frac{n}{100}. \quad (6)$$

Приклад роботи СПР

Розглянемо результат роботи системи на прикладі. Автоматична перевірка видає підозрювані акаунти, по яких надаються данні, наведені в табл. 1 (в таблиці наведена лише частинна). Для маскування від системи користувач намагався зареєструватися з мобільного пристрою.

Таблиця 1.

Результат роботи системи

Заголовок поля	Fiona	Iphone	+/-
Ім'я	Фіона Марківна	Телефон Фіони	-
Пароль	29bb6b6ec346e0291c9de68680feda31098d87f3	29bb6b6ec346e0291c9de68680feda31098d87f3	+
Телефон	933421212	505252411	-
Пошта	kiss@mail.ru	kiss@mail.ru	+
Держава	Україна	Росія	-
Місто	Київ	Москва	-
Адреса	Вул. Хрещатик	Червона площа	-
ОС	Safari	iPhone	-
Браузер	Safari 7.0	Safari 7.0	+
Останній IP	212.90.37.69	212.90.37.69	+
Ширина дисплею (px)	1440	320	-
Висота дисплею (px)	900	480	-
Глибина кольору дисплею (в бітах)	24	32	-
Платформа браузера	MacIntel	iPhone	-

За даними таблиці 1 ми можемо бачити, що в цих акаунтах було введено однакові паролі, що часто роблять особи, не здогадуючись, що їх паролі можна перевірити (паролі шифруються за однаковим алгоритмом, тому ми без порушення приватності можемо порівняти їх).

Також введено однакові електронні пошти, що вже надає нам велику імовірність того, що власниками цих акаунтів є одна особа. Однакова IP адреса мережі підтверджує наше припущення, щодо того, що це та сама людина, що зареєструвала акаунт спочатку зі свого

комп'ютера, а потім зі смартфона, під'єднаних до однієї мережі. Збіг браузера тут має лише випадковий характер. В даному випадку свою роль зіграє вагомість коефіцієнтів кожного поля інформації.

Основні висновки

В даній статті було продемонстровано багатофакторну систему прийняття рішень для виявлення декількох акаунтів одного і того самого користувача, для подальшого його

видалення. Запропонована система дає змогу адміністратору сайту визначити з певною мірою точності, чи відповідають 2 акаунти одному користувачу.

Аналізатор системи представлений автоматичною та адміністративною перевіркою. Прийняття рішення залишатиметься за людиною – система лише робитиме натяк на схожість 2-х акаунтів в системі.

Система має багатфакторну структуру. Фактори, що аналізуються – це як поля, що вводить користувач про себе при реєстрації, так і поля, що визначаються приховано від користувача з даних його браузера, комп'ютера,

програмного забезпечення, якихось поведінкових факторів самого користувача тощо.

Один з найголовніших плюсів розробленої системи – є можливість розширення. Чим більше факторів перевірки буде задіяно в системі, тим точніше буде результат перевірки. Також є можливість для надання ваги кожному з факторів, що підвищить точність оцінки.

Розроблену систему можна широко використовувати на практиці, наприклад для виявлення дублікатів сторінок користувачів в соціальних мережах чи форумах.

Список використаних джерел

1. В.А. Галатенко Основы информационной безопасности. / Галатенко В.А. – М.: ИНТУИТ.РУ «Интернет-Университет Информационных Технологий», 2006. – 208 с.
2. В.А.Галатенко Информационная безопасность. [Электронный ресурс] // Галатенко В.А. Открытые системы – 1996. – №4. – 10 с. – Режим доступа до журн.: <http://www.osp.ru/data/www2/os/1996/04/40.htm>.
3. Jodi L. Whitaker Keeping Children and Adolescents Safe. / Jodi L. Whitaker, Brad J. Bushman Online Dangers – 66 Wash. & Lee Rev., 2009. – 11 p.
4. Э. Гутманс PHP5. Профессиональное программирование / Э. Гутманс, С. Баккен, Д. Ретанс – СПб; М.: Символ-плюс, 2000. – 700 с.
5. Офіційний сайт CakePHP. Режим доступу: <http://cakephp.org/>
6. А. Нанда Oracle PL/SQL для администраторов баз данных / А. Нанда, С. Фейерштейн – СПб; М.: Символ-плюс, 2008. – 496 с.
7. М. Тим Джонс. Программирование искусственного интеллекта в приложениях. / М. Тим Джонс. – М.: ДМК Пресс, 2004. – 312 с.
8. Д. Рейсиг JavaScript. Профессиональные приемы программирования / Д. Рейсиг – СПб: Питер, 2008. – 351 с.

References

1. GALATENKO V.A. (2006) *Fundamentals of Information Security*, Ed: "Internet University of Information Technologies".
2. GALATENKO V.A. (1996) *Information security*. [Electronic resource] // Open systems #4. Avaliavle from: <http://www.osp.ru/data/www2/os/1996/04/40.htm>.
3. JODI L. WHITAKER, BRAD J. BUSHMAN (2009) *Online Dangers: Keeping Children and Adolescents Safe*. 66 Wash. & Lee L. Rev. 1053
4. E. GUTMANS, S. BAKKEN, D. RETANS (2000) *PHP5. Professional programming*, Ed: Prentice hall.
5. Official website CakePHP. Avaliavle from: <http://cakephp.org>.
6. A. NANDA, S. FERSTEIN (2005) *AJAX Oracle PL/SQL for DBA*. Ed: O'Reilly Media, Inc..
7. M. TIM JONES (2003) *AI Application Programmiring*. Ed: Charles River Media.
8. D. RASIG (2006) *Pro JavaScript techniques*, Ed: Apress.

Надійшла до редколегії 09.03.15

УДК 004.2:004.5

Коваль Ю.В., асистент

Iu. V. Koval, Assistant

**Протокол багатоточкового потоку:
означення та застосування в створенні
механізмів розпаралелювача та
аналізатора подій в системі жестового
інтерфейсу**

Київський національний університет імені
Тараса Шевченка, 83000, м. Київ, пр-т.
Глушкова 4д,
e-mail: kafedraTK@unicyb.kiev.ua

**Multipoint stream protocol: definition and
application for parallelizer and events
analyzer mechanisms in gestures interface
system creation**

Taras Shevchenko National University of
Kyiv, 83000, Kyiv, Glushkov ave., 4d,

e-mail: kafedraTK@unicyb.kiev.ua

В статті запропоновано протокол багатоточкового потоку. Визначено вимоги до такого потоку та запропоновано схему функціонування протоколу. В якості прикладу застосування багатоточкового протоколу розглянуто реалізацію механізмів розпаралелювача та аналізатора подій в системі інтерактивного жестового інтерфейсу комунікації людина-комп'ютер.

Ключові слова: протокол багатоточкового потоку, інтерфейс жестової мови.

Requirements for protocol to develop parallelizer and events analyzer mechanisms in gestures interface system are discussed in the article. A multipoint stream protocol is proposed. The complete set and semantics of protocol methods were proposed. Security key usage and strategy for next key selection proposed. Also additional techniques for security strengthening proposed such as using of key shift field and fields shift field. The next security improvement is the usage of cover fields. The complete scheme of state of one point that used the multipoint protocol presented. A different ways of using multipoint stream protocol were demonstrated on examples of realising a parallelizer mechanism and events analyzer mechanism of sign language human-computer interface system. A problem for commands set of parallelizer and events analyzer mechanisms was discussed and the solution proposed. The advantage of multistreaming of multipoint stream protocol for command and result stream was also demonstrated.

Key Words multipoint stream protocol, sign language interface.

Статтю представив д.ф.-м.н., проф. Анісімов А.В.

1. Вступ та постановка задачі

Схема інтерфейсу комунікації людина-комп'ютер на базі жестової мови та жестів, запропонована в (1), містить елемент розпаралелювач, призначенням якого є забезпечення одночасної передачі значень на обробку декільком процесам. Ці процеси є складовими віртуального процесу відповідно до (2) та керуються аналізатором подій, що також є складовим процесом. Серед сучасних технічних систем можливість одночасної передачі значень мають комп'ютерні мережі, побудовані за технологією broadcast. На сьогодні найбільш уживаною мережевою технологією є TCP/IP (3). Проте серед сучасних протоколів, реалізованих для мереж

у межах цієї технології, немає протоколу, для якого виконується такий набір вимог:

1. одночасна передача на декілька (можливо, сусідніх) комп'ютерів;
2. значення передаються потоком;
3. у потоці передаються окремі повідомлення;
4. є можливість приєднатися до потоку;
5. відсутність контролю отримування;
6. наявність системи захисту.

Вимоги 1, 2 та 3 відповідають задачі обробки жестової інформації.

Вимога 4 викликана потребою створення потоку на декілька отримувачів.

Вимога 5, традиційна для протоколу UDP, вважається недоречною для передачі потоку.

Причина полягає в неможливості передати потік значень без втрат та/або помилок навіть для невеликих за розміром потоків. Розбиття потоку на окремі повідомлення дозволяє передати їх безпомилково в більшості ситуацій, а у випадках помилок зберігати структуру потоку й правильно продовжувати його передачу. У випадку втрати отримувачем потоку відновлення його роботи відбувається завдяки виконанню вимоги 4. Такий підхід дозволяє відмовитися від технології TCP, що дуже доречно за умов, коли відновлений потік вже не буде актуальним із міркувань часу.

Вимога 6 необхідна для реалізації можливості роботи з незахищеними віддаленими мережами.

З вищесказаного випливає така постановка задачі: в межах технології TCP/IP розробити мережевий протокол, що відповідає вказаним вимогам. На базі запропонованого протоколу розробити механізми розпаралелювача та аналізатора подій в системі жестового інтерфейсу. В цій статті пропонується необхідний протокол та вказуються методи його реалізації.

2. Протокол багатоточкового потоку

Під потоком у цій статті будемо розуміти структуровану послідовність значень. Це означає: значення, що утворюють потік, об'єднані в блоки, що мають внутрішню структуру й об'єднані в послідовності. Передбачається, що блоки можуть бути різних розмірів, тому будемо називати їх повідомленнями. Послідовність повідомлень може утворювати складовий підпотік. В потоці завжди існує головний підпотік, в якому передаються команди на створення та знищення підпотоків. Головний підпотік також може використовуватися для передачі повідомлень. Обмін інформацією для програм організується з використанням вказаних структурних одиниць.

Одне повідомлення потоку при передачі в мережі розділяється на пакети, що передаються на нижчому мережевому рівні. Кожен такий пакет містить традиційні для протоколів сімейства TCP/IP поля – IP-адреси відправника та отримувача, числовий номер протоколу та розміру пакету. В якості

адреси отримувача можливе використання broadcast-адреси IP-підмережі. Це забезпечує одночасне отримувача пакета кількома комп'ютерами. Використання адреси конкретного комп'ютера означає, що процеси-отримувачі виконуються в межах цього комп'ютера або на комп'ютерах після нього в розумінні топології мережі та маршруту передачі.

На відміну від протоколів TCP та UDP, в яких вказуються порти відправника та отримувача (3), у протоколі багатоточкового потоку вказується номер порту потоку. Використання одного й того самого номеру порту для всіх точок, під'єднаних до потоку, необхідне для забезпечення одночасної багатоцільової передачі значень.

Для забезпечення захисту інформації в кожен пакет додаються поля упаковки, нового ключа, тасування ключа та тасування полів.

Поле з контрольною сумою пакета присутнє та використовуються як зазвичай.

Для узгодженості з блочною структурою інформації додаються поля з номерами та розмірами повідомлення та пакету в ньому.

Поле з числовим ідентифікатором підпотіку вказує відповідний ідентифікатор. Для запобігання колізії на створення підпотоків вважається, що у випадку створення існуючого підпотіку точка приєднується до нього.

Повна структура пакета така:

1. IP-адреса відправника;
2. IP-адреса отримувача;
3. числовий ідентифікатор протоколу;
4. розмір пакету;
5. порт потоку.

Подальші поля закриваються ключем:

6. упаковка;
7. тасування ключа;
8. тасування полів;
9. новий ключ;
10. контрольна сума;
11. числовий ідентифікатор підпотіку;
12. номер повідомлення;
13. розмір повідомлення;
14. номер пакету повідомлення;
15. розмір пакету повідомлення;

16. пакет повідомлення;
17. упаковка.

Під з'єднанням в описуваному протоколі розуміється сукупність точок, під'єднаних до потоку інформації, що формується процесами, які назвемо учасниками потоку. Кожен учасник потоку може передавати та приймати інформацію, використовуючи потік. Захист від одночасної передачі забезпечується технологією опорної мережі. Наприклад, для мережі Ethernet це є детектування колізії.

Кожен учасник потоку передає новий ключ, що може бути використаний у потоці. Пропонуються такі стратегії вибору ключа:

1. ключ від точки з найменшою IP-адресою;
2. ключ від точки з найбільшою IP-адресою;
3. ключ від точки з найбільшим об'ємом переданої інформації.

Вибір стратегії покладається на програміста при вказуванні параметрів з'єднання. Під'єднання точки до потоку можливе лише за умови правильно вказаної стратегії вибору наступного ключа.

Як вже вказувалося, рішення щодо відсутності механізму відновлення цілісності потоку прийнято з огляду на жорсткі часові вимоги. Незручності цього підходу мають компенсуватися як зовнішніми чинниками на кшталт надійності мережі в розумінні мінімальних втрат пакетів, що передаються, так і внутрішніми, як наприклад, обробка лише частини повідомлень для отримання результату. У випадку втрати з'єднання точка або заново під'єднується до потоку, або переходить до виконання інших дій.

Передбачено такі інтерфейсні функції для реалізації протоколу:

1. create – створення точки;
2. listen – очікувати на утворення потоку;
3. connect – об'єднання точок в потік;
4. join – приєднати точку до потоку;
5. close – закрити потік;
6. disconnect – від'єднати точку від потоку;
7. read – читати повідомлення;
8. write – надіслати повідомлення;

9. keepalive – підтримка потоку;
10. new – створити новий підпотік;
11. delete – закрити підпотік.

Схему станів роботи точки потоку зображено на рис. 1.

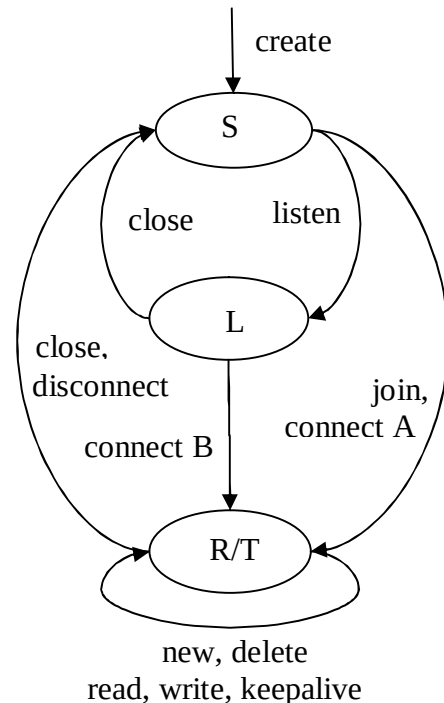


Рис. 1. Схема станів точки потоку

В наведеній схемі використані такі позначення:

S – початковий стан,

L – стан очікування з'єднання,

R/T – робочий стан потоку,

create, listen, close, join, new, delete, read, write, keepalive – відповідні дії точки, описані вище,

connect A – активне формування спільного потоку з окремих точок,

connect B – пасивне приєднання до потоку зі стану очікування.

3. Механізми розпаралелювача та аналізатора подій

Механізм розпаралелювача має з'єднувати вхідний потік, що передає інформацію, та вузли модулів, що отримують інформацію(1). Оскільки вхідні потоки можуть бути різними й наперед зреалізованими та з визначеними інтерфейсами, то модуль розпаралелювача з'єднується з вхідним потоком за традиційними протоколами. Зазвичай це TCP-сесія. Модулі, що здійснюють обробку, та сам модуль розпаралелювача утворюють

багатоточковий потік. Для формування потоку модулі обробки утворюють точки та очікують доти, доки розпаралелювач не з'єднає ці точки в єдиний потік.

Аналізатор подій виконує роль менеджера віртуального процесу для модулів функціональної обробки (2). На підставі рішень аналізатора подій запускаються або зупиняються окремі модулі. Для реалізації такого керування множиною підпроцесів запропоновано утворити багатоточковий потік, що створюється аналізатором подій, після чого до нього під'єднуються підпроцеси окремих модулів. В утвореному потоці передаються команди керування модулями та результати роботи модулів. Для розмежування цих інформаційних потоків в межах багатоточкового потоку створюються підпотоки.

Описана реалізація накладає певні вимоги на синтаксис команд, що використовуються для керування модулями. Синтаксично однакові команди повинні виконуватися в модулях одночасно. Прикладом такої команди є команда «перейти до обробки

наступного повідомлення». Що стосується різних команд, то вони мають бути синтаксично різними. Модулі, що не мають таких команд, повинні сприймати ці команди як такі, що не викликають ніяких дій. Прикладами таких команд є команди встановлення параметрів окремих модулів.

4. Висновки та подальші напрямки роботи

В результаті проведених досліджень було запропоновано протокол багатоточкового потоку, за допомогою якого реалізуються механізми розпаралелювача та аналізатора подій в системі інтерфейсу людина-комп'ютер на базі жестової мови та жестів.

Подальші дослідження будуть спрямовані на те, щоб реалізувати запропонований протокол, реалізувати на його основі модулі розпаралелювача та аналізатора подій, узгодити зовнішні функціональні модулі, створити загальну систему розпізнавання жестів та жестової мови та дослідити її роботу.

Список використаних джерел

1. Крак Ю.В., Коваль Ю.В., Тернов А.С. До розробки інтерактивного інтерфейсу моделювання та розпізнавання жестової інформації / Ю.В. Крак, Ю.В. Коваль, А.С. Тернов // Вісник Київського національного університету імені Тараса Шевченка Серія фізико-математичні науки. – 2014. – №4. – С.175-178.
2. Крак Ю.В., Коваль Ю.В., Ставровський А.Б. Віртуальний процес: означення та застосування в створенні системи жестового інтерфейсу / Ю.В. Крак, Ю.В. Коваль, А.Б. Ставровський // Вісник Київського національного університету імені Тараса Шевченка Серія фізико-математичні науки. – 2015. – №1. – С. 141-144.
3. Marsic I. Computer networks: Performance and quality of service/ I. Marsic // Department of Electrical and Computer Engineering, Rutgers University, New Brunswick, New Jersey - 2013.

References

1. KRAK IU.V., KOVAL IU.V., TERNOV A.S. (2014) *To creation of interactive interface for sign information modeling and recognition*. Bulletin of Taras Shevchenko National University of Kyiv Series Physics & Mathematics. Vol. 4. p.175-178.
2. KRAK IU.V., KOVAL IU.V., STAVROVSKIY A.B. (2015) *Virtual process: definition and application for gestures interface system creation*. Bulletin of Taras Shevchenko National University of Kyiv Series Physics & Mathematics. Vol. 1. p.141-144.
3. Marsic I. (2013) *Computer networks: Performance and quality of service* Department of Electrical and Computer Engineering, Rutgers University, New Brunswick, New Jersey.

Надійшла до редколегії 15.05.2015

УДК 519.9

Ковальов Д.І., аспірант

D.I. Kovaliov, post-graduate

Сервіс дистанційного навчання у форматі соціальної мережі

E-learning service in social network format

Київський національний університет імені Тараса Шевченка, 83000, м. Київ, пр-т. Глушкова 4д,
e-mail: daniil.kovaliov@gmail.com

Taras Shevchenko National University of Kyiv,
83000, Kyiv, Glushkova st., 4d,
e-mail: daniil.kovaliov@gmail.com

Описується стан електронної освіти в Україні відносно світових тенденцій. Розглядаються можливості впровадження сервісу для дистанційного або частково дистанційного навчання студентів у форматі соціальної мережі з обмеженою реєстрацією. Описуються мотиваційні механізми для студентів, переваги зворотного зв'язку для викладачів, використання мобільних пристроїв у навчальному процесі та інформуванні студентів і викладачів, система швидкої оцінки знань, що є частинами проекту та перспективи його розвитку.

Ключові слова: електронне навчання, соціальні мережі, адаптивні моделі, ігрофікація, мобільні пристрої.

Main subject of current article is an internal e-learning system that allows users to be aware of any urgent updates, useful information, and changes to schedule regarding ongoing learning activity in university and also maintain this information changes. E-learning situation in Kiev National University of Ukraine compared to world tendencies is briefly described in this article. Possibility of social network for distant or partly distant students learning service implementation as a social network with strict registration is reviewed. Motivational mechanisms in a gamification framework form, feedback advantages for teachers in a non-anonymous way, examples of mobile devices usage in educational process for immediate updates, system for fast assessment with adaptive evaluation as parts of the project oriented for education enhancement and its perspectives are described. Possibility of MOOC with video lectures that available for everyone on this platform is also mentioned and particular formulas, methodologies and situations are shown.

Keywords: e-learning, social networks, the adaptive model, gamification, mobile devices.

Статтю представив д.т.н. проф. Волошин О.Ф.

Вступ. Електронна освіта у двадцять першому столітті стала майже обов'язковим атрибутом будь-якого навчання. Реалізацію системи електронної освіти було втілено у методичній тестуюче-оцінюючій системі SMPR[1], що є додатком до курсу «Теорія прийняття рішень»[2]. Розробка цієї системи проходила протягом декількох років при залученні студентів третього курсу факультету кібернетики. У 2014 році було створено веб версію тестуючої системи, що успішно була застосована при екзаменаційному оцінюванні. Наступним кроком розвитку стала розробка специфікованого онлайн сервісу для комунікації

між викладачами та студентами в контексті навчання.

Актуальність. Більшість західних університетів мають свої власні внутрішні системи, що дозволяють проводити деякі або навіть більшість навчальних процесів через глобальну мережу інтернет. Це допомагає студентам і викладачам зберігати їхній час і надає можливість зосередитися саме на навчанні. Українські вищі навчальні заклади не мають централізованого сервісу з необхідними функціями, і, на думку авторів, його запровадження призведе до покращення якості та швидкості навчання.

Соціальна мережа. Основною частиною сервісу є соціальна мережа для студентів та викладачів. Реєстрація у мережі можлива лише за наявності студентського квитка або спеціального індивідуального коду викладача. Подібний підхід до реєстрації забезпечує неможливість комерційного використання мережі.

За тиждень до початку занять староста новосформованої групи отримує від адміністратора мережі доступ до редагування профіля групи та її складу. Староста власноруч заносить до соціальної мережі ПІБ та номер студентського квитка кожного студента його групи та заповнює розклад занять відповідно до офіційного розкладу. Назви предметів староста вибирає зі списку доступних для його курсу.

Сторінка групи у мережі має наступний функціонал:

- загальна стрічка повідомлень, що стосуються всіх членів групи;
- інтерактивний розклад занять (детальніше згодом);
- панель з документами, що стосуються групи;

Інтерактивний розклад має багато додаткових функцій:

- створення та редагування викладачами додаткових занять, контрольних робіт, консультацій для всіх слухачів або певних груп;
- створення та редагування студентами подій для групи, старостами – для потоку;
- можливість прикріплення документів до подій;
- створення та редагування викладачами подій та занять, що є обов'язковими для всіх студентів, з посланням на сторінки відповідних дисциплін.

Сторінки дисциплін надають можливість викладачам завантажувати та зберігати документи або відеозаписи, пов'язані з лекційними та практичними завданнями на огляд студентів певної групи або всіх відвідувачів сторінки. Студенти мають можливість задавати питання до певних лекцій чи документів, а також залишати коментарі та побажання до всього курсу загалом після його завершення.

Також на сторінках дисциплін та у розкладі можуть бути додані посилання на онлайн

тестування, що може бути використано викладачами для проведення швидких опитувань чи контрольних робіт.

Безпека мережі. Оскільки звичайна соціальна мережа є джерелом рекламних, а іноді й провокаційних інформаційних впливів, є потреба в ізолюванні навчальної соціальної мережі від анонімних та не авторизованих користувачів. Це можливо реалізувати шляхом попередньої реєстрації користувача в системі адміністратором чи іншим довіреним користувачем за короткими даними – ПІБ та номером студентського квитка. Заключні етапи реєстрації користувач проходить самостійно:

- підтвердження особи шляхом внесення номера студентського квитка;
- заповнення необхідної інформації
- підтвердження відповідності поштової скриньки шляхом переходу за посиланням, що на неї надіслано.

Авторизація на сайті здійснюється за допомогою логіну, що є номером студентського квитка та пароллю обраного користувачем.

Цей метод задовольняє декілька потреб: дозволяє вести актуальну статистичну інформацію по справжніх студентах та викладачах вищих навчальних закладів в цифровому вигляді (роки навчання, прослухані курси, бали, активність у групі тощо); захищає систему від несанкціонованого доступу людей, що мають на меті продаж, агітацію чи інші не пов'язані з навчанням цілі.

Відеолекції. Викладачі зможуть записувати свої лекції та редагувати їх перед публікацією безпосередньо на сайті. Це надає можливість вести лекції для студентів, що хворіють, або не мають можливості бути присутніми з інших причин а також брати участь у процесі прослуховування лекції тим студентам, що живуть в інших країнах. Аналогами подібних сервісів існують давно, але в контексті допомоги при очному навчанні вони використовуються не часто.

Мобільні пристрої. Цікавим та перспективним напрямком використання цього сервісу є застосування мобільного додатку.

Для викладачів це можливість відмічати присутніх натисканням однієї кнопки на екрані свого смартфона. При відповідних налаштуваннях географічна позиція пристрою

може бути вирахована з точністю, достатньою для визначення присутності студента на парі.

Аналогічним шляхом можливе ведення журналу через мобільний додаток. Достатньо лише вибрати прізвище студента зі списку присутніх і виставити йому оцінку.

Також у випадку затримок викладача, відміни занять, швидких контрольних робіт чи інших термінових подій викладач може надсилати повідомлення у групу студентів, яке вони отримають через додаток на мобільному пристрої, навіть перебуваючи далеко від монітору комп'ютера.

Також через мобільний додаток студенти можуть проходити тестування, а викладач отримувати результати студентів та автоматично обчислені оцінки.

Система тестування. Як частина мережі на сторінках предметів є окрема частина присвячена тестам з адаптивною оцінкою[3]. Викладачі можуть створювати тести та асоціювати їх з лекціями та практичними завданнями так само, як і звичайні файли. Відповідно, студенти можуть їх коментувати або задавати питання. Система тестування звільняє викладача від необхідності перевіряти власноруч відповіді студентів, позбавляє його можливості помилитись. Також у випадку, коли питання має нові або не передбачені викладачем відповіді, оцінка може бути перерахована для всіх студентів автоматично.

Крім цього, система дозволяє встановлювати чіткі часові межі проведення оцінювання.

Система тестування має можливість працювати з різними видами питань:

- питання з бінарними відповідями («так» чи «ні»);
- питання, де правильна відповідь вибирається з декількох, а правильність кожного варіанту відповіді оцінюється балом;
- питання, де декілька правильних відповідей вибираються із багатьох варіантів;
- питання, де відповідь це число. У цьому випадку до бази даних необхідно додати інформацію щодо правильної відповіді та діапазону можливих розбіжностей, на якому може знаходитись відповідь студента;
- питання, де відповіддю є інтервал;
- питання, відповіддю на які є нечіткі величини, що визначаються функцією

належності, що зводить їх до питань четвертого типу (відповідь це число);

- питання, відповіддю на які є слово або словосполучення. Через те, що відповідь може бути сформульована студентом правильно, але використовуючи термінологію, відмінну від термінології викладача, можливе виникнення неузгодженості. Для обробки такого типу питань необхідно використовувати словники з синонімічними рядами та алгоритми пошуку спорідненості текстів;

Після проведення тестування система перераховує вартість питань і записує нові значення у базу. Якщо в оцінці знань брало участь n студентів і було m питань, то корекція складності буде розрахована за формулою:

$$p_i^j = p_i^{j-1} + f_i(p_i^0, m, d^{j-1}, Z), j = \overline{1, n}, i = \overline{1, m},$$

де p_i^j - складність i -го завдання при проходженні j -го студента; f_i - функція згідно до якої відбувається корегування оцінки; d^{j-1} сумарна оцінка $j - 1$ студента; Z - деяка величина, що залежить від відхилення даної відповіді від правильної[4].

Оскільки під час роботи зі студентами можливі випадки «саботування» системи для отримання кращих результатів, система має реєструвати всі можливі зміни перед їх застосуванням та порівнювати з попередніми. У випадку надходження викладачу сповіщення про значну відмінність результатів попередніх проходжень тестового блоку та поточного, викладач має провести експертний аналіз результатів та підтвердити або відхилити внесення змін.

Мотиваційна система. На сьогоднішній день питання мотивації школярів, студентів та робітників має велике значення. Безумовно є самотивовані індивіди, що рухаються до мети без сторонньої допомоги. Але, нажаль, кількість таких людей невелика, через що так звана ігрофікація зазнає великого попиту.

У примітивній формі вона зустрічається у вигляді оцінок та балів звичної нам системи навчання. Але така система виключає індивідуальні досягнення кожного, як такі, що просто не можуть бути враховані через розподіленість балів між фіксованою кількістю контрольних та самостійних робіт. Також звична система оцінювання має лише епізодичний

характер, і не дає змогу студентові бачити свій постійний ріст чи його відсутність на фоні інших. Це призводить до того, що лише на передодні екзамену студент розуміє своє становище.

Система ігрофікації дасть викладачам додатковий інструмент, що дозволяє відмічати досягнення певних студентів за допомогою існуючих, або створених унікальних графічних відміток чи спеціальних завдань. Прикладом може бути відмітка «Всі пари відвідано», яка може буде надана студенту, що не пропустив жодної пари.

Крім цього, система в автоматичному режимі буде нараховувати умовні бали за відвідання лекцій, вчасність виконання лабораторних та домашніх робіт, активність (запитання викладачу, проведення заходів) тощо. Таким чином, у кожного студента буде свій унікальний рівень залученості в навчальний процес, згідно якого з'явиться можливість оцінити статус студента та рівень його знань.

Технології. Для створення цієї системи викристано платформу .NET та мови C#, JavaScript[5]. Ці технології були вибрані через їх розповсюдженість та легкість використання у будь якому середовищі. Для організації процесу розробки використовувалась методологія Scrum[6], що ґрунтується на періодичному підсумку результатів та створенні нових актуальних завдань.

Висновок. У цій статті розглянуто основний функціонал нового сервісу подібного до соціальної мережі, що спрямований на інформатизацію, оптимізацію та збільшення мотивації у навчальному процесі для студентів та викладачів. Система у багатоетапному режимі інтегрується для роботи на факультеті Інформаційних технологій як тестовому майданчику. За умови позитивних відгуків та доцільності використання системи, планується впровадження та розширення системи в масштабах інших факультетів та вищих навчальних закладів.

Список використаних джерел

1. Волошин О.Ф. Программная система поддержки курсов по теории принятия решений: стаття / Волошин О.Ф., Ковальов Д.І. [конференція: Problems of Computer Intellectualization. – Kyiv-Sofia: ITHEA], 2012. – 293-298 с.
2. Волошин О.Ф. Моделі та методи прийняття рішень / Волошин О.Ф., Машенко С.О. – К.: Видавничий центр «Київський університет», 2010. - 336 с.
3. O.F. Voloshyn Educational system support to decision making theory / O.F. Voloshyn, D.I. Kovaliov, [Computer Science And Information Technologies Conference], Yerevan 2013, 433-434 с.
4. Снитюк В.Е. Интеллектуальное управление оцениванием знаний / Снитюк В.Е., Юрченко К.М. – «Маклаут», Черкасы 2013. – 224 с.
5. Pro ASP. NET 4.5 in C#. 5th edition / A. Freeman [та ін.], Apress, 2013. – 1228 p.
6. H.Kniberg Scrum and XP from the Trenches / [C4Media, publisher of InfoQ.com], 2007 – 168 p.

References

1. VOLOSHYN, O.F., KOVALIOV, D.I. (2012) *Software system to support courses on the theory of decision-making*. Problems of Computer Intellectualization, Kyiv-Sofia: ITHEA. – p.293-298.
2. VOLOSHYN, O.F., MASHCHENKO, S.O. (2010) *Models and methods of decision-making*. Publishing center "Kiev University", Kiev.
3. VOLOSHYN, O.F., KOVALIOV, D.I. (2013) *Educational system support to decision making theory*. Computer Science and Information Technologies Conference, Yerevan. – p.433-434
4. SNYTYUK, V.E., YURCHENKO, K.M. (2013) *Intelligent control of knowledge evaluation*. Cherkasy: Maklout.
5. FREEMAN, A., MACDONALD M. and SZPUSZTA, M. (2013) *Pro ASP.NET 4.5 in C#*. 5th edition, Apress. – 1228 p.
6. KNIBERG, H. (2007) *Scrum and XP from the Trenches (Enterprise Software Development)*, C4Media, publisher of InfoQ.com – 168 p.

Надійшла до редакції 21.01.2015

УДК 681.3.062

Марченко О. О., д.ф.-м.н., доц.

O.O. Marchenko, Doctor of Sciences (Physics & Mathematics), Associate Professor

**Алгоритм виділення явних
семантичних зв'язків з векторів
семантико-синтаксичної валентності
концептів онтології**

**An algorithm for extraction of explicit
semantic relations from semantic-syntactic
valence vectors of ontology concepts**

Київський національний університет імені
Тараса Шевченка, 83000, м. Київ, пр-т.
Глушкова 4д, e-mail: rozenkrans@yandex.ua

Taras Shevchenko National University of Kyiv,
83000, Kyiv, Glushkova st., 4d,
e-mail: rozenkrans@yandex.ua

В роботі представлено метод перетворення векторів семантико-синтаксичної валентності синсетів WordNet у форму явних семантичних зв'язків-відношень між концептуальними вузлами онтологічної ієрархічної мережі. Вектори семантико-синтаксичної валентності синсетів, які є лише неявним представленням семантичних зв'язків між поняттями, методом «повторного читання» текстів навчального корпусу трансформуються у ребра семантичних відношень в графі онтології.

Ключові слова: штучний інтелект, комп'ютерна лінгвістика, семантика, здобування знань.

The work describes an algorithm for automatic detection of semantic relations between ontological knowledge base nodes by analyzing matrices of semantic-syntactic words valences. These matrices are obtained by means of non-negative factorization of syntactic words combinability tensor. The tensor has been generated in the course of frequency analysis of syntactic structures obtained from the large text corpora sentences taken from English Wikipedia and Simple English Wikipedia articles.

The vectors of semantic-syntactic synsets valences implicitly describe semantic relations between concept nodes in WordNet. The main purpose of the paper is developing of a method for transformation the implicit form of semantic-syntactic relations between concepts in ontology to explicit form of semantic relations structure. Presented method pretends to become a good basis for developing new effective tools for automating the process of building ontological linguistic knowledge bases. This perspective proves theoretical and practical value of the proposed algorithm.

Keywords: artificial intelligence, computational linguistics, semantics, knowledge extraction.

Статтю представив д.ф.-м.н., проф. Анісімов А.В.

У статті [1] описується метод автоматичного визначення семантичних відношень між концептами онтології. Дані зв'язки описуються так званими векторами семантико-синтаксичних валентностей концептів. В отриманій моделі можна легко перевірити наявність лінійного предикативного α - β -зв'язку між конкретними трьома концептами-вузлами WordNet [2] - a , b і c або кільцевого синтагматичного α - β -зв'язку між концептами a і b , що є достатнім для вирішення ряду задач в аналізі текстів природної мови. З іншого боку алгоритми, що використовують пошук найкоротших шляхів-ланцюжків між вузлами онтологій, не можуть безпосередньо

використовувати дану модель без явного опису відношень між концептами-вузлами у вигляді інцидентних ребер. Крім того, складно розглядати поповнення семантичної мережі новими відношеннями як процес збагачення онтологій без генерації явного опису семантичних зв'язків між понятійними вузлами.

З цього випливає необхідність побудови методів автоматичного виділення явних семантичних зв'язків з векторів семантико-синтаксичної валентності концептів онтології. Ідея методу виділення явних семантичних зв'язків предикативно-аргументного або кільцевого синтагматичного типу з векторної

моделі представлення відношень між вузлами-концептами онтології полягає в повторному «читанні» оброблених текстів корпусу за допомогою процедури синтаксичного аналізу, реалізованого на основі алгоритму Кока-Янгера-Касамі[3] на базі векторів семантико-синтаксичної валентності концептів, прив'язаних до синсетів WordNet.

Схематично можна представити цей процес наступним чином:

Алгоритм послідовно аналізує речення текстів вхідного корпусу і будує керуючі простори синтаксичних структур речень текстів.

Коли алгоритму потрібно визначити, чи можлива побудова кільцевого синтагматичного зв'язку між словами a та b , – переходимо до синсет-вузлів WordNet $\{A_i\}$, на які посиляється слово a , і до синсет-вузлів $\{B_i\}$, на які посиляється слово b .

Обчислюємо значення $(W_{a'}, H_{b'}^T)$, для слів $a' \in \{A_i\}$, $b' \in \{B_i\}$: якщо існують такі $a'' \in \{A_i\}$ і $b'' \in \{B_i\}$, що $(W_{a''}, H_{b''}^T) > T1$ (де $T1$ – пороговий рівень, що визначається емпірично; початково було прийнято $T1 \geq 1$), тоді кільцевий синтагматичний зв'язок утворюється та встановлюється семантичне відношення типу «об'єкт - атрибут» між концептом-синсетом A_k : ($a'' \in A_k$), та концептом-синсетом B_j : ($b'' \in B_j$) в мережі онтології.

Існування хоча б однієї такої пари ($a'' \in A_k$, $b'' \in B_j$) гарантує той факт, що дані для матриці D були отримані з тих же текстів та речень, які обробляються в даний момент. Відповідне синтаксичне відношення між словами a і b було занесено в матрицю D і потрапило в вектори її факторизованих матриць W і H . Під час розщеплення цих векторів семантико-синтаксичних валентностей відповідні значення були прив'язані до деяких двох синсетів, які й будуть знайдені цим алгоритмом, – між ними і встановиться дане відношення у явному вигляді. Алгоритм гарантовано коректно знаходить і прописує семантичне відношення між синсетами онтології при виконанні двох наступних умов:

1. При навчанні відповідний синтаксичний зв'язок був коректно прописаний в дерево підпорядкування і дерево виведення речення Стенфордським парсером. Тоді відповідні керуючі простори, на яких навчалася система, були складені коректно, і даний зв'язок був правильно прописаний в матриці

D і, отже, гарантовано присутній у векторах факторизованих матриць W і H .

2. Алгоритм розщеплення векторів семантико-синтаксичної валентності правильно розщепив відповідні даному зв'язку вектори з W і H і коректно виконав прив'язку до вірно встановлених вузлів – синсетів WordNet.

Проведені експерименти свідчать про високу точність роботи використовуваних алгоритмів, тому можна говорити про високу надійність даного методу прописування семантичних зв'язків між синсетами WordNet.

Для збільшення ступеня надійності алгоритму можна ввести додаткову перевірку на наявність подібних зв'язків у синів і/або батьків даних синсетів A_k та B_j . Якщо ці зв'язки існують, тобто $\exists A'_k$ – батьківський/синівський синсет A_k та $\exists B'_j$ – батьківський/синівський синсет B_j ; $\exists a' \in A'_k$ та $\exists b' \in B'_j$ такі, що $(W_{a'}, H_{b'}^T) > T1$, то імовірність побудови коректного семантичного зв'язку між синсетами A_k і B_j значно зростає.

Необхідно відзначити, що зв'язок між синсетами A_k та B_j буде знайдено безпосередньо відразу, якщо умови 1 і 2 виконуються. Якщо при навчанні зв'язок не був знайдений, або був знайдений, але неправильно прописаний до некоректної пари синсетів, то додаткові перевірки можуть збільшити надійність роботи методу визначення семантичних відношень між синсетами.

По-суті для «повторного читання» текстів корпусу використовується парсер-генератор керуючих просторів, описаний в роботі[4], з тією лише різницею, що там для визначення наявності зв'язків між словами використовувалися вектори семантико-синтаксичної валентності слів, отримані після факторизації матриці синтагматичних зв'язків D та тензору предикативних зв'язків F , а в даному випадку нам потрібно перейти від слів до їх значень – синсетів WordNet, так як в даній моделі вектори семантико-синтаксичної валентності прив'язані виключно до концептів-синсетів. Таким чином, парсер «вимушено» переходить від слів і словосполучень до їх семантичних значень, і дана фаза семантичного аналізу здійснюється паралельно з синтаксичним аналізом.

Окремого розгляду вимагає випадок, коли алгоритм виявляє більш ніж одну пару синсетів A_k і B_j ($a'' \in A_k$, $b'' \in B_j$): $(W_{a''}, H_{b''}^T) > T1$.

Сам алгоритм Кока-Янгера-Касамі являє собою процес динамічного збирання всіх можливих варіантів синтаксичної структури речення. На кожному рівні процесу побудови керуючого простору синтаксичної структури відбувається злиття двох структур (двох точок керуючих просторів) в одну більш велику структуру - точку, яка успадковує тим чи іншим чином своє лексичне значення у точок, які її утворюють. Подальше приєднання на більш високому структурному рівні відбуватиметься згідно векторів семантико-синтаксичної валентності тих синсетів, на які будуть посилатися нові лексичні значення утворених структур-точок керуючого простору. Потрібно зберігати в клітинках таблиці об'єднання всіх можливих пар A_k та B_j ($a'' \in A_k$, $b'' \in B_j$): $(W_{a'', H_{b''}^T}) > T_1$.

У процесі генерації структури алгоритмом Кока-Янгера-Касамі невірні варіанти відпадають внаслідок неможливості встановлення семантико-синтаксичного зв'язку на верхніх рівнях таблиці і неможливість побудови цілісної структури. У разі успіху повністю побудована структура керуючого простору в точках міститиме об'єднання відповідних коректних синсетів Wordnet з використанням правильних семантико-синтаксичних відношень-зв'язків між ними. Отже, можна внести ці відношення в семантичну базу WordNet шляхом додавання відповідних семантичних зв'язків між синсетами вже після закінчення побудови повної і цілісної структури керуючого простору для даного речення.

Розглянемо більш формально опис алгоритму

Bxiđ. WordNet з векторами семантико-синтаксичної валентності, прив'язаними до їх синсетів -вузлів, та вхідний ланцюжок речення $\omega = a_1 a_2 \dots a_n \in \Sigma^+$

Buxiđ. Таблиця семантико-синтаксичного аналізу T , що описує керуючий простір синтаксичної структури вхідного речення. Побудований КП містить у своїх точках значення слів - синсети WordNet, з'єднані у просторі семантичними α - β -відношеннями.

Алгоритм

1. Покладемо $t_{i1} = \{\{A_i\} | \{A_i\} - \text{синсети, на які посилаються } a_i \forall i=1..n\}$.

Якщо таких немає (наприклад для прийменників, сполучників і т.д.) – то замість синсетів t_{i1} містить просто лексеми, які також мають свої власні вектори семантико-синтаксичної валентності, завдяки яким вони на наступному етапі вступають у зв'язок з семантично значущими лексемами.

2. Припустимо, що вже обчислені t_{ij} для всіх $1 \leq i \leq n$ і всіх $1 \leq j' < j$.

Покладемо $t_{ij} = \{\{A_i\} \text{ для деякого } 1 < k \leq j, \{B_i\} \in t_{ik} \text{ та } \{C_i\} \in t_{i+k,j-k}, \xi(B_i, C_i) = A_i\}$. Оскільки $1 < k \leq j$, то k та $j-k$ менші за j . Таким чином, t_{ik} і $t_{i+k,j-k}$ обчислюються раніше, ніж

$A_i \Rightarrow +a_i a_{i+1} \dots a_{i+j-1}$. Після цього кроку з $A_i \in t_{ij}$ слідує $A_i \Rightarrow (B_i, C_i) \Rightarrow (+a_i \dots a_{i+k-1}, C) \Rightarrow +a_i \dots a_{i+k-1} a_{i+k} \dots a_{i+j-1}$.

3. Повторюємо крок (2) до тих пір, поки не стануть відомі t_{ij} для всіх $1 \leq i \leq n$ та $1 \leq j \leq n-i+1$.

Розглянемо роботу функції $\xi(B, C)$. Вона перевіряє точки B і C , чи можливо встановити між ними деякий зв'язок - кільцевий синтагматичний α - β -зв'язок, α -зв'язок лінійного предикативного відношення або β -зв'язок лінійного предикативного відношення.

Для визначення можливості встановлення кільцевого синтагматичного α - β -зв'язку між точками B та C перевірка здійснюється обчисленням операції скалярного добутку (W_B, H_C^T) (W_B – вектор концепту–синсету, який є семантичним значенням точки B , H_C – вектор концепту–синсету, який є семантичним значенням точки C): якщо $(W_B, H_C^T) \geq T_{\alpha\beta}$, то функція ξ встановлює даний кільцевий синтагматичний зв'язок, розміщує його в новій точці A , та обчислює її нове лексико-семантичне значення. Воно або успадковується з точки B , як головної точки пари, або в результаті об'єднання утворюється новий смисл («Теза» + «Чьорча» = «Теза Чьорча»). Точка A отримує знайдене функцією ξ лексичне значення і відповідний синсет (або набір синсетів у разі

неоднозначності) як семантичне значення. До синсету (або до синсетів) прив'язані його (їх) вектори семантико-синтаксичної валентності з різних матриць тензорної моделі мови.

Для визначення можливості встановлення β -зв'язку лінійного предикативного відношення між точками B та C перевірка здійснюється обчисленням операції скалярного добутку (Y_B, Z_C) (Y_B – вектор Y концепту–синсету, який є семантичним значенням точки B , Z_C – вектор Z концепту–синсету, який є семантичним значенням точки C): якщо $(Y_B, Z_C) \geq T_\beta$, то функція ξ встановлює даний β -зв'язок лінійного предикативного відношення, розміщує його в новій точці $A1$, та обчислює її лексико-семантичне значення. У $A1$, як правило, успадковуються обидва лексичних і обидва семантичних значення з точок B і C .

Для визначення можливості встановлення α -зв'язку лінійного предикативного відношення між точками D і $A1$ перевірка виконується

обчисленням операції $\sum_{i=1}^k X_D[i] * Y_B[i] * Z_C[i]$

(X_D – вектор концепту-синсету, що є значенням точки D , Y_B – вектор концепту-синсету, що є значенням точки B з $A1$, Z_C – вектор концепту-синсету, що є значенням точки C з $A1$): якщо

$\sum_{i=1}^k X_D[i] * Y_B[i] * Z_C[i] \geq T_\alpha$, то функція ξ

встановлює α -зв'язок лінійного предикативного відношення та завершує, таким чином, повну лінійну предикативну послідовність α - β -зв'язків керуючого простору даного речення.

Після того, як алгоритм успішно завершує збірку цілісного керуючого простору речення, повністю побудована структура КП в точках містить об'єднання відповідних коректних синсетів Wordnet з використанням правильних семантико-синтаксичних відношень між ними.

Далі йде етап обходу побудованої структури КП речення з перенесенням знайдених між синсетами семантичних зв'язків у базу даних WordNet у вигляді відповідних семантичних відношень між вузлами-концептами онтології.

Якщо між двома точками КП, що містять синсети A і B в якості їх семантичних значень, встановлений кільцевий синтагматичний α - β -зв'язок, то між A і B у WordNet прописується

явне відношення типу «об'єкт-атрибут» або «дія-атрибут».

Якщо між точками КП, що містять синсети A і B в якості їх семантичних значень, встановлений лінійний предикативний α -зв'язок, то між A і B у WordNet прописується явне відношення «суб'єкт-дія» або «актант-дія».

Якщо між точками КП, що містять синсети B і C в якості їх семантичних значень, встановлений лінійний предикативний β -зв'язок, то між B та C в WordNet прописується явне відношення «дія-об'єкт», або «дія-реципієнт», або «дія-інструмент» і т.д. - залежно від використання того чи іншого синтаксичного шаблону в реченні для вираження даного β -зв'язку. Семантичні рольові відмінки Філмора-Грубера [5,6], які можуть ефективно використовуватися в даному випадку, описують відповідність синтаксичних шаблонів типам семантичних відношень.

Одночасно з додаванням в онтологію набору семантичних зв'язків з одного речення фіксуються мітки про їх сумісне використання в рамках однієї функціональної аргументно-предикатної схеми відповідного концепту-дієслова.

Таким чином, після повторної обробки текстових корпусів всі неявні семантичні зв'язки синсетів, виражені векторами семантико-синтаксичних валентностей, прописуються в WordNet явним чином додаванням відповідних ребер семантичних відношень в граф онтології.

Висновки

У роботі був запропонований метод трансформації векторів семантико-синтаксичної валентності синсетів WordNet, які є неявною формою представлення семантичних зв'язків-відношень між концептуальними вузлами онтологічної ієрархічної мережі, у явну форму представлення в якості ребер семантичних відношень в онтологічному графі. Метод полягає у повторній обробці текстів навчального корпусу і з занесенням знайдених семантико-синтаксичних відношень «предикатно-аргументного» типу до бази даних онтології. В результаті відбувається поповнення ієрархічної мережі онтології семантичними зв'язками «горизонтального» типу, що становить теоретичну і практичну цінність даного методу.

Список використаних джерел

1. Анисимов А.В. Определение семантических валентностей концептов онтологий с помощью неотрицательной факторизации тензоров больших текстовых корпусов / А.В. Анисимов, А.А. Марченко, Т.Г. Вознюк // Кибернетика и Системный Анализ .– 2014. – № 3. – С. 3-16.
2. Miller G. A. WordNet: An online lexical database. / G. A. Miller, R. Beckwith, C. D. Fellbaum, D. Gross, K. Miller // International Journal of Lexicography – 1990. – 3(4). – pp. 235-244.
3. Kasami T. An efficient recognition and syntax-analysis algorithm for context-free languages/ Kasami T. // Scientific report AFCRL-1965-758. Air Force Cambridge Research Lab, Bedford, MA. – 1965.
4. Марченко О. О. Розробка семантико-синтаксичної моделі природної мови за допомогою методів невід'ємної тензорної і матричної факторизації / О. О. Марченко // Проблеми програмування.– №2-3. – 2014.– С. 263-272.
5. Fillmore Ch. The Case for Case /Ch. Fillmore // Universals in Linguistic Theory. Ed. by E. Bach and R. Tarms. – New York. – London. – Toronto, 1968.
6. Gruber J. S. Studies in Lexical Relations / J.S.Gruber// Ph.D. diss., MIT.- 1965. - Reprinted in Lexical Structure in Syntax and Semantics (1976). Amsterdam: North-Holland.

References

1. ANISIMOV A. V., MARCHENKO A. A., VOZNIUK T. G. (2014), *Determination of semantic valences of ontologies concepts using non-negative tensor factorization of large text corpora*, Cybernetics and System Analysis 3, pp. 3-16.
2. MILLER G. A., BECKWITH R., FELLBAUM C. D., GROSS D., MILLER K. (1990) *WordNet: An online lexical database*, International Journal of Lexicography: 3(4), pp. 235–244.
3. KASAMI T. (1965), *An efficient recognition and syntax-analysis algorithm for context-free languages*, Scientific report AFCRL-1965-758. Air Force Cambridge Research Lab, Bedford, MA.
4. MARCHENKO O.O.(2014), *Development of semantic and syntactic natural language model using the methods of non-negative matrix and tensor factorization*, Problems of Programming: 2-3, pp. 263-272.
5. FILLMORE C. (1968), *The Case for Case*, Universals in Linguistic Theory, Ed. by E. Bach and R. Tarms, New York, London, Toronto.
6. GRUBER J. S. (1976) *Studies in Lexical Relations*, Ph.D. dissertation, MIT. Reprinted in Lexical Structure in Syntax and Semantics, Amsterdam: North-Holland.

Надійшла до редколегії 30.06.15

УДК 519.8

Машченко С.О., д. ф.-м. н., проф.,
Моренець В.І., аспірант

**С-ядро кооперативної гри з нечіткою
множиною дозволених коаліцій**

Київський національний університет імені
Тараса Шевченка, 83000, м. Київ, пр-т.
Глушкова 4д,
e-mail: msomail@yandex.ua
e-mail: v.i.morenets@gmail.com

S.O. Mashchenko, D.Sci (Phys-Math.), Prof.,
Morenets V.I., Grad. Stud.

**C- core of cooperative game with the fuzzy set
of admissible coalitions**

Taras Shevchenko National University of Kyiv,
83000, Kyiv, Glushkova st., 4d,
e-mail: msomail@yandex.ua
e-mail: v.i.morenets@gmail.com

Розглядається кооперативна гра в характеристичній формі з нечіткою множиною допустимих коаліцій. Показано, що С-ядро цієї гри буде нечіткою множиною типу 2 (нечітка множина, функція належності якої приймає нечіткі значення), побудована її функція належності. Запропонована трьох-критеріальна задача пошуку раціонального розв'язку. Розглянутий підхід до отримання наближеного розв'язку цієї задачі.

Ключові слова: кооперативна гра, функція належності, нечітка множина типу 2, ядро гри.

The problem of the choice of an equitable distribution of profits among the players, if they agree to cooperation, but not clearly understand what the coalition may be admissible. This is due to the fact that in many real-life situations, not every group of players has the ability to cooperate and therefore all the coalition can not be formed. The well-known concept of C-core cooperative games generalized to the case of fuzzy sets of admissible coalition players. In this case, it can be described as the intersection of fuzzy set of crisp sets of distributions which satisfy to principle of separation. It is the fuzzy set of type 2 (a fuzzy set, membership function which takes a fuzzy values). This membership function is build. A three-criteria optimization problem of finding a rational solution is proposed. The approach to obtain an approximate solution of this problem is considered. The proposed approach extends the scope of the theory of fuzzy sets in case of cooperative games with fuzzy set of permissible coalitions and may provide a new approach to solving other problems of game theory.

Keywords: cooperative game, membership function, fuzzy set of type 2, core of game.

Статтю представив д. т. н., проф. Гаращенко Ф.Г.

Вступ. Кооперативна теорія ігор розв'язує задачу «справедливого» розподілу прибутку між гравцями у випадку їхньої співпраці. Тим не менш, у багатьох реальних життєвих ситуаціях, не кожна група гравців має можливість співпрацювати і тому всі коаліції не можуть утворюватися. В цьому випадку виникає кооперативна гра з обмеженою кооперацією. Причини обмежень на створення можливих коаліцій можуть бути різними, наприклад обмеження, які викликані правами; обмеження на максимальну кількість гравців, яким дозволено співпраця; обмеження, які пов'язані з відсутністю зв'язку між гравцями, або обмеження, які викликані потребою у згоді начальників гравців на формуванні коаліції з іншими.

Класична кооперативна гра у характеристичній формі $\langle N, v \rangle$ задається набором гравців $N = \{1, \dots, n\}$, $n \geq 2$, і характеристичною функцією $v: 2^N \rightarrow R^1$, яка для кожної коаліції гравців дорівнює її гарантованому виграшу, причому $v(\emptyset) = 0$, $v(S \cup T) \geq v(S) + v(T)$ для $\forall S, T \subset N$, $S \cap T = \emptyset$.

У більш загальному випадку, коли деякі коаліції можуть бути забороненими, кооперативна гра може бути описана трійкою $\langle N, v, \Omega \rangle$, де $\Omega \subseteq 2^N \setminus N$ – набір можливих коаліцій (в яких гравці можуть співпрацювати разом), і називається грою з обмеженою кооперацією. Знаходження справедливого

розподілу виграшу всієї спільноти гравців для гри з обмеженою кооперацією представляє собою більш загальну проблему, ніж для звичайної кооперативної гри. Через це будь-який розв'язок гри з обмеженою кооперацією є узагальненням деякого розв'язку класичної кооперативної гри. Узагальнення вектора Шеплі для ігор з обмеженою кооперацією розглядалися R. Myerson [1], пізніше їх аналізували G. Owen [2] і P. Borm, G. Owen і S. Tijs [3]. E. Algaba, J. Bilbao, P. Borm і J. Lopez [4, 5] вважають, що часткова модель співпраці на основі так званих союзних стійких систем є узагальненням ситуації спілкування.

В деяких випадках буває складно чітко вказати заборонені коаліції. В цій роботі розглядається С-ядро кооперативної гри з нечіткою множиною дозволених коаліцій.

С-ядро кооперативної гри. Спочатку розглянемо класичну кооперативну гру у характеристичній формі $\langle N, v \rangle$. Одним із методів розподілу виграшу $v(N)$ всієї спільноти гравців є С-ядро, яке складається з поділів (векторів $x = (x_1, \dots, x_n)$ таких, що $\sum_{i \in N} x_i = v(N)$ та $x_i \geq v(\{i\})$ для $\forall i \in N$), які задовольняють умовам $\sum_{i \in S} x_i \geq v(S)$ для $\forall S \subset N$. Як правило С-ядро містить більше одного елемента, тому виникає проблема вибору. Ця проблема може розв'язуватися різними методами, але всі вони зводяться до знаходження у певному сенсі «центра» С-ядра. Для цього розв'язується задача математичного програмування:

$$\begin{aligned} g(x) &\rightarrow \min, \\ \sum_{i \in S} x_i &\geq v(S), \quad S \in 2^N \setminus N, \\ \sum_{i \in N} x_i &= v(N), \end{aligned}$$

де $g(x)$ – деяка функція, за допомогою якої вибирають конкретний поділ з С-ядра. Вигляд цієї функції залежить від предметної області та домовленості між гравцями.

С-ядро кооперативної гри з нечіткою множиною дозволених коаліцій. Припустимо, що гравці не можуть чітко сказати, які коаліції з множини $\Omega = 2^N \setminus N$ усіх можливих, можуть бути дозволеними, а їм відома лише функція належності $\mu(S)$, $S \in \Omega$, нечіткої множини $\tilde{\Omega} \subseteq \Omega$ дозволених коаліцій. Тоді виникає кооперативна гра з нечіткою множиною

дозволених коаліцій, яка задається трійкою $\langle N, v, \tilde{\Omega} \rangle$. Якщо узагальнити на цей випадок метод розподілу виграшу всієї спільноти гравців згідно до концепції С-ядра, то ми одержимо задачу, яку можна сформулювати у так:

$$g(x) \rightarrow \min, \quad (1)$$

$$\sum_{i \in S} x_i \geq v(S), \quad S \in \tilde{\Omega}, \quad (2)$$

$$\sum_{i \in N} x_i = v(N), \quad (3)$$

де запис (2) слід розуміти таким чином. Нехай $F_S = \{x = (x_1, \dots, x_n) \mid \sum_{i \in N} x_i = v(N), \sum_{i \in S} x_i \geq v(S)\}$ –

множина векторів x , які задовольняють умові (2) для коаліції $S \in \Omega$ та умові (3). Тоді запис вигляду (2) означає, що множина допустимих розв'язків системи (2), (3) повинна задаватися множиною $\tilde{F} = \bigcap_{S \in \tilde{\Omega}} F_S$, де згідно [7] $\bigcap_{S \in \tilde{\Omega}} F_S$ – перетин нечіткої множини $\tilde{M}$ чітких множин F_S , який представляє собою нечітку множину типу 2.

Перетин нечіткої множини чітких множин. Визначимо це поняття відповідно до підходу, який був запропонований в [7].

В просторі R^n визначимо функцію належності (характеристичну функцію) чіткої множини F_S , $S \in \Omega$, яку позначимо

$$\phi_S(x) = \begin{cases} 1, & \sum_{i \in S} x_i \geq v(S), \\ 0, & \sum_{i \in S} x_i < v(S). \end{cases}$$

Для довільного $x \in R^n$ введемо відношення домінування на множині коаліцій Ω .

Говоритимемо, що коаліція $S \in \Omega$ домінує коаліцію $T \in \Omega$ для $x \in R^n$ і позначати це $S \succ_x T$, якщо справедливі такі нерівності: $\phi_S(x) \leq \phi_T(x)$, $\mu(S) \geq \mu(T)$ і хоча б одна з них строга.

Позначимо

$$\tilde{\mu}(x, S) = \begin{cases} \mu(S), & S \in PO(x), \\ 0, & S \notin PO(x), \end{cases}$$

де $PO(x) = \{S \in \Omega \mid T \not\succ_x S, \forall T \in \Omega\}$.

Перетином нечіткої множини $\tilde{\Omega}$ чітких множин F_S , $S \in \Omega$, називається [7] $\tilde{F} = \bigcap_{S \in \tilde{\Omega}} F_S$ – нечітка множина типу 2, яка задається трійками $(x, \psi(x, y))$, де

$\psi: R^n \times Y \rightarrow \{0,1\}$ – функція належності нечіткого відображення $\tilde{F}$, що виконує роль нечіткої функції належності, яка визначена таким чином:

$$\psi(x, y) = \max_{S \in \Omega} \{\mu(x, S) | \phi_S(x) = y\},$$

якщо $\exists S \in \Omega \quad \phi_S(x) = y$;

$$\psi(x, y) = 0,$$

якщо $\forall S \in \Omega \quad \phi_S(x) \neq y$;

x – елемент простору R^n ;

y – елемент універсальної множини $Y = \{0,1\}$

значень функції належності $\psi(x, y)$ нечіткої множини $\tilde{F}$ типу 2.

Значення $\psi(x^0, 1)$ можна розуміти як ступінь належності вектора $x^0 \in R^n$ множині $\tilde{F}$. Відповідно значення $\psi(x^0, 0)$ має сенс ступеню відсутності належності $x^0 \in R^n$ множині $\tilde{F}$.

Оскільки функція приналежності $\phi(x, y)$ однозначно визначає нечітке відображення $\tilde{F}$, яке, у свою чергу, однозначно визначає нечітку множину $\tilde{F}$ типу 2, то надалі називатимемо $\phi(x, y)$ функцією належності нечіткої множини $\tilde{F}$ типу 2, а її значення – достовірністю ступеня належності $y \in Y = [0,1]$ елемента простору R^n нечіткій множині $\tilde{F}$ типу 2.

Позначимо $S^* = \text{Arg} \max_{T \in \Omega} \mu(T)$ – множину коаліцій з максимальним ступенем належності до нечіткої множини допустимих коаліцій.

Спростити побудову функції належності $\psi(x, y)$ дозволяє наступна теорема [7].

Теорема. Нехай $F_S, S \in \Omega$, – чіткі множини, які задані в просторі R^n відповідними характеристичними функціями $\phi_S(x), x \in R^n, S \in \Omega$; $\mu(S), S \in \Omega$, – функція належності нечіткої множини коаліцій $\tilde{\Omega}$. Для того, щоб нечітка множина $\tilde{F}$ типу 2, яке задана функцією належності $\psi(x, y), x \in R^n, y \in \{0,1\}$, була перетином нечіткої множини $\tilde{\Omega}$ чітких множин $F_S, S \in \Omega$, тобто $\tilde{F} = \bigcap_{S \in \Omega} F_S$ необхідно і достатньо, що б для $x \in R^n$:

$$\psi(x, 0) = \begin{cases} \max_{\phi_S(x)=0} \mu(S), & \exists S \in \Omega \quad \phi_S(x) = 0, \\ 0, & \forall S \in \Omega \quad \phi_S(x) = 1, \end{cases}$$

$$\psi(x, 1) = \begin{cases} \max_{S \in \Omega} \mu(S), & \forall S \in S^* \quad \phi_S(x) = 1, \\ 0, & \exists S \in S^* \quad \phi_S(x) = 0. \end{cases}$$

З викладених вище міркувань стає зрозуміло, що множина $\tilde{F}$ є нечіткою множиною типу 2, яка задається функцією належності $\psi(x, y), x \in R^n, y \in \{0,1\}$, де

$$\psi(x, 0) = \begin{cases} \max_{\phi_S(x)=0} \mu(S), & \exists S \in \Omega \quad \sum_{i \in S} x_i < v(S), \\ 0, & \forall S \in \Omega \quad \sum_{i \in S} x_i \geq v(S), \end{cases} \quad (4)$$

- достовірність недопустимості $x \in R^n$;

$$\psi(x, 1) = \begin{cases} \max_{S \in \Omega} \mu(S), & \forall S \in S^* \quad \sum_{i \in S} x_i \geq v(S), \\ 0, & \exists S \in S^* \quad \sum_{i \in S} x_i < v(S), \end{cases} \quad (5)$$

- достовірність допустимості вектора $x \in R^n$.

Пошук раціональних поділів. В реальній ігровій ситуації гравці прагнуть мінімізувати окрім цільової функції $g(x)$ ще й достовірність $\psi(x, 0)$ недопустимості поділу, а також максимізувати достовірність $\psi(x, 1)$ його допустимості. Таким чином, перед гравцями постає наступна трьохкритеріальна задача:

$$\begin{aligned} g(x) &\rightarrow \min, \\ \psi(x, 0) &\rightarrow \min, \\ \psi(x, 1) &\rightarrow \max, \quad x \in R^n. \end{aligned} \quad (6)$$

Позначимо SO – множини оптимальних за Слейтером (слабо оптимальних за Парето) розв'язків цієї задачі. Нагадаємо, що розв'язок x^* називається оптимальним за Слейтером для задачі вигляду (6) якщо $\forall x \in X$, для якого мають місце нерівності: $g(x^*) > g(x), \psi(x^*, 1) < \psi(x, 1), \psi(x^*, 0) > \psi(x, 0)$.

Загальним раціональним розв'язком задачі (1) – (3) знаходження розподілу виграшу всієї спільноти гравців в кооперативній грі $\langle N, v, \tilde{\Omega} \rangle$ з нечіткою множиною дозволених коаліцій називається нечітка множина $\tilde{X}$ типу 2 з функцією належності

$$\psi_{\tilde{X}}(x, y) = \begin{cases} \psi(x, y), & x \in SO, \\ 0, & x \notin SO, \end{cases}$$

де $y \in Y = \{0,1\}$.

У разі, коли гравців цікавить конкретний раціональний розв'язок задачі (1) – (3), то його можна вибрати з множини SO за допомогою того або іншого методу багатокритеріальної оптимізації, вирішивши задачу (6).

Легко перевірити, що оптимальний розв'язок x^* «чіткого аналога» (1) – (3) належатиме множині SO із значенням достовірності його недопустимості $\psi(x^*, 0) = 0$. Назвемо цей розв'язок тривіальним розв'язком задачі (1) – (3). Нас, в першу чергу, цікавитимуть нетривіальні розв'язки задачі (1) – (3), які мають строго додатну достовірність недопустимості.

Оптимальний за Слейтером нетривіальний розв'язок x^* задачі (6) будемо називати раціональним розв'язком задачі (1) – (3) з достовірностями $\psi(x^*, 1)$ і $\psi(x^*, 0)$ відповідно його допустимості і недопустимості.

Позначимо $\Omega^\xi = \{S \in \Omega \mid \mu(S) \leq \xi\}$ множини коаліцій, які мають ступені належності нечіткій множині дозволених коаліцій $\tilde{\Omega}$ не більше за $\xi \in (0, 1)$. Спробуємо спростити задачу (6).

Твердження 1. Нехай $\xi \in (0, 1)$ – задане значення параметра, тоді якщо задача:

$$g(x) \rightarrow \min_{x \in R^T}, \quad (7)$$

$$0 < \psi(x, 0) \leq \xi, \sum_{i \in N} x_i = v(N), \quad (8)$$

$$\sum_{i \in S} x_i \geq v(S), S \in \Omega \setminus \Omega^\xi. \quad (9)$$

має оптимальний розв'язок, то він буде раціональним розв'язком задачі (1) – (3) з достовірністю допустимості $\mu^{\max} = \max_{S \in \Omega} \mu(S)$ і з достовірністю недопустимості не більше за ξ .

Доведення. Припустимо супротивне, що $x^* \notin SO$. Тоді $\exists \hat{x}$, який задовольняє умовам (8) і (9), для якого мають місце нерівності: $g(\hat{x}) < g(x^*)$, $\psi(\hat{x}, 1) > \psi(x^*, 1)$, $\psi(\hat{x}, 0) < \psi(x^*, 0)$.

Оскільки x^* – допустимий розв'язок задачі (7) – (9), то з (5) випливає, що $\psi(x^*, 1) = \max_{S \in \Omega} \mu(S) = \mu^{\max}$. Тоді з записаних вище нерівностей виходить, що $\psi(\hat{x}, 0) < \psi(x^*, 0) \leq \xi$ і $\psi(\hat{x}, 1) > \psi(x^*, 1) \geq \mu^{\max}$. Оскільки, згідно (5) $\psi(\hat{x}, 1) \leq \mu^{\max}$, то отримали суперечність.

Твердження доведене.

Спробуємо далі спростити задачу (6).

Твердження 2. Нехай $\xi \in (0, 1)$ – задане значення параметра, тоді якщо задача:

$$\min_{T \in \Omega^\xi} \min_{x \in R^n} g(x); \quad (10)$$

$$\sum_{i \in T} x_i < v(T); \sum_{i \in N} x_i = v(N); \quad (11)$$

$$\sum_{i \in S} x_i \geq v(S), S \in \Omega \setminus \Omega^\xi. \quad (12)$$

має оптимальний розв'язок, то він буде раціональним розв'язком задачі (1) – (3) з достовірністю допустимості $\mu^{\max} = \max_{S \in \Omega} \mu(S)$ і з достовірністю недопустимості не більше за ξ .

Доведення. Нехай $\xi \in (0, 1)$. Позначимо Q – множину допустимих розв'язків задачі (7) – (9), а P^T – множину допустимих розв'язків – (10) – (12). Спочатку покажемо, що

$$\bigcup_{T \in \Omega^\xi} P^T = Q. \quad (13)$$

Покажемо включення $\bigcup_{T \in \Omega^\xi} P^T \subseteq Q$. Якщо

$\bigcup_{T \in \Omega^\xi} P^T = \emptyset$, то воно є очевидним. Нехай $x^* \in \bigcup_{T \in \Omega^\xi} P^T \neq \emptyset$. Припустимо супротивне, що $x^* \notin Q$. Тоді можливі два випадки. В першому

випадку не виконується умова (8), тобто $\psi(x^*, 0) > \xi$. Звідси з формули (4) отримаємо $\psi(x^*, 0) = \max_{S \in \Omega} \{\mu(S) \mid \sum_{i \in S} x_i < v(S)\} > \xi$. Тоді

$\exists S \in \Omega \setminus \Omega^\xi$, для якого $\sum_{i \in S} x_i^* < v(S)$, що суперечить умові (12). В другому випадку не виконується умова (9), тобто $\exists S \in \Omega \setminus \Omega^\xi$, для якого $\sum_{i \in S} x_i^* < v(S)$. Тоді не виконується

відповідна умова (12) і тому $x^* \notin P^T \forall T \in \Omega^\xi$ і як наслідок $x^* \notin \bigcup_{T \in \Omega^\xi} P^T$. Отримали суперечність.

Таким чином, $x^* \in Q$ і тому $\bigcup_{T \in \Omega^\xi} P^T \subseteq Q$.

Покажемо включення $Q \subseteq \bigcup_{T \in \Omega^\xi} P^T$. Якщо

$Q = \emptyset$, то воно очевидне. Нехай $x^* \in Q \neq \emptyset$. Припустимо супротивне, що $x^* \notin \bigcup_{T \in \Omega^\xi} P^T$. Тоді

$x^* \notin P^T \forall T \in \Omega^\xi$. Оскільки умови (9) і (12) співпадають, то $\sum_{i \in S} x_i^* \geq v(S) \forall S \in \Omega^\xi$. Оскільки

$\Omega^\xi = \{S \in \Omega \mid \mu(S) \leq \xi\}$, то з формул (4) випливає або $\psi(x^*, 0) = \max_{S \in \Omega} \{\mu(S) \mid \sum_{i \in S} x_i^* < v(S)\} > \xi$, або

$\psi(x^*, 0) = 0$. В обох випадках одержуємо суперечність з умовою (8). Таким чином,

$x^* \notin \bigcup_{T \in \Omega^\xi} P^T$, звідси випливає $Q \subseteq \bigcup_{T \in \Omega^\xi} P^T$ і

справедлива рівність (13).

Завершимо доведення твердження. З рівності (13) очевидно випливає еквівалентність задач (7) – (9) і (10) – (12), чого достатньо на підставі твердження 1 для доведення.

Твердження доведене.

Таким чином, згідно до запропонованого підходу для знаходження раціонального розв'язку задачі (1) – (3) достатньо зробити наступне:

- вибрати число $\xi \in (0,1)$, яке згідно твердженню 2 характеризує максимальну для гравців достовірність недопустимості розв'язку;

- побудувати множину коаліцій $\Omega^\xi = \{S \in \Omega \mid \mu(S) \leq \xi\}$, які мають ступінь належності до допустимих не більше заданого числа $\xi \in (0,1)$;

- для кожної коаліції $T \in \Omega^\xi$ розв'язати задачу:

$$\min_{x \in R^n} g(x);$$

$$\sum_{i \in T} x_i < v(T); \sum_{i \in N} x_i = v(N);$$

$$\sum_{i \in S} x_i \geq v(S), S \in \Omega \setminus \Omega^\xi.$$

якщо вона має розв'язок, то позначимо його $x^{(T)}$, в протилежному випадку вибираємо наступну коаліцію $T \in \Omega^\xi$;

- з отриманих розв'язків $x^{(T)}$, $T \in \Omega^\xi$, вибираємо рекордне $\hat{x}$ за значенням цільової функції, тобто $g(\hat{x}) = \arg \min_{T \in \Omega^\xi} g(x^{(T)})$.

Висновок. На закінчення слід зазначити, що запропонований підхід розширює область застосування теорії нечітких множин на випадок кооперативних ігор з нечіткою множиною допустимих коаліцій і може дати новий підхід до розв'язання інших постановок ігрових задач з нечіткою структурою.

Список використаних джерел

1. Myerson R.B. Graphs and cooperation in games / R.B. Myerson // Math. Oper. Res. – 1977. – 2. – pp. 225–229.
2. Owen G. Values of graph-restricted games / G. Owen // SIAM J. Algebraic and Discrete Methods. – 1986. – 7. – pp. 210–220.
3. Borm P. On the position value for communication situations / Borm P., Owen G., Tijs S. // SIAM J. Disc. Math. – 1992. – 5. – pp. 305–320.
4. Algaba E. The position value for union stable systems / Algaba E., Bilbao J.M., Borm, P., Lopez J.J. // Math. Meth. Oper. Res. – 2000. – 52. – pp. 221–236.
5. Algaba E. The Myerson value for union stable structures / Algaba E., Bilbao J.M., Borm, P., Lopez J.J. // Math. Meth. Oper. Res. 54. – 2001. pp. 359–371.
6. Мащенко С.О. Задача математического программирования с нечетким множеством индексов ограничений / Мащенко С.О. // Кибернетика и системный анализ. – 2013. – № 1. – С. 62 – 68.
7. Орловский С.А. Проблемы принятия решений при нечеткой исходной информации / С.А. Орловский. – М. : Наука, 1981.

References

1. MYERSON, R.B. (1977) Graphs and cooperation in games. Math. Oper. Res. 2. pp. 225–229.
2. OWEN, G. (1986) Values of graph-restricted games. SIAM J. Algebraic and Discrete Methods 7. pp. 210–220.
3. BORM, P.; OWEN, G.; TIJS, S. (1992) On the position value for communication situations. SIAM J. Disc. Math. 5. pp. 305–320.
4. ALGABA, E.; BILBAO, J.M.; BORM, P.; LOPEZ, J.J. (2000) The position value for union stable systems. Math. Meth. Oper. Res. 52. pp. 221–236.
5. ALGABA, E.; BILBAO, J.M.; BORM, P.; LOPEZ, J.J. (2001) The Myerson value for union stable structures. Math. Meth. Oper. Res. 54. pp. 359–371.
6. MASHCHENKO, S. (2013) The problem of the mathematical programming with the fuzzy set of indexes of constraints. Cybernetics and systems analysis. 49(1). pp. 62 – 68.
7. ORLOVSKY, S. (1981) Problems of decision making at fuzzy initial information. Moskva: Nauka.

Надійшла до редколегії 11.06.2015.

УДК 681.3.062

Еміл М. Насіров, аспірант

**Блочно-діагональний підхід до невід'ємної
факторизації розріджених матриць і тензорів
надвеликої розмірності**

Київський національний університет імені
Тараса Шевченка, 03680, м. Київ, пр-т.
Глушкова 4д, e-mail: enasirov@gmail.com

Emil M. Nasirov, Postgraduate Student

**Blocks approach of non-negative huge sparse
matrix and tensors factorization**

Taras Shevchenko National University of Kyiv,
03680, Kyiv, Glushkova ave., 4d,
e-mail: enasirov@gmail.com

В роботі описано алгоритми невід'ємної факторизації розріджених матриць та тензорів, популярну технологію в штучному інтелекті взагалі, та в комп'ютерній лінгвістиці зокрема. Оброблювана матриця – матриця зв'язків Слова×Статті, яка містила частотну оцінку вживання слів в текстах статей Вікіпедії. Наведено існуючі підходи до факторизації матриць та тензорів. Запропоновано блочно-діагональний підхід до невід'ємної факторизації матриць та тензорів надвеликої розмірності. Використання пам'яті та мережових операцій було проаналізовано та оптимізовано.

Ключові слова: штучний інтелект, комп'ютерна лінгвістика, паралельні обчислення

This paper proposes parallel method of non-negative large sparse matrix factorization - very popular technique in computational linguistics. The initial matrix contains frequency estimations of using terms in Wikipedia articles.

The main purpose of the paper is developing of a parallel method of factorization. Proposed blocks factorization approach, which can speedup non-negative sparse extra-large matrix and tensor factorization. Memory usage and data transmitting necessity of factorization algorithm was analysed and optimized. The described algorithms were compared by means of large sparse matrices and tensors processing. Presented method pretends to become a good basis for developing new effective tools for non-negative huge sparse matrix and tensors factorization. It can be used to perform the tasks of industrial scale to factorize sparse matrices of large dimension with an acceptable time using available computing resources. This perspective proves theoretical and practical value of the proposed algorithm.

Keywords: artificial intelligence, computational linguistics, parallel computations

Статтю представив чл.-кор. НАН України, д.ф.-м.н., проф. Анісімов А. В.

Сьогодні невід'ємна факторизація матриць та тензорів є дуже популярною технологією в штучному інтелекті взагалі, та в комп'ютерній лінгвістиці зокрема. Використовуючи невід'ємну факторизацію в рамках парадигми латентно-семантичного аналізу, комп'ютерні лінгвісти застосовують даний підхід для розв'язання таких прикладних задач, як класифікація, кластеризація текстів та термінів[1, 2], побудова мір семантичної близькості[3], автоматичне виділення лінгвістичних структур та відношень

(*Selectional Preferences* [4] та *Verb Sub-Categorization Frames* [5]) і багато інших.

Дана робота описує побудову моделі паралелізації обчислення невід'ємної факторизації розріджених матриць надвеликої розмірності, що представляє особливий інтерес та актуальність з точки зору її застосування у великих NLP системах загального тематичного призначення, не обмежених використанням лише для вузьких предметних областей.

Задача невід'ємної факторизації розріджених матриць надвеликої розмірності постала в процесі розробки системи визначення міри семантичної близькості-зв'язності за технологією Латентного Семантичного Аналізу[6]. Для побудови системи був оброблений великий текстовий корпус статей англomовної Вікіпедії. Обробка корпусу полягала у лексичному аналізі та лематизації лексем речень статей та в обчисленні частот вживання множини слів та словосполучень англійської мови в складі різних статей англomовної Вікіпедії. В результаті була побудована велика матриця Слова×Статті яка містила частотну оцінку вживання слів в текстах статей Вікіпедії. Розмірність матриці дорівнює 2,437,234 слів-словосполучень на 4,475,180 статей англomовної Вікіпедії. Після цього для усунення з матриці випадкових даних був встановлений пороговий рівень $T=3$, частотні оцінки в матриці, які менші за пороговий рівень T , були обнулені). В результаті була побудована розріджена матриця надвеликого розміру - 156,236,043 ненульових елементів при розмірі $2,437,234 \times 4,475,180$. Для невід'ємної факторизації розрідженої матриці такого розміру знадобилася розробка спеціальної моделі паралелізації матричних обчислень, яка була реалізована із застосуванням паралельних розподілених обчислень та обчислень на GPU.

Алгоритм невід'ємної матричної факторизації

Алгоритм невід'ємної матричної факторизації виконує декомпозицію невід'ємної матриці V на невід'ємні матриці W та H таким чином, щоб

$$V \approx WH$$

В якості функції оцінки може бути використана функція виміру відстані між двома невід'ємними матрицями. Однією з таких мір є квадрат Евклідової метрики:

$$\mu = \|A - B\|^2 = \sum_{ij} (A_{ij} - B_{ij})^2$$

Така цільова функція обмежена знизу. Нижня границя 0 досягається тоді і тільки тоді коли $A = B$.

Отже, при використанні Евклідової метрики, факторизація матриці полягає в мінімізації $\|V - WH\|^2$ при умові невід'ємності W та H .

Така цільова функція не зростаюча при наступних правилах:

$$H_{ij} \leftarrow H_{ij} \frac{(W^T V)_{ij}}{(W^T W H)_{ij}} \quad (1)$$

$$W_{ij} \leftarrow W_{ij} \frac{(V H^T)_{ij}}{(W H H^T)_{ij}} \quad (2)$$

Виконання ітерацій алгоритму продовжується до тих пір, поки не буде досягнута стаціонарна точка, або не буде виконана максимальна кількість ітерацій[7].

Аналіз моделі

В Таблиці 1 представлено необхідні об'єми пам'яті для зберігання матриць W та H при різних k при використанні типу даних float (32bit).

K	100	200	300
W	0.98Gb	1.95Gb	2.92Gb
H	1.79Gb	3.58Gb	5.37Gb
Всього	2.76Gb	5.53Gb	8.29Gb

Таблиця 1. Необхідні об'єми пам'яті для зберігання матриць W та H при різних k

Описаний алгоритм на кожній ітерації потребує в 2 рази більше пам'яті для збережень матриць (не враховуючи потреби на збереження початкової матриці V). Враховуючи такі потреби в пам'яті, для виконання алгоритму на одному комп'ютері необхідно проводити збереження даних на жорсткий диск. Далі буде розглянуто та порівняно 2 підходи до паралелізації алгоритму: локальний та розподілений.

Застосування моделі до факторизації тензорів

Для того, щоб обчислити невід'ємні матриці-компоненти $\{A, B, C\}$ зазвичай застосовується обмежений оптимізаційний підхід для мінімізації підходящої функції оцінки. Зазвичай мінімізують наступну функцію:

$$D_F(Y \| [A, B, C]) = \|Y - [A, B, C]\|_F^2 + \alpha_A \|A\|_F^2 + \alpha_B \|B\|_F^2 + \alpha_C \|C\|_F^2$$

, де $\alpha_A, \alpha_B, \alpha_C$ – невід'ємні регуляційні параметри.

Існує щонайменше три різні підходи до розв'язання такої оптимізаційної задачі. Перший підхід полягає у використанні векторної форми функції оцінки: $J(X) = \text{vec}(Y - [A, B, C]) = 0$. Для розв'язання використовується метод найменших квадратів. Цей метод для факторизації тензорів вперше використав Паатеро [7]. Якоб'ян такої функції може мати розмір $ITQJ \times (I + T + Q)$, а отже такий підхід потребує значних розрахункових витрат.

У другому підході Акар, Колда та Дунлаві запропонували штучно оптимізувати функцію оцінки використовуючи техніку нелінійної зв'язної градієнтної оптимізації.[1]

Третім, і найпоширенішим підходом є використання техніки ALS. В цьому підході підраховується градієнт функції оцінки для кожної матриці окремо.

$$\begin{aligned}\nabla_A D_F &= -Y_1(C \odot B) + A[(C^T C) * (B^T B) + \alpha_A I], \\ \nabla_B D_F &= -Y_2(C \odot A) + B[(C^T C) * (A^T A) + \alpha_B I], \\ \nabla_C D_F &= -Y_3(B \odot A) + C[(B^T B) * (A^T A) + \alpha_C I].\end{aligned}$$

Прирівнюючи кожен компоненту градієнта до 0 та накладаючи умову невід'ємності можна отримати ефективні та прості правила оновлення матриць:

$$\begin{aligned}A &\leftarrow [Y_{(1)}(C \odot B) + [(C^T C) * (B^T B) + \alpha_A I]^{-1}]_+, \\ B &\leftarrow [Y_{(2)}(C \odot A) + [(C^T C) * (A^T A) + \alpha_B I]^{-1}]_+, \\ C &\leftarrow [Y_{(3)}(B \odot A) + [(B^T B) * (A^T A) + \alpha_C I]^{-1}]_+.\end{aligned}$$

Основними перевагами такого підходу є висока швидкість збіжності та можливість розподілення для задач великої розмірності.

Блочно-діагональний підхід до факторизації матриць та тензорів

Так як розглядувана розріджена матриця є матрицею зв'язності Слова×Статті, отже можливі перестановки як рядків так і стовпців початкової матриці.

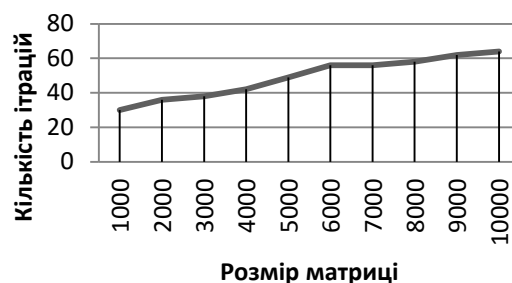
Розглядувана розріджена матриця має 156,236,043 ненульових елементів з 10,907,060,852,120 що складає лише 0.014%. Така розрідженість дозволяє привести матрицю до блочно-діагональної форми за допомогою перестановок рядків та стовпчиків. У випадку, коли слово може належати декільком блокам, рядок може бути розділений на два рядки, що відповідатимуть одному слову.

Аналогічно, до блочно-діагонального вигляду можуть бути приведені тензори за допомогою перестановок шарів.

Розміри блоків найкраще обирати такими, що можуть бути повністю завантаженні в пам'ять. В результаті буде отримано N блоків.

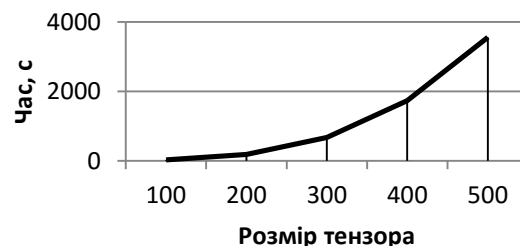
Переваги такого підходу:

1. Не потрібно мережевих операцій на кожній ітерації. Необхідне лише початкове розподілення блоків матриці або тензора між вузлами.



Графік 1.

2. Швидша збіжність. На графіку 1 та графіку 2 показано час необхідний для факторизації квадратних матриць та тензорів різних розмірів відповідно, в один потік без використання GPU. На графіку 3 показано необхідну кількість ітерацій для невід'ємної факторизації квадратних матриць різних розмірів.

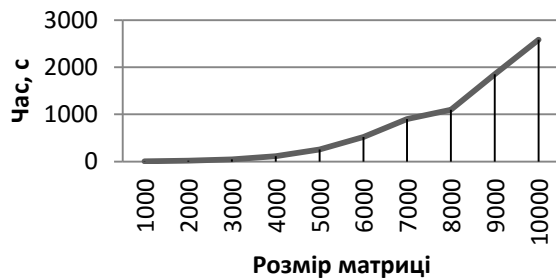


Графік 2.

3. Зменшення k, а отже і зменшення розмірів результатів та об'ємів необхідних даних для зберігання в пам'яті під час ітерацій.

4. Швидший підрахунок функції оцінки μ . Швидший підрахунок добутку $W * H$ так як менше k. Кількість елементів для підрахунку функції оцінки менша в N разів(для матриці), кількість необхідних перевірок менша так як зменшується загальна кількість ітерацій.

На графіку 4 показано залежність якості результатів факторизації від k та кількості ітерацій для звичайного алгоритму факторизації матриць.

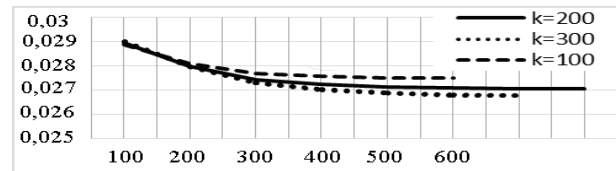


Графік 3.

У випадку зведеної до блочно-діагональної матриці, в результаті факторизації окремих блоків для кожного слова буде отримано один або декілька рядків та ідентифікатор блоку в якому це слово знаходиться. При відновленні, якщо слово та стаття знаходяться в різних блоках – результатом буде 0, інакше результатом буде добуток відповідних рядка та стовпчика.

Список використаних джерел

1. Xu W. Document-clustering based on n-negative matrix factorization / Xu W., Liu X., Gong Y. // In Proceedings of SIGIR'2003.–2003.– p. 267–273.
2. Document clustering using nonnegative matrix factorization / [F.Shahnaz and other] // Information Processing and Management.–2006. – vol. 42. –p. 373 – 386
3. Ukrainian WordNet: Creation and Filling / [A. Anisimov and other] // FQAS-2013. –2013. – p. 649-660
4. T. Van de Cruys A Non-negative Tensor Factorization Model for Selectional Preference Induction / Tim Van de Cruys // Journal of Natural Language Engineering. –2010.– vol.16(4).– p. 417–437.
5. T. Van de Cruys. Multi-way Tensor Factorization for Unsupervised Lexical Acquisition / T. Van de Cruys [and other] // Proceedings of COLING-2012. –2012. –p. 2703–2720.
6. Indexing by Latent Semantic Analysis / [Scott Deerwester and other] // Journal of the American Society for Information Science – 1990 – vol.41 (6) – p. 391–407.
7. D. D. Lee. Algorithms for non-negative matrix factorization / D.D. Lee and H. S. Seung // In NIPS. –2000. –p. 556-562.



Графік 4

Висновки

В роботі описано алгоритм невід'ємної факторизації матриць та алгоритм невід'ємної факторизації тензорів. Запропоновано блочно-діагональний підхід до невід'ємної факторизації матриць та тензорів. Наведено переваги блочно-діагонального підходу щодо швидкодії.

References

1. XU W., LIU X., GONG Y. (2003) *Document-clustering based on n-negative matrix factorization*. In Proceedings of SIGIR'2003. pp. 267–273.
2. SHAHNAZ F. et al. (2006) *Document clustering using nonnegative matrix factorization*. Information Processing and Management. Vol 42, pp. 649-660.
3. ANISIMOV A. et al. (2013) *Ukrainian WordNet: Creation and Filling*. FQAS-2013. pp. 649-660.
4. VAN DE CRUYS T. (2010) *A Non-negative Tensor Factorization Model for Selectional Preference Induction*. Journal of Natural Language Engineering, Vol. 16(4), pp. 417–437.
5. VAN DE CRUYS T. et al. (2012) *Multi-way Tensor Factorization for Unsupervised Lexical Acquisition*. Proceedings of COLING-2012. pp. 2703–2720.
6. DEERWESTER S. et al. (1990) *Indexing by Latent Semantic Analysis*. Journal of the American Society for Information Science. Vol. 41 (6), pp. 391–407.
7. LEE D.D., SEUNG H. S. (2000) *Algorithms for non-negative matrix factorization*. In NIPS, MIT Press. pp. 556-562.

Надійшла до редколегії 30.06.15

УДК 681.3, 004.05

Панченко Т. В., к.ф.-м.н., доц.

**Метод доведення коректності
паралельних програм з використанням
спрощеного стану та його застосування**

Київський національний університет імені
Тараса Шевченка, Україна, 01601, місто Київ,
вул. Володимирська, 64/13
e-mail: pantaras@ukr.net

T. V. Panchenko, PhD, docent.

**Parallel Programs Correctness Proof
Method Using Simplified State with
Applications**

Taras Shevchenko National University of Kyiv,
64/13, Volodymyrska Street, City of Kyiv, Ukraine,
01601
e-mail: pantaras@ukr.net

Доведення коректності паралельних програм є складною задачею, оскільки потоки виконання можуть впливати один на інший через спільні змінні або механізм повідомлень. Класичні методи типу Флойда-Хоара не застосовні у цьому випадку напряму і необхідні нові нетривіальні підходи, які враховують це ускладнення. Розроблений метод для доведення коректності програм (для визначеного спеціального класу програм – а саме серверного програмного забезпечення в архітектурі SMP як-то сервер баз даних або веб-сервер) – є предметом даної статті. Його простота у випадку відсутності локальних даних підпрограм і застосування до класичної задачі-прикладу паралельного додавання до спільної змінної подано у роботі.

Ключові слова: коректність програм, паралельні програми, композиційно-номінативні методи.

The correctness proof for programs with parallelism is complicated problem due to possible inference of one execution path on another via shared data or messaging mechanics. Classical methods like Floyd-Hoare cannot be applied in this situation directly and new non-trivial methods are required to cope with such complexity. Developed method for software correctness proof (for defined custom class of programs – namely server software for SMP architecture like DB-server or Web-server) is the subject of this article. Its simplicity in case of local subroutine data absence and application to classical sample task of shared variable concurrent increment are shown here.

Key Words: software correctness, concurrent programs, composition-nominative methods.

Статтю представив д.ф.-м.н., проф. Буй Д.Б.

Introduction

The methodology for software correctness proof was developed in [1] and presented there in details. Using this methodology one can formulate and prove properties of programs developed in Interleaving Parallel Composition Language (IPCL) [1]. The syntax and operational semantics of IPCL presented *ibid.* Here we concentrate on simplified method for program properties (including correctness) proof for some subclass of all IPCL programs. The notion of simplified state for this kind of reasoning (introduced in [1]) as well as clarification on specific IPCL subclass mentioned above are also discussed here.

IPCL Subclass and Method Description

Following [2], the composition languages class [3] IPCL and methodology proposed by the author for programs properties (including correctness) proof in composition languages IPCL [2] are well suited for the specification and verification of the server-side software in client-server environments.

But it should be noted that in the absence of local data in programs $A, B, \dots, C$ (which are subroutines of program P), it is advisable to consider the concept of a **Simplified State** (here we use the notation and follow the terms according to [2], in particular, the program $P = A^n \parallel B^m \parallel \dots \parallel C^k \in SeqILProgs$, where

SeqILProgs is the sequential programs subclass of all IPCL programs). Simplified state is an aggregated state of the following form $N^{Pmq} \times D$, where $Pmq = \|A_{marks}\| + \|B_{marks}\| + \dots + \|C_{marks}\|$, i.e. the sum of all tags amounts (A_{marks} , B_{marks} , ..., C_{marks}) for programs A , B , ..., C (which are subroutines of P), where A_{marks} is the set of labels for labeled program A – i.e. when we labeled every particular atomic operator (or operation, function call etc.) in the text notation of program A , $\|A\| = \text{card}(A)$ is the power of the set A , and D contains the global (shared, common) data for all routines in the form of *nominative data* [3], which means a set of pairs $\text{name} \mapsto \text{value}$.

First $\|A_{marks}\|$ components of such a state include the number of programs that are executing now operators at appropriate labels of program A in this state, and the sum of these components in each state for the program P is equal to n – i.e. the number of instances (copies) of A program in P (in terms of program P). The following $\|B_{marks}\|$ components contain the number of programs that are executing now operators at appropriate labels of B program in this state, and the sum of these components in each state for the program P is equal to m (the number of instances of the program B in P) in terms of P programs and so on. The last component D contains global shared data for all routines of P .

Simplified state operates with the number of routines which are executing now at appropriate label instead of label identification for each instance of every individual routine (A , B , ..., C). In fact, routines are not distinguished from each other accurate to the label of current execution if they do not change local data (actually – do not have them at all), but only operate with shared global data.

One can obtain a simplified state for the state $S \in A_{marks}^n \times B_{marks}^m \times \dots \times C_{marks}^k \times D \times D^{numprocs}$, $numprocs = n + m + \dots + k$, in the following way. As sets A_{marks} , B_{marks} , ..., C_{marks} are finite, let $A_{marks} = \{A_1, \dots, A_a\}$, ..., $C_{marks} = \{C_1, \dots, C_c\}$. Let $Pr_i(S)$ be the i -th component of a tuple S . Then for any regular state S the corresponding simplified state will be $SS = (a_1, \dots, a_a, \dots, c_1, \dots, c_c, d)$, where $a_j = \|\{i \mid Pr_i(S) = A_j, i \in N_{Pmq}\}\| = P_{S[A_j]}, \quad \forall j \in N_a, \dots,$
 $c_j = \|\{i \mid Pr_i(S) = C_j, i \in N_{Pmq}\}\| = P_{S[C_j]}, \quad \forall j \in N_c,$
 $d = P_{numprocs+1}(S).$

The set of simplified states *SStates* introduced accordingly. The same goes with simplified initial states *SStartStates* (they all have the structure $(n, 0, \dots, 0, \dots, k, 0, \dots, 0, d)$, which means all the routines are at their appropriate start labels), simplified final states *SStopStates* (their structure is $(0, \dots, 0, n, \dots, 0,$

$\dots, 0, k, d)$, which means all the routines have stopped execution at their exit (“after-program”) label) and the transition function (one step execution of P program) over the simplified states $SStep: SStates \rightarrow SStates$. For every particular possible transition of execution control for each routine this new transition function *SStep* decreases the value of some components of the *SStates* vector by 1 (eg. first component) and at the same time increases the value of some other component by 1 (eg. second component), which means the transfer of execution control from one label (in this case A_1) to another label (in this case A_2) for one of the routines (in this case A), and also changes the value of the last components (global data) if the current function is in subclass *Oper* of class *F*. All these objects could be easily obtained by transferring states into their simplified states appropriately or by direct construction.

Sample Correctness Task: Parallel Increment To Shared Variable. Discussion

To prove the properties of both types (defined in [1,2]) one could apply the methodology given in [1,2]. To demonstrate the simplicity of the method [1,2] and the convenience of simplified state model usage, let us discuss a sample that – de facto – (see. [4], [5] and others) become “standard” to check the “efficiency” of parallel programs and methods concerning modeling, execution, model checking and correctness proof. Let the program twice (separately, independently) increments some shared (global) variable concurrently. In the IPCL notation program P will look like $x := x + 1 \parallel x := x + 1$, where x is a shared (global) variable and the set $F = \{f_1\}$, where $f_1(d) \equiv d \nabla [x \mapsto (x \nabla (d) + 1)]$ (here the semantic function f_1 has syntactic form of $x := x + 1$). It is clear that every action is atomic (read, add +1, write back the variable value). The problem is formulated as follows: to prove that when the initial value of shared (global) variable x is 0, the final value of x (after the program stops) will be equal to 2.

Without going into details, “pure” Owicki-Gries method [4] requires program properties to be formulated and tested for each operator taking into account their potential interference and “cross”-interference effects. Recall that inspections quantity is quadratic with respect to the program operator count.

The extended version of the rely-guarantee Owicki-Gries method modification [5] requires two additional variables (for such a trivial program!)

introduction and formulating of nontrivial (!) rely- and guarantee- conditions for application of the method to this task. The proof itself takes more than one page space [5].

TLA [6] offers (for the very similar task) to build a model that is not much easier than the first two, and the formulation of the model and, in fact, the proof itself (with explanations) is again at least the page of formulas in space.

Sample Correctness Task: Parallel Increment To Shared Variable. Proof

Let us now consider a *detailed* proof of *generalization* of this property in IPCL. Explore generalized version of the task first. Thus, let us have a program $P = Inc^n$ (for some fixed $n \in \mathbb{N}$), where $Inc = x := x + 1$. The semantic model described above with $F = \{f_i\}$. We consider a simplified model, as we have no local variables for routines (just shared global x). Labelling algorithm [1] (obviously) will result to the next for Inc : $Inc = [M1] x := x + 1 [M2]$, $SStates = \{(s_1, s_2, d) \mid s_1 \in \mathbb{N}, s_2 \in \mathbb{N}, d \in D\}$, where s_1 is the number of programs (all performed in parallel) executing currently at the label [M1], s_2 is the number of programs executing at the label [M2] (in fact, finished currently), and d is the shared (global) data (containing variable x and its value).

It is clear that $SStep = \{((s_1, s_2, d), (s_1 - 1, s_2 + 1, f_1(d))) \mid s_1 > 0 \text{ \& } s_1 \in \mathbb{N} \text{ \& } s_2 \in \mathbb{N}\}$, $SStartStates = \{(n, 0, [x \mapsto 0])\}$, $SStopStates = \{s \mid s = (0, n, d') \text{ \& } \exists s_1, s_2, \dots, s_l \bullet (s_1 \in SStartStates \text{ \& } s_l = s \text{ \& } (\forall i \in \mathbb{N}_{1..l} \bullet (s_i, s_{i+1}) \in SStep))\}$ – by definition.

Let $PreCond(SS) \equiv (x \Rightarrow (d) = 0)$ and $PostCond(SS) \equiv (x \Rightarrow (d) = n)$, where $SS = (s_1, s_2, d) \in SStates$ – (simplified) state. Consider the invariant $Inv(SS) \equiv (s_2 = x \Rightarrow (d))$. Let us verify $InvCond(Inv, PreCond, PostCond)$: $\forall S \in SStartStates \bullet (PreCond(S) \rightarrow Inv(S)) = PreCond((n, 0, [x \mapsto 0])) \rightarrow Inv((n, 0, [x \mapsto 0])) = True \rightarrow True = True$, $\forall S \in SStopStates \bullet (Inv(S) \rightarrow PostCond(S)) = \forall S = (s_1, s_2, d) = (0, n, d') \in SStopStates \bullet ((s_2 = x \Rightarrow (d)) \rightarrow (x \Rightarrow (d) = n))$, since $s_2 = n$, then the implication is True, and the entire predicate is True, $\forall (S, S') \in SStep \bullet (Inv(S) \rightarrow Inv(S')) = \forall (S, S') = ((s_1, s_2, d), (s_1 - 1, s_2 + 1, f_1(d))) \in SStep \bullet ((s_2 = x \Rightarrow (d)) \rightarrow (s_2 + 1 = x \Rightarrow (f_1(d)))) = \forall (S, S') \in SStep \bullet True$, since it is obvious that $x \Rightarrow (f_1(d)) = x \Rightarrow (d) + 1$.

Hence, Inv indeed is invariant for the program P , besides at the starting states it is a logical

consequence of the pre-condition and it implies post-condition at the final states. Q.E.D.

Sample Correctness Task: Parallel Increment To Shared Variable. Conclusions

Thus, we have proved a *generalized* property. Originally formulated (correctness) property can be derived from a more general just letting $n = 2$.

Note that some of sources mentioned here are considered a variant of the example $P' = x := x + 1 \parallel x := x + 2$ with pre-condition $x = 0$. To solve this problem with the method proposed one should use another (but obvious) invariant $Inv(SS) \equiv (s_2 + 2 * s_4 = x \Rightarrow (d))$, where $SS = (s_1, s_2, s_3, s_4, d) \in SStates$ is simplified state, for generalized program $P'' = (x := x + 1)^n \parallel (x := x + 2)^m$ apart with labels (labelled program P): [M1] $x := x + 1$ [M2], [M3] $x := x + 2$ [M4], and initial states $SStartStates = \{(n, 0, m, 0, [x \mapsto 0])\}$ (with $n = m = 1$).

Note also that the number of inspections to prove is linear regarding the number of program operators in both examples (in general, for every program). Details on this and other examples as well as detailed theory can be found in [1].

Conclusions

The simplified method with simplified state introduced for program properties proof (including program correctness checking) is appropriate for parallel programs without local variables which use only shared ones. This case is appropriate model of server software (in SMP architecture) like Web-servers and Database-server applications and such a kind of computing. The method and the model as well as Interleaving Parallel Compositional Languages class (IPCL) are described in details in [1] and were first presented in [2] and then in [7] with modifications (simplified state).

All reasoning presented here are in terms of composition-nominative approach [8,3], but the main formalism for concurrent behavior modelling is Abstract State Machines [9,10], which in line with compositional semantics gives us enough means to describe interleaving parallel execution within operational approach indeed.

Список використаних джерел

References

1. Панченко Т.В. Композиційні методи специфікації та верифікації програмних систем. Дисертація на здобуття наукового ступеня кандидата фізико-математичних наук.: 01.05.03 / Панченко Тарас Володимирович – К., 2006. – 177 с.
2. Панченко Т.В. Методологія доведення властивостей програм в композиційних мовах IPCL / Тарас Володимирович Панченко // Доповіді Міжнародної конференції “Теоретичні та прикладні аспекти побудови програмних систем” (TAAPSD’2004). – К., 2004. – С. 62–67.
3. Nikitchenko N. A Composition Nominative Approach to Program Semantics / N. Nikitchenko. – Technical Report IT-TR: 1998-020. – Technical University of Denmark, 1998. – 103 p.
4. Owicki S. An Axiomatic Proof Technique for Parallel Programs / S. Owicki, D. Gries // *Acta Informatica*. – 1976. – Vol. 6, № 4. – P. 319–340.
5. Xu Q. The Rely-Guarantee Method for Verifying Shared Variable Concurrent Programs / Q. Xu, W.-P. de Roever, J. He // *Formal Aspects of Computing*. – 1997. – Vol. 9, № 2. – P. 149–174.
6. Lamport L. Verification and Specification of Concurrent Programs / L. Lamport // deBakker J., deRoever W., Rozenberg G. (eds.) *A Decade of Concurrency*, Vol. 803. – Berlin: Springer-Verlag, 1993. – P. 347–374.
7. Панченко Т.В. Модель спрощеного стану для методу доведення властивостей в мовах IPCL та її застосування і переваги / Тарас Володимирович Панченко // Доповіді міжнародної наукової конференції TAAPSD’2007. – Berdyansk, 2007. – С. 319–322.
8. Редько В.Н. Композиции программ и композиционное программирование / Владимир Никифорович Редько // Программирование. – 1978. – № 5. – С. 3–24.
9. Reisig W. The Expressive Power of Abstract-State Machines / W. Reisig // *Computing and Informatics*. – 2003. – Vol. 22, № 3–4. – P. 209–219.
10. Gurevich Y. Reconsidering Turing's thesis (toward more realistic semantics of programs) / Y. Gurevich. – Technical report CRL-TR-36-84. – University of Michigan, 1984. – 13 p.
1. PANCHENKO, T. (2006) *Compositional Methods for Software Systems Specification and Verification* (PhD Thesis). Kyiv. 177 p.
2. PANCHENKO, T. (2004) The Methodology for Program Properties Proof in Compositional Languages IPCL. In *Proceedings of the International Conference "Theoretical and Applied Aspects of Program Systems Development" (TAAPSD'2004)*. Kyiv. pp. 62–67.
3. NIKITCHENKO, N. (1998) *A Composition Nominative Approach to Program Semantics*. Technical Report IT-TR: 1998-020. Technical University of Denmark. 103 p.
4. OWICKI, S., GRIES, D. (1976) An Axiomatic Proof Technique for Parallel Programs. *Acta Informatica*. Vol. 6, No 4. pp. 319–340.
5. XU, Q., DE ROEVER, W.-P. and HE, J. (1997) The Rely-Guarantee Method for Verifying Shared Variable Concurrent Programs. *Formal Aspects of Computing*. Vol. 9, No 2. pp. 149–174.
6. LAMPORT, L. (1993) Verification and Specification of Concurrent Programs. deBakker, J., deRoever, W. and Rozenberg, G. (eds.) *A Decade of Concurrency*. Vol. 803. Berlin: Springer-Verlag. pp. 347–374.
7. PANCHENKO, T. (2007) The Simplified State Model for Properties Proof Method in IPCL Languages and its use and advantages. In *Proceedings of the International Conference "Theoretical and Applied Aspects of Program Systems Development" (TAAPSD'2007)*. Berdyansk. pp. 319–322.
8. REDKO, V. (1978) Compositions of programs and composition programming. *Programming*. 5. pp. 3–24.
9. REISIG, W. (2003) The Expressive Power of Abstract-State Machines. *Computing and Informatics*. Vol. 22, No. 3–4. pp. 209–219.
10. GUREVICH, Y. (1984) *Reconsidering Turing's thesis (toward more realistic semantics of programs)*. Technical report CRL-TR-36-84. University of Michigan. 13 p.

Надійшла до редколегії 05.05.2015

УДК 004.93

Скотаренко Ф.М., аспірант

F.M. Skotarenko, postgraduate student

**Алгоритм лінійної дискримінації для
матричних просторів та його основи**

**Algorithm of linear discrimination problem
for matrix spaces: base principals**

Київський національний університет імені
Тараса Шевченка, 03680, м. Київ, пр.-т
Глушкова 4д,

Taras Shevchenko National University of Kyiv,
03680, Kyiv, Glushkova av., 4d,

e-mail: fred.unicyb@gmail.com

e-mail: fred.unicyb@gmail.com

У роботі наведено алгоритм задачі лінійної дискримінації як один з варіантів розв'язання проблеми групування інформації (ПГІ) у варіанті задачі класифікації чи кластеризації. Загалом ПГІ, що є важливою складовою прикладних досліджень, виступає в одному з проявів: як задача відновлення функції, представленої своїми спостереженнями, чи – у вигляді задачі класифікації, кластеризації, розпізнавання образів. Задача лінійної дискримінації в R^n є важливим інструментом групування в прикладних дослідженнях, коли досліджувані об'єкти представлені векторами ознак. В той же час у багатьох важливих областях дослідження: розпізнавання мовних сигналів, обробка зображень, динамічні символи жестової мови, – природними представниками досліджуваних об'єктів є матриці. Робота присвячена розробці та обґрунтуванню алгоритму розпізнавання саме останньої задачі на основі лінійної дискримінації в матричних евклідових просторах, розвитку математичного апарату для оперування з матричними об'єктами. Як і в евклідових просторах числових векторів, необхідні математичні засоби будуються на основі сингулярного подання та техніки псевдо обернення за Муром - Пенроузом для слушних лінійних операторів над слушними евклідовими просторами. В роботі наведені основні результати теорії кортежних операторів, які дозволяють сформулювати та конструктивно реалізувати у вигляді алгоритму задачу лінійної дискримінації для матриць.

Ключові слова: матричні евклідові простори, кортежні оператори, сингулярне подання кортежних операторів, лінійна дискримінація в матричних евклідових просторах.

In this paper, represented algorithm of a linear discrimination problem as a variant of the problem of grouping information (GIP). In general GIP which is an important component of applied research, appears in one of the manifestations: a problem of recovery function, presented his observations or - in the form of classification, clustering, pattern recognition, or recognition dynamic symbols sign language. Linear discrimination of two classes is an important representative of the latest version of PIP. Linear discrimination inserted and considered for elements of Euclidean space matrices of a fixed dimension. The importance of clustering or classification tasks defined by the recognition that in many important applied problems of representative natural objects are grouped perform matrix. Matrix, in particular, presented image. Just - spectrogram in processing audio signals, dynamic symbols sign language, too, is the image matrix. Formulation and solution of linear discrimination based on the use of a special class of linear operators introduced by the authors in [Donchenko, Zinko, Skotarenko, 2012] - cortege operators. The work the proposed are necessary facts from the theory cortege operators.

Key Words: Matrix Euclidean spaces, cortege operators, matrix corteges operators, Single Valued Decomposition for cortege linear operators, linear discrimination of Matrix Euclidean Spaces.

Статтю представив д.т.н., професор Гаращенко Ф.Г

Вступ

Одним з надійних засобів розв'язування задачі класифікації в R^n є класифікація на основі лінійної дискримінації (ЛД) класів. В той же час у багатьох важливих прикладних областях: мовні сигнали, зображення, динамічні символи жестової мови, - природними представниками досліджуваних об'єктів є матриці. Тому важливим є розробка методів, які б дозволяли розв'язування ЛД-задач для матричних об'єктів. Нижче представлені основні складові математичного апарату, який дозволяє формувати та конструктивно розв'язувати ЛД-задачу як задачу розпізнавання для матричних представників. Як і в евклідових просторах числових векторів, необхідні математичні засоби будуються на основі сингулярного подання та техніки псевдо обернення за Муром - Пенроузом (ПдО) для слухних лінійних операторів над слухними евклідовими просторами. Такі слухні складові пропонує концепція кортежних операторів для матричних кортежів. В роботі наведені основні результати із статті [Донченко, Зінко, Скотаренко, 2012] в якій сформульована та реалізована концепція кортежних операторів які дозволяють сформулювати та конструктивно реалізувати у вигляді алгоритму розв'язання ЛД-задачі лінійної дискримінації для матриць. Основні визначення та необхідні результати наводитимуться саме за цитованою вище роботою.

1 Кортежні оператори

Під "кортежним оператором" (КорО) за матричним кортежем $\alpha = (X_1, \dots, X_K)$, $X_k \in R^{m \times n}$, $k = \overline{1, K}$ будемо розуміти лінійний оператор $\wp_\alpha : R^K \rightarrow R^{m \times n}$, який визначається співвідношенням:

$$\wp_\alpha z = \sum_{i=1}^K z_i X_i, \quad z \in R^K : z^T = (z_1, \dots, z_K).$$

Наступна теорема технічною, але є важливою для реалізації програми концепції кортежних

операторів: конструктивної побудови сингулярний подання (SVD) КорО та побудови теорії псевдо обернення для них. В ній $\wp_\alpha^*$ - позначає спряжений до $\wp_\alpha$.

Теорема 1. Добутком двох операторів $\wp_\alpha^* \wp_\alpha : R^K \rightarrow R^K$ є лінійний оператор, визначений звичайною матрицею F :

$$\wp_\alpha^* \wp_\alpha = \begin{pmatrix} \text{tr} X_1^T X_1, \dots, \text{tr} X_1^T X_K \\ \dots \\ \text{tr} X_K^T X_1, \dots, \text{tr} X_K^T X_K \end{pmatrix} = F,$$

яка є матрицею Грама для $X_1, \dots, X_K$ в евклідовому просторі $R^{m \times n}$ із слідовим скалярним добутком: $(X_i, X_j) = \text{tr} X_i^T X_j, i, j = \overline{1, K}$

Сингулярний розклад для матриці F це спектральний розклад звичайної симетричної невід'ємно визначеної матриці. Це означає, що задача побудови сингулярного розкладу для кортежного оператора зводиться до класичної задачі сингулярного розкладу для операторів між Евклідовими просторами числових векторів. Позначатимемо в подальшому (v_k, λ_k^2) , $k = \overline{1, r}$, $r = \text{rank} F$ всі ненульові сингулярності матриці F (пари власний вектор - власне число), які відповідають додатнім власним числам, $\lambda_1^2 \geq \dots \geq \lambda_r^2 > 0$.

Теорема 2. Набір пар (U_k, λ_k^2) , $k = \overline{1, r}$ в якій

$$U_k = \frac{1}{\lambda_k} \wp_\alpha v_k, \quad k = \overline{1, r} \quad \text{є набором ненульових}$$

сингулярностей оператора $\wp_\alpha \wp_\alpha^*$. Крім того

$$v_k = \frac{1}{\lambda_k} \wp_\alpha^* U_k, \quad k = \overline{1, r}.$$

Теорема 3. (SVD-подання кортежного оператора) Будь який кортежний оператор $\wp_\alpha$ та спряжений до нього $\wp_\alpha^*$ можуть бути подані у вигляді:

$$\wp_\alpha = \sum_{j=1}^r \lambda_j U_j v_j^T = \sum_{j=1}^r \lambda_j U_j (v_j, \cdot)_{R^K},$$

$$\wp_\alpha^* = \sum_{j=1}^r \lambda_j v_j (U_j, \cdot)_{R^{m \times n}}$$

Теорема 4. ПДО $\wp_\alpha^+$ для КорО $\wp_\alpha$ визначається співвідношенням

$$\wp_\alpha^+ X = \sum_{k=1}^K \frac{1}{\lambda_k} v_k (U_k, X), X \in R^{m \times n},$$

а для його спряженого -

$$\wp_\alpha^{+*} = \sum_{k=1}^K \frac{1}{\lambda_k} U_k v_k^T$$

В подальшому будуть потрібні аналоги Z-операторів в евклідових просторах числових векторів, що є ортогональними проекторами на ортогональні доповнення до множини можливих значень лінійного оператора та спряженого до нього. Цими операторами для кортежних операторів із збереженням змісту є Z - оператори, що визначатимуться співвідношеннями:

$$Z(\wp_\alpha) = E_K - \sum_{j=1}^r v_j v_j^T, \quad E_K - \text{одинична в } R^K,$$

$$Z(\wp_\alpha^*) X = \sum_{k=1}^K U_k (U_k, X)_{R^{m \times n}}, X \in R^{m \times n}$$

2 Лінійна дискримінація в матричних Евклідових просторах

ЛД-проблема полягає в розділенні двох наборів векторів в евклідових просторах слухною гіперплощиною. Для R^n в робастній постановці ЛД-задача застосуванням техніки ПДО була розв'язана в [Кириченко, Лепеха, 2002]. Концепція кортежних операторів дозволяє перенести згаданий результат на матриці.

2.1 Постановка задачі

Нехай $X_1, \dots, X_K \in R^{m \times n}$ множина матриць з навчальної вибірки представляє два класи Cl_1, Cl_2 об'єктів: для

$$J_1 \cap J_2 = \emptyset, J_1 \cup J_2 = \{1, 2, \dots, K\}$$

$$X_j \in Cl_1, j \in J_1, X_j \in Cl_2, j \in J_2 : .$$

Необхідно знайти $\Delta > 0$ і визначити лінійний функціонал $A: R^{m \times n} \rightarrow R^1$ таким чином, щоб $(A, X_j) > \Delta, j \in J_1, (A, X_j) < -\Delta, j \in J_2$ (1)

Позначимо через $\Omega(\Delta)$ множину числових векторів $y^T = (y_1, \dots, y_K)$ з R^K , які задовольняють обмеженням

$$y_j > \Delta, j \in J_1, y_j < -\Delta, j \in J_2$$

2.2 Розв'язок ЛД-задачі для матриць

Зауважимо, що набір $(A, X_1)_{tr}, \dots, (A, X_K)_{tr}$ значень лінійних форм на матрицях, що підлягають дискримінації, визначає значення спряженого $\wp_\alpha^*$ оператора до кортежного оператора $\wp_\alpha$ для аргументу A :

$$(\wp_\alpha^* A)^T = ((X_1, A), \dots, (X_K, A)), \quad \wp_\alpha^* A \in R^K$$

Має місце наступна теорема

Теорема 5. ЛД-задача для матриць в формі, сформульованій вище, є рівносильною розв'язності лінійного рівняння

$$\wp_\alpha^* X = y \quad (2)$$

в області $\Omega(\Delta)$: існуванню $y \in \Omega(\Delta)$, для якого (2) розв'язне.

Доведення. Дійсно, виконання (1) означає, що вектор

$$y \in R^K : y^T = ((A, X_1), \dots, (A, X_K))$$

належить до $\Omega(\Delta)$. Навпаки, якщо для деякого $y \in \Omega(\Delta)$ розв'язок рівняння (2) існує, то цей розв'язок і є розв'язком ЛД-задачі.

Теорема 6. ЛД-задача має розв'язок тоді і тільки тоді, коли існує $y_* \in \Omega(\Delta) \subseteq R^K$, для якого

$$y_*^T Z(\wp_\alpha) y_* = y_*^T (E_K - \sum_{k=1}^r v_k v_k^T) y_* = 0.$$

За таким y_* ЛД-розв'язок визначається рівнянням

$$A = \wp_\alpha^{*+} y_* \quad (3)$$

Наслідок. Матрична ЛД розв'язна задача якщо існує $y_* \in \Omega(\Delta) \subseteq R^K$ для якого виконуються такі умови

$$y_*^T (E_K - \sum_{k=1}^r v_k v_k^T) y_* = 0$$

і відповідні розв'язки раніше визначені (2).

Теорема 7. Матрична ЛД рівносильне задачі оптимізації в області $\Omega(\Delta) \subseteq R^K$ для квадратичної форми

$$y^T Z(\varphi_\alpha) y = y^T \left(E_K - \sum_{k=1}^r v_k v_k^T \right) y.$$

Якщо розв'язок y_* в оптимізаційній задачі дає мінімум, який дорівнює нулю, то ЛД- розв'язок визначається співвідношенням

$$A = \varphi_\alpha^{**} y_*.$$

3. Матрична ЛД - задача: алгоритм

Теорема 7 є основою в алгоритмі розв'язку матричної ЛД-задачі для двох матричних класів.

Він складається із шести наступних кроків

1-й крок: Обчислення матриці Грама F для набору матриць $X_1, \dots, X_K \in R^{m \times n}$:

$$F = \begin{pmatrix} \text{tr} X(1)^T X(1), \dots, \text{tr} X(1)^T X(K) \\ \dots \\ \text{tr} X(K)^T X(1), \dots, \text{tr} X(K)^T X(K) \end{pmatrix} = \begin{pmatrix} (X(1), X(1))_{tr}, \dots, (X(1), X(K))_{tr} \\ \dots \\ (X(K), X(1))_{tr}, \dots, (X(K), X(K))_{tr} \end{pmatrix}$$

2-й крок: Обчислення ненульових сингулярностей матриці F :

$$(v_k, \lambda_k^2), \lambda_k^2 > 0, \quad k = \overline{1, r}, \quad r = \text{rank} F - \text{набору пар}$$

власний вектор - власне число для ненульових власних чисел ;

3-й крок: Обчислення матриці $E_K - \sum_{k=1}^r v_k v_k^T$ квадратичної форми ;

4-й крок: Розв'язання задачі квадратичної оптимізації в області $\Omega(\Delta)$ для квадратичної форми з матрицею попереднього кроку:

- мінімізація квадратичної форми

$$y^T \left(E_K - \sum_{k=1}^r v_k v_k^T \right) y \text{ в області } \Omega(\Delta) \subseteq R^K$$

числовий методами . Розв'язок задачі позначається y_* ;

5-й крок: Перевірку умови рівності мінімуму в точці y_* мінімуму нулю:

$$y_*^T \left(E_K - \sum_{k=1}^r v_k v_k^T \right) y_* = 0; \quad (4)$$

6-й крок:

- якщо задовольняє умові (4) : тоді будується розв'язок матричної ЛД- задачі за формулою,

$$A = \varphi_\alpha^{**} y_*$$

за співвідношенням

$$\varphi_\alpha^{**} = \sum_{k=1}^K \frac{1}{\lambda_k} U_k v_k^T ;$$

- якщо умова (4) не задовольняється, робиться висновок, що матрична ЛД - задача не розв'язна.

Висновок

Наведений алгоритм лінійної дискримінації переносить дослідження від числових векторів евклідового простору до евклідового простору матриць фіксованої розмірності. В свою чергу це дозволяє досліджувати динамічні символи жестової мови. Як і в евклідових просторах числових векторів, необхідні математичні засоби будуються на основі сингулярного подання та техніки псевдо обернення. Цей алгоритм може бути використаний у багатьох важливих областях дослідження: розпізнавання мовних сигналів, обробка зображень, динамічні символи жестової мови, - природними представниками досліджуваних об'єктів є матриці.

Список використаних джерел

1. Кириченко Н.Ф. Применение псевдо-обратных и проекционных матриц к исследованию задач управления, наблюдения и идентификации / Н.Ф. Кириченко, Н.П. Лепеха // Кибернетика и системный анализ. – 2002. – № 4. – С. 107 - 124.
2. Кириченко Н.Ф. Нелинейные рекурсивные регрессионные преобразователи: динамические системы и оптимизация/ Н.Ф. Кириченко, В.С. Донченко, Д.П. Сербаев // Кибернетика и системный анализ. – №3, 2005. – С. 58 - 68.
3. Кириченко Н.Ф. Синтез систем нейрофункциональных преобразователей в решении задач классификации/ Н.Ф. Кириченко, Ю.Г. Кривонос, Н.П. Лепеха // Кибернетика и системный анализ. – 2007. – №3. С. 47 – 57
4. Донченко В.С. Евклидовы пространства числовых векторов и матриц: конструктивные методы описания базовых структур и их использование. //International Journal “Information technologies & Knowledge”. - 2011. - Vol. 5.- Number 3. - С. 203 - 216.
5. Донченко В. “Вектори ознак” в задачі групування інформації: вектори і матриці/ В. Донченко, Т. Зінько, Ф. Скотаренко// Problems of Computer Intellectualization.- Institute of Cybernetics NASU, ITHEA.-Kyiv, Ukraine -Sofia, Bulgaria. - 2012. - С. 111 – 124.

References

1. KIRICHENKO, M., LEPEHA, M. (2002) *The use of pseudo-and projection matrices to research problems of control, surveillance and identification*. Cybernetics and Systems Analysis. Vol.4. P. 107 - 124.
2. KIRICHENKO, M., DONCHENKO, V. & SERBAEV, D (2005). *Nonlinear regression recursive converters: dynamical systems and optimization*. Cybernetics and Systems Analysis. Vol.3. P. 58 - 68.
3. KIRICHENKO, M., KRIVONOS, Y., LEPEHA, M. (2007) *Synthesis systems neurofunctional converters in solving classification problems*. Cybernetics and Systems Analysis. Vol.3. P. 47 – 57.
4. DONCHENKO, V. (2011) *Euclidean spaces numerical vectors and matrices: design methods for describing the basic structures and their use*. International Journal “Information technologies & Knowledge”. Vol. 5. - Number 3. - P. 203 - 216.
5. DONCHENKO, V., ZINKO, T., SKOTARENKO, F. (2012) *“Feature Vectors” in Grouping Information Problem in Applied Mathematics: Vectors and Matrixes*. - Problems of Computer Intellectualization.- Institute of Cybernetics NASU, ITHEA.-Kyiv, Ukraine -Sofia, Bulgaria. - 2012. - P. 111 - 124.

Received 07.06.2015

УДК 004.925.8, 004.272.2

Терещенко В.М., д. ф. -м. н., проф.,
Осадчий Б.О., студент
Петльована М.В., студентка,
Селезньов А.І., студент,
Тара А.М., студент.

Архітектура моделі єдиного алгоритмічного середовища (МЄАС) для розв'язання задач обчислювальної геометрії

Київський національний університет імені
Тараса Шевченка, 83000, м. Київ, Глушкова 4д
e-mail: v_ter@ukr.net
e-mail: bodyaosadchiy@gmail.com
e-mail: mariapetlovana@gmail.com
e-mail: anatoliy.seleznyov@gmail.com
e-mail: derubioo@gmail.com

V.M. Tereshchenko, doctor of ph.-mat. sc., prof.,
B.O. Osadchyi, student,
M.V. Petlovana, student,
A.I. Seleznov, student,
A.M. Tara, student.

Architecture of model of the united algorithmic environment (MUAЕ) for solving computational geometry problems

Taras Shevchenko National University of Kyiv,
83000, Kyiv, Glushkova st., 4d
e-mail: v_ter@ukr.net
e-mail: bodyaosadchiy@gmail.com
e-mail: mariapetlovana@gmail.com
e-mail: anatoliy.seleznyov@gmail.com
e-mail: derubioo@gmail.com

В статті розглядається архітектура моделі єдиного алгоритмічного середовища (МЄАС) для розв'язання задач обчислювальної геометрії та запропонована її практична реалізація.

Ключові слова: МЄАС, метод “розділай та пануй”.

Architecture of model of the united algorithmic environment (MUAЕ) for solving computational geometry problems is considered in the article. Implementation of MUAЕ is proposed. Implementation description includes: problem statement, architecture requirements, description of chosen approach, which satisfies all key features of MUEA. Description of the universal data structure, which is used to maintain all final and intermediate results for all algorithms, is provided. “Divide and conquer” method is a base approach for considered algorithms. In order to algorithms used functions and primitives only from the united base of geometric functions, a separate geometric library was created. Key feature of the implementation is that every algorithm has its own merge method, which connects two already processed parts with possible usage of intermediate results of other algorithms. Algorithm dependency graph that used to verify, if the given algorithm computational sequence is valid, is described.

Key Words: MUAЕ, “Divide and Conquer” method.

Статтю представив д. ф. -м. н., проф. Анісімов А.В.

1. Вступ

На сьогодні існує безліч окремих ефективних алгоритмів для розв'язання основних класів задач обчислювальної геометрії. В той же час, при розв'язанні задачі візуального моделювання, виникає необхідність у використанні цілої сукупності алгоритмів, які, зазвичай, пов'язані між собою: результати роботи одного алгоритму є вхідними даними для іншого. Якщо кожен алгоритм з цієї сукупності реалізований окремо, то загальна ефективність розв'язання загальної задачі зменшується, а саме:

- 1) на етапах попередньої обробки даних і виконання алгоритмів дублювання одних і тих же операцій для кожного алгоритму (наприклад, сортування, побудова дерева розбиття, створення необхідних структур даних) суттєво уповільнює загальний час виконання;
- 2) використання структур даних для кожного з алгоритмів збільшує об'єм загальної пам'яті і додатковий час для доступу до елементів даних;

- 3) конвертація даних під необхідний формат, яка може призвести до втрати точності обчислень і теж уповільнює час виконання.

Таким чином, постає питання розробки єдиного алгоритмічного середовища для ефективного розв'язання сукупності пов'язаних між собою задач обчислювальної геометрії, яке позбавлено цих проблем.

В роботі [1] була запропонована базова модель єдиного алгоритмічного середовища (МСАС), яка вирішує поставлену задачу. Перевагою цієї моделі є те, що вона підтримує паралельне виконання алгоритмів, які базуються на методі «розділяй та пануй». В роботах [2,3] основна увага приділяється математичному опису моделі та реалізації паралелізму, в той час як архітектурне рішення самої моделі залишається відкритим питанням.

В цій роботі пропонується можливий варіант архітектури МСАС, оснований на ідеях моделі [4], яка вирішує такі проблеми:

- 1) Масштабованість моделі на велику кількість алгоритмів;
- 2) Автоматизована перевірка середовища на коректність;
- 3) Забезпечення гнучкості програмного коду до змін.

Постановка задачі. Нехай задано множину S з N точок у просторі R^2 . Розробити архітектуру єдиного алгоритмічного середовища для ефективної реалізації алгоритмів сукупності задач, що визначені на множині S , методом «розділяй та пануй» з нижньою оцінкою складності $\Omega(N \log N)$, яка буде розв'язувати перераховані вище проблеми.

2. Проектування архітектури

Зважаючи на ті задачі, які має вирішувати реалізація МСАС, на етапі проектування були сформульовані такі основні вимоги до архітектури середовища:

1. Підтримка сукупності окремих геометричних алгоритмів, які можна виконувати на довільних вхідних даних в довільному порядку. Кожен алгоритм в середовищі має реалізовувати метод «розділяй та пануй». Алгоритми можуть використовувати результати роботи інших алгоритмів середовища, тобто є залежними від них.

2. Масштабованість за кількістю алгоритмів. Додавання або вилучення нового алгоритму з середовища потребує мінімальних та інтуїтивних змін коду.
3. Забезпечення коректності середовища: відсутність циклічних залежностей між алгоритмами, перевірка того, що всі залежні алгоритми дійсно присутні в середовищі.
4. Архітектура має бути побудована таким чином, щоб кількість спільного коду, структур даних, функцій була максимальна. Це забезпечує легкість в підтримці коду середовища, внесенні змін, а також зменшує час виконання алгоритмів.
5. Попередня обробка даних повинна мінімізувати кількість однакових дій для різних алгоритмів (наприклад, сортування, розбиття вхідних даних тощо).
6. Кожен алгоритм має виконуватись тільки в разі необхідності та не більше одного разу на одних і тих самих вхідних даних.
7. Забезпечення візуалізації результатів роботи кожного алгоритму.

3. Реалізація МСАС

Запропонована реалізація відповідає всім сформульованим вище вимогам. Діаграми, які описують архітектурне рішення наведені на рис.1-4.

3.1. Попередня обробка та дерево розбиття

На кроці попередньої обробки відбувається зчитування вхідних даних, сортування та розбиття для методу «розділяй та пануй». Розбиття вхідних точок зберігається у вигляді дерева, де кожен вузол відповідає деякій множині точок, і будується лише один раз на початку роботи. Дерево розбиття реалізується класом DACTree (Divide and Conquer tree). У вузлах цього дерева, за які відповідає клас DACNode (рис.1), зберігаються результати роботи кожного з алгоритмів на відповідній множині точок в полі data. Клас дерева розбиття DACTree (рис.2) також підтримує список усіх алгоритмів, які можна виконати на цьому розбитті вхідних даних – AlgorithmContainer (див. пункт 3.2).

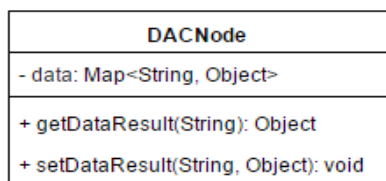


Рис.1. Вузол DACNode дерева розбиття DACTree.

Згідно з вимогами 1, 6 алгоритми можуть виконуватись в довільному порядку, але можуть залежати один від одного. Тому виникає необхідність перевірки, чи є певний алгоритм вже виконаним. Для цього в дереві для кожного алгоритму зберігаються прапорці PROCESSED, NOT_PROCESSED. Відмітимо, що алгоритм не може бути порахованим частково (це порушило б цілісність структур даних), тобто він завжди є або виконаним для всього дерева, або невиконаним для всього дерева.

Виконання алгоритмів відбувається за допомогою функції processAlgorithm, яка спочатку виконує всі алгоритми, від яких залежить поточний алгоритм (якщо вони ще не виконані), і тільки потім виконує поточний.

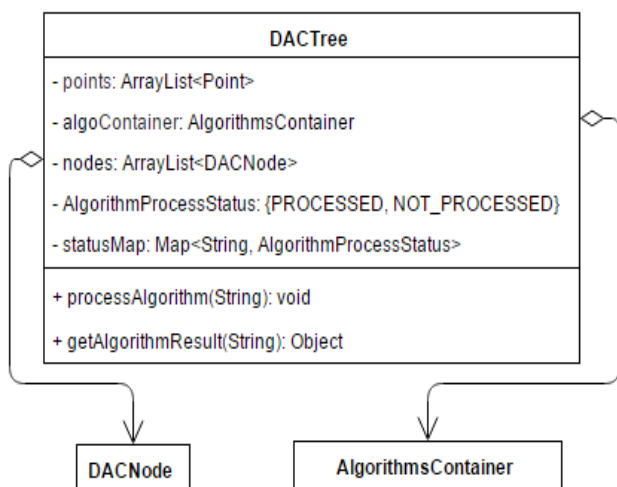


Рис.2. Дерево розбиття DACTree.

3.2. Представлення алгоритмів

Кожен алгоритм в середовищі представлений класом який реалізує сам алгоритм, та класом який представляє собою дані – результат роботи алгоритму (наприклад, діаграма Вороного представлена як реберний список з подвійними зв'язками).

Виходячи з пункту 1 пропонується мати абстрактний клас Algorithm (рис.3), в який винесені спільні для всіх алгоритмів функції.

Оскільки усі вони реалізують метод “розділяй та пануй” і використовують спільне розбиття вхідних даних, то для кожного конкретного алгоритму достатньо описати тільки свою функцію merge(), яка реалізує крок злиття результатів, та свої функції обробки тривіальних випадків (зазначимо, що для кожного алгоритму тривіальною може бути різна кількість точок, яка задається константою TrivialCount).

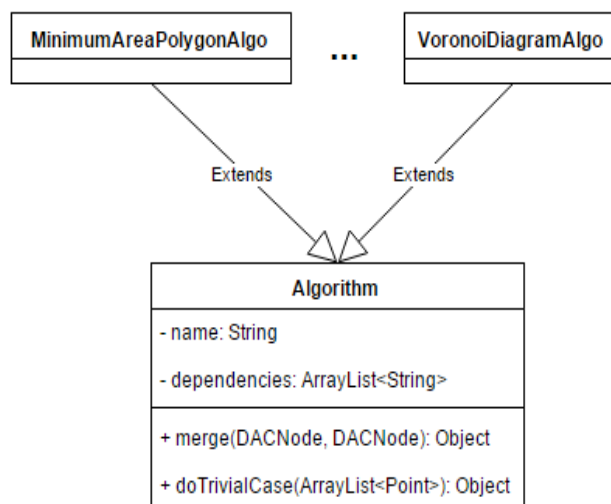


Рис.3. Абстрактний клас Algorithm.

Функція merge() приймає DACNode-и синів в дереві розбиття, що дозволяє отримувати результати роботи алгоритмів, від яких залежить алгоритм, що виконується, і повертає об'єкт даних алгоритму. Окрім цього, клас Algorithm зберігає список тих алгоритмів, від яких він залежить.

Вимога пункту 2 про масштабованість реалізується через клас AlgorithmContainer – контейнер для алгоритмів, через який відбувається додавання та видалення алгоритмів. Ці операції можуть порушити коректність середовища, тому що алгоритми є залежними один від одного. Для того, щоб автоматизувати перевірку залежностей між алгоритмами (пункт 3 вимог), використовується шаблон проектування Builder. Клас AlgorithmContainerBuilder (рис.4) дозволяє “будувати” об'єкт AlgorithmContainer (який використовується для додавання та видалення алгоритмів), а перед тим, як повернути готовий об'єкт, виконує перевірку середовища (вимога 3).

Перевірка середовища здійснюється шляхом побудови орієнтованого графу залежностей алгоритмів. За допомогою обходу графу в глибину перевіряється, чи не містить граф

невідомих для середовища алгоритмів, а також чи немає циклічних залежностей між алгоритмами.

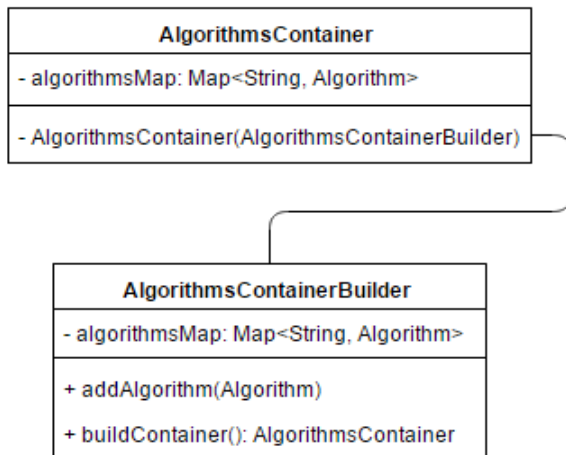


Рис.4. Класи AlgorithmContainer та AlgorithmContainerBuilder.

3.3. Спільна геометрична бібліотека

Для того, щоб алгоритми в середовищі працювали з єдиною базою геометричних примітивів та функцій, було створено окрему геометричну бібліотеку. Це вирішує питання перетворення даних від одного представлення в інше, а також спрощує визначення обчислювальної точності та похибки алгоритмів. Бібліотека побудована таким чином, що алгоритми можна писати на високому рівні абстрактності, оперуючи звичайними геометричними функціями та об'єктами – наприклад, знайти точку перетину між прямою та сегментом. Також в бібліотеці є складені об'єкти, як, наприклад, планарне розбиття у вигляді РСПЗ, яке може використовуватись алгоритмами незалежно один від одного.

4. Список реалізованих алгоритмів

Розроблене за запропонованою архітектурою середовище підтримує ряд алгоритмів, що базуються на методі «розділяй та пануй»:

- 1) Побудова опуклої оболонки [5];
- 2) Побудова діаграми Вороного [6,7];
- 3) Пошук усіх найближчих сусідів [2,3,6];
- 4) Побудова Евклідового мінімального кістякового дерева [8-10];

- 5) Пошук простого охоплюючого багатокутника мінімальної площі [11-16]. Для цього алгоритму була запропонована функція злиття, яка базується на пошуці суміжного простого багатокутника мінімальної площі.

Таким чином, граф залежностей алгоритмів, який підтримується середовищем, буде мати вигляд, що зображено на рис.5.

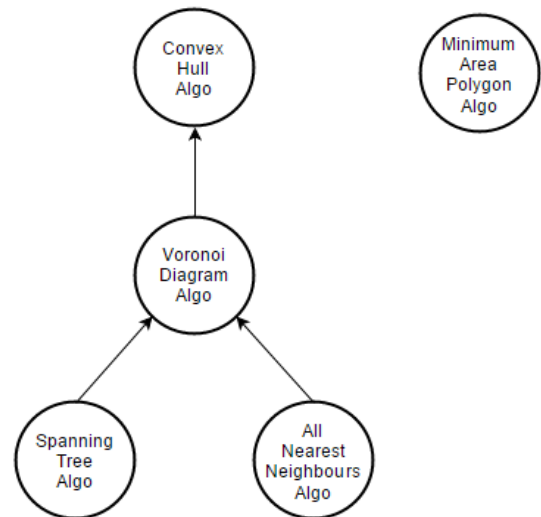


Рис.5. Граф залежностей алгоритмів.

5. Висновок

У статті запропоновано архітектуру та реалізацію моделі єдиного алгоритмічного середовища, яка має єдину універсальну структуру даних для задач обчислювальної геометрії, що розв'язуються методом «розділяй та пануй». Таким чином, можна уникнути окремих структур даних для кожного алгоритму, а отже зайвого споживання пам'яті. Необхідність у процедурах попередньої обробки та рекурсивного спуску для кожного алгоритму зникає, оскільки виконується лише одна процедура, яка готує необхідну структуру вхідних даних для алгоритмів. Отже, це зменшує час виконання алгоритмів.

Також варто зауважити, що наведена в цій роботі реалізація дуже добре масштабується на велику кількість алгоритмів, які можуть бути імплементовані в цьому середовищі. При цьому програмний код залишається гнучким до змін. Відмітимо, що інваріантом даного середовища є те, що дерево залежностей алгоритмів буде коректним після кожної операції вставки або видалення алгоритму. Перевірка на коректність відбувається автоматично.

Список використаних джерел

1. Терещенко В.М. Побудова єдиного алгоритмічного середовища для розв'язування комплексу задач обчислювальної геометрії: автореф. дис. канд. техн. наук: 01.05.01/ Терещенко В. М. // – м. Київ, 2011. – 30 с
2. Терещенко В.Н. Подход к решению взаимосвязанных задач геометрического моделирования / В.Н. Терещенко // Штучний інтелект, 4'2009. – 2009. – с. 161-167.
3. Tereshchenko V.N. Recursion and parallel algorithms in geometric modeling problems / V.N. Tereshchenko, A.V. Anisimov // Cybernetics and Systems Analysis, Springer. Vol. 46(2). – 2010. – pp. 173-184.
4. Tereshchenko V.N. The Unified Algorithmic Platform for Solving Complex Problems of Computational Geometry / V. N. Tereshchenko, I. Budjak, A. Fisunen // 12th International Conference. Lecture Notes in Computer Science, 7979. St. Petersburg, Russia, September 30 – October 4, 2013. Springer Berlin Heidelberg. – 2013. – pp. 424-428.
5. Терещенко В.Н. Построение обобщенной стратегии решения некоторых задач вычислительной геометрии / В.Н. Терещенко, В.В. Сапсай, И.С. Статкевич, А. Федоров // Параллельные вычислительные технологии. – Киев, 2009. – с. 736-737.
6. Препарата Ф. Вычислительная геометрия: введение / Ф. Препарата, М. Шеймос // . – пер. с англ., Москва: Мир. – 1989. – 478 с.
7. Dobrin A. A Review of Properties and Variations of Voronoi Diagrams / A. Dobrin // Whitman College, WA. – 2005. – pp. 2-10.
8. Sedgewick R. Minimum Spanning Tree lecture notes / R. Sedgewick, K. Wayne // Computer Science: Algorithms & Data Structures, Princeton, Princeton University Press. – 2007. – p. 226.
9. Eades P. The realization problem for Euclidean minimum spanning trees is NP-hard / P. Eades, S. Whitesides // Proceeding of the 10th ACM Symposium on Computational Geometry, New York, USA. – 1994. – pp. 49-56.

References

1. TERESCHENKO, V. (2011) Pobudova algoritmychnogo seredovysha dlya rozvyazuvannya kompleksu zadach obchislualnoi geometrii (Doctoral dissertation). Retrieved from Library of Taras Shevchenko National University of Kyiv.
2. TERESCHENKO, V.N. (2009) Podhod k resheniyu vzaimosvyazannyih zadach geometricheskogo modelirovaniya. *Shtuchniy intelekt*. 4'2009. pp. 161-167.
3. TERESCHENKO, V.N. and ANISIMOV, A.V. (2010) Recursion and parallel algorithms in geometric modeling problems. *Cybernetics and Systems Analysis*. 46(2). pp. 173-184.
4. TERESHCHENKO, V.N., BUDJAK, I. and FISUNENKO, A. (2013) The Unified Algorithmic Platform for Solving Complex Problems of Computational Geometry. In *12th International Conference. Lecture Notes in Computer Science*, 7979. St. Petersburg, Russia, September 30 – October 4, 2013. Springer Berlin Heidelberg. pp. 424-428.
5. TERESHCHENKO, V.N. et al. (2009) Postroenie obobshchennoi strategii reshenia zadach vychislitelnoi geometrii. *Parallel computational technologies*. Kyiv. pp. 736-737.
6. PREPARATA, F. and SHAMOS, M. (1989) Computational geometry: an introduction. – translation, Moscow: Mir.
7. DOBRIN, A. (2005) A Review of Properties and Variations of Voronoi Diagrams. Whitman College, WA. pp. 2-10.
8. SEDGEWICK, R. and WAYNE, K. (2007) Minimum Spanning Tree lecture notes. *Computer Science: Algorithms & Data Structures*. Princeton: Princeton University Press. p. 226.
9. EADES, P. and WHITESIDES, S. (1994) The realization problem for Euclidean minimum spanning trees is NP-hard. In *Proceeding of the 10th ACM Symposium on Computational Geometry*. New York, USA. ACM. pp. 49-56.
10. FORTUNE, S. (1992) Voronoi diagrams and Delaunay triangulations. *Computing in Euclidean geometry*. 1. pp. 193-233.

10. Fortune S. Voronoi diagrams and Delaunay triangulations / S. Fortune // *Computing in Euclidean Geometry*. – 1992. – с. 193-233.
11. Chong Z. Generating random polygons with given vertices. / Z. Chong, S. Gopalakrishnan, J. Snoeyink, J. S. B. Mitchell // *Computational Geometry: Theory and Applications*, Elsevier. Vol. 6, Issue 5. – 1996. – pp. 277-290.
12. Melkemi M. Computing the shape of a planar points set / M. Melkemi, M. Djebali // *Pattern Recognition*, Elsevier. Vol. 33(9). – 2000. – pp. 1423–1436.
13. Garai G. A split and merge procedure for polygonal border detection of dot pattern / G. Garai, B. Chaudhuri // *Image and Vision Computing*, Elsevier. Vol. 17(1). – 1999. – pp. 75–82.
14. Auer T. Heuristics for the generation of random polygons / T. Auer, M. Held // *Proceeding. of the 8th Canadian Conference on Computational Geometry*, August 12-15, 1996, Carleton University, Ottawa, Canada. – 1996. – pp. 38-43.
15. Galton A. What is the region occupied by a set of points? / A. Galton, M. Duckham // *Geographic Information Science, 4th International Conference, Proceeding, ser. Lecture Notes in Computer Science*, September 20-23, 2006, Munster, Germany. – 2006. – pp. 81-98.
16. Tereshchenko V. Constructing a Simple Polygonalizations / V. Tereshchenko, V. Muravitskiy // *World Academy of Science, Engineering and Technology, International Scholarly and Scientific Research & Innovation*, Vol. 5(5), June 24-26, 2011, Paris, France. – 2011. – pp. 563-566.
11. CHONG, Z. et al. (1996) Generating random polygons with given vertices. *Computational Geometry: Theory and Applications*. 6(5). pp. 277-290.
12. MELKEMI, M. and DJEBALI, M. (2000) Computing the shape of a planar points set. *Pattern Recognition*. 33(9). pp. 1423–1436.
13. GARAI, G. and CHAUDHURI, B. (1999) A split and merge procedure for polygonal border detection of dot pattern. *Image and Vision Computing*. 17(1). pp. 75-82.
14. AUER, T. and HELD, M. (1996) Heuristics for the generation of random polygons. In *Proceeding. of the 8th Canadian Conference on Computational Geometry*. Ottawa, August 12-15, 1996. Ottawa: Carleton University Press. pp. 38-43.
15. GALTON, A. and DUCKHAM, M. (2006) What is the region occupied by a set of points? In *Geographic Information Science, 4th International Conference, Proceeding, ser. Lecture Notes in Computer Science*. Munster, September 20-23, 2006. Springer. pp. 81-98.
16. TERESCHENKO, V. and MURAVITSKIY, V. (2011) Constructing a Simple Polygonalizations. In *World Academy of Science, Engineering and Technology, International Scholarly and Scientific Research & Innovation*. 5(5). Paris, June 24-26, 2011. pp. 563-566.

Надійшла до редколегії 30.04.2015

УДК 626.13

Хомицький В.В.¹, к.т.н., пров.н.с.
Терещенко Л.М.², к.ф.-м.н., с.н.с.
Фомін В.В.³, д.ф.-м.н., проф.
Абрамова Л.П.⁴, провідний гідротехнік

**Застосування чисельної спектральної
моделі SWAN для розрахунків
параметрів вітрового хвилювання в
районі Кривої коси**

¹ Інститут гідромеханіки НАНУ, вул.
Желябова, 8/4, Київ, 03680, Україна
e-mail: homickey@ukr.net

² Інститут гідромеханіки НАНУ, вул.
Желябова, 8/4, Київ, 03680, Україна
e-mail: lida_tereschenko@mail.ru

³ Морське відділення Українського НІГМІ, вул.
Советська, 61, м. Севастополь, 99011, Україна

⁴ Інститут гідромеханіки НАНУ, вул.
Желябова, 8/4, Київ, 03680, Україна

V. V. Khomicky¹, Senior Scientific Associate
L.N. Tereshchenko², Principle Scientific Associate
V.V. Fomin³, Doctor of Sciences
L. P. Abramova⁴, leading Hydraulic Engineering

**The use of numerical models SWAN spectral
parameters for the calculation of wind waves
in the area Curve spit**

¹ Institute of Hydromechanics, Academy of Sciences
of Ukraine, 8/4 Zheliabov str., 03680 Kiev, Ukraine,
e-mail: homickey@ukr.net

² Institute of Hydromechanics, Academy of Sciences
of Ukraine, 8/4 Zheliabov str., 03680 Kiev, Ukraine
e-mail: lida_tereschenko@mail.ru

³ Marine Branch of the Ukrainian
Hydrometeorological Institute, 61, Soviet street,
Sevastopol, 99011, Ukraine

⁴ Institute of Hydromechanics, Academy of Sciences
of Ukraine, 8/4 Zheliabov str., 03680 Kiev, Ukraine,

Для моделювання вітрового хвилювання та штормового нагону в районі коси Кривої розташованої в Таганрозькій затоці Азовського моря використовувалася модель SWAN, основана на чисельному розв'язку рівняння балансу хвильової енергії в спектральній формі. У розрахунках використовувалася SWAN версії 40.85 в стаціонарному режимі. Кутівий дозвіл моделі складав 10°. По частотній координаті, що змінюється від 0,04 до 2 Гц, використовувалася нерівномірна сітка з 71 вузлом. Джерело генерації хвиль вітром визначалося на основі квазілінійної теорії Янсена. Була виконана серія з 20-ти чисельних експериментів по розрахунку хвильових характеристик для швидкостей і напрямків вітру представлених. У кожному експерименті глибина моря в 3-й області коректувалася шляхом додавання до неї максимальних підйомів рівня моря. Для побудови рельєфу дна розрахункових областей проводилася оцифровка навігаційних карт.

Ключові слова: вітрове хвилювання, модель SWAN, нагін, берегова зона.

To simulate wind waves and storm surge in the area of spit Curve located in Taganrog Bay of the Azov Sea used model SWAN, based on the numerical solution of wave energy balance equation in spectral form. The calculations used SWAN version 40.85 in the steady state. The angular resolution of the model was 10°. For coordinate frequency varying from 0.04 to 2 Hz was used irregular grid of 71 node. Source waves wind generation was determined on the basis of quasi-linear theory Janssen. For a description of bottom friction applied the theory of the relation of the Grant – Madsen. It was made a series of 20 experiments on numerical calculation of wave characteristics of wind speed and direction. In each experiment, the depth of the sea in the 3rd region corrected by adding to it the maximum sea level rise. For the construction of the bottom topography of settlement areas digitize navigational charts. The bottom relief was obtained by digitizing the navigation map 3434 scale 1:50 000. The results of calculations of the wave fields in the area Crooked braids show that the intensity of the wind waves on the wind direction is the same as for the storm surge, dangerous wind direction in order of importance are: the south- west, south, west, south-east.

Key Words: wind waves, model SWAN, assault, coastal zone.

Статтю представив д.ф.-м.н., проф. Анісімов А.В.

Азовське море за морфологічними ознаками відноситься до плоских морів і є мілководною водоймою з невисокими береговими схилами. Площа моря без Сивашу і лиманів складає 37 802 км². Площа Сивашу - 2 540 км², а водойм лиманового типу - 4 200 км². Найбільша його довжина від коси Арабатська стрілка до гирла р.Дон дорівнює 380 км, найбільша ширина між вершинами Темрюкської і Белосарайської заток - 200 км. Об'єм води при середній глибині 8,5 м дорівнює 320 км³.

Підводний рельєф Азовського моря порівняно простий: з віддалення від берега глибини повільно і плавно нарастають, досягаючи в центральній частині моря 14,4 м. Основна площа характеризується глибиною 5 - 13 м. Область найбільших глибин знаходиться в центрі моря. Близьке до симетричного розташування ізобат, порушується витягнутістю Таганрозької затоки. Ізобата 5м розташовується приблизно в 2км від берега, віддаляючись від нього в затоках і лиманах.

У рельєфі дна Азовського моря відзначаються системи підводних височин, витягнутих уздовж східного (банка Железинська) і західного (банки Морська і Арабатська) узбереж'я моря, глибини над якими зменшуються від 8-9м до 3-5м. Для підводного берегового схилу північного узбережжя характерне широке мілководдя (20.30км) з глибинами 6.7м, для південного узбережжя - крутий підводний схил до глибин 11-12м.

Північно-східна ділянка моря врізається в сушу на 140 км, утворюючи Таганрозьку затоку. Високі і обривисті береги північної та південної ділянок затоки прорізають ярами і балками. Від обох берегів виступають коси. На північному березі затоки найбільш великі коси - Беглицька і Крива.

Характеристики штормових нагонів і вітрового хвилювання в районі Кривої коси, розташованої в Таганрозькій затоці Азовського моря, визначаються: рельєфом дна і конфігурацією берегової лінії; швидкістю і напрямом вітру; величинами ефективних розгонів.

Берегова зона с. Обрив, що примикає зі сходу до основи Кривої коси, є мілководною акваторією з глибинами 1-2м (ізобата 2м розташовується на відстані 500-750м від урізу води).

Для моделювання вітрового хвилювання використовувалася модель SWAN, основана на

чисельному розв'язку рівняння балансу хвильової енергії в спектральній формі:

$$\frac{\partial}{\partial t} N + \frac{\partial}{\partial x} c_x N + \frac{\partial}{\partial y} c_y N + \frac{\partial}{\partial \sigma} c_\sigma N + \frac{\partial}{\partial \theta} c_\theta N = S / \sigma \quad (1)$$

Тут S - функція джерела, що описує механізми формування хвильового поля; $c_x, c_y, c_\theta, c_\sigma$ - швидкості перенесення хвильової енергії, що визначаються на основі методу геометричної оптики:

$$c_x = c_g \frac{k_x}{k} + U, \quad c_y = c_g \frac{k_y}{k} + V,$$

$$c_\theta = -\frac{1}{k} \left(\frac{\partial \sigma}{\partial H} \frac{\partial H}{\partial m} + k_x \frac{\partial U}{\partial m} + k_y \frac{\partial V}{\partial m} \right), \quad (2)$$

$$c_\sigma = \frac{\partial \sigma}{\partial H} \left(\frac{\partial H}{\partial t} + U \frac{\partial H}{\partial x} + V \frac{\partial H}{\partial y} \right) - \frac{c_g}{k} \left(k_x \frac{\partial U}{\partial s} + k_y \frac{\partial V}{\partial s} \right), \quad (3)$$

де c_g - групова швидкість; k_x, k_y - компоненти хвильового вектору; $k = \sqrt{k_x^2 + k_y^2}$; $\sigma = \sqrt{gk \tanh(kH)}$ - власна частота; s - просторова координата, паралельна напрямку поширення хвиль; m - просторова координата перпендикулярна напрямку поширення хвиль. При виведенні співвідношень (2), (3) використовуються дисперсійне співвідношення і рівняння збереження хвильових гребнів.

В результаті чисельного розв'язку рівняння (1) виходить просторово-часовий розподіл частотно-кутового спектру хвиль F . Якщо ввести операцію подвійної інтеграції по частоті від $\sigma_{\min}$ до $\sigma_{\max}$ і по напрямку від 0 до 2π :

$$\langle \Phi \rangle = \int_0^{2\pi} \int_{\sigma_{\min}}^{\sigma_{\max}} \Phi(\theta, \sigma) d\sigma d\theta, \quad (4)$$

Знаючи середню висоту хвиль, можна знайти висоти хвиль різних забезпеченостей на основі інтегральної функції розподілу для моря кінцевої глибини.

У розрахунках використовувалася модель SWAN версії 40.85 в стаціонарному режимі. Кутовий дозвіл моделі складав 10°. По частотній координаті, що змінюється від 0,04 до 2 Гц, використовувалася нерівномірна сітка з 71 вузлом. Джерело генерації хвиль вітром визначалося на основі квазілінійної теорії

Янсена. Для опису донного тертя застосовувалися співвідношення теорії Гранта-Мадсена з параметром шорсткості дна $k_N = 0,01\text{ м}$. З метою підвищення точності розрахунків використовувався 3-х кроковий метод вкладених сіток. На першому кроці моделювалися поля хвилювання в усьому Чорноморському басейні на сітці $\sim 4,5 \times 4,5\text{ км}$ (246×160 вузлів). На другому кроці на сітці $\sim 1,5 \times 1,5\text{ км}$ (246×205 вузлів) хвилювання розраховувалося для акваторії, що включає Азовське море, Керченську протоку і північно-східний шельф Чорного моря.

На третьому кроці з дозволом $\sim 50 \times 50\text{ м}$ (669×518 вузлів) розраховувалися поля вітрових хвиль в районі Кривої коси. На кроках 2 і 3 параметри хвиль на рідких межах визначалися інтерполяцією модельних даних, отриманих на попередніх кроках. На третьому кроці враховувалося вплив зміни рівня моря на вітрове хвилювання. Для цього глибина моря в SWAN задавалася у виді: $H = h + \eta$, де h - рельєф дна басейну; η - рівень моря, отриманий за допомогою моделі ADCIRC.

Для побудови рельєфу дна розрахункових областей проводилося оцифрування навігаційних карт. Рельєф дна третьої області показаний на рис. 1 Він був отриманий шляхом оцифрування навігаційної карти 3434 масштаби 1:50 000.

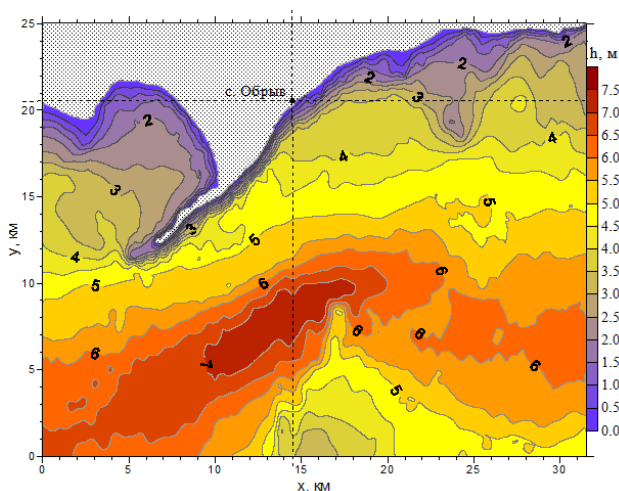


Рис.1 Модельний рельєф дна (м) в районі Кривої коси. Ізобати проведені через 0,5м.

Була виконана серія з 20-ти чисельних експериментів за розрахунком хвилевих характеристик для різних швидкостей і напрямів вітру. У кожному експерименті глибина моря в

3-ій області коригувалася шляхом додавання до неї максимальних підйомів рівня моря.

Результати розрахунків хвилевих полів в районі Кривої коси показують, що залежність інтенсивності вітрового хвилювання від напрямів вітру та ж, що і для штормових нагонів, тобто хвиленебезпечні напрями вітру в порядку їх значущості наступні: південний захід, південь, захід, південний схід. Модельні поля хвилювання для східного вітру слід сприймати з урахуванням вищевикладених вище зауважень, що стосуються ефектів обміління і осушення дна. Результати моделювання дають основу припускати, що при східних вітрах виникають згони, що перешкоджають виникненню інтенсивного хвилювання в досліджуваному районі.

Проаналізуємо основні особливості модельних полів хвилювання в районі Кривої коси. На рис. 2 приведені просторові розподіли висот хвиль 1% -ої забезпеченості ($h_{1\%}$) і середнього напрям хвилювання ($\bar{\theta}$) для усіх 20-ти варіантів розрахунків. А на рис.3 амплітуда швидкості придонних орбітальних рухів.

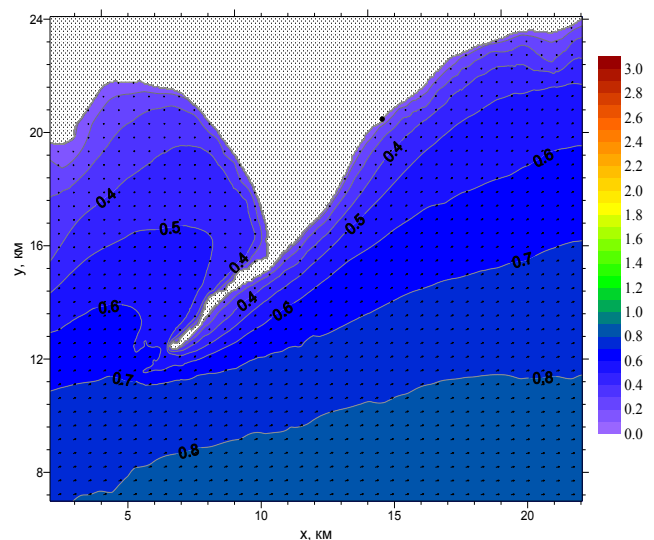


Рис.2 Висота хвиль 1% -ої забезпеченості (м) і середній напрям хвилювання в районі Кривої коси при західному вітрі 8 м/с.

Як видно, розрахунковий нагін призводить до збільшення висот хвиль та із зменшенням глибин стає все більш значимим. Так якщо в районі улоговини висоти хвиль збільшуються на 0,25-0,3 м ($\sim 15\%$), то у береговій зоні, що примикає до Кривої коси, збільшення висот хвиль досягає 0,7-0,9 м ($\sim 75-90\%$).

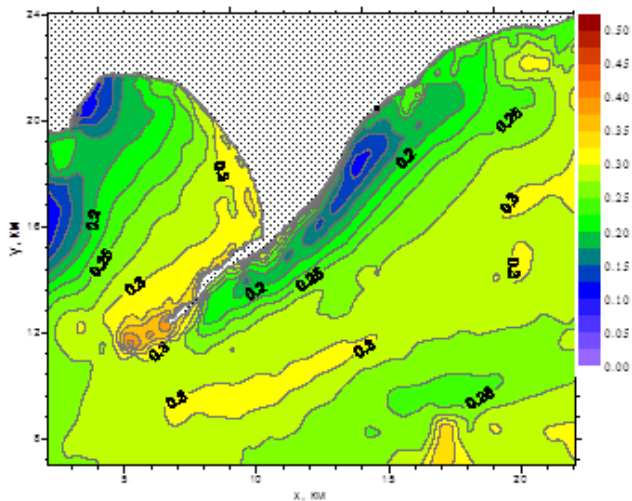


Рис.3 Амплітуда швидкості придонних орбітальних рухів (м/с) в районі Кривої коси при західному вітрі 20,3 м/с.

Висновки

Отриманий результат демонструє важливість розрахунку штормових нагонів при моделюванні вітрового хвилювання в мілководних районах. Проведені чисельні експерименти показують поєднання сильних штормових нагонів інтенсивних вітрових хвиль, що виникають на їх фоні, може призводити до виникнення особливо небезпечних явищ у береговій зоні с. Обрив.

Список використаних джерел

1. Luettich, R.A. ADCIRC: an advanced three-dimensional circulation model for shelves coasts and estuaries, report 1: theory and methodology of ADCIRC-2DDI and ADCIRC-3DL / R.A. Luettich, J.J.Westerink, N.W Scheffner.- Vicksburg, Dredging Research Program Technical Report DRP-92-6, U.S. Army Engineers Waterways Experiment Station, MS.- 1992.- 137p.
2. Luettich R.A. Formulation and Numerical Implementation of the 2D/3D ADCIRC / R.A. Luettich, J.J.Westerink - University of Notre Dame.-2004.-74 p
3. Kerr P.C. Surge Generation Mechanisms in the Lower Mississippi River and Discharge Dependency./ P.C. Kerr, J.J.Westerink, J.C. Dietrich, R.C.Martyr // Journal of Waterway, Port, Coastal, and Ocean Engineering.- 2013, 139(4). - P. 326-335.
4. Booij. N. A third-generation wave model for coastal regions. Model description and

validation. / N. Booij., R.C. Ris, L.H.Holthuijsen // J. Geophys. Res.- 1999, 104(C4).- P. 7649-7666.

5. Holthuijsen L.H. Waves in oceanic and coastal waters. / L.H Holthuijsen. - Cambridge University Press: Cambridge.-2007. - 387p.

References

1. LUETTICH R.A., WESTERINK J.J., SCHEFFNER N.W. (1992) ADCIRC: an advanced three-dimensional circulation model for shelves coasts and estuaries, report 1: theory and methodology of ADCIRC-2DDI and ADCIRC-3DL. Vicksburg: Dredging Research Program Technical Report DRP-92-6, U.S. Army Engineers Waterways Experiment Station, MS. 137p.
2. LUETTICH R.A., WESTERINK J.J. (2004) Formulation and Numerical Implementation of the 2D/3D ADCIRC. Notre Dame: University of Notre Dame. 74 p.
3. KERR P.C., WESTERINK J.J., DIETRICH J.C., MARTYR R.C. (2013) Surge Generation Mechanisms in the Lower Mississippi River and Discharge Dependency. // *Journal of Waterway, Port, Coastal, and Ocean Engineering*, - 139(4).- P. 326-335.
4. BOOIJ. N., RIS. R.C., HOLTHUIJSEN L.H (1999). A third-generation wave model for coastal regions. Model description and validation. // *J. Geophys. Res.*- 104(C4).- P. 7649-7666.
5. HOLTHUIJSEN L.H. (2007) Waves in oceanic and coastal waters. Cambridge: University Press Cambridge XVI. 387p.

Надійшла до редколегії 24.03.15

УДК 519.9

Царук В. І., аспірант

Про оптимізацію транспортних потоків в моделі Бекмана

Київський національний університет імені
Тараса Шевченка, факультет кібернетики,
03680, Україна, м. Київ, пр. Академіка
Глушкова, 4д,
e-mail: volodymyr.tsaruk@gmail.com

V. I. Tsaruk, post-graduate student

On transport flow optimization in Backman model

Taras Shevchenko National University of Kyiv,
cybernetics faculty, 03680, Kyiv, Glushkova st.,
4d,
e-mail: volodymyr.tsaruk@gmail.com

В роботі досліджується задача оптимізації транспортних потоків з метою мінімізації загальних затрат системи за допомогою зовнішніх обмежень або, навпаки, заохочень. Розглянуто зведення моделі Бекмана до задачі дворівневої оптимізації і задачі з обмеженнями вигляду варіаційної нерівності. Запропоновано підхід до її розв'язання методом Каруша-Куна-Такера. Детально розглянуто конкретну реалізацію даної задачі, а саме задачу керування громадським транспортом. Продемонстровано зведення її до набору задач опуклої оптимізації на опуклій множині.

Ключові слова: транспортні потоки, принцип Вордроп, рівновага за Нешем, дворівнева оптимізація

Transport flow congestion problem is relevant for big cities. In case of uncontrolled development of transport flow the equilibrium of traffic, which is far from effective, is quickly set in the city. Different methods of road network management are used to overcome this problem: limitation or extension of the road traffic, regulation by means of traffic lights, or development of public transport. Generally, mathematical background of this management leads us to a highly complicated problem. This research studies the problem of optimizing transport flows to minimize the total cost of the system by means of external constraints or, on the contrary, incentives. Its implementation for the Beckman model is considered, and reduction to a bilevel optimization problem, or mathematical problem with variational inequality constraints. Its solution by Karush-Kuhn-Tucker method is suggested. Specific implementation of this problem is considered in detail, namely the problem of public transport control. Its reduction to a set of convex optimization problems for convex sets is demonstrated.

Key Words: transport flows, Wardrop principle, Nash equilibrium, bilevel optimization

Статтю представив д.т.н., п.н.с. Кудін В.І.

Вступ

В великих містах особливо актуальною є проблема перевантаженості транспортної мережі. Для боротьби з нею при плануванні міста намагаються проектувати дороги на кілька кроків наперед, а з існуючими проблемами розбираються, використовуючи різні методи: розвиток громадського транспорту, обмеження руху, регулювання частот світлофорів, тощо.

Задача моделювання транспортних потоків складається з двох підзадач: розрахунок матриці кореспонденції [1], що визначає попит на перевезення, і пошук точки рівноваги, що відповідає даній матриці. В даній роботі матриця кореспонденцій вважається відомою і

фіксованою, і досліджується керування точкою рівноваги.

При пошуку точки рівноваги потоків індивідуального транспорту використовують перший принцип Вордроп [2]: кожен користувач транспортної мережі незалежно від інших вибирає маршрут слідування в точку призначення, що відповідає найменшим затратам для нього. В результаті цього вибору користувачі утворюють потоки на відповідних маршрутах, пошук значень яких і є ціллю їх моделювання. Найбільш вивченою моделлю є модель Бекмана [3]. Вона припускає, що витрати на користування дорогами є неперервними і монотонними функціями від їх завантаженості, і зводить задачу пошуку точки рівноваги до розв'язання варіаційної нерівності. В [4] доводиться, що до розв'язку цієї ж варіаційної нерівності збігається

еволюційний алгоритм за тих же функцій витрат, а отже задача може моделювати реальні процеси.

Очевидно, що отримана рівновага потоків не буде точкою, в якій сумарні затрати є найменшими. Більше того, при неграмотному плануванні доріг різниця між оптимальними і наявними затратами тільки зростатиме. В [5] наводиться приклад мережі, де один з сегментів доріг тільки погіршує загальний транспортний стан, причому такі сегменти зустрічались і в реальних системах доріг різних міст.

В роботі розглянуто метод оптимізації роботи транспортної мережі, впливаючи на затрати користування нею. Цей вплив може бути викликаний звуженням або розширенням проїжджої частини, введенням плати за користування, чи просто збільшенням кількості автобусів на маршруті. Задача зводиться до задачі дворівневої оптимізації [6].

Модель Бекмана

Використовуватимемо опис моделі з [1]. Нехай $\Gamma(V, E)$ – орієнтований граф, де V – множина вершин, а E – множина дуг. Позначимо $S \subset E$ – множина джерел і $D \subset E$ – множина стоків. Між елементами множин S і D існує попит на перевезення. Позначимо множину потокотворчих пар:

$$W = \{w = (i, j) : i \in S, j \in D\}.$$

Кожній парі $w \in W$ відповідає число $\rho_w \geq 0$, яке називається попитом на перевезення.

Позначимо: P_w – множина маршрутів, що ведуть з i в j ; $P = \bigcup_{w \in W} P_w$ – множина всіх можливих маршрутів; $x_p, p \in P$ – потік через відповідний маршрут. Для кожної пари $w \in W$ виконується наступна умова:

$$X_w = \left\{ x_p \geq 0 : p \in P_w, \sum_{p \in P_w} x_p = \rho_w \right\}.$$

Позначимо також $x = (x_p, p \in P)$ – вектор потоків з допустимою множиною:

$$x \in X = \prod_{w \in W} X_w.$$

Нехай $\delta_{p,e} \in \{0,1\}$, де $p \in P, e \in E$ – індекс належності ребра e маршруту p . Тоді кожній дузі $e \in E$ відповідає число – вартість подорожі через неї $t_e = g_e(x_e)$, де $x_e = \sum_{p \in P} \delta_{p,e} x_p \in \mathbb{R}$ – потік через відповідну дугу. Тоді вартість подорожі маршрутом $p \in P$ визначається

функцією $G_p = G_p(x) = \sum_{e \in E} \delta_{p,e} g_e(x_e)$, де $x \in X$, а вектор вартостей маршрутів – вектор-функцією $G(x)$.

Оскільки кожен водій мінімізує власні затрати, то ненульовими будуть тільки ті потоки, які є мінімальними для відповідної пари «джерело-витік»:

$$x_p > 0 \Leftrightarrow G_p(x) = \min_{k \in P_w} G_k(x), p \in P_w. \quad (1)$$

В такому випадку транспортна рівновага, що встановлюється в місті, визначається вектором $x^\dagger$, що є розв'язком варіаційної нерівності:

$$G(x^\dagger)(x - x^\dagger) \geq 0 \quad \forall x \in X. \quad (2)$$

Керування потоками

Міське керівництво зацікавлене в мінімізації загальних транспортних витрат. Впливати на ці витрати воно може, змінюючи пропускну спроможність доріг, тобто впливаючи на функції витрат.

Нехай $\{R_i \subset E, i \in I \subset \mathbb{N}\}$ – сімейство множин ребер, керованих за одним принципом, таких що $\forall i, j \in I, R_i \cap R_j = \emptyset$. Позначимо тоді $r = (r_i, i \in I)$ – вектор керування ребрами. Вартість проїзду заданим ребром буде подано в вигляді функції з параметром:

$$t_p = g_e(x_e, r) = g_e(x_e, r_i). \quad (3)$$

Тоді нерівність, що задає рівновагу, набуває вигляду:

$$G(x^\dagger, r)(x' - x^\dagger) \geq 0 \quad \forall x' \in X. \quad (4)$$

Нехай також вартість застосування керування задається строго опуклою функцією $K(r) = \sum_{i \in I} K_i(r_i)$. За таких умов отримуємо математичну задачу з обмеженням вигляду варіаційної нерівності:

$$\begin{aligned} G(x^\dagger, r)x^\dagger + K(r) &\xrightarrow{x^\dagger, r} \min, \\ G(x^\dagger, r)(z - x^\dagger) &\geq 0 \quad \forall z \in X. \end{aligned} \quad (5)$$

Підходи до розв'язання

Дворівнева задача оптимізації. Нехай функції вартості проїзду по ребру $G_e(x, r) = g_e(x_e, r) = g_e(\sum_{p \in P} \delta_{p,e} x_p, r)$ неперервно диференційовні, строго монотонно зростаючі по x_e . Тоді для довільних $p, k \in P'$ маємо:

$$\begin{aligned}\frac{\partial G_p(x, r)}{\partial x_k} &= \sum_{e \in E} \delta_{p,e} \frac{\partial g_e(\sum_{q \in P} \delta_{q,e} x_q, r)}{\partial x_k} = \\ &= \sum_{e \in E} \delta_{p,e} \delta_{k,e} \frac{dg_e(\sum_{q \in P} \delta_{q,e} x_q, r)}{dx_e} = \\ &= \frac{\partial G_k(x, r)}{\partial x_p}.\end{aligned}$$

Звідси, відповідно до теореми Шварца, впливає існування такої функції $F(x, r)$, що $G_p(x, r) = \frac{\partial F(x, r)}{\partial x_p}$. Вона може бути представлена у вигляді:

$$F(x, r) = \sum_{e \in E} \int_0^{\sum_{p \in P} \delta_{p,e} x_p} g_e(z, r) dz. \quad (6)$$

В такому випадку задача (5) зводиться до дворівневої задачі мінімізації:

$$\begin{aligned}G(x^\dagger, r)x + K(r) &\xrightarrow{x^\dagger, r} \min, \\ x^\dagger &\in \arg \min_{x^\dagger \in X} F(x, r).\end{aligned} \quad (7)$$

Оскільки функція $G(x, r)$ – строго монотонна, то, відповідно, функція $F(x, r)$ буде строго опуклою.

Метод Каруша-Куна-Такера. Для розв'язання цієї задачі застосуємо метод Каруша-Куна-Такера [6]. Розглянемо структуру множини X . Кожній парі «джерело-стік» $w \in W$ відповідає обмеження вигляду

$$v_w(x) = \sum_{p \in P_w} x_p - \rho_w = 0.$$

Позначимо $N_w = |P_w|$ – кількість маршрутів на даній парі.

На кожен потік накладається обмеження невід'ємності:

$$h_p(x) = x_p \geq 0, p \in P.$$

Тоді функція Лагранжа для задачі нижнього рівня матиме вигляд:

$$L(x, \lambda, \mu) = F(x, r) - \lambda v(x) - \mu h(x).$$

Тут $\lambda \in \mathbb{R}^m, m = |W|$, а $\mu \in \mathbb{R}^n, n = |P|$. У такому випадку отримаємо необхідну умову мінімуму:

$$\nabla_x F(\bar{x}, r) = \nabla_x F(x, r) - \lambda \nabla_x v(x) - \mu \nabla_x h(x).$$

Розпишемо її:

$$\left\{ \begin{aligned} G_p(\bar{x}, r) - \lambda_w N_w - \mu_p &= 0, p \in P_w, \\ \sum_{p \in P_w} x_p - \rho_w &= 0, \\ \lambda_w &\geq 0, \\ x_p &\geq 0, \\ \mu_p &\geq 0, \\ \mu_p x_p &= 0. \end{aligned} \right. \quad (8)$$

Тоді задача (7) набуває вигляду оптимізаційної задачі з обмеженнями взаємодоповнення:

$$\left\{ \begin{aligned} G(x^\dagger, r)x^\dagger + K(r) &\xrightarrow{x^\dagger, r, \mu, \lambda} \min \\ G_p(x^\dagger, r) - \lambda_w N_w - \mu_p &= 0, p \in P_w, \\ \sum_{p \in P_w} x_p - \rho_w &= 0, w \in W, \\ \lambda_w &\geq 0, w \in W, \\ x_p^\dagger &\geq 0, p \in P, \\ \mu_p &\geq 0, p \in P, \\ \mu_p x_p^\dagger &= 0, p \in P, \\ r_i &\geq 0, i \in I. \end{aligned} \right. \quad (9)$$

Розглянемо величину $\lambda_w N_w = \lambda_w^0$. Для кожного маршруту $p \in P_w$, якщо потік через нього ненульовий, виконується рівність:

$$G_p(x, r) = \lambda_w^0.$$

Тому λ_w^0 – найменша вартість подорожі за парою $w \in W$. Тоді задача (9) може бути представлена у вигляді:

$$\left\{ \begin{aligned} \sum_{w \in W} \sum_{p \in P_w} \lambda_w^0 x_p + K(r) &\xrightarrow{x^\dagger, r, \mu, \lambda} \min \\ G_p(x^\dagger, r) - \lambda_w N_w - \mu_p &= 0, p \in P_w, \\ \sum_{p \in P_w} x_p - \rho_w &= 0, w \in W, \\ \lambda_w &\geq 0, w \in W, \\ x_p^\dagger &\geq 0, p \in P, \\ \mu_p &\geq 0, p \in P, \\ \mu_p x_p^\dagger &= 0, p \in P, \\ r_i &\geq 0, i \in I. \end{aligned} \right. \quad (9')$$

Спрощена гібридна транспортна модель

Дана модель об'єднує одночасно громадський та індивідуальний транспорт. Вона використовує наступні припущення:

- Маршрути громадського транспорту прокладені заздалегідь;

- Кожен маршрут громадського транспорту має лише початкову і кінцеву зупинку;
- Громадський транспорт їздить лише своїми виділеними смугами і ніяк не перетинається з індивідуальним;
- Кожен пасажир приймає рішення: щодо використання типу транспорту. При цьому має повну інформацію щодо всіх можливих маршрутів.

Доповнимо ребра графа E ребрами $e_w^* \in E^*$, $w \in W$, що відповідають маршрутам громадського транспорту. Відповідно, кожне з цих ребер породжує новий маршрут $p^* \in P^*$ на $w \in W$. Кожному маршруту $p \in P^*$ відповідає його інтенсивність $r_p \geq 0$, а також $K_p > 0$ – вартість обслуговування одиничної інтенсивності. Тоді ці величини утворюють вектори $r \in K$. Загальна множина всіх маршрутів складає $P' = P \cup P^*$.

Позначимо x' – вектор потоків на кожному з маршрутів $p \in P'$, і, відповідно, $G'(x', r)$ – вектор вартостей проїзду, причому

$$G'_p(x', r) = \begin{cases} G_p(x'), p \in P, \\ g_p(x_p, r_p), p \in P^*, \end{cases}$$

тобто вартість поїздки маршрутом громадського транспорту залежить тільки від завантаженості цього маршруту і його інтенсивності.

Тоді транспортна рівновага визначається розв'язком $x^\dagger \in X'$ варіаційної нерівності:

$$G'(x^\dagger, r)(z - x^\dagger) \geq 0 \quad \forall z \in X'. \quad (10)$$

В цьому випадку загальна задача набуває відповідно оптимізаційної задачі з обмеженнями вигляду варіаційних нерівностей.

$$\begin{cases} \sum_{p \in P} (C_p x + b_p) x_p + \sum_{p \in P^*} (c_p \hat{x}_p + b_p) + \sum_{p \in P^*} K_p \frac{\hat{x}_p}{x_p} \xrightarrow{x, \hat{x}, \mu, \lambda} \min \\ \begin{cases} C_p x + b_p - \lambda_w N_w - \mu_p = 0, p \in P_w, \\ c_p \hat{x}_p + b_p - \lambda_w N_w - \mu_p = 0, p \in P_w^*, \\ \sum_{p \in P'_w} x_p - \rho_w = 0, w \in W, \\ \lambda_w \geq 0, w \in W, \\ x_p \geq 0, p \in P, \\ \mu_p \geq 0, p \in P, \\ \mu_p x_p = 0, p \in P, \\ \hat{x}_p \geq 0, p \in P'. \end{cases} \end{cases} \quad (13)$$

$$\begin{aligned} G'(x^\dagger, r) x + K r \xrightarrow{x^\dagger, r} \min, \\ G'(x^\dagger, r)(z - x^\dagger) \geq 0 \quad \forall z \in X'. \end{aligned} \quad (11)$$

Лінійний випадок. Нехай функції вартості проїзду маршрутами мають вигляд:

$$G'_p(x', r) = \begin{cases} C_p x + b_p, p \in P, \\ \frac{c_p x_p}{r_p} + b_p, p \in P^*, \end{cases} \quad (12)$$

$$C_p \in \mathbb{R}^n, n = |P'|.$$

Тоді, підставивши (11) і (12) в (9), отримаємо:

$$\begin{cases} G'(x^\dagger, r) x^\dagger + K r \xrightarrow{x^\dagger, r, \mu, \lambda} \min \\ \begin{cases} C_p x + b_p - \lambda_w N_w - \mu_p = 0, p \in P_w, \\ \frac{c_p x_p}{r_p} + b_p - \lambda_w N_w - \mu_p = 0, p \in P_w^*, \\ \sum_{p \in P'_w} x_p - \rho_w = 0, w \in W, \\ \lambda_w \geq 0, w \in W, \\ x_p \geq 0, p \in P, \\ \mu_p \geq 0, p \in P, \\ \mu_p x_p = 0, p \in P, \\ r_p > 0, p \in P'. \end{cases} \end{cases}$$

Довизначимо дану систему при $r_p = 0$. Зробимо заміну:

$$\hat{x}_p = \begin{cases} \frac{x_p}{r_p}, p \in P^*, r_p > 0, \\ 0, p \in P^*, r_p = 0. \end{cases}$$

Тоді $r_p = \frac{\hat{x}_p}{x_p}$ при $x_p > 0$. Також покладемо вектор $\hat{x} = (\hat{x}_p, p \in P^*)$. В цьому випадку роль керування виконуватимуть компоненти x_p вектора x при $p \in P^*$.

На ній задача (13) набуває вигляду:

Зафіксуємо деяку множину індексів

$$\Lambda = \{p \in P' | \mu_p = 0\}.$$

$$\sum_{p \in P} (C_p x + b_p) x_p + \sum_{p \in P^*} (c_p \hat{x}_p + b_p) + \sum_{p \in P^* \cap \Lambda} K_p \frac{\hat{x}_p}{x_p} \xrightarrow{x, \hat{x}, \mu, \lambda} \min$$

$$\left\{ \begin{array}{l} C_p x + b_p - \lambda_w N_w = 0, p \in P_w \cap \Lambda, \\ C_p x + b_p - \lambda_w N_w - \mu_p = 0, x_p = 0, p \in P_w \cap (P \setminus \Lambda), \\ c_p \hat{x}_p + b_p - \lambda_w N_w = 0, p \in P_w^* \cap \Lambda, \\ b_p - \lambda_w N_w - \mu_p = 0, p \in P_w^* \cap (P \setminus \Lambda), \\ \sum_{p \in P'_w} x_p - \rho_w = 0, w \in W, \\ \lambda_w \geq 0, w \in W, \\ x_p \geq 0, p \in P \cap (P \setminus \Lambda), \\ \mu_p \geq 0, p \in P \cap \Lambda, \\ \hat{x}_p \geq 0, p \in P'. \end{array} \right. \quad (14)$$

Зафіксуємо $w \in W$. Тоді для довільного $\lambda_w \geq 0$ система рівнянь

$$C_p x + b_p - \lambda_w N_w = 0, p \in P_w \cap \Lambda$$

має єдиний розв'язок $x^{(w)}(\lambda_w) = \lambda_w a^{(w)} + d^{(w)}$. Підставивши його в рівняння $\sum_{p \in P'_w} x_p - \rho_w = 0$, отримаємо:

$$\sum_{P_w^* \cap \Lambda} x_p = \rho_w - \sum_{P_w \cap \Lambda} \lambda_w a^{(w)} + d^{(w)}$$

Позначивши вектор $\tilde{x} = (x_p, p \in P^*)$ вектором керування, вказавши, що вектор μ , не впливає на розв'язок і виконавши перетворення, аналогічні (9'), ми отримаємо задачу оптимізації опуклої функції на опуклій множині:

$$\lambda \rho + \sum_{w \in W} \sum_{p \in P_w^* \cap \Lambda} K_p \frac{b_p - \lambda_w N_w}{c_p x_p} \xrightarrow{\tilde{x}, \lambda} \min$$

$$\left\{ \begin{array}{l} x^{(w)}(\lambda_w) = \lambda_w a^{(w)} + d^{(w)}, P_w \cap \Lambda, \\ \hat{x}_p = \frac{b_p - \lambda_w N_w}{c_p}, 0, p \in P_w^* \cap \Lambda, \\ \sum_{P_w^* \cap \Lambda} x_p = \rho_w - \sum_{P_w \cap \Lambda} (\lambda_w a_p^{(w)} + d_p^{(w)}), w \in W, \\ \lambda_w \geq 0, w \in W, \\ x_p \geq 0, p \in P \cap (P \setminus \Lambda), \\ \lambda_w \leq \min\{b_p | p \in P_w \cap \Lambda\}, \\ \rho_w - \sum_{P_w \cap \Lambda} (\lambda_w a_p^{(w)} + d_p^{(w)}) \geq 0, \\ \hat{x}_p \geq 0, p \in P'. \end{array} \right. \quad (15)$$

Очевидно, що вона має множину розв'язків, яким відповідає множина векторів $\{(x, r)_\Lambda\}$ і число $G_\Lambda((x, r)_\Lambda)$ – значення цільової функції. Відповідно, розв'язком задачі (13) буде $\arg \min_{\Lambda} G_\Lambda((x, r)_\Lambda)$.

Висновки

В статті розглянуто задачу оптимізації транспортних потоків на прикладі моделі Бекмана. Її зведено до задачі дворівневої оптимізації, а також досліджено її для окремого

випадку, коли функції виплат лінійні за значеннями потоків.

Задача оптимізації транспортних потоків має високу обчислювальну складність. Навіть в найпростішому випадку, де функції витрат є лінійними, вона розбивається на 2^N задач опуклої оптимізації, де N – кількість маршрутів, а враховуючи, що кількість

маршрутів теж зростає експоненційно з ростом міста, проблеми з обчисленням виникають досить швидко.

Для подолання цієї проблеми можна накласти обмеження на кількість маршрутів для кожної пари стік-витік або використати модель, відмінну від моделі Бекмана.

Список використаних джерел

1. Гасников А. В. Введение в математическое моделирование транспортных потоков / Гасников А.В., Кленов С.Л., др. — Москва, МФТИ, 2010. — 362 с.
2. Wardrop J. G. Some theoretical aspects of road traffic research. / J. G. Wardrop, J. I. Whitehead. — ICE Proceedings: Engineering Divisions, Volume 1, Issue 5, 1952. — pp. 72-73.
3. M. Beckman Studies in Economics of Transportation. / M. Beckman, C. McGuire, and C. Winsten. — Yale University Press, New Haven, 1956. — pp. 767–768.
4. Sandholm W H. Evolutionary Implementation and Congestion / William H. Sandholm. — Oxford, Review of Economic Studies, 69, 2002. — pp. 667-689.
5. Braess D. Über ein Paradoxon aus der Verkehrsplanung. / D. Braess. — Unternehmensforschung 12, 1969. — pp. 258–268.
6. Allende G. B. Mathematical Programs with Equilibrium Constraints: Solution Techniques from Parametric Optimization / Gemayqzel Bouza Allende. — University of Twente, 2006. — 161 p.

References

1. GASNIKOV A. V., KLENOV S. L., etc. (2010). *Introduction into transport flows mathematical modelling*. Moscow, MPhTI.
2. J. G. WARDROP, J. I. WHITEHEAD (1952). *Some theoretical aspects of road traffic research*. ICE Proceedings: Engineering Divisions, Volume 1, Issue 5. pp. 72-73.
3. M. BECKMAN, C. MCGUIRE, and C. WINSTEN (1956). *Studies in Economics of Transportation*. New Haven: Yale University Press. pp. 767–768.
4. W. H. SANDHOLM (2002). *Evolutionary implementation and congestion pricing*. Oxford: Review of Economic Studies, 69. pp. 667-689.
5. D. BRAESS (1969). *Über ein Paradoxon aus der Verkehrsplanung*. Unternehmensforschung 12. pp. 258–268.
6. GEMAYZEL BUOZA ALLENDE (2006). *Mathematical Programs with Equilibrium Constraints: Solution Techniques from Parametric Optimization*. University of Twente.

Надійшла до редколегії 22.05.2015

УДК 519.25, 378.146

Шарапов М.М., к.ф.-м. н., доцент

Статистичне корегування результатів тестів

Київський національний університет імені
Тараса Шевченка, 83000, м. Київ, пр-т.
Глушкова 4д,
e-mail: sharapov@unicyb.kiev.ua.

M.M. Sharapov, Cand.Sci.(Phys.-Math.)

Statistical correction of test results

Taras Shevchenko National University of Kyiv,
83000, Kyiv, Glushkova st., 4d,
e-mail: sharapov@unicyb.kiev.ua.

В статті розглядається статистичне оцінювання кількості тих питань тестів, на які було свідомо дано правильні відповіді. Також пропонується новий підхід до тестування, за якого опитуваний може явно вказувати на своє незнання відповіді. Показано, що запропонований підхід не змінює кількісне значення статистичної оцінки, але дозволяє ввести до розгляду якісно нові показники оцінювання знань.

Ключові слова: оцінювання тестів, свідомі відповіді, статистичне корегування, нові показники якості знань.

Here we consider a statistical estimation of a number of consciously given correct answers during testing. If a testee doesn't know a correct answer in the test then he can try to guess the correct variant. The intuitive approach can be used to estimate the number of non-guessed correct answers while statistical approach can be used too. Furthermore, the statistical approach gives a higher estimated grade. A new approach to testing is considered as well. Under new conditions a student has opportunity to mark the question as unbeknown without trying to guess the correct answer. It is shown that new method doesn't change the estimation of consciously given correct answers under statistical approach but makes it possible to introduce new ratings for knowledge estimation. Moreover, the introduced sub-indicator of knowledge strength makes it possible to build new grades, estimation and indicators using the described technique.

Key Words: tests assessment, consciously given correct answers, statistical correction, new ratings for knowledge estimation.

Статтю представив д. т. н., професор Гаращенко Ф.Г.

The main questions we are going to deal with are "What is statistical correction of test results (SCTR)?" and "Why/when do we need it?". Well, the SCTR is a method (algorithm) to make test results (TR) closer to a real skill level (SL) of a student (testee) because sometimes TR and SL can differ seriously. So we do need SCTR to decrease this difference. Suppose the test consists of 40 questions with 4 variants of answer each while student gave 28 correct answers and 12 incorrect ones. It is logically to suppose that he guessed some answers when he didn't know the right answer. So, the real result should be 24 instead of 28.

The aspects that should be taken onto account:

- (A) The purpose of test.
- (B) The kind of the test.

(A) If the purpose is just a ranking then we do not need a SCTR because it will not change the result (ranks of students). But if the purpose is

to obtain the SL over some standard scale then the SCTR is needed.

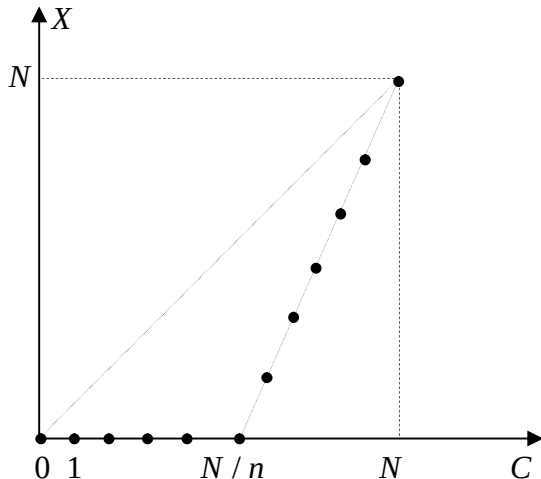
(B) There are many different kinds of tests and we will start with the simplest ones when test consists questions under three conditions:

- (i) All questions have the same difficulty level.
- (ii) All questions have the same number of variants of answer.
- (iii) All variants of answers have the same (equal) probability to be guessed.

Under conditions (i)-(iii) it is easy to generalize the result of mentioned example: if test consists of N questions with $n > 1$ variants of answer each while student gave C correct answers and I incorrect ones ($I + C = N$) then the corrected result should be $X = C - Y$ where Y is a number of guessed answers. Well, $X \leq C$ means that the corrected result supposed to be less or equal to C .

We state that X could be obtained (estimated) as follows

$$X_0 = \begin{cases} \frac{Cn-N}{n-1}, & \frac{N}{n} < C \leq N \\ 0, & C \leq \frac{N}{n} \end{cases}. \quad (1)$$



In troth, the average number of guessed answers should be $Y = \frac{N-X}{n}$ (each n -th of all unknown ones) and $X = C - Y = C - \frac{N}{n} + \frac{X}{n}$ yield (1). By the way, the case $X = 0 \left(C \leq \frac{N}{n} \right)$ means that unfortunately student couldn't even guess the sufficient answers.

Now let's try to use common probabilistic approach to obtain the desired estimation for X . Let $P(C|X)$ be a probability to give C correct answers while X of them have been guessed ($C \geq X$). Then

$$\begin{aligned} P(C|X) &= P\{\text{guessed } Y - C - X \text{ of } N - X\} = \\ &= C_{N-X}^Y p^Y (1-p)^{N-X-Y} \end{aligned}$$

is nothing but binomial probability where $p = \frac{1}{n}$ is a probability of success and $C_a^b = \frac{a!}{b!(a-b)!}$ is a binomial coefficient. In order to obtain the argmaximum (X) for $P(C|X)$ we'll deal with the inequality

$$\frac{P(C|X+1)}{P(C|X)} = \frac{N-C}{(N-X)(1-p)} \geq 1 \quad (2)$$

because for the corresponding values of X $P(C|X+1) \geq P(C|X)$. Inequality (2) has a plane solution $X \leq \frac{Cn-N}{n-1}$ so the desired argmaximum (see the numerator of LHS at (2)) is $X+1 = \frac{Cn-N}{n-1} + 1$. So, we get the estimation

$$X'_0 = X_0 + 1 = \begin{cases} \frac{Cn-N}{n-1} + 1, & \frac{N}{n} < C \leq N \\ 0, & C \leq \frac{N}{n} \end{cases} \quad (3)$$

and it differs from (1). Why? What's the problem? Well, let's try to substitute X'_0 into (2) and we'll see that

$$\frac{P(C|X'_0+1)}{P(C|X'_0)} = \frac{N-C}{(N-X'_0)(1-p)} = 1.$$

This equality means that both $X_0 = \frac{Cn-N}{n-1}$ and

$X'_0 = X_0 + 1 = \frac{Cn-N}{n-1} + 1$ give the same maximum for $P(C|X)$. SCTR decreases amount of earned points so it'll be honestly to deal with X'_0 instead of X_0 . For example, if $N=12$, $n=4$, $C=6$ then $X_0=4$ but we'll propose $X'_0=5$ instead (we're going to steal only one point instead of two).

The next interesting idea is to add to every question one more variant of answer – "I don't know" (IDK). Well, the testee can be honest (choosing IDK every time he doesn't know the correct answer) or can try to guess the correct answer anyway.

If the testee is honest then we have N questions with $n+1$ variants of answer each (n real variants and IDK one), the probability to guess is $p = \frac{1}{n}$ (when the honest testee really thinks that he knows the correct answer while he doesn't), C is a number of given correct answers, D is a number of IDK answers, X is a number of questions when testee really knew the correct answer, Y – the number of guessed correct answers (when the honest testee really thought that he knew the correct answer while he didn't). $C = X + Y$. Well, this Y is the same one we used at the beginning, so the SCTR will be

$$X_1 = \begin{cases} \frac{Cn - (N - D)}{n - 1} + 1, & \frac{N - D}{n} < C \leq N - D \\ 0, & 0 \leq C \leq \frac{N - D}{n} \end{cases} \quad (4)$$

The number C in (4) differs from C in (3) because now $C = X + Y$ and $Y = \frac{N - D - X}{n}$. But the numerical values of X'_0 and X_1 are the same because of

$$C = X + Y = X + \frac{N - D - X}{n}$$

that is $Cn + D = N - X$ doesn't depend upon D ; SCTR doesn't depend upon chosen schema or testee's honesty because (4) turns into (3) when $D = 0$). E.g. $n = 4$, testee knew 40 questions, mistakenly assumed he knew 40 questions and didn't know 20 questions ($N = 100$). Then in case of (3)

we have $C = 40 + \frac{60}{4} = 55$, $X'_0 = \frac{55 \cdot 4 - 100}{3} + 1 = 41$.

And in case of (4) $C = 40 + \frac{40}{4} = 50$ and

$X_1 = \frac{50 \cdot 4 - (100 - 20)}{3} + 1 = 41$. So, **if the testee is**

dishonest, will he get the same statistical result? In statistical sense the answer is positive but unfortunately sometimes he will get an even better result just guessing correct answers.

In order to remedy the situation we introduce a **new sub-indicator of knowledge strength** (SIKS) in case of IDK-test

$$k = \frac{X_1 - 1}{N - D} =$$

$$= \begin{cases} \frac{Cn - (N - D)}{(n - 1)(N - D)}, & \frac{N - D}{n} < C \leq N - D \\ 0, & 0 \leq C \leq \frac{N - D}{n} \end{cases} \quad (5)$$

It is time to explain why we reduced the numerator. Wasn't it logical to deal with $\frac{X_1}{N - D}$?

Well, it was. But we're not going to obtain a new SCTR, we're going to introduce SIKS with comfy properties.

First, $k \in [0; 1]$ monotonically while

$$C \in \left[\frac{N - D}{n}; N - D \right].$$

Second, SIKS k depends upon D – the greater D , the greater k . Third, no matter what kind of user we have – a cheater or just an unsure one – SIKS k will describe the strength of his knowledge. Under conditions of mentioned example in case of (3) we have

$$k = \frac{X_0}{N} = \frac{40}{100} = 0.4$$

while in case (4)

$$k = \frac{X_1 - 1}{N - D} = \frac{41 - 1}{100 - 20} = 0.5.$$

In practice it will be useful and interesting to include in IDK-test several questions without corrects answers (dead question) and analyze student's choice. For example it is interesting to investigate the correlation between SIKS k and answer to the dead question.

One more interesting question is how SIKS k depends on the level of honesty. Obviously, the testee can behave honestly with some probability $p \in [0; 1]$. $p = 0$ means the dishonest behaviour (trying to guess correct answer by all means), $p = 1$ means the honest behaviour (choosing IDK in appropriate cases). Let $\sim$ be a part of those episodes when student erroneously supposes that he knows the correct answer, in other words $N = X + \sim(N - X) + (1 - \sim)(N - X)$ where student really knows X questions and really doesn't know $(1 - \sim)(N - X)$ ones. Then he will guess approximately

$$Y = \frac{\sim(N - X)}{n} + \frac{(1 - \sim)(N - X)(1 - p)}{n} = \frac{(N - X)(\sim + (1 - \sim)(1 - p))}{n}$$

correct answers. If he knows the formula (4) he will try to maximize the value of $Cn + D = C(p)n + D(p) \rightarrow \max$ but he'll not be able to because

$$\begin{aligned} C(p)n + D(p) &= (X + Y)n + (1 - \sim)(N - X)p = \\ &= Xn + (N - X)(\sim + (1 - \sim)(1 - p)) + \end{aligned}$$

$$+(1-\sim)(N-X)p = X(n-1) + N$$

doesn't depend upon p . Well, this effect we saw earlier – the honesty doesn't affect the estimated grade. On the other hand, the honesty affects the SIKS k . Really, according to independency X_1 of p we obtain

$$k = k(p) = \frac{X_1 - 1}{N - D(p)} = \frac{X_1 - 1}{N - (1-\sim)(N-X)p}.$$

This means that honesty is a good strategy to increase the SIKS k though if $\sim \rightarrow 1$ the honesty loses its effect: if one cannot distinguish white from black no matter how honest he is. And on the contrary, if $\sim \rightarrow 0$ (and may be in addition $X \rightarrow 0$ in sense of first line at (4)) then honesty affects SIKS k strongly. E.g. $N = 40$, $n = 4$, $\sim = 0$, $X = 2$. Then

$$D = D(p) = (1-\sim)(N-X)p = 38p,$$

$$\begin{aligned} C &= C(p) = X + Y(p) = \\ &= X + \frac{(1-\sim)(N-X)(1-p)}{n} = 2 + \frac{38(1-p)}{4}, \end{aligned}$$

Список використаних джерел

1. Булах І. Є. Створюємо якісний тест: навч. посіб./ І. Є. Булах, М. Р. Мруга. – К.: Майстер-клас, 2006 – 160 с.
2. Білоусова Л. І. Потенціал комп'ютерного тестування / Л. І. Білоусова // Вісник ТІМО. – 2008. – № 10. – С. 40 – 44.

$$\begin{aligned} X_1 &= \frac{Cn - (N - D)}{n} + 1 = \\ &= \frac{8 + 338(1-p) - 40 + 38p}{3} + 1 = 3 \end{aligned}$$

(here is our bonus grade – compare $X_1 = 3$ vs $X = 2$).

$$k = k(p) = \frac{X_1 - 1}{N - D(p)} = \frac{2}{40 - 38p} \in \left[\frac{1}{20}; 1 \right].$$

Well, the effect is rather great: $k_{\min} = k(p) = \frac{1}{20}$ and

$k_{\max} = k(1) = 1$ (the biggest possible value for SIKS k).

Well, we have introduced so-called point estimation while interval estimation can be examined too. The standard confidence interval seems to have a doubtful practical value but it will be usefully to obtain the estimate that could be guaranteed (like “not less than” or “not more than”) with some given probability. This can be done both in cases with IDK variant and without it. Some kinds of confidence intervals can be built for SIKS too. The described technique splits up quantitative and qualitative grades but makes it possible to obtain new grades and ratings as functions of introduced estimates (3), (4) and (5).

References

1. BULAH I.Ye. (2006) *Creating high quality tests*/ Kiev: Maister-klas.
2. BILOUSOVA L.I. (2008) *Computer-based testing potential*. Visnyk TIMO, 10.

Надійшла до редколегії 30.05.2015

УДК 519.9

Шликов М. П., аспірант

Про алгоритм побудови довірчого еліпсоїда Петуніна

Київський національний університет імені Тараса
Шевченка, факультет кібернетики, 03680, Україна,
м. Київ, пр. Академіка Глушкова, 4д.
e-mail: maksym_shlykov@ukr.net

M. P. Shlykov, post-graduate student

On the algorithm of Petunin confidence ellipsoid construction

Taras Shevchenko National University of Kyiv,
cybernetics faculty 03680, Kyiv, Glushkova st., 4d,
e-mail: maksym_shlykov@ukr.net

Довірчі еліпсоїди та сфери мають широке застосування у сучасному машинному навчанні. Їх використовують для багатовимірного ранжування, знаходження медіани, знаходження викидів та оцінки квантилів певних рівнів. На сьогодні існує чимало алгоритмів, що використовують еліпсоїди та сфери. Зокрема така конструкція як еліпсоїд мінімального об'єму, що містить дану множину точок використовується для побудови класифікатора, що обмежує область входних даних (gap-tolerant classifier). Ще одним прикладом є застосування сфер для виявлення викидів, запропоноване Таксом та Дуйном. Більшість існуючих алгоритмів мають статистичну природу або ж їх створення було пов'язане з широковідомим методом опорних векторів. Їх основною метою є побудова структури, яка б якомога щільніше містила або відділяла тренувальні дані. Існує довірчий еліпсоїд, запропонований Ю. І. Петуніним. Його головною перевагою є відомий рівень значущості. У роботі запропоновано новий алгоритм побудови еліпсоїда Петуніна. Оригінальний алгоритм, запропонований автором, важко застосовувати у просторах із розмірністю більшою за три. Дана роботи містить модифікацію вихідного алгоритму, що дозволяє легко будувати довірчий еліпсоїд Петуніна у довільних просторах скінченної розмірності.

Ключові слова: довірчі еліпсоїди, багатовимірне ранжування даних, класифікація

Confidence ellipsoids and spheres have wide application in modern machine learning. They can be used for multivariate ranking, determining median, detecting outliers, determining quantile of prescribed level, resolving uncertainties etc. A lot of algorithms that use ellipsoids and spheres have been introduced. Such construction as minimum volume enclosing ellipsoid is used to build gap-tolerant classifier. Another example is support vector data description introduced by Tax and Duin that uses a hypersphere to detect outliers. Most of existing algorithm have statistical origin or were inspired by well-known support vector machine algorithm. Their purpose is to build a structure that tightly contains or separates training data. There exists a confidence ellipsoid introduced by Y. I. Petunin. Its main advantage is known confidence level. In this article we introduce a new algorithm of Petunin ellipsoid construction. The original algorithm has some constraints in its application for finite-dimensional space whose dimension is higher than three. This article contains a modification of the algorithm that allows easy construction of Petunin ellipsoid construction in any finite-dimensional space.

Key Words: confidence ellipsoids, data-ranking, classification

Статтю представив д.ф.-м.н., проф. Буй Д.Б.

Вступ

Машинне навчання сьогодні динамічно розвивається. Це зумовлено можливістю накопичувати та зберігати великі об'єми даних. Однією з актуальних задач машинного навчання є задача класифікації. Зокрема набирає популярності так звана задача однокласової

класифікації (one-class classification), тобто задача відсіювання аномальних об'єктів. Одним з найпопулярніших методів класифікації є метод опорних векторів (support vector machine) [1]. Він будує гіперплощини у просторі ознак, що забезпечують найбільшу ширину проміжку, що розділяє об'єкти тренувальної вибірки. Своєю популярністю метод опорних векторів завдячує

відносній простоті. Задача оптимізації, що лежить в його основі, зводиться до задачі квадратичного програмування.

Метод опорних векторів будує «просту» лінійну поверхню. Сьогодні зроблено чимало спроб ускладнити форму цієї поверхні. Так у роботах [2] та [8] розглядаються методи відсіювання викидів за допомогою сфер. Автори назвали свій алгоритм Support vector data description через його подібність до методу опорних векторів. Основною ідеєю методу є побудова сфери найменшого об'єму, що містила б усі точки, що відмічені як точки заданої множини. При цьому усі тренувальні об'єкти, що є викидами, мають знаходитися ззовні побудованої сфери.

Ще однією широко застосовною конструкцією є еліпсоїд мінімального об'єму, що містить дану множину точок (minimum volume enclosing ellipsoid). У [6] досліджено ймовірність того, що об'єкт, який має бути віднесений до множини, буде невірно класифікований даним еліпсоїдом. Ще одним застосуванням є обмеження області, в якій лежать дані. У роботі [7] еліпсоїд мінімального об'єму застосовується разом із методом опорних векторів для побудови так званого gap-tolerant classifier.

Взагалі застосування еліпсоїдів дуже широке: їх використовують для оцінки медіани, оцінки квантилів певних рівнів а також для побудови багатовимірного ранжування даних. Автори роботи [5] пропонують алгоритм побудови еліпсоїда з найменшим ефективним радіусом (least effective radius) на основі коваріаційної матриці отриманих даних та застосовують його для виявлення викидів у бездротовій сенсорній мережі. Зокрема цікавою є процедура оновлення коваріаційної матриці та переходу до нового еліпсоїда на основі даних, що надходять.

Існує алгоритм, запропонований Ю. І. Петуніним та наведений у роботах [3] та [4], який дозволяє побудувати еліпсоїд, що містить задану множину об'єктів. Спочатку він розглядався як наближення еліпсоїда мінімального об'єму, проте пізніше було виявлено, що даний еліпсоїд має фіксовану ймовірність попадання в нього, яка залежить лише від об'єму вибірки, за якою його будували:

$\frac{n}{n+1}$. Основними застосуваннями даного еліпсоїда на даний момент є багатовимірне ранжування даних ([3]) та еліптичний пілінг ([4]).

Оригінальний алгоритм, запропонований Ю. І. Петуніним, має суттєвий недолік: його важко застосовувати у просторах із розмірністю більшою за три.

Метою даної роботи є запропонувати модифікацію вихідного алгоритму, що дозволяє будувати еліпсоїд Петуніна у просторах довільної скінченної розмірності. Асимптотична складність модифікованого алгоритму співпадає зі складністю оригінального.

Структура статті є наступною. У другій частині ми наведемо оригінальний алгоритм та вкажемо його недоліки. Третя частина буде присвячена модифікації даного алгоритму. Далі буде оцінено складність отриманого алгоритму ті наведено його порівняння з вихідним.

Еліпсоїд Петуніна

У даній частині ми наведемо алгоритм побудови еліпсоїда Петуніна. Як і в роботах [3] та [4] для наочності опишемо побудову спочатку для простору R^2 , а потім сформулюємо загальний алгоритм для довільного скінченновимірного простору. Отже, нехай ми маємо набір точок простору R^m (позначимо його $M_n = \{\vec{x}_1, \dots, \vec{x}_n\}$).

Випадок $m=2$. Знайдемо діаметр множини M_n . Нехай його визначають точки (x_k, y_k) і (x_l, y_l) . Проведемо через точки (x_k, y_k) і (x_l, y_l) пряму L . Знайдемо вершини опуклої оболонки (x_r, y_r) і (x_q, y_q) , що є найвіддаленішими від прямої L та лежать по різні сторони від неї. Проведемо через точки (x_r, y_r) і (x_q, y_q) прямі L_1 і L_2 , паралельні до L . Через точки (x_k, y_k) і (x_l, y_l) будемо дві прямі L_3 і L_4 , перпендикулярні до L . Перетин прямих L_1 , L_2 , L_3 і L_4 утворюють прямокутник, сторони якого мають довжини a і b (нехай, для визначеності, $a \leq b$). Здійснимо поворот та перенос системи координат, так щоб лівий нижній кут прямокутника був розташований на початку нової системи координат з осями Ox' і Oy' , а точки (x_1, y_1) , (x_2, y_2) , ..., (x_n, y_n) перейшли в точки (x'_1, y'_1) , (x'_2, y'_2) , ..., (x'_n, y'_n) . Виконаємо стискання абсцис усіх точок (x'_1, y'_1) , (x'_2, y'_2) , ..., (x'_n, y'_n) з коефіцієнтом $\alpha = \frac{a}{b}$ і одержимо сукупність точок $(\alpha x'_1, y'_1)$, $(\alpha x'_2, y'_2)$, ...,

$(\alpha x'_n, y'_n)$, що лежать у квадраті S . Знайдемо центр (x'_0, y'_0) квадрата S й обчислимо відстані від нього до кожної точки $r_1, r_2, \dots, r_n$. Знайдемо найбільше число $R = \max(r_1, r_2, \dots, r_n)$. Побудуємо коло з центром у точці (x'_0, y'_0) і радіусом R . (Усі точки $(\alpha x'_1, y'_1), (\alpha x'_2, y'_2), \dots, (\alpha x'_n, y'_n)$ тепер лежать всередині цього кола.) Застосовуємо до цього кола операцію розтягування уздовж осі Ox' з коефіцієнтом $\beta = \frac{1}{\alpha}$ і зворотні перетворення повороту і переносу, одержуючи шуканий довірчий еліпс. Легко бачити, що середня складність цього алгоритму визначається складністю побудови опуклої оболонки і дорівнює $O(n \lg n)$.

Випадок $m > 2$. Знайдемо діаметр множини M_n . Нехай $\bar{x}_k$ і $\bar{x}_l$ — дві вершини, що лежать на ньому. Проведемо через $\bar{x}_k$ та $\bar{x}_l$ пряму L . Здійснимо поворот і перенос системи координат, так щоб діаметр опуклої оболонки лежав на осі Ox'_1 . Спроектуємо точки на підпростір, перпендикулярний Ox'_1 . У даному підпросторі здійснимо аналогічні дії (побудову опуклої оболонки, перенос та поворот). Повторюватимемо дані кроки, доки розмірність підпростору, на який необхідно проектувати точки, не стане рівною одиниці. Після цього побудуємо найменший прямокутний паралелепіпед, сторони якого напрямлен вздовж координатних вісей, що містить точки $\bar{x}'_1, \dots, \bar{x}'_n$. Стиснемо простір, перетворюючи прямокутний паралелепіпед у гіперкуб. Знайдемо центр $\bar{x}_0$ гіперкуба й обчислимо відстані $r_1, r_2, \dots, r_n$ від нього до кожної точки. Знайдемо найбільше число $R = \max(r_1, r_2, \dots, r_n)$. Побудуємо гіперкулю з центром у точці $\bar{x}_0$ і радіусом R . Застосуємо до цієї гіперкулі зворотні операції розтягування, повороту і переносу, одержуючи необхідний еліпсоїд у вихідному просторі.

Основним недоліком даного алгоритму є складність його застосування у випадку $m > 3$. Так зокрема складною є побудова повороту на кожному кроці. У зв'язку з цим зменшення розмірності підпростору також створює складнощі при реалізації алгоритму. Практичне застосування алгоритму є зручним лише у дво- та тривимірному просторах.

Альтернативний алгоритм

У цій частині ми дещо модифікуємо вихідний алгоритм побудови довірчого еліпсоїда з метою його зручного застосування для точок простору R^m , $m > 2$. Помітимо, що основною метою алгоритму в цьому випадку є побудова такого лінійного перетворення простору, при якому обрані на кожному кроці діаметри множин (або їх опуклих оболонок) будуть паралельними до векторів базису простору. Після цього все зводиться до простого стискання та знаходження найбільш віддаленої від центру точки.

Нашою основною метою є чітко описати побудову зазначеного вище лінійного перетворення. Уявімо, що в процесі побудови ми на кожному кроці не виконували повороти, а просто проектували та зсували наші вектори. Тоді в кінці нам відомі усі вектори діаметрів опуклих оболонок. З побудови очевидно, що вони утворюють ортогональну систему векторів. Нормуємо її та позначимо $B_m = \{\bar{a}_1, \bar{a}_2, \dots, \bar{a}_m\}$.

Після цього нашим завданням є знайти унітарне перетворення простору, за якого дана ортонормована система векторів стане базисом. Ця задача стає тривіальною, якщо врахувати, що при оберненому перетворенні образами базисних векторів будуть вектори нашої системи B_m . Тобто фактично нам відома матриця оберненого перетворення:

$$U^{-1} = (\bar{a}_1 | \bar{a}_2 | \dots | \bar{a}_m)$$

або ж

$$U = (\bar{a}_1 | \bar{a}_2 | \dots | \bar{a}_m)^{-1}$$

На цьому кроці дуже зручним є те, що U — унітарний оператор, а тому $U^{-1} = U^T$. Наступним кроком до спрощення процесу реалізації нашого алгоритму буде модифікація переходу до підпростору меншої розмірності. Нашою метою тут є зробити так, щоб зсув простору був одним. Для цього розглянемо перший крок алгоритму. Нехай ми знайшли вектори діаметра опуклої оболонки нашої множини точок (позначимо їх $\bar{x}_k$ та $\bar{x}_l$). Якщо не робити поворот та зсув, на першому кроці нам необхідно спроектувати наші точки на лінійний многовид $\bar{x}_k + L$, де L — ортогональне доповнення прямої, визначеної вектором $\bar{x}_k - \bar{x}_l$. Фактично це гіперплощина, що проходить через $\bar{x}_k$ та ортогональна до $\bar{x}_k - \bar{x}_l$. Позначимо точки, отримані після проектування $M_n^{(1)} = \{\bar{x}_i^{(1)}, i = 1, n-1\}$. Зауважимо, що кількість різних точок після проектування зменшиться на

одну, бо за побудовою проекцією точки $\vec{x}_k$ буде $\vec{x}_l$. Покажемо, що для подальших кроків ми можемо не переходити до підпросторів меншої розмірності. Замість цього нам достатньо на кожному кроці виконувати проектування на лінійний многовид, підпростір якого є гіперплощиною, ортогональною до вектора діаметру опуклої множини. Нехай на деякому кроці ми маємо множину $M_n^{(k)} = \{\vec{x}_i^{(k)}, i = 1, n-k\}$ векторів простору R^m , що належать лінійному многовиду $u + L_1$, розмірність якого дорівнює p . Припустимо також, що ми знайшли вектор діаметру даної множини і він дорівнює $\vec{x}_l^{(k)} - \vec{x}_t^{(k)}$. Розглянемо тепер гіперплощину, ортогональну до цього вектора. Її рівняння має вигляд:

$$(\vec{x}, \vec{x}_l^{(k)} - \vec{x}_t^{(k)}) = (\vec{x}_l^{(k)}, \vec{x}_l^{(k)} - \vec{x}_t^{(k)})$$

При цьому проекція точки на дану гіперплощину має вигляд

$$P\vec{x} = \vec{x} - \frac{(\vec{x}, \vec{x}_l^{(k)} - \vec{x}_t^{(k)}) - (\vec{x}_l^{(k)}, \vec{x}_l^{(k)} - \vec{x}_t^{(k)})}{\|\vec{x}_l^{(k)} - \vec{x}_t^{(k)}\|^2} (\vec{x}_l^{(k)} - \vec{x}_t^{(k)})$$

Оскільки $\vec{x}_t^{(k)}$ лежить в L_1 , то $\vec{y} + L_1 = \vec{x}_l^{(k)} + L_1$.

Тому якщо $\vec{x} \in \vec{x}_t^{(k)} + L_1$, то $\vec{x} = \vec{x}_t^{(k)} + \vec{y}, \vec{y} \in L_1$

$$P\vec{x} = \vec{x} - \alpha(\vec{x}_l^{(k)} - \vec{x}_t^{(k)}) = \vec{x}_t^{(k)} + \vec{y} - \alpha(\vec{x}_l^{(k)} - \vec{x}_t^{(k)})$$

де $\alpha = \frac{(\vec{x}, \vec{x}_l^{(k)} - \vec{x}_t^{(k)}) - (\vec{x}_t^{(k)}, \vec{x}_l^{(k)} - \vec{x}_t^{(k)})}{\|\vec{x}_l^{(k)} - \vec{x}_t^{(k)}\|^2}$. Також

ми знаємо, що $\alpha(\vec{x}_l^{(k)} - \vec{x}_t^{(k)}) \in L_1$, а тому

$P\vec{x} \in \vec{x}_t^{(k)} + L_1$. При цьому за означенням проекції

$P\vec{x}$ належить гіперплощині $(\vec{x}, \vec{x}_l^{(k)} - \vec{x}_t^{(k)}) = (\vec{x}_l^{(k)}, \vec{x}_l^{(k)} - \vec{x}_t^{(k)})$, яка теж є лінійним многовидом. Тому проекції наших точок лежатимуть у перетині двох лінійних многовидів, причому наша гіперплощина є лінійним многовидом $\vec{x}_t^{(k)} + L_2$, де L_2 — це підпростір розмірності $m-1$. Перетин цих лінійних многовидів має вигляд $\vec{x}_t^{(k)} + L_1 \cap L_2$

При цьому оскільки $\vec{x}_l^{(k)} - \vec{x}_t^{(k)}$ ортогональний до

L_2 , то $L_1 + L_2 = R^m$. Запишемо тепер формулу

Грасмана, щоб знайти розмірність $L_1 \cap L_2$:

$$m = m-1 + p - \dim L_1 \cap L_2$$

звідки

$$\dim L_1 \cap L_2 = p-1$$

З урахуванням цього ми на кожному кроці можемо виконувати проекцію на гіперплощину перпендикулярну вектору діаметра, що проходить через одну з точок цього діаметра. При цьому насправді на кожному кроці розмірність многовида, у якому лежатимуть точки, зменшуватиметься. В оригінальному алгоритмі це перехід до підпростору, що перпендикулярний черговій осі координат, отриманій після повороту.

Таким чином у підсумку наш алгоритм набуває вигляду:

Вхідні дані: $M_n = \{\vec{x}_1, \dots, \vec{x}_n\}$

Алгоритм

1. $M_n^{(0)} \leftarrow M_n, B = \emptyset, k = 0$;

2. Доки $k < m-1$ виконувати:

2.1. Знайти $\vec{x}_l^{(k)}$ та $\vec{x}_t^{(k)}$ — точки, що визначають діаметр множини $M_n^{(k)}$;

2.2. $B \leftarrow B \cup \left\{ \frac{\vec{x}_l^{(k)} - \vec{x}_t^{(k)}}{\|\vec{x}_l^{(k)} - \vec{x}_t^{(k)}\|} \right\}$;

2.3. $M_n^{(k+1)} \leftarrow P_L M_n^{(k)}$, де L — це гіперплощина, визначена рівнянням $(\vec{x}, \vec{x}_l^{(k)} - \vec{x}_t^{(k)}) = (\vec{x}_l^{(k)}, \vec{x}_l^{(k)} - \vec{x}_t^{(k)})$;

2.4. $k \leftarrow k+1$;

3. Сформувати матрицю U , стовпчиками якої є вектори множини B ;

4. $M_n' \leftarrow U^T M_n$;

5. Знайти мінімальний паралелепіпед зі сторонами, напрямленими вздовж координатних осей, що містить множину M_n' . Нехай $\vec{c}$ — центр цього паралелепіпеда;

6. Побудувати перетворення S , що переводить цей паралелепіпед в куб;

7. $M_n'' \leftarrow S M_n', \vec{c}' \leftarrow S \vec{c}$;

8. $R \leftarrow \max \{ \|\vec{x} - \vec{c}'\|, x \in M_n'' \}$;

9. $E \leftarrow \frac{1}{R^2} \begin{pmatrix} 1 & \dots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \dots & 1 \end{pmatrix}$;

10. $E \leftarrow U S^T E S U^T, \vec{c} \leftarrow U \vec{c}'$;

Вихідні дані: E — матриця вихідного еліпсоїда, $\vec{c}$ — центр.

Застосування запропонованого алгоритму можливе лише у випадку, коли шуканий еліпсоїд існує та є не виродженим. Іншими словами не

має існувати підпростору вихідного простору R^m , що містить вхідну множину точок.

Складність

У даній частині ми вважатимемо операції проектування точки на афінний підпростір, лінійного перетворення простору та порівняння двох точок елементарними, оскільки

Асимптотично найскладнішою частиною обох алгоритмів є знаходження діаметра скінченної множини точок. Кількість операцій, необхідна для пошуку діаметра перебором, дорівнює $\frac{n(n-1)}{2}$. Проте цей результат може

бути покращений для дво- та тривимірних просторів за допомогою опуклої оболонки. Можна побудувати опуклу оболонку та здійснювати перебір по її вершинах. Асимптотична складність таких дій буде $O(n \lg n + \frac{z(z-1)}{2})$, де z — кількість вершин

опуклої оболонки. В найгіршому випадку складність залишиться квадратичною, проте для більшості вибірок кількість вершин опуклої оболонки є незначно у порівнянні з розміром вибірки, тому складність для таких вибірок буде $O(n \lg n)$. Для просторів R^m , $m > 3$ складність алгоритмів буде $O(n^2)$

Порівняння на практиці

Порівняння алгоритмів проводилося для двотривимірного випадку. Обидва алгоритми було реалізовано мовою R. Було згенеровано 100 вибірок, кожна з яких містила по 200 векторів. Кожна вибірка містила по 100 векторів з двох нормальних розподілів з однаковими діагональними матрицями коваріації та різними векторами математичних сподівань. Для кожного алгоритму розглядався сумарний час роботи на всіх 100 вибірках. Приймавши сумарний час роботи оригінального алгоритму у двовимірному випадку за одиницю, результати мають наступний вигляд:

	Оригінальний алгоритм	Модифікований алгоритм
R^2	1	2,0082
R^3	1,9951	2,9404

Таблиця 1. Порівняння часу роботи алгоритмів.

Слід зауважити, що для тривимірного простору після першого проектування в оригінальному алгоритмі задача зводиться до випадку $m=2$. Це означає, що після повороту ми можемо на площині, перпендикулярній, Ox'_1 застосувати алгоритм для випадку $m=2$ та побудувати шуканий прямокутник, а потім за ним побудувати паралелепіпед у просторі. Це було враховано при програмній реалізації оригінального алгоритму. Таке спрощення зменшує кількість необхідних пошуків діаметру множини на 1 і нам необхідно буде їх здійснювати $m-1$ разів. Модифікований алгоритм при цьому вимагає здійснювати пошук діаметру множини m разів. В іншому алгоритми здійснюють подібні дії. Таким чином співвідношення між часом роботи двох алгоритмів на одних і тих самих вхідних даних має бути близьким до $\frac{m-1}{m}$, що і відображено у результатах.

Як впливає з результатів порівняння, модифікований алгоритм має дещо більший час роботи. Проте зі збільшенням розмірності простору дана різниця нівелюється. Крім того, як ми вже казали раніше, застосування оригінального алгоритму у просторах з розмірністю, більшою за 3, є проблематичним у зв'язку зі складністю побудови поворотів

Висновки

У статті запропоновано новий алгоритм побудови еліпсоїда Петуніна. Даний еліпсоїд застосовується для багатовимірного ранжування та пілінгу даних для оцінки медіани. Наведено складність отриманого алгоритму та проведено порівняння з оригінальним алгоритмом. Основною його перевагою є легкість застосування для просторів довільної скінченної розмірності. Недоліком є дещо більший час роботи порівняно з оригінальним алгоритмом на просторах розмірності 2 і 3. Подальшим покращенням могло б стати вдосконалення процесу знаходження діаметру множини.

Ще однією проблемою, яка виникає при побудові даного еліпсоїда є досягнення наперед визначеного рівня значущості. Зі зростанням кількості тренувальних даних рівень значущості згідно ймовірності потрапляння до еліпсоїда буде прямувати до нуля. Для отримання, наприклад, рівня значущості 0,05 необхідно 19 тренувальних точок. Тому постає питання вибору цих 19 точок так, щоб отриманий еліпсоїд містив тренувальну

вибірку та деякою мірою відображав сукупність,
з якої було взято дані.

Список використаних джерел

1. Herbrich R., Learning Kernel Classifiers: Theory and Algorithms. / R. Herbrich — MIT Press — Cambridge, MA, 2001 — 382 P.
2. Tax D.M.J., Support vector data description. / D.M.J. Tax, R.P.W. Duin // Machine learning — 2004. — 54(1). — P. 45-66.
3. Ключин Д.А., Еліптична функція статистичної глибини даних. / Д.А. Ключин, М.В. Присяжна // Вісник Київського національного університету імені Тараса Шевченка, Серія фізико-математичні науки. — 2014. — № 1. — С.179-184.
4. Lyashko S.I., Multivariate ranking using elliptical peeling. / S.I. Lyashko, D.A. Klyushin, V.V. Alexeyenko // Cybernetics and Systems Analysis — 2013. — 49(4). — P. 511-516.
5. Zhang Y., Distributed online outlier detection in wireless sensor networks using ellipsoidal support vector machine. / Y. Zhang, N. Meratnia, P.J.M. Havinga // Ad hoc networks. — 2012. — 11(3). — P. 1062-1074.
6. Dolia A.N., The minimum volume covering ellipsoid estimation in kernel-defined feature spaces. / A.N. Dolia, T. Bie, C.J. Harris // ECML, 18-22 September 2006, Berlin, Germany. — 2006. — P. 630-637.
7. Shivaswamy, P., Ellipsoidal kernel machines. / P.Shivaswamy, T. Jebara // AISTATS, 21-24 March 2007, San Juan, Puerto Rico. — 2007 — P. 484-491.
8. Wang J., Pattern classification via single spheres. / J. Wang, P. Neskovic, L.N. Cooper // Discovery Science, 8-11 October 2005, Singapore, — 2005 — P. 241-252.

References

1. HERBRICH R. (2001) *Learning Kernel Classifiers: Theory and Algorithms* MIT Press, Cambridge, MA
2. TAX D.M.J., DUIN R.P.W. (2004) Support vector data description. *Machine learning* 54(1). P. 45-66.
3. KLYUSHIN D.A., PRYSIAZHNA M.V. (2014) Elliptical statistical data depth function. *Bulletin of Taras Shevchenko National University of Kyiv. Series Physics & Mathematics*, 1. P.179-184.
4. LYASHKO S.I., KLYUSHIN D.A., ALEXEYENKO V.V. (2013) Multivariate ranking using elliptical peeling. *Cybernetics and Systems Analysis* 49(4). P. 511-516.
5. ZHANG Y., MERATNIA N., HAVINGA P.J.M. (2012) Distributed online outlier detection in wireless sensor networks using ellipsoidal support vector machine. *Ad hoc networks*. 11(3). P. 1062-1074.
6. DOLIA A. N., BIE T., HARRIS C. J. (2006) The minimum volume covering ellipsoid estimation in kernel-defined feature spaces. In *ECML*. Berlin, Germany, 18th to 22nd September 2006. P. 630-637.
7. SHIVASWAMY P., JEBARA T. (2007). Ellipsoidal kernel machines. In *AISTATS*. San Juan, Puerto Rico, 21st to 24th March 2007 . P. 484-491.
8. WANG J., NESKOVIC P., COOPER L. N. (2005) Pattern classification via single spheres. In *Discovery Science*. Singapore, 8th to 11th October 2005. P. 241-252.

Надійшла до редколегії 29.05.2015

УДК 519.6

Яковенко О.А., аспірант

Про прикладне програмне забезпечення для задач нелінійного програмування

Київський національний університет імені
Тараса Шевченка, 83000, м. Київ,
пр-т. Глушкова 4д,
e-mail: sandra.yakov@gmail.com

O. Iakovenko, graduate student

Nonlinear programming software

Taras Shevchenko National University of Kyiv,
83000, Kyiv, Glushkova st., 4d,
e-mail: sandra.yakov@gmail.com

У статті проводиться порівняльний аналіз можливостей сучасних пакетів математичного програмування для задачі нелінійної оптимізації. Окремлено коло базових чисельних методів оптимізації розв'язання нелінійних задач. Подається перелік існуючих оптимізаторів та мов математичного програмування. Наводяться основні напрямки можливого розвитку існуючих інструментаріїв.

Ключові слова: алгоритм, розв'язувач, мова алгебраїчного моделювання, чисельний метод, нелінійна оптимізація.

Nonlinear programming problems are in general more difficult to solve than linear programming problems. Solving nonlinear problems without modern technologies would be complicated and laborious. Effective application of a particular algorithm for solving a nonlinear constrained optimization problem primarily depends on the optimization algorithm practical realization quality.

Most frequently used solvers (for solving constrained optimization problems) and algebraic modelling languages are represented. Point is also made of software packages operational functionality. It is concluded that there is no universal algorithm for solving nonlinear optimization problem.

An article deals with most common nonlinear optimization problem solvers and methods employed by this solvers. Numerical methods which are mostly used in mathematical software are the following: GRG method, interior point method, quasi newton method, branch and bound method, Frank–Wolfe algorithm (also known as conditional gradient method) and trust region method. Programming languages used to develop optimization algorithms were also analyzed.

It is also stated that nonlinear systems properties analysis is carried out only within used optimization method. Therefore, the question of instrumental analysis (especially for systems where the mathematical model changes are required) is not resolved yet.

Keywords: algorithm, solver, numerical method, nonlinear optimization.

Статтю представив д.т.н., проф. Волошин О.Ф.

Розвиток моделей і методів оптимізації стимулювався значним збільшенням розмірності і складності оптимізаційних задач, викликаним істотним технологічним підйомом після другої світової війни [1, с.13].

Найбільш поширені пакети математичного програмування, що надають можливість проведення аналізу властивостей лінійних моделей, були розглянуті у попередній статті [2].

Проте, процеси різної природи формалізуються задачами нелінійного програмування різної складності. Переважна більшість прикладних задач оптимізації є

нелінійними. Чисельне розв'язання задачі нелінійної оптимізації без застосування сучасних технологій бачиться складним та обтяжливим. Тому сучасні процедури аналізу та алгоритми нелінійної оптимізації, що використовуються для їх розв'язання у пакетах прикладного математичного програмування, викликають жвавий науковий інтерес як базовий інструментарій розв'язання практичних задач.

Загальна теорія розв'язання задач нелінійного програмування на даний час знаходиться в стані становлення. Для розв'язання таких задач на практиці використовуються алгоритмічно запрограмовані чисельні методи,

які розв'язують задачу за допомогою ЕОМ. Теоретична та практична складність проблеми обмежує можливості для створення загальноприйнятого універсального чисельного методу, який дозволяв би отримувати оптимальний розв'язок для будь-якої задачі нелінійної оптимізації. Через це під час розв'язання оптимізаційної задачі, що була сформульована як задача нелінійного програмування, можуть знадобитись декілька методів. Тому вибір та ефективне застосування кожного з них у процесі розв'язання задачі нелінійної оптимізації залежить насамперед від якості практичної реалізації алгоритмів оптимізації, які є в розпорядженні у дослідника.

Формально задача нелінійного програмування може бути сформульована таким чином:

Мінімізувати

$$f(x), x \in E^n, \quad (1)$$

за m лінійних та(або) нелінійних обмежень у вигляді рівностей

$$h_1(x) = 0, j = 1, \dots, m, \quad (2)$$

і $(p - m)$ лінійних та(або) нелінійних обмежень у вигляді нерівностей:

$$g_1(x) \geq 0, j = m + 1, \dots, p. \quad (3)$$

[3, с. 19].

Математичні методи, які були розроблені дослідниками в минулому, дозволяють розв'язати (в багатьох випадках) задачі вигляду (1)-(3) багатьма способами.

Методи їх розв'язання прийнято поділяти на прямі та двоїсті, у залежності від того, чи розв'язують вони вихідну задачу, або ж знаходять відповідь із допомогою двоїстої. Програмна реалізація методів дозволила підняти на якісно новий рівень швидкість вирішення проблем оптимізації [1, 4].

В результаті бурхливого розвитку прикладного програмного забезпечення у сфері математичних комплексів з'явилась велика кількість математичних пакетів, що дозволяють розв'язувати задачі математичної оптимізації. З одного боку, це було "спровоковане" ростом та можливостями обчислювальних потужностей комп'ютерних систем, і з іншого, нагальною виробничою необхідністю (формулюються вимоги до інструментарію проведення обчислень). Тому на сьогоднішній день ми маємо досить широкий діапазон для вибору інструментарію, що використовується.

Таке розмаїття породжує резонне запитання – то який саме математичний пакет

краще враховує потреби "замовника" та придатний для розв'язання конкретної задачі?

У сучасних наукових дослідженнях ретельно розглянуті питання використання прикладних програмних комплексів та наявних алгоритмів для розв'язання конкретної прикладної задачі нелінійної оптимізації. У таких дослідженнях робляться висновки про продуктивність та доцільність використання конкретного алгоритму для розв'язання типової задачі. Нам здається необхідним зробити акцент і на огляді та порівнянні існуючого інструментарію не лише на базі конкретної задачі, а за допомогою аналізу і висвітлення спектру функціональних можливостей у програмних продуктах, що використовуються найчастіше.

В цілому арсенал для розв'язання задачі оптимізації можна поділити на мови алгебраїчного моделювання та на розв'язувачі.

Серед мов алгебраїчного моделювання можна виділити шість найбільш вживаних [5]. Такі системи використовують для опису та розв'язання проблем високої складності великомасштабних математичних розрахунків.

Цей список містить такі мови:

1. AIMMS (Advanced Interactive Multidimensional Modeling System);
2. AMPL (A Mathematical Programming Language);
3. GAMS (General Algebraic Modelling System);
4. LINDO/LINGO;
5. MPL (Mathematical Programming Language);
6. MathLab.

Будучи системами моделювання високого рівня, вони підтримують велику кількість і комерційних розв'язувачів, і таких що мають відкритий код. Перевага використання мови алгебраїчного моделювання як середовища програмування полягає у тому, що у разі її використання оптимізаційна модель записується лише один раз, а для того, щоб змінити алгоритм (тобто розв'язувач, який реалізовує такий алгоритм), достатньо змінити декілька рядків коду і параметри моделі будуть передаватися у зазначеному напрямку. Це не тільки звільняє від необхідності переписувати модель декілька разів, але й надає можливість порівняти розв'язання тієї ж самої моделі різними методами.

Також існують дослідження, що порівнюють конкретні розв'язувачі, які

використовує дослідник під час розв'язання своєї задачі у рамках використання конкретної мови математичного програмування [6], наприклад, AMPL [7] або GAMS [8].

Серед розв'язувачів найбільш вживаними для розв'язання задач нелінійної умовної оптимізації визначають такі: CONOPT, SNOPT, Gurobi, MOSEK, KNITRO, MINOS, LOQO, IPOPT, BARON, XPRESS.

Варто також зазначити, що серед вище перелічених мов, LINDO та MathLab (Optimization Toolbox) являють собою і мову алгебраїчного моделювання і оптимізатор одночасно. Саме тому вважаємо за необхідне включити їх у список розв'язувачів, які будемо досліджувати.

Таблиця 1

Розв'язувач	Алгоритмізований метод
LINDO	GRG method (Метод умовного градієнта)
CONOPT	GRG method (Метод умовного градієнта)
LOQO	Метод внутрішньої точки
SNOPT	Квазіньютонівський метод
KNITRO	Метод довірчих інтервалів
MOSEK	Метод внутрішньої точки
MINOS	Метод умовного градієнта
BARON	Метод гілок і меж
XPRESS	Метод гілок і меж
IPOPT	Метод внутрішньої точки
Mathlab Optimization Toolbox	Метод довірчих інтервалів

У таблиці 1 наведені найбільш поширені пакети прикладного програмування, що оснащені розв'язувачами нелінійних задач математичного програмування та методи, які вони використовують для розв'язання.

Керуючись інформацією, отриманою під час проведення дослідження, можемо констатувати, що найбільш згадуваними у літературі і поширеними вважаються оптимізатори: KNITRO, SNOPT та MINOS.

Порівняльна таблиця 1 наочно демонструє, що розв'язувачі використовують різноманітні методи. Це підтверджує зроблене на початку статті зауваження, що нелінійне програмування, на відміну від лінійного, не має універсального методу розв'язання оптимізаційних задач. З іншого боку, це надає можливість для отримання кращого результату за рахунок того, що можна вирішити проблему

оптимізації декількома методами і порівняти результати, або навіть заздалегідь подивитись на спеціалізованих сайтах, який саме метод використовується для розв'язання конкретної задачі, що дозволяє користувачу самостійно визначити, який саме розв'язувач використовувати. Адже кожна задача унікальна, а найефективніший алгоритм може варіюватись.

Мови, на яких написані алгоритми оптимізації, належать до ешелону перших мов програмування: C/C++ та FORTRAN (його різні модифікації(версії), як то Fortran77, Fortran 90, Fortran95 тощо). Це пов'язано з тим, що теоретична база цих алгоритмів була розроблена раніше, аніж почали бурхливо розвиватися комп'ютерні технології загалом і мови програмування зокрема. Як тільки з'явилась фундаментальна практична база і можливість запрограмувати існуючі алгоритми оптимізації, починаючи з 70-х років з'являлися і перші версії оптимізаторів, які були згодом покращені.

У той час, як у лінійному програмуванні переважна більшість задач розв'язується за допомогою програмного забезпечення, що розв'язує задачі загальновизначим ефективним симплекс-методом [2], та робляться активні спроби застосування релаксаційних схем та структурного аналізу властивостей лінійних систем [9], нелінійне програмування все ще залишається областю дослідження на стадії становлення, включає в себе елементи експериментування.

Можна констатувати, що нелінійні задачі оптимізації все ще не мають універсального алгоритму розв'язання. Більшою мірою на сьогоднішній день процес розв'язання задачі нелінійного програмування зводиться до пошуку найкращого алгоритму для розв'язання конкретної задачі шляхом порівняння результатів використання різних методів. Аналіз властивостей нелінійних систем проводиться лише у рамках методу оптимізації, що використовується. І тому питання інструментального аналізу властивостей таких систем (особливо при змінах параметрів математичної моделі) комплексно не вирішено. Це залишає широке поле для подальших наукових досліджень.

Список використаних джерел

1. Базара М. Нелинейное программирование. Теория и алгоритмы / М. Базара, К. Шетти, Х. Шерали. – М.: Мир, 1982. – 583 с.
2. Яковенко О.А. Про інструментарій аналізу властивостей лінійних моделей у пакетах математичного програмування / О.А. Яковенко, В.І. Кудін // Вісник Київського національного університету імені Т.Шевченка. Серія физ.-мат. науки. – 2014. – 4. – С. 210-215.
3. Химмельблау Д. Прикладное нелинейное программирование: [пер. с англ. И.М. Быховской, Б.Т. Вавилова] / Д. Химмельблау. – М.: Мир, 1975. – 534 с.
4. Numerical methods for constrained optimization / Edited by P. E. Gill and W. Murray. – Academic Press, London, 1974. – 283 p.
5. Kallrath J. Modeling Languages in Mathematical Optimization / Josef Kallrath. – Springer, 2004. – 408 p.
6. Neumaier A. A comparison of complete global optimization solvers / Arnold Neumaier, Oleg Shcherbina, Waltraud Huyer, Tamás Vinkó // Mathematical Programming. – June 2005. – Volume 103, Issue 2. – pp. 335-356.
7. Benson H. Y. A Comparative Study of Large-Scale Nonlinear Optimization Algorithms / Hande Y. Benson, David F. Shanno, Robert J. Vanderbei // High Performance Algorithms and Software for Nonlinear Optimization, Applied Optimization. – 2003. – Volume 82. – pp. 95-127.
8. Lastusilta T. Comparison of some high-performance MINLP solvers / T. Lastusilta, M. Bussieck, T. Westerlund // Chemical Engineering Transactions, 2007. – pp. 125-130.
9. Кудин В.И. Релаксационная система решения задачи линейного программирования с использованием метода последовательного анализа / В.И. Кудин // Автоматика. – 1989. – №3. – С. 42-47.

References

1. BAZARAA, M. S., SHERALI, C., SHETTY, K. (2006) *Nonlinear programming: theory and algorithms*. New York : Wiley.
2. IAKOVENKO, O.A., KUDIN, V.I. *About existing linear models properties analysis tools in mathematical software packages*. Bulletin of Taras Shevchenko National University of Kyiv. Series Physics & Mathematics. – 2014. – 4. – pp. 210-215.
3. HIMMELBLAU, D. (1972) *Applied nonlinear programming*. New York : Mcgraw-Hill.
4. GILL, P. E., MURRAY, W., ed. (1974) *Numerical methods for constrained optimization*. London : Academic Press.
5. KALLRATH, J. (2004) *Modeling Languages in Mathematical Optimization*. New York : Springer.
6. NEUMAIER, A., SHCHERBINA, O., HUYER, W., VINKÓ, T. *A comparison of complete global optimization solvers*, Mathematical Programming. – June 2005, Volume 103, Issue 2. – pp. 335-356.
7. BENSON, H. Y., SHANNO, D. F., VANDERBEI, R. J. *A Comparative Study of Large-Scale Nonlinear Optimization Algorithms*. High Performance Algorithms and Software for Nonlinear Optimization, Applied Optimization. – 2003. – Volume 82. – pp. 95-127.
8. LASTUSILTA, T., BUSSIECK, M., WESTERLUND, T. *Comparison of some high-performance MINLP solvers*. Chemical Engineering Transactions, 2007. – pp. 125-130.
9. KYDIN, V.I. *Relaxation system for solving the linear programming problem using sequential analysis*. Avtomatika. – 1989. – №3. – pp. 42-47.

Надійшла до редколегії 20.04.15

РАДІОФІЗИКА

УДК 537.52:533.9.07:538.911

Веклич А.М., д.ф.-м.н., проф.,
Клешич М.М., асп.,
Борецький В.Ф., к.ф.-м.н., доц.,
Фесенко С.О., асист.,
Телега В.М., к.ф.-м.н., зав. лаб. НДІ ФЕ

Спектроскопія плазми електричної дуги між Cu-Ni електродами в потоці аргону

Київський національний університет імені
Тараса Шевченка, 03022, просп. Глушкова 4Г,
Київ, Україна
e-mail: van@univ.kiev.ua

Veklich A.N., D.Sci., Professor,
Kleshich M.M., Ph.D. stud.,
Boretskij V.F., Ph.D., Associate Professor,
Fesenko S.O., Assistant Professor,
Telega V.N., Ph.D., head of R&D Phys. El. Lab.

Plasma spectroscopy of electric arc between Cu-Ni electrodes in argon flow

Taras Shevchenko National University of Kyiv,
03022, Glushkova Av., 4G, Kyiv, Ukraine
e-mail: van@univ.kiev.ua

Досліджено плазму електродугового розряду між Cu – Ni електродами в потоці аргону за допомогою оптичної емісійної спектроскопії у припущенні локальної термодинамічної рівноваги. Методом відносних інтенсивностей одержані радіальні профілі температури у стовпі плазми. Для визначення радіального розподілу температури використані спектральні лінії міді та нікелю. Рекомендовані спектральні лінії Ni I для діагностики плазми.

Ключові слова: плазма електродугового розряду, оптична емісійна спектроскопія, метод відносних інтенсивностей.

Plasma of electric arc discharge between Cu–Ni electrodes in argon flow at arc current 3.5 A in the assumption of local thermodynamic equilibrium was investigated by optical emission spectroscopy. The qualitative difference between emission spectrum of the investigated plasma source in comparison with the free-burning electric arc in air was found. Namely, its spectrum contains a much larger number of nickel atomic spectral lines. Selection of Ni I spectral lines for purposes of plasma diagnostics was carried out. The radial profiles of temperature in discharge column were obtained by method of relative intensities of spectral lines. Copper and nickel atomic spectral lines were used with this aim. The total intensity of three spectral lines Ni I 508.0, 508.1, 508.4 nm in combination with line Ni I 547.6 nm is proposed to determine the temperature by method the relative intensities in diagnostics of plasma with nickel vapour impurities.

Key words: plasma of electric arc discharge, optical emission spectroscopy, method of relative intensities.

Статтю представив д.ф.-м.н., проф. Анісімов І.О.

Вступ

На сьогодні при виготовленні електродів або контактів приладів, які використовуються в електротехнічній галузі для комутації електричного струму, часто використовуються композитні матеріали [1]. Перевагою таких матеріалів є поєднання високої ерозійної стійкості та задовільної теплопровідності і електропровідності. Зокрема, прийнятні ерозійні властивості композитів забезпечуються завдяки тугоплавкій компоненті в цій композиції (наприклад, вольфраму, молібдену або оксидів металів). Наявність легкоплавкої компоненти у складі композита (наприклад, міді або срібла)

забезпечує необхідну електропровідність та теплопровідність. Такі композитні матеріали, як правило, виготовляються методами порошкової металургії [2, 3].

Очевидно, розробка композитних матеріалів не може бути покращена без досконалого вивчення впливу електричної дуги на робочий шар поверхні електродів. Отже, доцільно дослідити вплив параметрів плазми дугового розряду на ерозійні властивості таких композитів. Тому, першочерговим завданням такого типу досліджень є розробка засобів контролю параметрів плазми таких розрядів [4]. Раніше для діагностики електричної дуги, яка

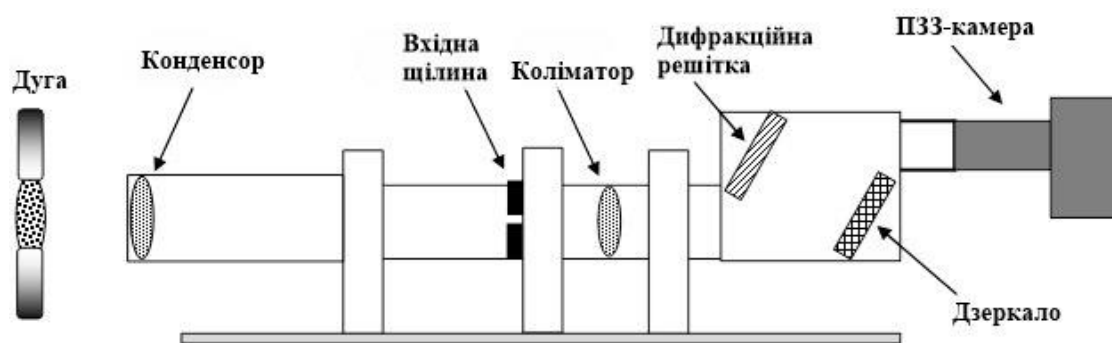


Рис. 1 Оптична схема експериментальної установки [8].

горить у повітрі [5-7] або в потоці аргону [8] між композитними електродами на основі міді або срібла, розроблені методи оптичної емісійної спектроскопії плазми та лазерної абсорбційної спектроскопії. У припущенні локальної термодинамічної рівноваги (ЛТР) в плазмі досліджено дуговий розряд силою струму 3.5 або 30 А між плоскими торцевими поверхнями неохолоджуваних стрижневих електродів. Такі дуги можуть бути використані як модельне джерело реальних дугових розрядів у комутуючих приладах електротехнічної промисловості.

Мета цієї роботи – розробка методики оптичної емісійної спектроскопії плазми електричної дуги з домішками парів нікелю. А саме, в рамках цього дослідження, виконується:

- порівняння спектрів випромінювання плазми вільно існуючого розряду між Cu–Ni електродами у повітрі та в потоці аргону;
- селекція спектральних ліній та спектроскопічних констант атому нікелю для використання їх в оптичній емісійній спектроскопії при визначенні температури плазми.

Отже, застосування перевіреної в цій роботі спектроскопічної методики дозволить дослідити на наступних етапах проблему взаємодії плазми дуги і поверхні композитних Cu–Ni електродів.

Експериментальна установка

Електрична дуга запалювалась у потоці аргону (6.4 стандартних літрів на хвилину) між торцевими поверхнями неохолоджуваних електродів. Розрядний проміжок складав 8 мм, діаметр електродів – 6 мм, струм дуги – 3.5 А. Електроди розташовані вертикально: верхній електрод – Ni (катод), нижній електрод – Cu

(анод). Такого типу розряд можна розглядати як спрощену початкову модель електричної дуги між композитними електродами Cu–Ni. Потік газу (аргону) у цьому джерелі спрямовувався з боку верхнього електроду співвісно до розряду.

З метою порівняння одержаних результатів виконували дослідження вільно існуючої у повітрі електричної дуги у такому ж режимі роботи розряду.

Для одночасної реєстрації спектральних та просторових розподілів інтенсивності випромінювання електричної дуги застосували розроблену раніше методику діагностики [8]. У цьому випадку використали дифракційний спектрометр та цифрову камеру на основі приладу із зарядовим зв'язком (ПЗЗ) (рис. 1).

Для обробки отриманих зображень спектрів випромінювання плазми використано графічний інтерфейс користувача [8].

Реєстрацію просторових розподілів інтенсивності випромінювання електродугового розряду виконували у спектральному діапазоні 400 – 660 нм.

Через бічне спостереження неоднорідного плазмового об'єкта (як було реалізовано в даній експериментальній установці) для одержання локальних значень інтенсивності необхідно застосовувати перетворення Абеля. В графічному інтерфейсі користувача ця процедура виконується відповідно до алгоритму, запропонованому в роботі [9], у припущенні осової симетрії джерела оптичного випромінювання [8].

В процесі реєстрації спектрів випромінювання електродугової плазми врахована нерівномірність спектральної чутливості ПЗЗ матриці. Максимальна чутливість використаної в експериментальних дослідженнях

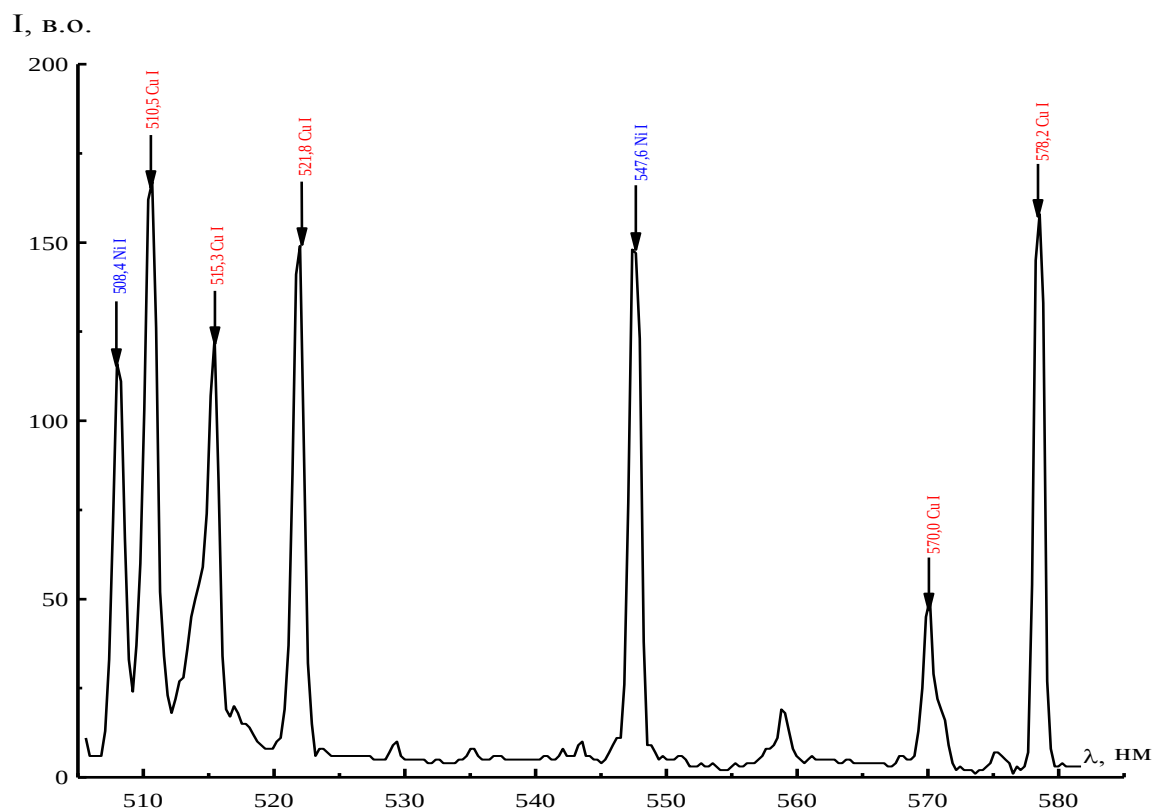


Рис. 2. Спектр випромінювання електродугового розряду між Cu-Ni електродами в повітрі.

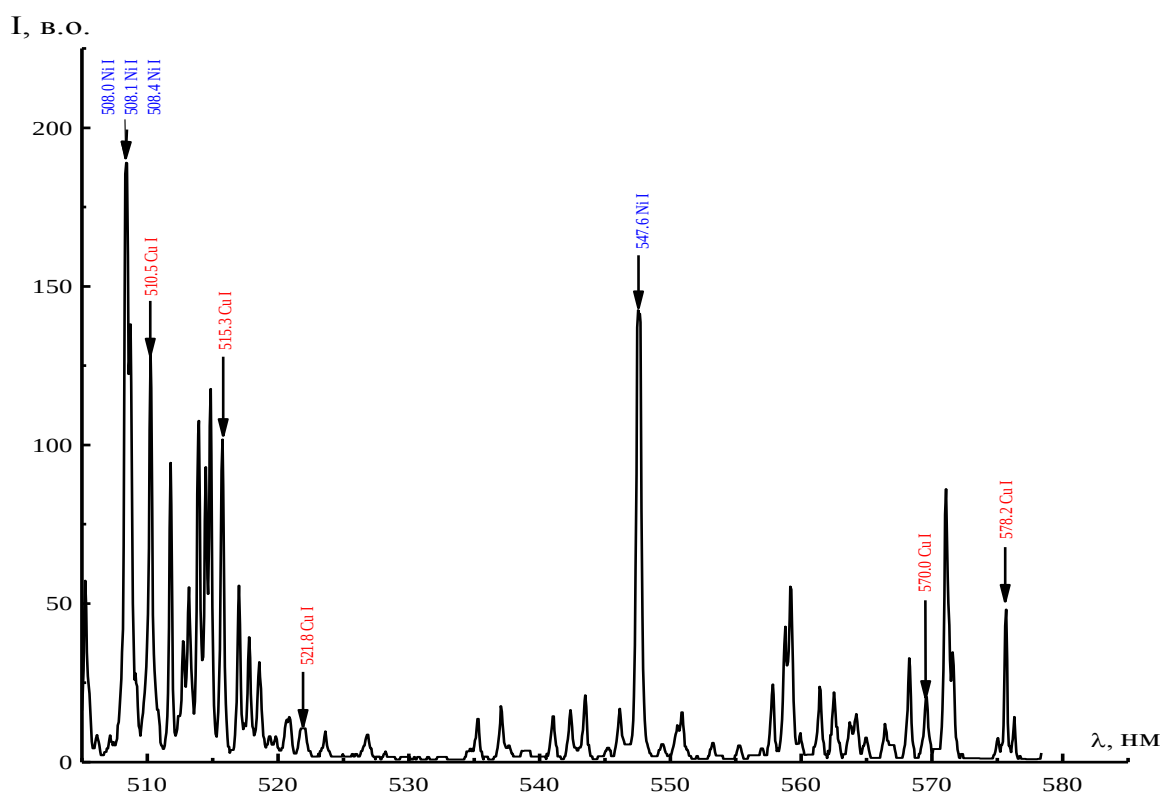


Рис. 3. Спектр випромінювання електродугового розряду між Cu-Ni електродами в потоці аргону.

Таблиця 1

Спектроскопічні дані Cu I та Ni I.

Елемент	λ , нм	Перехід $i \rightarrow j$	g_j	g_i	E_j , eB	E_i , eB	$g_j f_{ji}$ [10]	$g_j f_{ji}$ [11]	$g_j f_{ji}$ [5]
Cu I	510.5	$3d^9 4s^2 \rightarrow 3d^{10} 4p$	6	4	1.38	3.81	0.0312	0.02	0.0197
	515.3	$3d^{10} 4p \rightarrow 3d^{10} 4d$	2	4	3.78	6.19	0.96	1.9	1.6466
	521.8	$3d^{10} 4p \rightarrow 3d^{10} 4d$	4	6	3.81	6.19	1.84	2.4	1.9717
	570.0	$3d^9 4s^2 \rightarrow 3d^{10} 4p$	4	4	1.64	3.81	0.0048	0.0069	0.0057
	578.2	$3d^9 4s^2 \rightarrow 3d^{10} 4p$	4	2	1.64	3.78	0.01656	0.027	0.0130
Ni I	508.0	$3d^9(^2D)4p \rightarrow 3d^9(^2D_{5/2})4d$	11	9	3.65	6.09	1.35	8.0	
	508.1	$3d^9(2D)4p \rightarrow 3d^9(^2D_{3/2})4d$	9	7	3.84	6.28	1.96	6.7	
	508.4	$3d^9(^2D)4p \rightarrow 3d^9(^2D_{5/2})4d$	7	9	3.67	6.11	1.05	2.1	
	547.6	$3d^{10} \rightarrow 3d^9(^2D)4p$	1	3	1.82	4.08	0.13	0.21	

матриці перебуває в діапазоні довжин хвиль 500 – 600 нм. Для корекції нерівномірності її спектральної чутливості (в досліджуваному діапазоні 400 – 600 нм) використали еталонну вольфрамову лампу.

Додатковою проблемою у процесі реєстрації спектрів є недостатній динамічний діапазон цього типу матриці. З метою оптимального вимірювання інтенсивності спектральних ліній в кожному експерименті підбирали відповідний час експозиції. З цією ж метою, в разі потреби, використовували нейтральні світлофільтри.

Результати та обговорення

На першому етапі виконували спектроскопічні дослідження плазми вільно існуючого електродугового розряду у повітрі. На рис. 2 показано спектр випромінювання плазми, в якому чітко спостерігаються спектральні лінії як атома міді, так і атома нікелю. На цьому етапі також виконали попередню селекцію спектральних ліній та спектроскопічних констант атома нікелю.

На наступному етапі досліджень зареєстровано спектр плазми електродугового розряду в потоці аргону (рис. 3). Порівняння спектрів, представлених на рис. 2 та 3, дозволяє зробити висновок, що спектр випромінювання плазми електродугового розряду у потоці аргону

містить значно більшу кількість спектральних ліній атома нікелю. Цей ефект потребує подальшого, більш детального, дослідження.

Для визначення температури плазми (у припущенні ЛТР) використали метод відносних інтенсивностей із залученням спектральних ліній Ni I 508.4 та 547.6 нм.

Проте, дослідженнями встановлено, що поруч зі спектральною лінією 508.4 нм розташовані ще лінії Ni I 508.0 та 508.1 нм. Тобто, фактично зареєстрований просторовий розподіл інтенсивності випромінювання в околі 508 нм складається із трьох спектральних компонент. Зазначимо, що енергії збудження верхніх рівнів спектральних переходів досліджуваних ліній (508.0, 508.1, 508.4 нм) складають, відповідно, 6.09, 6.28, 6.11 еВ. Крім того, внесок випромінювання кожної спектральної лінії в результуючий просторовий розподіл інтенсивності є співвимірним. Тому, при визначенні радіального розподілу температури запропоновано використати сумарну інтенсивність цих ліній.

У розрахунках профілю температури використали середню енергію збудження зазначених трьох спектральних ліній, а добуток $g_j \cdot f_{ji}$ необхідно взяти як суму трьох доданків (табл. 1):

$$(g_j \cdot f_{ji})_{508.0} + (g_j \cdot f_{ji})_{508.1} + (g_j \cdot f_{ji})_{508.4} \quad (1)$$

Крім того, з метою перевірки доцільності використання такого підходу, при визначенні радіального розподілу температури плазми методом відносних інтенсивностей додатково залучили такі спектральні лінії атома міді – Cu I 510.5 та 515.3 нм (табл. 1). В досліджуваному режимі роботи дуги спектральні лінії Cu I 521.8, 570.0, 578.2 нм використати практично неможливо через їх слабку інтенсивність випромінювання.

На рис. 4 показано, одержаний із залученням спектральних ліній Ni I та Cu I, радіальний розподіл температури плазми в середньому поперечному перерізі електродугового розряду між Cu–Ni електродами силою струму 3.5 А в потоці аргону. В розрахунках температури використали спектроскопічні константи атома нікелю та міді із джерел [11] та [5], відповідно. Радіальні профілі температури, визначені із залученням спектральних ліній Cu I та Ni I, збігаються в межах похибки вимірювань.

Таким чином, запропонований у цьому дослідженні спосіб використання сумарної інтенсивності спектральних ліній Ni I 508.0, 508.1, 508.4 нм для визначення температури методом відносних інтенсивностей у поєднанні з лінією Ni I 547.6 нм можна рекомендувати для діагностики плазми з домішками нікелю.

Крім того, на основі аналізу одержаного розподілу температури (рис. 4), можна рекомендувати для подальшого використання у дослідженнях значень спектроскопічних констант атому нікелю саме із роботи [11] (на відміну від наведених в базі даних NIST [10]).

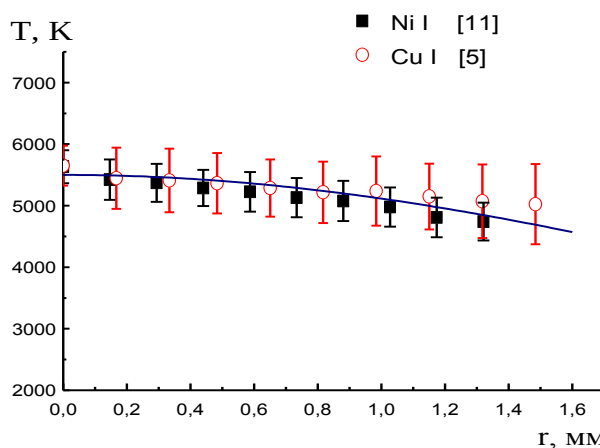


Рис. 4. Радіальний розподіл температури плазми електродугового розряду між Cu–Ni електродами силою струму 3.5 А із залученням спектральних ліній Ni I та Cu I.

Висновки

Спектроскопічними дослідженнями плазми електродугового розряду в потоці аргону між Cu–Ni електродами силою струму 3.5 А встановлено:

- якісну відмінність спектру випромінювання такого плазмового джерела у порівнянні із вільно існуючим розрядом у повітрі, а саме: його спектр містить значно більшу кількість спектральних ліній атома нікелю;

- можливість використання для діагностики плазми з домішками нікелю сумарної інтенсивності трьох спектральних ліній Ni I 508.0, 508.1, 508.4 нм для визначення температури методом відносних інтенсивностей у поєднанні з лінією Ni I 547.6 нм.

Список використаних джерел

1. Буткевич Г.В. Электрическая эрозия сильноточных контактов и электродов / Буткевич Г.В., Белкин Г.С. – Москва: Энергия, 1978, – 253 с.
2. Минакова Р.В. Тенденции развития технологий изготовления композиционных материалов и контактов из них // Минакова Р.В., Кресанова А.П., Чураков М.М., Хоменко Е.В. // Электрические контакты и электроды. – Киев: Ин-т пробл. Материаловедения НАН Украины. – 1998. – С. 5-19.
3. Minakova R.V. Contact interaction, W (Mo, Cr)-Cu structure and properties of composites with additives// Minakova R.V., Lesnik N.D., Kresanova A.P., Flis A.A., Khomenko E.V. //

References

1. BUTKEVICH G.V. (1978) *Electrical erosion of high-current contacts and electrodes*. Moscow: Energiya.
2. MINAKOVA R.V. KRESANOVA A.P., CHURAKOV M.M., KHOMENKO E.V. (1998) The development tendencies of manufacturing technologies of composite materials and their contacts. *Electric Contacts and Electrodes*. Kiev, Frantsevich Institute for Problems of Materials Science. p. 5-19.
3. MINAKOVA R.V., LESNIK N.D., KRESANOVA A.P., FLIS A.A., KHOMENKO E.V. (1996) Contact interaction, W (Mo, Cr)-Cu structure and properties of composites with additives. *Powder Metallurgy and Metal Ceramics*

- Powder Metallurgy and Metal Ceramics. – Kiev: Frantsevich Institute for Problems of Materials Science. – 1996. – **35**. – P.7-8.
4. Лохте-Хольтгревен В. Методы исследования плазмы. – Москва: Мир, 1971, – 552 с.
 5. Babich I.L. Spectroscopic data and Stark broadening of Cu I and Ag I spectral lines: Selection and analysis // Babich I.L., Boretskij V.F., Veklich A.N., Semenyshyn R.V. // *Advances in Space Research*. – 2014. – **54**. – P.1254-1263.
 6. Semenyshyn R.V.. Spectroscopy peculiarities of thermal plasma of electric arc discharge between electrodes with Zn admixtures // Semenyshyn R.V., Veklich A.N., Babich I.L., Boretskij V.F. // *Advances in Space Research*. – 2014. – **54**. – P.1235-1241.
 7. Fesenko S. Properties of thermal air plasma with admixing of copper and carbon // Fesenko S., Veklich A., Boretskij V., Cressault Y., Gleizes A., Teulet Ph. // *Journal of Physics: Conference Series*. – 2014. – **550**. – 012008 (8pp.).
 8. Veklich A.N. Investigations of thermal plasma with metal impurities. Part II: Peculiarities of spectroscopy by W I, Mo I, Cu I spectral lines// Veklich A.N., Lebid A.V., Soroka P.V., Boretskij V.F., Babich I.L. // *Problems of Atomic Science and Technology. Series: Plasma Physics*. – 2013. – №1. – **19**. – P.213-215.
 9. Bockasten. K. Transformation of observed radiances into radial distribution of the emission of a plasma // *Journ. of Opt. Soc. of Am.* – 1961. – **51**. – №9. – P.943-947.
 10. Kramida A. NIST. Atomic Spectra Database [База даних] / Kramida A., Ralchenko Yu., Reader J. and NIST ASD Team (2014) // Gaithersburg, MD: National Institute of Standards and Technology. – accessed 27.06.2015. <http://physics.nist.gov/asd>
 11. Корлисс Ч. Вероятности переходов и силы осцилляторов 70 элементов // Корлисс Ч., Бозман У. – Москва: Мир, 1968, – 552 с.
 - Kiev, Frantsevich Institute for Problems of Materials Science. 35. p.7-8.
 4. LOCHTE-HOLTGREVEN W. (1971) *Plasma diagnostics*. Moscow: Mir.
 5. BABICH I.L., BORETSKIJ V.F., VEKLICH A.N., SEMENYSHYN R.V. (2014) Spectroscopic data and Stark broadening of Cu I and Ag I spectral lines: Selection and analysis *Advances in Space Research*. 54. p.1254-1263.
 6. R.V. SEMENYSHYN. VEKLICH A.N., BABICH I.L., BORETSKIJ V.F. (2014) Spectroscopy peculiarities of thermal plasma of electric arc discharge between electrodes with Zn admixtures. *Advances in Space Research*. 54. p.1235-1241.
 7. FESENKO S., VEKLICH A., BORETSKIJ V., CRESSAULT Y., GLEIZES A., TEULET PH. (2014) Properties of thermal air plasma with admixing of copper and carbon. *Journal of Physics: Conference Series*. 550. 012008 (8pp.)
 8. VEKLICH A.N., LEBID A.V., SOROKA P.V., BORETSKIJ V.F., BABICH I.L. (2013) Investigations of thermal plasma with metal impurities. Part II: Peculiarities of spectroscopy by W I, Mo I, Cu I spectral lines. *Problems of Atomic Science and Technology. Series: Plasma Physics*. 19(1). p.213-215.
 9. K. BOCKASTEN. (1961) Transformation of observed radiances into radial distribution of the emission of a plasma. *Journ. of Opt. Soc. of Am.* 51(9). p.943-947.
 10. KRAMIDA A., RALCHENKO YU., READER J. AND NIST ASD TEAM NIST. (2014) *Atomic Spectra Database* [Online]. Gaithersburg, MD: National Institute of Standards and Technology. – accessed 28.06.2015. <http://physics.nist.gov/asd>
 11. CORLISS C.H., BOZMAN W.R. (1968) *The transition probabilities and oscillator strengths of 70 elements*. Moscow: Mir.

Надійшла до редколегії 30.06.15

УДК 621.383

Козинець О.В., к.ф.-м.н.

**Розрахунок напруги холостого ходу
кремнієвого сонячного елемента з
модифікованим шаром в області
просторового заряду**

Інститут високих технологій, Київський
національний університет імені Тараса
Шевченка, 03680, м. Київ, просп. Глушкова 4г,
e-mail: alk@univ.kiev.ua

A.V. Kozinetz, Ph.D.

**Calculation of open circuit voltage for silicon
solar cell with modified layer in space charge
region**

Institute of High Technologies, Taras Shevchenko
National University of Kyiv, 03680, Kyiv, Glushkova
St., 4g
e-mail: alk@univ.kiev.ua

Розраховано напругу холостого ходу кремнієвого сонячного елемента на основі асиметричного n^+-p переходу з модифікованим шаром в області просторового заряду. Розглянуто надбар'єрний та тунельний механізми переносу струму через модифікований шар. Для випадку надбар'єрного переносу зменшення струму насичення може бути співрозмірне із випадком тунельного переносу. Перший випадок є імовірним для модифікованих шарів більшої товщини. Показано що відносна зростання напруги холостого ходу складає близько десяти відсотків, якщо струм короткого замикання вихідної структури зберігається, принаймні, незмінним. В свою чергу більша товщина модифікованого шару дозволяє отримувати вищі значення додаткового струму внаслідок додаткового інфрачервоного поглинання.

Ключові слова: сонячний елемент, модифікований шар, асиметричний n^+-p перехід.

An open circuit voltage of n^+-p silicon solar cell with amorphized layer inserted in space charge region was calculated. Insertion of optical dopant into space charge region can be accompanied by partial amorphization of silicon. This leads to local energy band gap increase and potential barriers in conduction and valence bands formation. As efficiency of holes injection is very low for asymmetrical junctions, saturation current is defined by injection of electron in p-base completely. It was analyzed two principal cases of current flow through space charge region: first one corresponds to situation when electrons from emitter pass the barrier with assistance of thermionic emission, second one is tunneling. It was shown that an increase of open circuit voltage consists nearly ten per cent for both cases, if short circuit current remains, at least, unchanged. However, amorphized layer can change current flows of initial cell significantly and precise analysis of fill factor and photocurrent is needed to obtain ultimate result. The relative increase of open circuit current is comparable for both cases. In turn, larger modified layers allow more efficiently improve value of additional photocurrent due to infrared absorption.

Key Words: solar cell, modified layer, asymmetrical n^+-p junction.

Статтю представив д. ф.-м. н., проф. Скришевський В.А.

1. Вступ

Коефіцієнт корисної дії кремнієвих фотоперетворювачів на основі монокристалічного n^+-p переходу на сьогоднішній день не перевищує 16-18%. Дослідження можливостей підвищити коефіцієнт корисної дії є важливим напрямком сучасної фізики напівпровідникових приладів. Одним з перспективних підходів до розв'язання такої

проблеми є застосування домішкового фотовольтаїчного ефекта для використання квантів з енергією меншою ніж ширина забороненої зони монокристалічного кремнію 1.12 еВ. Суть цього підходу полягає у введенні в кремнієву ґратку "глибоких" домішок (як наприклад *In*, *Er*, *S*, *Te*), які створюють рівні всередині забороненої зони [3,4]. Зазвичай для цього застосовують іонну імплантацію з наступними циклами

термічного відпалу [4]. Позитивний ефект полягає у створенні можливості для двоступеневої фото генерації при поглинанні квантів з відповідними енергіями. Однак, в процесі іонної імплантації можлива небажана аморфізація монокристалічного матеріалу в певному об'ємі та, як наслідок, збільшення ширини забороненої зони до 1.6-1.8 еВ.

Можливий варіант ефективного застосування фотовольтаїчного ефекту є іонна імплантація в область просторового заряду. По-перше в такому випадку є можливість контролювати рекомбінаційні характеристики інтерфейсів аморфізованої області, по-друге інтенсивність небажаних рекомбінаційних процесів по типу Шоклі-Ріда-Хола через додатково створений глибокий рівень в забороненій зоні. Зазначимо, що параметри аморфізованого шару суттєво залежать від технологічних умов проведення іонної імплантації та характеристик відпалу. Наявність додаткової бар'єрної області може суттєво впливати на процеси переносу струму та основні характеристики фотоелектричного перетворення. Отже, необхідно проаналізувати особливості формування таких характеристик та визначити допустимі параметри аморфізованої області. В роботі [2] розглянуто особливості її впливу на формування густини струму короткого замикання. Розрахунок напруги холостого ходу, яка є іншим важливим параметром сонячного елемента, запропоновано в даній роботі. Для адекватного опису аморфізованої області з оптичними домішками використано теоретичну модель, яка передбачає наявність тонкої (50-100 нм) бар'єрної структури всередині області просторового заряду. Метою роботи є вивчення закономірностей формування напруги холостого ходу в системі із складною зонною структурою та визначення параметрів аморфізованої області, які забезпечують збільшення напруги холостого ходу за умови незмінного значення густини струму короткого замикання. Отримані результати можуть бути корисними для практичних досліджень подібних структур.

2. Теоретична модель

Для розрахунку використаємо діодну модель, в рамках якої припускають, що носії проходять область просторового заряду без розсіяння. Зазвичай концентрація донорних домішок в емітері на два-три порядки величини більша ніж акцепторної домішки в базовій області. Як наслідок, область просторового заряду виникає лише в базовій області а інжекція електронів з емітера визначає струм насичення такої структури. Припустимо, що в деякій точці x області просторового заряду асиметричного кремнієвого $n^+ - p$ переходу існує аморфізована структура товщиною d , яка створює потенціальний бар'єр в зоні провідності величиною ϕ_b (відлік енергії від рівня зони провідності в точці x_{AL}). Потенціал в області просторового заряду шириною L з висотою бар'єра ϕ_0 можна апроксимувати виразом

$$\varphi = \varphi_0 \left(1 - \frac{x}{L}\right)^2.$$

Струм на межі поділу область просторового заряду-нейтральна область бази можна подати або як градієнт концентрації неосновних носіїв, або як різницю потоків електронів з емітерної та базової областей (з урахуванням параметрів аморфізованого шару та механізму переносу струму). Прирівнявши ці величини, можна знайти розподіл концентрації електронів в базовій області та густину темного струму [2]:

$$j(V) = (j_0 / (1 + j_0 / f)) (\exp(eV / kT) - 1), \quad (1)$$

де j_0 - струм насичення вихідної структури, $f_n = \frac{1}{4} e v_n n_n \exp(-(\phi + \phi_b + eV_2) / kT)$ - для надбар'єрного $f_n = \frac{1}{4} e v_n n_n P_n(V) \exp(-\phi_0 / kT)$ - для тунельного механізму. В цих виразах $\phi = \phi_0 - \phi(x_{AL})$, V_2 - частина напруги зправа від аморфізованого шару, $P_n(V)$ - коефіцієнт тунельної прозорості, v_n - теплова швидкості електронів, n_n - рівноважна концентрація електронів в емітері, k - стала Больцмана, T - температура, e - заряд електрона.

Напругу холостого ходу можна оцінити за відомою формулою [1]

$$V_{xx} = nkTj_{kз} / j_{of} \ln(j_{kз} / j_{of} + 1), \quad (2)$$

де $j_{of} = j_o / (1 + j_o / f)$ - ефективна густина струму насичення, $j_{kз}$ - густина струму короткого замикання, n -коефіцієнт не ідеальності.

3. Результати та їх обговорення

Струм насичення базової області більше ніж струм насичення емітерної області на три порядки величини для асиметричного n^+-p переходу. Як випливає з аналізу формули (2), струм насичення вихідної структури зменшується у випадку створення додаткової бар'єрної структури в області просторового заряду. Наступні розрахунки виконаємо для типових параметрів кремнієвих фотоперетворювачів: $j_{kз} = 33 \text{ мА/см}^2$, $N_d = 10^{19} \text{ см}^{-3}$, $N_A = 10^{16} \text{ см}^{-3}$, $D_n = 30 \text{ см}^2/\text{с}$, $D_p = 13 \text{ см}^2/\text{с}$, $v_p = v_n = 10^7 \text{ см/с}$, $\phi_0 = 0.8 \text{ еВ}$, $L = 0.5 \text{ мкм}$. Ефективна висота бар'єра $\phi_f = \phi_b - \phi$ вибрана для параметризації розрахункових кривих. Рис. 1 ілюструє випадок тунелювання електронів через аморфізований шар. Як відомо, такий механізм може бути актуальним у випадку її невеликих товщин коли $d < 10 \text{ нм}$. Збільшення товщини шару суттєво зменшує струм насичення вже для невеликих $d = 3-5 \text{ нм}$. Як показують розрахунки, у випадку незмінної густини

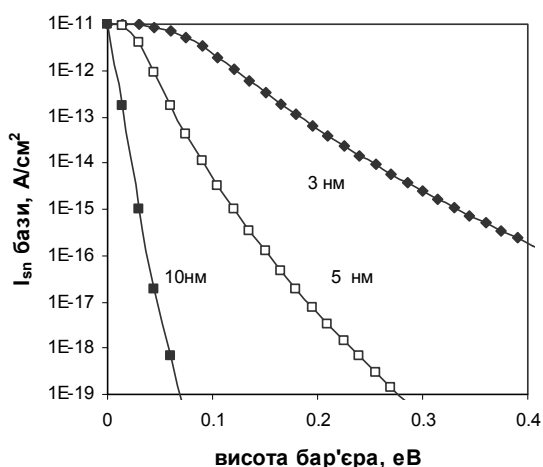


Рис. 1. Залежність густини струму насичення бази від ефективного бар'єра для тунельного механізму

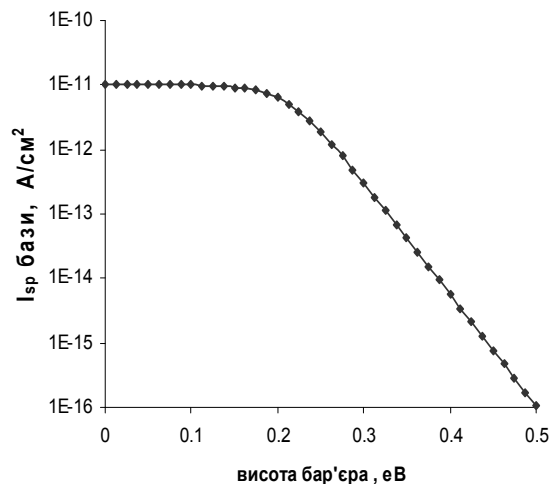


Рис. 2 Залежність густини струму насичення бази від ефективного бар'єра для термоемісійного механізму

струму короткого замикання, зростання напруги холостого ходу може складати 10-15% від початкової величини. У випадку зростання товщини аморфізованого шару імовірність термоемісійного механізму збільшується. Причому для фіксованого значення x_{AL} , чим більша величина бар'єра ϕ_b , тим більш суттєво зменшується струм насичення вихідної структури.

Навпаки, якщо ϕ_b фіксовано, більші x_{AL}

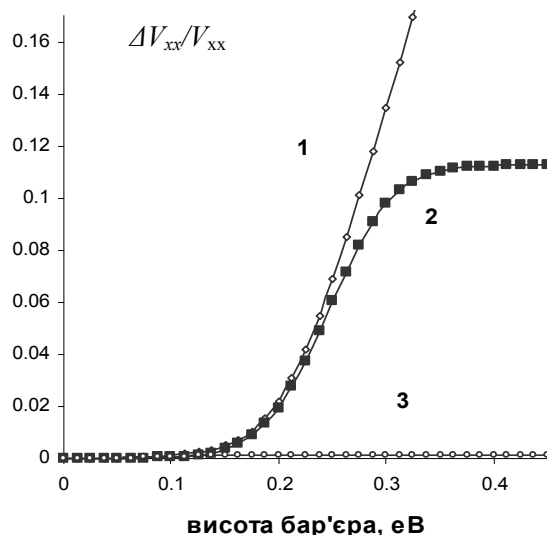


Рис. 3 Відносна зміна напруги холостого ходу від висоти ефективного бар'єра у випадку інжекції електронів та дірок (1), тільки електронів (2), тільки дірок (3).

(аморфізований шар розміщено ближче до границі область просторового заряду-нейтральна область бази) відповідають меншим значенням струму насичення. Рис.2 ілюструє вплив величини ϕ_f на струм насичення базової області. Розрахунки свідчать що і у випадку термоемісійного механізму відносно зростання напруги холостого ходу може складати 10-15%, Рис.3. Плато в області величини параметра $\phi_f > 0.3$ еВ (крива (2), рис.3) відповідає випадку коли струм насичення бази зменшується на два порядки величини і його величина стає співмірним із струмом насичення емітера. Якщо аморфізований шар створює потенціальний бар'єр у валентній зоні аналогічні міркування можуть бути застосовані. Обмеження інжекції дірок в емітер без зміни умов для переносу електронів через область просторового заряду не призводить до суттєвої змін напруги холостого ходу, крива 3. Якщо аморфізований шар створює одночасно бар'єри у зоні провідності та у валентній зоні, напруга холостого ходу збільшується внаслідок впливу

аморфізованого шару на обидві компоненти струму насичення. Слід зазначити, що припущення незмінності густини струму короткого замикання є ідеалізованим наближенням, оскільки фізична природа впливу бар'єрного шару на світлові та темнові струми однакова і необхідні додаткові числові розрахунки в рамках самоузгодженої моделі. Зазначимо що шари більшої товщини дозволяють отримувати більшу величину додаткового фотоструму в області енергій квантів менше ніж 1.12 еВ.

4. Висновки

На основі діодної моделі розраховано напругу холостого ходу кремнієвого фотоперетворювача з аморфізованим шаром в області просторового заряду. Розглянуто термоемісійний та тунельний механізми переносу струму через аморфізований шар. Показано що збільшення напруги холостого ходу може складати 10-12% в обох випадках, якщо густина струму короткого замикання залишається, принаймні, незмінною.

Список використаних джерел

1. Sze S. M. Physics of semiconductor device.- John Wiley & Sons, Inc.: New York, 1969-456 p.
2. Kozinetz A. V. Physical properties of silicon photodetectors with inserted delta- and psi- layers.-Manuscript. Thesis for a physical and mathematical sciences candidate's degree-2007-153 p.
3. Kasai H Study for improvement of solar cell efficiency by impurity photovoltaic effect/ H. Kasai, H. Matsumura // Solar Energy Materials & Solar Cells. – 1997. – 48. – P. 93–100.
4. Ley M. Experimental and theoretical investigation of a new potential barriers due to sharp α -Si-c-Si heterointerfaces buried in the solar cell emitter/ M. Ley, Z.T. Kuznicki // Solar Energy Materials & Solar Cells. – 2002. – 72. – P. 613–619.

References

1. SZE S. (1969) *Physics of semiconductor device*. John Wiley & Sons, Inc.: New Jersey.
2. KOZINETZ A. (2007) Physical properties of silicon photodetectors with inserted delta- and psi- layers.- Manuscript. Thesis for candidate's degree, 153 p.
3. KASAI, H. MATSAMJURA H (1997) Study for improvement of solar cell efficiency by impurity photovoltaic effect. *Solar Energy Materials & Solar Cells*. 48. p. 93–100.
4. LEY, M., KUZNICKI, Z. T. (2002) Experimental investigation of a new potential barriers due to sharp α -Si/c-Si heterointerfaces buried in the solar cell emitter *Solar Energy Materials & Solar Cells*. 48. p. 613–619.

Надійшла до редколегії 10.07.15

УДК 681.3

Ляковський В.В.¹, студент
Баужа О.С.², к.ф.-м.н., асистент

**Розробка макету мобільного
уніфікованого комплексу реєстрації
радіаційного випромінювання**

Київський національний університет імені
Тараса Шевченка, 83000, м. Київ, пр-т.
Глушкова 4г,

¹ e-mail: lwv2006@gmail.com

² e-mail: asb@mail.univ.kiev.ua

V. V. Liaskovskyi¹, degree,
O.S. Bauzha², PhD.

**Development of model unified
mobile complex for registration of radioactive
emissions**

Taras Shevchenko National University of Kyiv,
83000, Kyiv, Glushkova st., 4g,

¹ e-mail: lwv2006@gmail.com

² e-mail: asb@mail.univ.kiev.ua

В роботі пропонується метод об'єднання в автоматизований вимірювальний комплекс лічильників Гейгера-Мюллера з мобільними пристроями (ноутбуком, телефоном, планшетом). Результатом роботи є розробка універсального комплексу реєстрації радіаційного випромінювання. Даний комплекс може фіксувати дозу випромінювання небезпечної рівня для людини рівня та сигналізувати про ступені небезпеки.

Ключові слова: дозиметр, лічильник Гейгера-Мюллера, Bluetooth, Arduino

A method of combination of the Geiger-Muller counters with the mobile devices (a laptop, a smartphone, a tablet computer) into the automated measuring system is offered in this work. The result of the work is the development of a universal set of registration of radiation based on the microcontroller Atmel Atmega 328. This set can register the dose of gamma radiation levels dangerous for humans and can signal the degrees of danger. The device also has a built-in counter of the accumulated dose. The adjustment and work with the device occurs with the help of a wireless interface Bluetooth, or with the help of a Universal Serial Bus USB. Bluetooth interface realized with the help of microcircuit BlueCore 417 and USB with the help of CH340G. It is not necessary to use a special set software for the measuring set for the work of a mobile device. The set is compact, portable, has an integrated power supply and can operate without a mobile device as a signal of the dangerous zones. The built-in flashing alarm system and audible alarm system are implemented in the device. This device can be used in the industry, for work on the facilities with high radiation background, and at home, because it is not expensive and it is also affordable.

Key Words: dosimeter, Geiger-Muller counter, Bluetooth, Arduino.

Статтю представив д.т.н. Погорілий С.Д.

Вступ

В наш час, після катастрофи на Чорнобильській АЕС вимірювання зовнішнього радіаційного фону є актуальною метрологічною задачею, адже велика кількість забруднених речовин потрапляє із зони відчуження до решти території України [1,4]. Також, на території України є велика кількість об'єктів, які несуть у собі потенційну загрозу для радіаційної безпеки.

Радіаційна безпека - це забезпечення захисту від іонізуючого опромінення окремих осіб, їх нащадків і людства в цілому. З іншого боку - це необхідність створення відповідних умов для

нормальної повсякденної діяльності людини, під час якої люди можуть потрапляти під дію іонізуючого випромінювання [1,3,7]. Дозиметричний контроль територій проводиться для своєчасного отримання даних про дози опромінення людей та ступінь зараження місцевості, будинків, об'єктів народного господарства, техніки тощо, для вживання своєчасних заходів щодо зменшення небезпеки радіаційного ураження [1,7].

Проникаюча радіація небезпечна своїми наслідками для здоров'я людини. Маючи велику енергію, радіоактивні промені глибоко

приникають в тканини організму людини, іонізують їх та збуджують променеву хворобу.

В статті пропонується створення нового універсального приладу для проведення регулярного радіологічного контролю навколишнього середовища. Необхідно створити універсальний прилад, який дозволяв би вести радіологічний контроль.

Універсальність полягає в наступному:

1. Прилад об'єднує в собі функції дозиметра, радіометра та сигналізатора небезпечних зон.

2. Прилад має широкий діапазон вимірювання, а отже повинен застосовуватись водночас на промислових об'єктах з високим радіаційним фоном і в побуті.

3. Прилад повинен мати низьку собівартість, малі розміри, невелику вагу.

Основні принципи роботи приладу

Радіаційний контроль – це комплекс заходів по дотриманню норм радіаційної безпеки й основних санітарних правил роботи з радіоактивними речовинами та іншими джерелами іонізуючого випромінювання. Основним документом, який регулює радіаційний контроль є Закон України «Про захист людини від впливу іонізуючого випромінювання» від 14 січня 1998 року №15/98-ВР, ст. 9 та 15).

З гігієнічної точки зору, своєчасно проведений радіаційний контроль та усунення

чи зниження до нормативного рівня іонізуючого випромінювання захистить здоров'я, так як знаходження тривалий час у приміщеннях з підвищеним радіаційним фоном чи рівнем радону в повітрі може привести до негативного впливу на стан здоров'я людей.

Найбільш поширеним методом радіаційного контролю є використання дозиметрів, які ґрунтуються на лічильниках Гейгера-Мюллера. Є багато реалізацій лічильників Гейгера-Мюллера, але найбільшого поширення на території СНД здобув лічильник СБМ-20.

Причиною широко використання є простота виробництва та невелика вартість, що дозволяє використовувати даний лічильник водночас в побутових і професійних приладах радіаційного контролю. СБМ-20 має габарити трохи більше за олівець, що дозволяє створювати на основі нього як великі стаціонарні, так і портативні дозиметри. Дозиметрам називають сімейство приладів фіксації потужності радіоактивного випромінювання, вимірювальний прилад для вимірювання дози або потужності дози іонізуючого випромінювання отриманої приладом за деякий проміжок часу, наприклад, за період перебування на деякій території або за робочу зміну. Вимірювання вищезгаданих величин називається дозиметрією.

Компонентна база

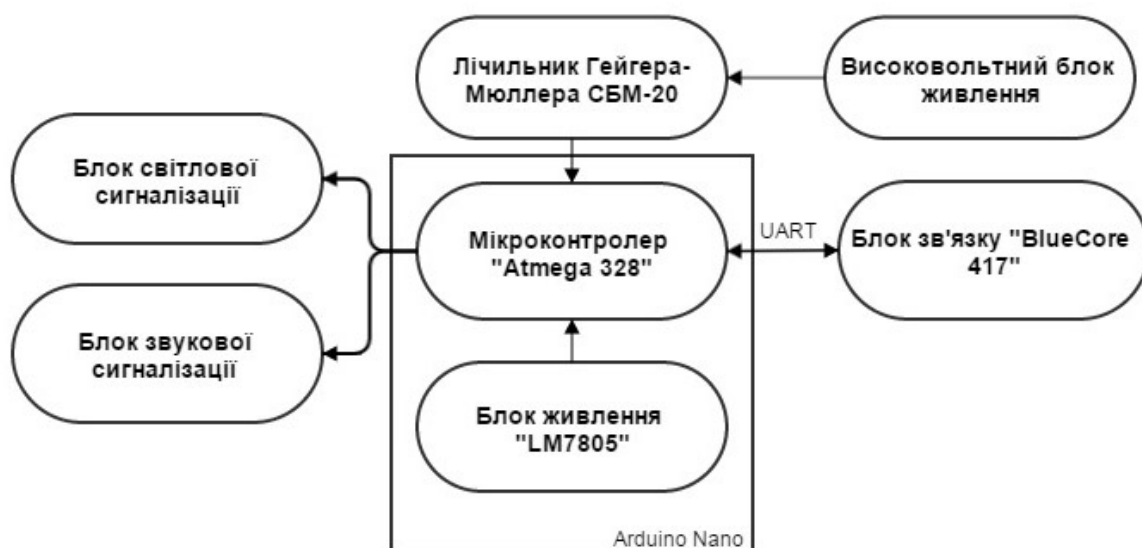


Рис. 1 Блок-схема дозиметричного комплексу

Виходячи з поставленої задачі та представлених на ринку датчиків та мікроконтролерів, було підібрано такі компоненти:

- Мікроконтролер Arduino Nano на процесорі Atmel Atmega328 [5].
- Лічильник Гейгера-Мюллера СБМ-20;
- Модуль зв'язку BC 417 [6];

Принципова схема пристрою зображена на рис.1.

Мікроконтролери сімейства Arduino [5] є мініатюрною апаратно-обчислювальною платформою, оснащеною процесором, оперативною пам'яттю, енергонезалежною флеш-пам'яттю та інтерфейсом вводу-виводу. Програмне забезпечення для цієї платформи може бути розроблене мовами Асемблер або С-подібною мовою Processing/Wiring. Сімейство мікроконтролерів Arduino може використовуватися або створення автономних інтерактивних об'єктів, або для комунікаційного з'єднання до програмного забезпечення, що працює на інших комп'ютерах локальної мережі, зокрема Adobe Flash, Processing, Max/MSP, Pure Data, SuperCollider тощо. Інформація про схемотехнічну реалізацію мікроконтролерних систем Arduino знаходиться у відкритому доступі..

Мікроконтролер Arduino Nano складається з процесора фірми Atmel архітектури AVR, а також елементів обв'язки для програмування та інтеграції з іншими пристроями. На багатьох платах наявний лінійний стабілізатор напруги +5В та +3,3В. Тактування здійснюється на частоті 16 або 8 МГц за допомогою зовнішнього кварцового резонатора. У мікроконтролер записано завантажувач (bootloader), тому для зміни програмного забезпечення мікроконтролера, використання зовнішнього програматора не є обов'язковим.

Плати Arduino дозволяють використовувати значну кількість I/O виводів мікроконтролера у зовнішніх схемах. У платі Arduino Nano доступно 14 цифрових входів/виходів, 6 із яких можуть видавати широко-імпульсний сигнал, і 8 аналогових входів.

Лічильники Гейгера-Мюллера СБМ-20 представляють із себе сталеву трубку наповнену спеціальною газовою сумішшю, довжиною

близько 10 сантиметрів і діаметром близько одного сантиметра. З двох сторін лічильник Гейгера СБМ-20 має контакти під цокольне з'єднання.

Лічильники Гейгера-Мюллера СБМ-20 (Рис.1) широко використовувалися в популярних в СРСР дозиметрах побутового та військового призначення.

Лічильник СБМ-20 виробляється з кінця 70х років в без істотних змін конструкції і характеристик.

BC417 [6] однокристальна радіосмугова інтегральна мікросхема яка реалізує протокол Bluetooth. Дана мікросхема підтримує підвищену швидкість передачі даних (EDR), що забезпечує передачу даних на швидкості до 3 Мбіт. Дозволяє оперативно з'єднувати пристрій з будь-яким мобільним телефоном, планшетним комп'ютером, або ноутбуком за допомогою протоколу Bluetooth. BC417 (Рис.1) має інтерфейс зовнішньої Flash пам'яті, зі швидкістю до 8Mbit. При використанні з Bluetooth КСВ програмний стек забезпечує повну сумісний Bluetooth специфікації 2.0 для передачі даних і голосу. BC417 може бути використаний як конвертер RS232-to-Bluetooth.

До мікроконтролера Atmega за допомогою цифрових входів під'єднано лічильник Гейгера-Мюллера СБМ-20 (рис.1). При кожному спрацюванні лічильника, відбувається переривання на мікроконтролері.

Живлення мікроконтролера забезпечує мікросхема LM7805, яка з обв'язкою, забезпечує 5 Вольт на виході. Цієї напруги достатньо для живлення всього приладу. Виключенням є лічильник Гейгера-Мюллера СБМ-20. Даний лічильник входить в стан детектування при напрузі живлення в межах від 260 до 320 вольт.

Для живлення лічильника застосовується високовольтний блок живлення (Рис.2), який видає потрібну напругу при 6 вольтах на вході. Схема високовольтного блоку живлення розроблена ПО «Горизонт» та опублікована в вільному доступі.

На даній схемі зображено три резистори. R1 та R3 по 100 кОм, а R2 – 5,1 МОм. Конденсатори C3 та C4 мають ємність 4700 пФ, C1 – 50 мкФ, а C2 – 0.1 мкФ. В схемі присутні діоди D1 – D2 серії КД102Б та діод D3 серії КД522. Транзистор Q1 серії КТ315Г.

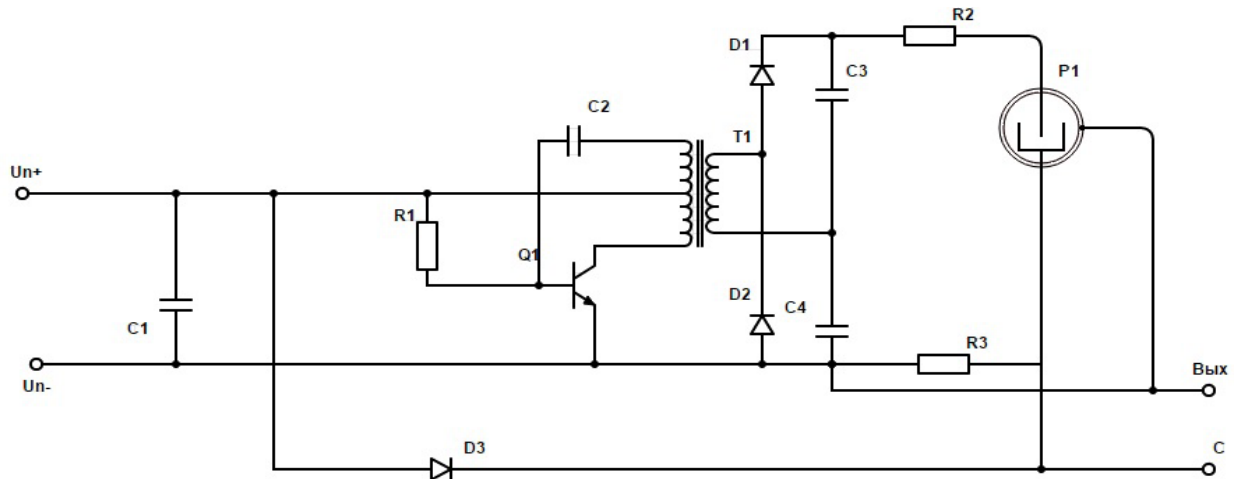


Рис. 2 Принципова схема високовольтного блоку живлення

Трансформатор є основою даного блоку живлення і складається з двох обмоток. Перша обмотка складається із двох блоків по 9 витків, а друга – з 400 витків. В трансформаторі використовується магнітопровід М2000НМ. Р1 позначає лічильник Гейгера-Мюллера СБМ-20.

В приладі присутні блоки звукової та світлової сигналізації (рис.1). Вони призначені

для попередження користувача про потенційну небезпеку. Світлова та звукова сигналізація вмикається при перевищенні поточного порогу значення радіаційного фону, або якщо накопичена доза перевищує задану за день норму.

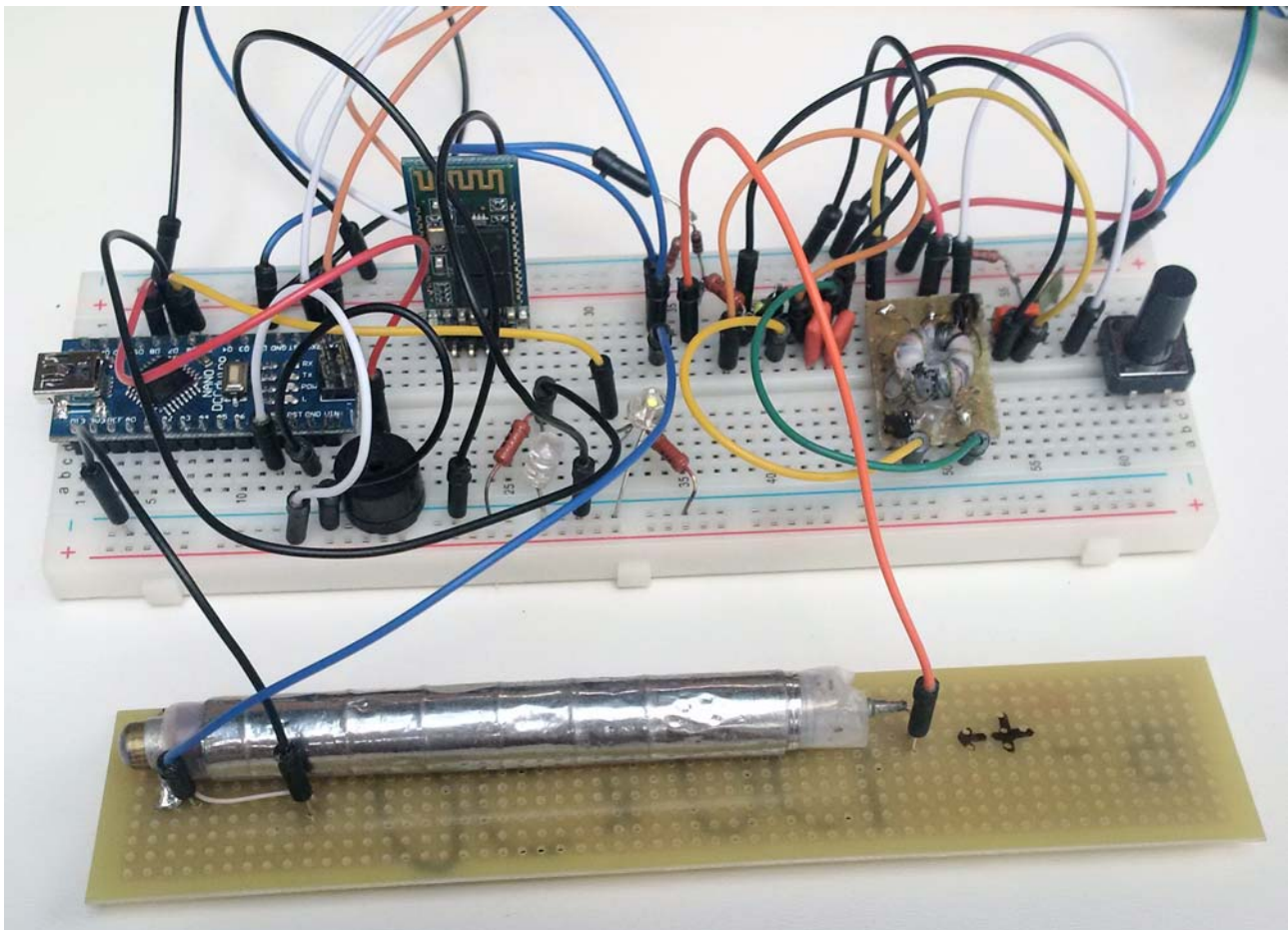


Рис. 3 Стендовий експериментальний зразок дозиметричного комплексу

Є два пороги сигналізування: попередження та тривога. Попередження спрацьовує, якщо поточне значення радіаційного фону наближається до критичного значення, але не досягає його. У випадку подальшого росту показника поточного радіаційного фону, вмикається режим тривоги, який продовжує працювати, поки користувач не покине небезпечну зону.

Пороги попередження та тривоги, як для поточного радіаційного фону, так і для накопиченої дози, користувач може встановити за допомогою терміналу.

Модуль зв'язку BC 417 (Рис.1), дозволяє під'єднатися до пристрою за допомогою протоколу Bluetooth. Сам модуль зв'язку під'єднано до мікроконтролера за допомогою інтерфейсу UART [2,6].

Модуль BC 417 дозволяє використовувати будь-який пристрій, який підтримує протокол Bluetooth для перегляду даних з приладу та налаштування його. В якості терміналів до дозиметра можуть бути використані, як персональні комп'ютери на системах (Windows, Linux, Mac OS), так і планшетні ПК та смартфони (Windows Mobile, Windows Phone, Android, iOS, BlackBerry OS).

Для наглядної демонстрації простоти схеми, на Рис. 3 зображено прототип дозиметричного комплексу, зібраного на платі безпаячного монтажу.

Розроблений пристрій є простим у повторенні, адже було використано популярні та дешеві компоненти. Вартість одного приладу при малосерійному виготовленні близька до 14-15 \$.

Прилад можна живити, як від акумуляторів, батарейок, так і від зовнішнього джерела живлення з напругою від 6 – 18 вольт. Габарити пристрою є співрозмірними з пультом до побутової техніки, у випадку збирання на базі одношарової односторонньої плати.

Висновки:

Важкий багаж технологічних та соціальних проблем, який успадкувала Україна в результаті

аварії на ЧАЕС, а також нещодавня аварія на ядерному об'єкті в японській префектурі Фукусіма дають наочні докази того, що своєчасна розробка і впровадження високоточних, портативних і доступних для широкого використання індивідуальних засобів радіаційного контролю є важливим і актуальним питанням для будь-якої розвиненої країни з ядерною енергетикою.

Серед різних проаналізованих методів реєстрації радіаційного забруднення перевагу надано використанню класичного лічильника Гейгера-Мюллера високоенергетичних елементарних частинок моделі СБМ-20 українського виробництва як інтегруючого датчика для побудованого приладу контролю радіаційної безпеки.

Як апаратний обчислювальний пристрій обрано яскравого представника моделі Arduino Nano популярного сімейства сучасних мікроконтролерів, оснащених процесорами корпорації Atmel з власною архітектурою команд AVR через мініатюрні розміри та низьке енергоспоживання цієї моделі.

Оскільки прилад контролю радіаційної безпеки має вбудований емулятор терміналу з власною системою керуючих команд, то цифровий зв'язок з приладом організовано за допомогою універсального асинхронного прийомопередавача UART, для якого фізичним передаючим середовищем є поширений інтерфейс Bluetooth. Це дає можливість практично будь-якому мобільному пристрою або стаціонарному комп'ютеру, оснащеному Bluetooth-інтерфейсом з'єднатись з приладом і виконати керуючі команди.

Даний прилад може застосовуватися як в промисловості, для роботи на об'єктах з високим радіаційним фоном, так і в побуті, адже є не дорогим та доступним.

Список використаних джерел

1. Ильин Л.А. Радиационная безопасность и защита: Справочник / Л.А. Ильин, В.Ф. Кириллов, И.П. Коренков - Москва: Медицина, 1996. – 336с.

References

1. ILYIN L., KIRILLOV V. and KORENKOV I. (1996) *Radiacionnaya bezopasnost' i zaschita: spravochnik* Moskva: Medicina.

2. Гольдберг Л.М. Цифровая обработка сигналов. [изд. 2-е перераб. и доп.] / Л.М. Гольдберг, Б.Д. Матюшкин, М.Н. Поляк – Москва.: Радио и связь, 1990. – 256 с.
3. Шарапов В.М. Датчики: Справочное пособие / В.М. Шарапов, Е.С. Полищук, Н.Д. Кошевой, Г.Г. Ишанин, И.Г. Минаев, А.С. Совлукон – Москва.: Техносфера, 2012. – 624 с.
4. Гвай А.С. Современные методы и средства дозиметрии ионизирующих излучений в медицине / А.С. Гвай, Л.А. Аверьянова, О.Ю. Шалёпа // Східно-Європейський журнал передових технологій. – 2012. 57, №9. – С. 40-44.
5. ATMEL 8-BIT MICROCONTROLLER WITH 4/8/16/32KBYTES IN-SYSTEM PROGRAMMABLE FLASH [документація, електронний ресурс] – 2014, 650p. – доступ до документації: – http://www.atmel.com/images/Atmel-8271-8-bit-AVR-Microcontroller-ATmega48A-48PA-88A-88PA-168A-168PA-328-328P_datasheet_Complete.pdf
6. Blue CoreTM4-External, Single Chip Bluetooth® v2.0+EDR System Production Information Data Sheet For BC417143B-IQN-E4 BC417143B-IRN-E4 [документація, електронний ресурс] – 2005, 116 p. – доступ до документації: <https://cdn.sparkfun.com/datasheets/Wireless/Bluetooth/CSR-BC417-datasheet.pdf>
7. Закон України. Про використання ядерної енергії та радіаційну безпеку. Закон від 08.02.1995 № 39/95-ВР- [електронний ресурс] доступ до закону: <http://zakon4.rada.gov.ua/laws/show/39/95-вр>
2. HOLNDBERH L., MATYUSHKINA B. and POLYAK M. (1990) *Cifrovaya obrabotka signalov*. 2 Ed. Moskva: Radio i svyaz'.
3. SHARAPOV V., POLYSCHUK E., KOSHEVOY N., YSHANYN G., MINAEV I. and SOVLUKOV A. (2012) *Datchiki: Spravochnoe posobie*. Moskva: Tehnosfera.
4. HVAY A., AVERYANOVA L. & SHALYEPА O. (2012) *Sovremennye metody i sredstva dozimetrii ionisuyuschih szluchenij v medisine*. *Eastern-European Journal of Enterprise Technologies* 57 (9). p. 40-44.
5. ATMEL 8-BIT MICROCONTROLLER WITH 4/8/16/32KBYTES IN-SYSTEM PROGRAMMABLE FLASH [Datasheet, online] –2014, 650p. – Available from: http://www.atmel.com/images/Atmel-8271-8-bit-AVR-Microcontroller-ATmega48A-48PA-88A-88PA-168A-168PA-328-328P_datasheet_Complete.pdf
6. Blue CoreTM4-External, Single Chip Bluetooth® v2.0+EDR System Production Information Data Sheet For BC417143B-IQN-E4 BC417143B-IRN-E4 [Datasheet, online] –2005, 116 p. – Available from: <https://cdn.sparkfun.com/datasheets/Wireless/Bluetooth/CSR-BC417-datasheet.pdf>
7. Zakon Ukrainy. Pro Viktoristannya Yadernoy energii ta radiaciynu bezpeku. Zakon vid 08.02.1995 № 39/95 VR [Online] Available from: <http://zakon4.rada.gov.ua/laws/show/39/95-вр>

Надійшла до редколегії 05.06.15

УДК 53.088

Манілов А.І., к.ф.-м.н.

**Вдосконалення математичної обробки
фотоелектричного відгуку на
модульоване освітлення**

Інститут високих технологій, Київський
національний університет імені Тараса
Шевченка, 03680, м. Київ, просп. Глушкова 4г,
e-mail: anmanilov@univ.kiev.ua

A.I. Manilov, Ph.D.

**Improved mathematical processing of
photoelectric response on modulated
illumination**

Institute of High Technologies, Taras Shevchenko
National University of Kyiv, 03680, Kyiv, Glushkova
St., 4g
e-mail: anmanilov@univ.kiev.ua

Розроблено вдосконалений метод математичної обробки фотоелектричного відгуку сонячного елемента на модульоване освітлення. Розглянуто два варіанти наближень, подібності елементів і зсуву вольт-амперної характеристики, та розширення відповідних обчислювальних схем на випадок довільного фотоелектричного модуля. Запропоновані вдосконалення дозволяють тестувати сонячні батареї з неоднаковими параметрами компонентів.

Ключові слова: сонячний елемент, фотоелектричний сигнал, тестування.

The improved mathematical processing of photoelectric signal obtained from solar cell under modulated illumination is proposed. Suggested method extends the scope of the original testing approach, which basic assumption is constant proportionality coefficient between the photoelectric response and derivative of current-voltage characteristics of the element. So, the more flexible assumption methods are proposed. The first method is based on supposition that the solar module elements have identical parameters but different responses for different values of current. The second method requires a fixed bias of current at light, but allows for differences in the components of the device. The corresponding modifications of calculation schemes for current-voltage characteristics restoration are listed. The advantages and disadvantages of used calculation approaches are discussed. It is possible to combine these methods in the algorithm, which does not require any of the previous assumptions. This combination provides a tool for research and diagnostics of semiconductor solar cells with heterogeneous parameters.

Key Words: solar cell, photoelectric signal, testing.

Статтю представив д. ф.-м. н., проф. Скришевський В.А.

1. Вступ

Дослідження сонячних елементів займає помітне місце у сучасній фізиці. На сьогоднішній день напівпровідникова промисловість виробляє кремнієві фотоелектричні модулі з коефіцієнтом корисної дії біля 18-19% для публічного використання. Найкращі показники ефективності сучасних пристроїв сягають 25% [1-2]. При цьому, ціна на сонячні батареї знизилась на 70% протягом останніх 5 років. Системи на основі напівпровідникових модулів демонструють автономність та здатність повністю забезпечити потреби споживача у енергії [3-4].

Важливим напрямком досліджень сонячних елементів є розробка засобів їх тестування. Класичними способами тестування фоточутливих пристроїв є вимірювання їх вольт-амперних характеристик (ВАХ) та реєстрація спектральних залежностей наведеного фотоелектричного сигналу [1-2]. Вони дозволяють визначити експлуатаційні показники пристроїв. Ці методи постійно вдосконалюються новими особливостями в освітленні, попередній обробці зразків та конструкції приладів [5].

Недоліком класичних способів є здатність тестування лише модуля в цілому, а не його

компонент. Сучасні сонячні батареї складаються зі зв'язаних між собою фотоелектричних елементів, кожен з яких робить свій внесок у загальну ефективність пристрою. Відповідно, постають задачі визначення характеристик окремих елементів, виявлення дефектних складових.

Одним з перспективних засобів тестування є метод діагностики сонячних елементів у модулях шляхом вимірювання електричного відгуку на модульованому освітленні [6]. Він дозволяє отримати ВАХ складових пристрою без необхідності у його руйнуванні чи створенні додаткових контактів. Такий підхід має ряд обмежень, покладених в основу обробки результатів вимірювань. Область застосування методу можна розширити за рахунок вдосконалення математичної обробки даних [7]. У статті запропоновано шляхи вирішення цієї проблеми.

2. Вихідний метод

2.1. Засади методу

Вихідний метод полягає у вимірюванні і наступній математичній обробці малого сигналу, наведеного при освітленні кожного елемента сонячної батареї модульованим світлом [6].

Зразок підключається до електричного кола, в якому реалізується гальваностатичний режим. За таких умов опромінення слабким модульованим світлом призводить до генерації змінної напруги, необхідної для того, щоб наскрізний струм залишився незмінним.

Модульоване світло малої інтенсивності фокусується на окремому n -му елементі модуля. Відповідно, фотоелектричний сигнал δv від освітленого елемента виділяється на частоті модуляції. Цей фотоелектричний сигнал є пропорційним похідній ВАХ на елементі:

$$\delta v(l, n) = K \cdot (dv/dl)_n \quad (1)$$

Відгук реєструється для кожного елемента при різних значеннях струму. Коефіцієнт пропорційності K розраховується з

експериментальних значень фотоелектричного сигналу та похідної загальної ВАХ модуля на основі адитивності характеристик елементів. Відновлення ВАХ n -го елемента здійснюється інтегруванням похідної:

$$V_n(I) = \int_0^I \frac{\delta v(i, n)}{K} di + V_{xx} \quad (2)$$

де V_{xx} – постійна інтегрування, яка дорівнює нулю для темнових ВАХ і напрузі холостого ходу при постійному освітленні.

Тобто, одержання темної ВАХ елемента виконується у два етапи: (а) модульоване освітлення елемента із вимірюванням електричного відгуку, пропорційного похідній від ВАХ, (б) обчислення коефіцієнтів пропорційності та відновлення ВАХ інтегруванням.

2.2. Обмеження методу

Розрахунки вихідного методу базуються на припущенні, що коефіцієнт пропорційності K між значенням фотоелектричного відгуку елемента та похідною його ВАХ є постійним. Але у загальному випадку, цей коефіцієнт, визначаючий зсув ВАХ при освітленні, залежить від струму I , на якому проводять вимірювання, а також від елемента n , що тестується. Тобто, справедливості подальших розрахунків у методі базується на переході:

$$K(I, n) = K = const \quad (3)$$

Що ж робити, якщо таке спрощення не працює? Задача є досить складною для розрахунку. Для полегшення завдання можна виділити 2 варіанти вдосконалення методу, які припускають незалежність K від однієї зі змінних. Це спрощує розрахунки щодо загального випадку, але в той же час розширює область застосовності вихідного методу.

3. Запропонований метод

3.1. Наближення подібності елементів

Перше наближення базується на ідентичності параметрів елементів сонячного

модуля. В той же час допускається нерівномірний зсув ВАХ по осі струму. У цьому випадку коефіцієнт пропорційності має вигляд:

$$K(I, n) = K(I) \quad (4)$$

Відповідно, фотоелектричний відгук n -го елемента при значенні струму I буде визначатися за формулою:

$$\delta v(I, n) = K(I) \cdot (dV/dI)_n \quad (5)$$

З адитивності ВАХ елементів модуля, $K(I)$ може бути обчислений зі значень фотоелектричного відгуку елементів та похідної повної ВАХ пристрою:

$$K(I) = \sum_n \delta v(I, n) / (dV/dI)_{\text{модуля}} \quad (6)$$

Отримавши набір коефіцієнтів на різних струмах, можна розрахувати ВАХ елемента за формулою (2). Тобто при цьому наближенні залишається вихідний математичний апарат, але коефіцієнти розраховуються при різних струмах із врахуванням залежності $K(I)$ при інтегруванні.

3.2. Наближення зсуву ВАХ

Другий варіант наближення полягає у здатності ВАХ пристрою зсуватися на фіксовану величину по осі струмів при освітленні. Це наближення застосовується у вихідному методі. Вдосконалимо цей підхід дозволом елементам сонячного модуля мати відмінні один від одного параметри. В такому випадку коефіцієнт пропорційності K не залежить від сили струму, проте має своє значення для кожного елемента:

$$K(I, n) = K(n) \quad (7)$$

Тоді фотоелектричний відгук задається співвідношенням:

$$\delta v(I, n) = K(n) \cdot (dV/dI)_n \quad (8)$$

Склавши ВАХ всіх елементів модуля отримаємо рівняння:

$$(dV/dI)_{\text{модуля}} = \sum_n \delta v(I, n) / K(n) \quad (9)$$

Невідомими в цьому рівнянні виступають

коефіцієнти $K(n)$. Відповідно, кількість невідомих дорівнює кількості елементів модуля. Набір рівнянь виду (9) можна отримати, вимірюючи фотоелектричний відгук при різних струмах.

Нехай у модулі N елементів, а виміри проводяться при M значеннях струму. Тоді результатом буде система M рівнянь із N невідомими, розв'язання якої дасть значення всіх $K(n)$. Для розв'язності системи необхідно, щоб $M \geq N$. Узгодити кількість робочих точок на ВАХ із кількістю елементів у модулі досить просто. Відновлення ВАХ кожного елемента здійснюється за допомогою формули (2).

3.3. Загальний випадок

Розглянемо загальний випадок, коли коефіцієнт пропорційності $K(I, n)$ залежить від струму i має різне значення для кожного елемента модуля. Відповідно, фотоелектричний відгук n -го елемента матиме вигляд:

$$\delta v(I, n) = K(I, n) \cdot (dV/dI)_n \quad (10)$$

У цьому випадку формули (6) і (9) не дозволяють визначити коефіцієнти пропорційності для повної ВАХ елемента. Але існує можливість застосування розглянутих наближень для ділянки ВАХ.

Для цього необхідне виконання наступної умови: кількість значень сили струму, при яких проводились вимірювання, має значно перевищувати кількість елементів у модулі: $M \gg N$. Тоді можна розбити ВАХ на ділянки, кількість точок на кожній з яких рівна N : $(I_1, I_2, \dots, I_N)$, $(I_{N+1}, I_{N+2}, \dots, I_{2N})$, ..., $(I_{M-N}, I_{M-N+1}, \dots, I_M)$. Введемо на кожній ділянці наближення зсуву ВАХ (7). В такому випадку коефіцієнти пропорційності на ділянках можуть бути обчислені з систем рівнянь (9). Це дасть можливість відновити їх ВАХ.

Поєднавши ВАХ всіх ділянок можна отримати загальну характеристику кожного елемента модуля. Звуження ділянок і відповідне зростання кількості точок вимірювань сприятиме більшій точності методу.

4. Висновки

Запропоновані способи вдосконалення математичної обробки фотоелектричного сигналу дозволяють розширити область застосування методу тестування сонячних батарей за допомогою модульованого освітлення.

Перший спосіб заснований на припущенні, що елементи сонячного модуля мають ідентичні параметри, але різні відгуки для різних значень струму. Другий спосіб

вимагає фіксованого зсуву струму при освітленні, але допускає неоднаковість компонентів пристрою.

Поєднання запропонованих методів дає інструмент для дослідження і діагностики напівпровідникових сонячних батарей з неоднорідним розподілом параметрів.

Список використаних джерел

1. *Fraas L.* Solar cells and their applications / L. Fraas, L. Partain. - John Wiley & Sons, Inc.: Hoboken, New Jersey, 2010. – 627 p.
2. *El Chaar L.* Review of photovoltaic technologies / L. El Chaar, L.A. Lamont, N. El Zein // *Renewable and Sustainable Energy Reviews*. – 2011. – **15**. – P. 2165–2175.
3. *Krauter S.C.W.* Development of an integrated solar home system / S.C.W. Krauter // *Solar Energy Materials & Solar Cells*. – 2004. – **82**. – P. 119–130.
4. *Wisseem Z.* Modeling and technical-economic optimization of an autonomous photovoltaic system / Z. Wisseem, K. Gueorgui, K. Hédi // *Energy*. – 2012. – **37**. – P. 263–272.
5. *Agroui K.* Indoor and outdoor characterizations of photovoltaic module based on multicrystalline solar cells / K. Agroui // *Energy Procedia*. – 2012. – **18**. – P. 857–866.
6. *Litvinenko S.V.* Nondestructive diagnostics of solar cells in modules and batteries by means of modulated optical beam-induced photovoltaic signal / S.V. Litvinenko // *Solar Energy Materials & Solar Cells*. – 2003. – **77**. – P. 369–376.
7. *Manilov A.I.* Improved method of diagnostics for multi-component structures / A.I. Manilov // *International Conference "Functional Materials"*, 29 September - 5 October 2013, Gaspra, Ukraine: abstracts. – Gaspra. – 2013. – P. 482.

References

1. FRAAS, L., PARTAIN, L. (2010) *Solar cells and their applications*. John Wiley & Sons, Inc.: Hoboken, New Jersey.
2. EL CHAAR, L., LAMONT, L.A., EL ZEIN, N. (2011) Review of photovoltaic technologies. *Renewable and Sustainable Energy Reviews*. 15. p. 2165–2175.
3. KRAUTER, S.C.W. (2004) Development of an integrated solar home system. *Solar Energy Materials & Solar Cells*. 82. p. 119–130.
4. WISSEM, Z., GUEORGUI, K., HÉDI, K. (2012) Modeling and technical-economic optimization of an autonomous photovoltaic system. *Energy*. 37. p. 263–272.
5. AGROUI, K. (2012) Indoor and outdoor characterizations of photovoltaic module based on multicrystalline solar cells. *Energy Procedia*. 18. p. 857–866.
6. LITVINENKO, S.V. (2003) Nondestructive diagnostics of solar cells in modules and batteries by means of modulated optical beam-induced photovoltaic signal. *Solar Energy Materials & Solar Cells*. 77. p. 369–376.
7. MANILOV, A.I. (2013) Improved method of diagnostics for multi-component structures. In *International Conference "Functional Materials"*, Sunday 29th September to Saturday 05th October 2013. Gaspra. p. 482.

Надійшла до редколегії 15.06.15

УДК 535.345.67

Первак Ю.О., д.ф.-м.н., проф.

Особливості структури дисперсійних дзеркал з негативною дисперсією групової затримки

Київський національний університет імені Тараса Шевченка, 03680, м. Київ, пр-т. Глушкова 4г,
e-mail: yupervak@univ.kiev.ua

Yu.O. Pervak, Dr. Sci.

Features of the structure of dispersive mirrors with negative group delay

Taras Shevchenko National University of Kyiv,
03680, Kyiv, Glushkova st., 4g,
e-mail: yupervak@univ.kiev.ua

Досліджено особливості структури дисперсійного дзеркала з робочим діапазоном від 750 до 850 нм, в якому дзеркало має лінійну залежність групової затримки (GD) і негативну дисперсію групової затримки (GDD). Середній рівень GDD дорівнює -600 фс^2 . Дзеркало утворено нанесенням на скляну підкладку багатошарової структури (73 шари) з оксидів ніобію і кремнію, що чергуються. Повна фізична товщина багатошарової структури дорівнює 9331,7 нм. Аналіз просторово-спектрального розподілу електричного поля свідчить про наявність в структурі семи резонаторів, кожен з яких утворений трьома сусідніми шарами. Центральний шар в двох резонаторах має високий показник заломлення, а в п'яти – низький. Електричне поле завжди концентрується в центральних і бічних шарах з низьким показником заломлення.

Ключові слова: дисперсійне дзеркало, негативна дисперсія групової затримки

The features of structure of dispersive mirror with the operating range from 750 to 850 nm were investigated. A mirror has linear dependence of group delay (GD) and negative group delay dispersion (GDD). The middle level of GDD equals -600 fs^2 . A mirror consists deposited multi-layer structure (73 layers) from the oxides of niobium and silicon that alternate on the glass substrate. The complete physical thickness of multi-layer structure equals 9331,7 nm. The analysis of spatially spectral distribution of electric field testifies to the presence in the structure of seven resonators, each of which consists of three adjacent layers. In two resonators, the central layer has a high refractive index and in five – low. Always electric field is concentrated in the central and lateral layers with low refractive index. The presence of half-wave resonators in ordinary quarter wave structure produces due to the effect of energy storage in the resonator the highest negative group delay dispersion. All half wave resonators are formed a combination of three layers. Key words: dispersive mirror, negative group delay dispersion. The total optical thickness combinations just under half the wavelength of the mid-operating spectral range of the dispersion mirror.

Статтю представив д.техн.н., проф. Погорілий С.Д.

Прогрес в технології дисперсійних дзеркал (ДД) [1-5] дозволив подолати проблему компенсації дисперсії при генерації фемто- і аттосекундних імпульсів [6, 7]. Оптичні елементи на основі ДД забезпечують точний контроль GD і GDD у різних частотних діапазонах, що дуже важливо при створенні фемтосекундних лазерів. Однак, отримання необхідного для компенсації дисперсії високого рівня негативної GDD часто вимагає використовувати схеми компенсаторів, що налічують десятки ДД. Для генерації надкоротких світлових імпульсів потрібно суттєво розширювати спектральний діапазон точного контролю значень GD/GDD. Спектральний діапазон дов-

жин хвиль в дві октави, дозволяє генерувати імпульси тривалістю менше 2 фс у видимому діапазоні [4]. Саме тому, проектування ДД з бажаними спектральними характеристиками (відбиття та GDD), дуже важливе у фізиці надкоротких імпульсів.

Широкодіапазонна оптика на основі ДД з великими від'ємними значеннями GDD використовується для компенсації GDD, накопиченої імпульсом при проходженні через дисперсійні середовища, оптичні підсилювачі та ін. Чим ширший робочий спектральний діапазон, тим менше середнє значення GDD може бути отримано. Це обумовлено принципом невизначеності і тісно

пов'язане з фазовими властивостями багатошарових структур [5]. Нещодавно отримані результати [8] дозволили оцінити максимально досяжні значення GDD, що відповідають одному відбиттю. Проте поза увагою залишилися особливості будови таких ДД.

Тому в даній роботі досліджено особливості структури ДД з максимально високою негативною GDD на прикладі ДД з робочим діапазоном від 750 до 850 нм, в якому ДД має лінійну залеж-

ність групової затримки (GD) і максимальну негативну дисперсію групової затримки (GDD). Дзеркало утворено нанесенням на скляну підкладку багатошарової структури (73 шарів) з оксидів ніобію (H) і кремнію (L), що чергуються. Повна фізична товщина структури дорівнює 9331,7 нм. Оптичні товщини на довжині хвилі середини робочого діапазону ($\lambda_0 = 800$ нм) усіх 73 шарів структури показано на рис.1.

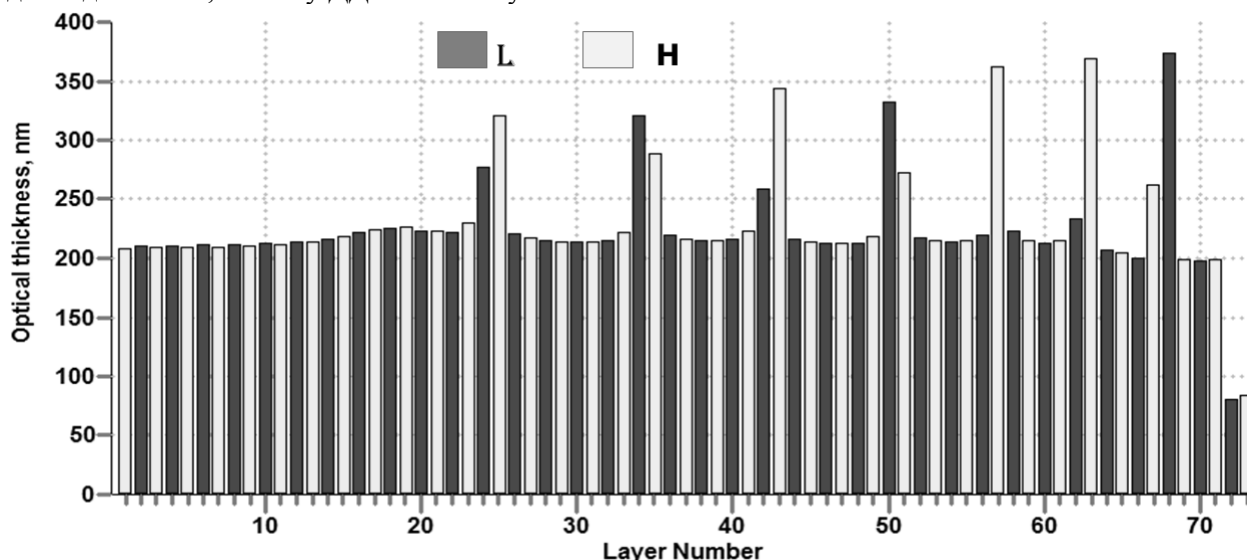


Рис.1. Оптичні товщини шарів (нумерація шарів починається від підкладки).

Процедура отримання конструкції ДД з використанням програмного пакету Optilayer [9, 10] детально описана раніше [11]. Критерії пошуку конструкції визначалися умовами отримання максимального негативного значення GDD і відбиван-

ня (середнє) не менше 99% для p -поляризованого світла при куті падіння 10° . Спектральна залежність відбивання R_p досліджуваного ДД показана на рис.2, а відповідні спектральні залежності GD і GDD на рис.3.

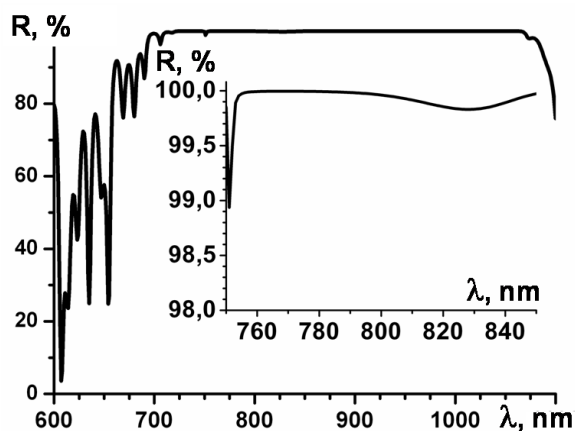


Рис.2. Спектральна залежність відбивання R_p при куті падіння 10° .

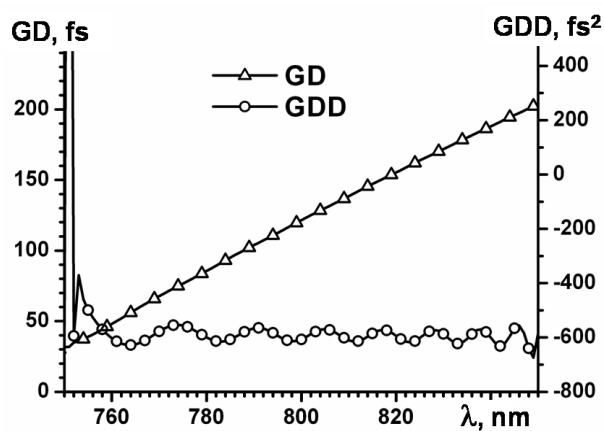


Рис.3. Спектральна залежність GD і GDD при куті падіння 10° .

Аналіз просторово-спектрального розподілу електричного поля всередині багат шарової структури (рис.4) свідчить про наявність в ній семи резонаторів, в яких поле концентрується.

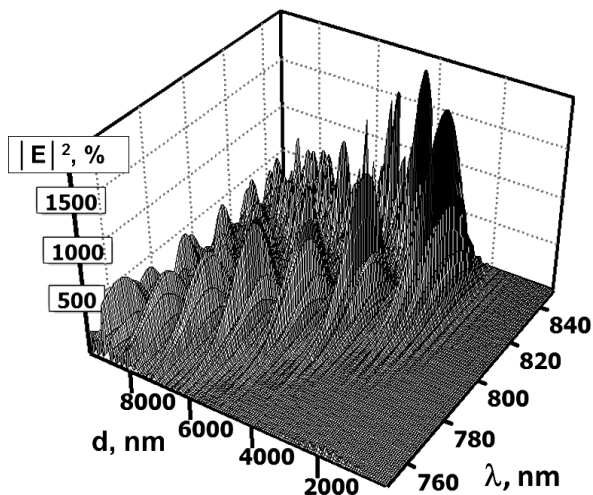


Рис.4. Просторово-спектральний розподіл електричного поля ДД (просторова координата d відраховується від підкладки)

Таблиця 1

Характеристики резонаторів

Номер резонатору	Номер шару	Матеріал	Оптична товщина шару, нм	Оптична товщина резонатору, нм
1	69	Nb ₂ O ₅	200	838
	68	SiO ₂	374	
	67	Nb ₂ O ₅	264	
2	64	SiO ₂	207	813
	63	Nb ₂ O ₅	372	
	62	SiO ₂	234	
3	58	SiO ₂	224	809
	57	Nb ₂ O ₅	365	
	56	SiO ₂	220	
4	51	Nb ₂ O ₅	274	828
	50	SiO ₂	333	
	49	Nb ₂ O ₅	221	
5	43	Nb ₂ O ₅	347	831
	42	SiO ₂	258	
	41	Nb ₂ O ₅	224	
6	35	Nb ₂ O ₅	291	835
	34	SiO ₂	321	
	33	Nb ₂ O ₅	223	
7	25	Nb ₂ O ₅	323	832
	24	SiO ₂	277	
	23	Nb ₂ O ₅	232	

Перший резонатор (нумерація резонаторів починається від зовнішнього середовища), найближчий до поверхні ДД, утворений 67, 68 і 69 шарами. З рис.1 видно, що усі парні шари мають нижчий показник заломлення (SiO₂), ніж непарні (Nb₂O₅). Отже центральний шар першого резонатору має низький показник заломлення. Оптична товщина 67, 68 і 69 шарів дорівнює 264; 374 і 200 нм, відповідно. Сумарна товщина цих трьох шарів дорівнює 838 нм, що трохи більше довжини хвилі середини робочої області ДД. Комбінацію 67, 68 і 69 шарів можна розглядати як деякий еквівалентний шар з відносно низьким показником заломлення. Це підтверджується характером просторово-спектрального розподілу електричного поля. Електричне поле концентрується в 68 шарі.

Другий резонатор утворений 67, 68 і 69 шарами. Тут центральний шар має високий показник заломлення. Комбінація 67, 68 і 69 шарів еквівалентна одному шару з високим показником заломлення. Сумарна товщина цих трьох шарів дорівнює 813 нм, що теж трохи більше довжини хвилі середини робочої області ДД. Відповідні значення параметрів для усіх резонаторів зведено в таблиці.

Також утворені з трьох шарів і інші резонатори. Вони також можуть бути представлені еквівалентними шарами з низьким або високим показником заломлення. Третій резонатор утворений 56, 57 і 58 шарами. Електричне поле концентрується в 56 і 58 шарах і мінімальне в 57 шарі.

Четвертий, п'ятий, шостий і сьомий резонатори можна представити еквівалентними шарами з відносно низьким показником заломлення. Електричне поле концентрується в центральному шарі кожного резонатора. Оптична товщина резонаторів трохи перевищує довжину хвилі середини робочої області ДД.

Електричне поле довгохвильових компонентів випромінювання концентрується в усіх семи резонаторах, в той час як електричне поле короткохвильових компонентів випромінювання концентрується в резонаторах, найближчих до поверхні, що контактує з зовнішнім середовищем. Це забезпечує більшу групову затримку довгохвильових компонентів електромагнітного поля відносно короткохвильових і негативну дисперсію групової затримки (див. рис.3).

Потрібно відзначити, що принциповою відмінністю досліджуваного ДД є саме наявність в

його конструкції пів хвильових резонаторів. І саме тому такі конструкції дзеркал називають дисперсійними дзеркалами, на відміну від традиційних чирпованих дзеркал [1], які вперше дозволили отримати негативну дисперсію групової затримки і вирішити проблему компенсації дисперсії надкоротких світлових імпульсів. В традиційних чирпованих дзеркалах не було півхвильових резонаторів, а збільшення GD з довжиною хвилі і значно менша, ніж в ДД негативна GDD досягалися за рахунок поступового збільшення

оптичної товщини шарів від зовнішнього середовища до підкладки.

Таким чином, півхвильові резонатори, введені в звичайну чверть хвильову стопу дозволяють, завдяки ефекту накопичення енергії в резонаторах, отримувати максимальну високу негативну дисперсію групової затримки. При цьому півхвильовий резонатор утворюється комбінацією з трьох шарів. А сумарна оптична товщина комбінації трохи менша половини довжини хвилі середини робочого спектрального діапазону дисперсійного дзеркала.

Список використаних джерел

1. Szipocs R. Chirped multilayer coatings for broadband dispersion control in femtosecond lasers / R. Szipocs, K. Ferencz, C. Spielmann, F. Krausz // *Opt. Lett.*—1994—**19**— N.3.— P.201–203.
2. Pervak V. High-dispersive mirrors for femtosecond lasers/ V.Pervak, C.Teisset, A.Sugita, et al.//*Opt.Express.* — 2008.—**16** — N.14. — P. 10220–10233.
3. Pervak V. High-dispersive mirrors for high power applications / V. Pervak, O. Pronin, O. Razskazovskaya, et al. // *Opt. Express.* — 2012.—**20**. — N.4. — P.4503–4508.
4. Pervak V. Recent development and new ideas in the field of dispersive multilayer optics / V. Pervak // *Appl.Opt.*—2011.—**50**. — N.9. — P.C55–C61.
5. Nohadani O., Birge J.R., Kartner F.X., Bertsimas D.J. Robust chirped mirrors / O. Nohadani, J.R. Birge, F.X. Kartner, D.J. Bertsimas // *Appl. Opt.* — 2008.—**47**. — N.14. — P.2630–2636.
6. Pervak V. Robust synthesis of dispersive mirrors / V. Pervak, M.K. Trubetskov, A.V. Tikhonravov // *Opt.Express.*—2011.—**19**.—N.3.—P.2371–2380.
7. Krausz F. Attosecond physics / F. Krausz // *Rev. Mod. Phys.* — 2009.—**81**. — N.1. — P.163–234.
8. Wirth A. Synthesized light transients / A. Wirth, M.T. Hassan, V. Pervak, et al. // *Science.* — 2011.— **334**. — P.195–200.
9. Tikhonravov A.V. Phase properties of multilayers / A.V. Tikhonravov, B.W. Baumeister, K.V. Popov // *Appl. Opt.* — 1997.—**36**. — N.19. — P.4382–4392.
10. Tikhonravov A.V. OptiLayer software / A.V. Tikhonravov., M.K. Trubetskov// <http://www.optilayer.com>.
11. Pervak V. Empirical study of the group delay dispersion achievable with multilayer mirrors / V. Pervak, V. Fedorov, Yu.A. Pervak, M.K. Trubetskov // *Opt. Express.* —2013.—**21**. — N.15.— P.18311–18316.

References

1. SZIPOCS, R. et al.(1994) Chirped multilayer coatings for broadband dispersion control in femtosecond lasers. *Opt. Lett.* 19 (3). p. 201–203.
2. PERVAK, V. et al. (2008) High-dispersive mirrors for femtosecond lasers. *Opt. Express.* 16 (14). p.10220–10233.
3. PERVAK, V. et al.(2012) High-dispersive mirrors for high power applications. *Opt. Express.* 20 (4). p.4503–4508.
4. PERVAK, V. (2011) Recent development and new ideas in the field of dispersive multilayer optics. *Appl. Opt.* 50 (9). p. C55–C61.
5. NOHADANI, O. et al. (2008) Robust chirped mirrors. *Appl. Opt.* 47 (14). p.2630–2636.
6. PERVAK, V. & Trubetskov, M., Tikhonravov, A. (2011) Robust synthesis of dispersive mirrors. *Opt. Express.* 19. — N.3. — P.2371–2380.
7. KRAUSZ, F. (2009) Attosecond physics. *Rev. Mod. Phys.* 81 (1). p.163–234.
8. WIRTH, A. et al. (2011) Synthesized light transients. *Science.* 334. p.195–200.
9. TIKHONRAVOV, A.V. & BAUMEISTER, B.W. POPOV, K.V. Phase properties of multilayers. *Appl. Opt.* 36 (19). p.4382–4392.
10. TIKHONRAVOV A.V. & TRUBETSKOV M.K. OptiLayer software. <http://www.optilayer.com>.
11. PERVAK, V. et al.(2013) Empirical study of the group delay dispersion achievable with multilayer mirrors. *Opt. Express.* 21 (15). p.18311–18316.

Надійшла до редколегії 20.06.15

УДК 537.86/.87

Слободянюк Д. В.¹, к. ф.-м. н.,
Попов М. О.², к. ф.-м. н.,
Малишев В. Ю.³, асистент,
Мойсеєнко В. А.⁴, провідний інженер

D. V. Slobodianiuk¹, Ph. D.,
M. O. Popov², Ph. D.,
V. Yu. Malyshev³, Assist. Prof.,
V. A. Moiseienko⁴, Leading Engineer

Спотворення профілю вихідного сигналу параметричним підсилювачем в багатомодовому режимі

Signal profile distortion in multimode parametric amplifier

^{1,2,3,4} Київський національний університет імені
Тараса Шевченка, 03022, м. Київ, пр-т.
Глушкова 4г.

^{1,2,3,4} Taras Shevchenko National University of
Kyiv, 03022, Kyiv, Glushkova st., 4g.

¹ e-mail: denslobod@ukr.net

¹ e-mail: denslobod@ukr.net

² e-mail: maxim_popov@univ.kiev.ua

² e-mail: maxim_popov@univ.kiev.ua

³ e-mail: forstain@i.ua

³ e-mail: forstain@i.ua

⁴ e-mail: v_a_m@univ.kiev.ua

⁴ e-mail: v_a_m@univ.kiev.ua

В роботі експериментально та за допомогою числових методів досліджено роботу параметричного підсилювача на основі тонкої плівки залізо-ітрієвого гранату. Для теоретичного опису такої системи застосовувалася S-теорія. В ході аналізу використовувалась модель багатомодової системи з паралельною накачкою, особливість якої полягає у врахуванні присутності короткохвильових спінових збуджень (т. з. домінантної групи хвиль), що вироджені по частоті з сигнальною хвилею та мають мінімальний параметр релаксації.

В результаті було встановлено, що багатомодовість призводить до спотворення часових профілів вихідних сигналів, підсиленних під дією повздовжньої параметричної накачки. При цьому затримка підсиленого сигналу як функція потужності накачки має екстремальний характер, а тривалість самого сигналу зменшується приблизно в 50 разів, від 1 мкс до значень порядку 20 нс.

Ключові слова: параметричний підсилювач, залізо-ітрієвий гранат, спінові хвилі, S-теорія.

In this paper the operation of the parametrical amplifier based on thin yttrium-iron garnet film was investigated both theoretically and experimentally. S-theory approach was used for the theoretical description of the given system. During the analysis the multimode model of parametrical amplifier with parallel pumping was utilized. The distinctive feature of such model is the presence of short-wave spin excitations (so-called dominant waves) that have the lowest dissipation parameter and are frequency degenerated with signal waves. Consequently, a competition takes place between two simultaneous processes, namely, parametric amplification of signal wave and parametric generation of dominant waves.

It was shown that multimode nature of this system leads to the distortion of output amplified signal time profiles. The parametrically amplified signal delay time versus pump power dependence demonstrates nonlinear behavior with clear maximum, whereas duration of amplified signal decreases monotonously with power increase in more than 50 times, from 1 μ s down to 20 ns value.

Key words: parametrical amplifier, yttrium-iron garnet, spin waves, S-theory.

Статтю представив д.ф.-м.н., доц. Савенков С.М.

Вступ

Вже більш ніж 50 років відомі та широко використовуються радіоелектронні пристрої, в яких підсилення сигналу по потужності відбувається за рахунок енергії зовнішнього джерела (генератора накачки), що періодично

змінює індуктивність або ємність нелінійного реактивного елемента електричного кола підсилювача. Такий тип приладів отримав назву „параметричні підсилювачі”, оскільки їх принциповою особливістю є модуляція параметрів коливальної системи [1]. Параметричні підсилювачі знайшли своє

застосування в радіоастрономії, в дальньому космічному та супутниковому зв'язку та радіолокації, як малопотужні підсилювачі для слабких сигналів, що потрапляють на вхід радіоприймального пристрою. Частіше за все такі пристрої використовуються в НВЧ діапазоні довжин хвиль. В якості реактивного елемента параметричних підсилювачів переважно використовується напівпровідниковий параметричний діод, а в НВЧ діапазоні застосовуються електронно-променеві лампи. Генератори накачки створюються на основі лавино-пролітних діодів, діодів Ганна, відбиваючих клістронів. На сьогоднішній день реалізовані параметричні підсилювачі з коефіцієнтом підсилення сигналу по потужності на рівні 10-30 дБ, смугою пропускання, що складає 10-20 % від несучої частоти сигналу та значенням шумової температури, що не перевищує 100 К [1].

Проте, як і інші типи підсилювачів, вони мають недоліки, серед яких окреме місце займають нелінійні спотворення. Нелінійними спотвореннями називають зміни форми огинаючої сигналу, причина виникнення яких полягає в нелінійності характеристик електронних елементів, з яких побудований підсилювач. До типових елементів електричних схем, яким притаманні нелінійні властивості, відносяться транзистори, вакуумні лампи, діоди та трансформатори.

При наявності нелінійних спотворень в підсилювачі (а в реальних підсилювачах вони є завжди) у вихідному сигналі виникають нові частоти, котрі були відсутні у вхідному. Загальний рівень нелінійних спотворень кількісно оцінюється коефіцієнтом нелінійних спотворень [2], який зазвичай виражається у відсотках. Практично мають значення тільки друга та третя гармоніки. Наприклад, для підсилювачів низької частоти простої побутової радіотехніки максимальним рівнем можна вважати 15-20 %, а для високоякісних підсилювачів сучасної стереоапаратури коефіцієнт нелінійних спотворень складає десяти або й навіть соті долі відсотка.

Метою даної роботи було дослідження параметричного підсилювача на базі тонких плівок залізо-ітрієвого гранату, який функціонує в багатомодовому режимі. Вхідний сигнал такого підсилювача являв собою спінову хвилю, а для накачування використовувалася електромагнітна параметрична накачка на подвоєній частоті вхідного сигналу. Основна увага в роботі

приділялась дослідженню впливу багатомодовості на спотворення часових профілів вихідних сигналів під дією повздовжньої параметричної накачки.

Теоретична модель

Для аналізу багатомодової взаємодії при параметричному підсиленні спінових хвиль будемо використовувати модифіковану систему рівнянь Бломбергена [3, 4], що має наступний вигляд:

$$\left(\frac{\partial}{\partial t} + \Gamma_i + \nu_i \frac{\partial}{\partial z} \right) A_i + i P_i A_i^* = 0,$$
$$P_i = h(z, t) V_i + S_{i,i} A_i A_{-i} + S_{i,j} A_i A_j, \quad (1)$$
$$i, j = s, d$$

і описує сигнальну A_s та домінуючу A_d групу спінових хвиль, причому від'ємний індекс відповідає хвилі, що рухається в зворотному напрямку (т.з. холоста хвиля). В загальному випадку, враховуючи особливості спіно-хвильового спектру, існує нескінченна кількість хвиль, вироджених по частоті з сигнальною. Однак в рамках запропонованої нами моделі ми обмежимося врахуванням лише однієї хвилі, що має найменший параметр дисипації. Питання про знаходження вказаного параметра буде обговорено надалі.

Наявність домінуючої спінової хвилі суттєвим чином впливає на еволюцію сигнальної групи хвиль, оскільки в системі виникає конкуренція між процесами параметричного підсилення сигнальної групи та параметричного збудження домінуючої. Так, в найпростішому випадку, внаслідок перерозподілу енергії накачки в системі сигнальна група спінових хвиль може подавлюватись [5]. Далі цей висновок буде проаналізовано детальніше.

В рівняннях системи (1) ν_i та Γ_i - це, відповідно, групова швидкість та параметр дисипації відповідних груп хвиль, $S_{i,i}, S_{i,j}$ - нелінійні параметри, що відповідають за взаємодію хвильових пакетів в рамках S-теорії [5], а V_i - величина, що характеризує зв'язок відповідних хвиль з накачкою і визначається, як:

$$V_i = V_{k_i, k'_i} = \gamma \frac{\omega_{k_i} \omega_{k'_i} - \omega_H^2}{2 \sqrt{\omega_{k_i} \omega_{k'_i} \omega_H}} \quad (2)$$

де γ – гіромагнітне відношення, $\omega_H = \gamma H_0$ – характерна частота спектру спінових хвиль (СХ), H_0 – внутрішнє магнітне поле.

Профіль сигналу накачки вважався однорідним в межах області $0 < z < L$ і прямокутним у часі, з тривалістю τ_p :

$$h(z, t) = \begin{cases} 0, z < 0 \cup z > L \\ \Theta(t) \Theta(\tau_p - t) h_p, 0 < z < L \end{cases} \quad (3)$$

Характерний розмір області накачки L задається фактичними параметрами системи, наприклад, розміром діелектричного резонатора, який використовується в експерименті, або ефективною товщиною мікросмужкової антени.

Рівняння (1) необхідно доповнити початковими та граничними умовами для сигнальної, холостої та домінантної груп хвиль. Вважаючи швидкість домінантної групи близькою до нуля та поклавши $A_d = A_{-d}$ матимемо:

$$\begin{aligned} A_s(0, t) &= A_0, \\ A_s(z, 0) &= A_0 \exp(-\Gamma_s z / v_s) \\ A_s(L, 0) &= A_{-s}(L, 0) = A_T, \\ A_d(L, 0) &= A_T \end{aligned} \quad (4)$$

такий вибір початкових умов відповідає випадку неперервного вхідного сигналу амплітуди A_0 , що взаємодіє з домінантною групою теплової амплітуди A_T .

Система рівнянь (1) розв'язувалась для випадку дотично намагніченої плівки залізо-ітрієвого гранату (ЗІГ) числовими методами. В ході розв'язку використовувались параметри, ідентичні до зазначених у роботі [6]: $v_s = 3.6 \cdot 10^6$ см/с, $\Gamma_s / 2\pi = 0.7$ МГц, $\Gamma_d = 0.9\Gamma_s$, всі коефіцієнти $S_{i,j} = 2\pi \cdot 200$ МГц, $\tau_p = 1$ мкс.

Для спрощення вважалося, що швидкість поширення домінантних хвиль близька до нуля.

Тут необхідно зазначити, що параметр дисипації домінантної групи, в загальному випадку, не може бути знайдений точно. Однак можна стверджувати, що, через ряд причин, дисипація домінантної групи буде мінімальною. Головним чином це пов'язано з тим, що, по-перше, швидкість поширення коротких спінових хвиль є близькою до нуля, відповідно, такий механізм затухання хвиль, як винос енергії з області дії накачки, для домінантної групи буде малоефективним. По-друге, для

короткохвильових спінових збуджень практично відсутній такий додатковий вклад в параметр дисипації, як двамагнотний механізм розсіяння на неоднорідностях кристалу, що теж зменшує дисипацію домінантної групи.

В результаті розв'язку системи (1) було одержано просторово-часові профілі вихідних підсиленних сигналів в багатомодовій системі, а також основні їх характеристики: напівширину, тривалість та час затримки, при різних рівнях потужності накачки або, що те ж саме, при різних коефіцієнтах підсилення. На рис. 1 наведено теоретично розраховані профілі вихідних сигналів для різних потужностей накачки. Як бачимо, при зростанні потужності накачки форма вихідного сигналу зазнає спотворень. Так, максимум огинаючої сигналу спочатку зсувається вправо, в область більших часів затримки, а з подальшим ростом потужності накачки починається звуження сигналу із зсувом його максимуму в сторону менших затримок. Тривалість сигналу (визначена як інтервал між точками, в яких амплітуда сигналу вдвічі менша за максимальну) теж зменшується, від початкового значення порядку тривалості накачки до тривалості близько 100 нс.

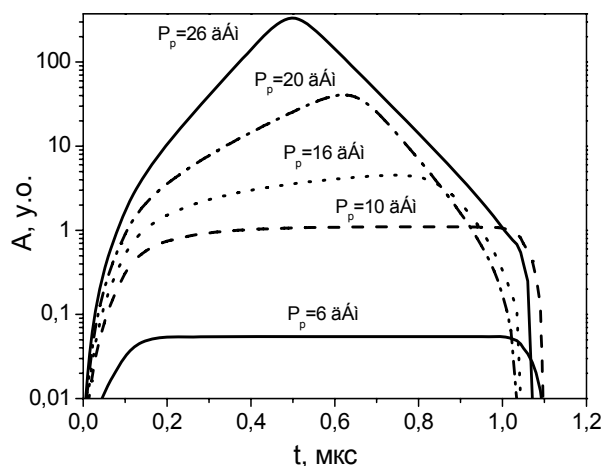


Рис. 1 Еволюція профілів вихідних сигналів при зростанні потужності накачки

Експеримент та обговорення результатів

Для перевірки результатів числового розрахунку були проведені відповідні експериментальні вимірювання. Дослідження виконувались на установці, основною частиною, якої була пасивна лінія затримки (ПЛЗ) на основі тонкої плівки ЗІГ. Типову схему експериментальної секції на ПЛЗ наведено на

рис. 2. Вона складалася з дотично намагніченої плівки ЗІГ 3, товщиною 5 мкм, розміщеної всередині відкритого діелектричного резонатор (ВДР) 2. ВДР використовувався для концентрації електромагнітного поля накачки в об'ємі плівки ЗІГ і модуляції на подвоєній частоті сигналу внутрішнього магнітного поля H_0 . В свою чергу, модуляція магнітного поля призводила до періодичної зміни величини зв'язку хвиль з накачкою (див. (2)), що робило можливим процес параметричного підсилення. Для підведення та відведення мікрохвильового сигналу використовувались дві мідні мікросмужкові антени-перетворювачі 1 (вхідна та вихідна) шириною 25 мкм, що розміщувались поблизу поверхні плівки.

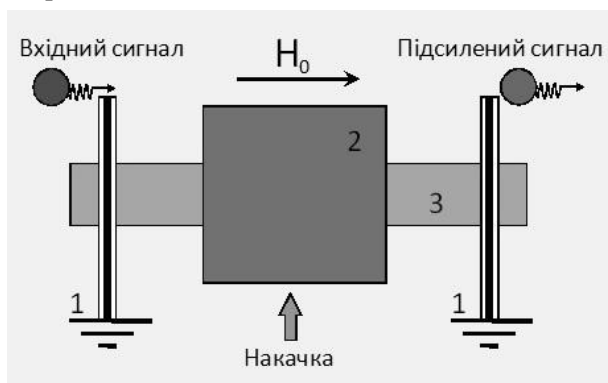


Рис. 2 Схема експериментальної секції

В результаті дії мікрохвильового поля антени на дотично намагнічену плівку ЗІГ в ній збуджувалися спінові хвилі, котрі рухались до вихідної антени з певною груповою швидкістю v_g , причому вектор фазової швидкості був спрямований протилежно відносно напрямку поширення спін-хвильового пакету. В той момент часу, коли хвилі змінної намагніченості, збуджені сигнальним генератором, потрапляли в область дії накачки (область розташування ВДР), на ВДР подавався імпульс параметричної накачки на подвійній частоті вхідного сигналу, який призводив до параметричного підсилення всіх тих хвиль, що задовольняли умовам параметричного резонансу, тобто для яких одночасно виконувалися закони збереження енергії $\omega_p = \omega_r + \omega_s$ (де $\omega_p, \omega_r, \omega_s$ – частота накачки, холостої та сигнальної хвиль, відповідно) та імпульсу. З вихідної антени НВЧ сигнал проходив на два каскади підсилювачів (коефіцієнт підсилення 30 дБ, шумова температура – 90 К). З підсилювачів сигнал потрапляв на квадратичний детектор. Продетектований сигнал спостерігався за

допомогою цифрового осцилографа Tektronix TDS 3032B. Для запису осцилограм було реалізовано спряження осцилографа з персональним комп'ютером по протоколу Ethernet.

При цьому такий параметричний підсилювач характеризувався певним коефіцієнтом підсилення та смугою підсилення частот. Коефіцієнт підсилення параметричного підсилювача визначався потужністю накачки [7], а смуга підсилення залежала як від потужності, так і від тривалості накачки [8].

В даній роботі досліджувалася еволюція форми вихідного підсиленого сигналу в багатомодовому параметричному підсилювачі при збільшенні потужності параметричної накачки. Для цього використовувалися сигнали накачки з прямокутним профілем огинаючої, що створювались за допомогою швидкісних (час перемикання 2 нс) дискретних модуляторів амплітуди НВЧ коливань. Частота вхідного сигналу $\omega_s / 2\pi$ складала 4700 МГц, а частота накачки $\omega_p / 2\pi$ була 9400 МГц. Плівка намагнічувалась до насичення зовнішнім магнітним полем в 990 Е. На вхідну антену-перетворювач подавався неперервний НВЧ сигнал ($\tau_s = \infty$), а тривалість накачки τ_p у всій множині експериментів варіювалась від 1 до 12 мікросекунд. Максимальна усереднена за період потужність накачки P_p не перевищувала 1 Вт. Зміни часового профілю огинаючої підсиленого сигналу реєструвались та записувались за допомогою осцилографа.

Детальне порівняння експериментальних та теоретичних результатів щодо еволюції вихідного сигналу з ростом потужності накачки, при тривалості накачки $\tau_p = 1$ мкс, наведено на рис. 3. Тривалість підсиленого сигналу визначалась на половині висоти останнього. Час затримки вимірювався відносно переднього фронту поданого в систему вхідного сигналу. Числове моделювання параметричного підсилення сигналів прямокутного профілю проводилось з використанням моделі параметричного підсилювача в багатомодовому режимі, як описано у попередньому розділі.

Як бачимо з рисунку, тривалість сигналу монотонно спадає, починаючи з величини порядку 1 мкс, яка природно відповідає тривалості накачки. Цікавим є той факт, що при великих рівнях потужності накачки тривалість підсиленого сигналу суттєво зменшується і може

складати десятки наносекунд. Це проілюстровано асимптотичним наближенням кривої тривалості сигналу до рівня 15-20 нс при зростанні потужності накачки. Причиною цього є те, що параметрично збуджена домінантна група спінових хвиль з ростом потужності накачки першою досягає стаціонарного рівня, пригнічуючи тим самим сигнальну групу. Внаслідок цього підсилення вихідного сигналу припиняється і він набуває характерного трикутного профілю, що супроводжується відповідним зменшенням напівширини.

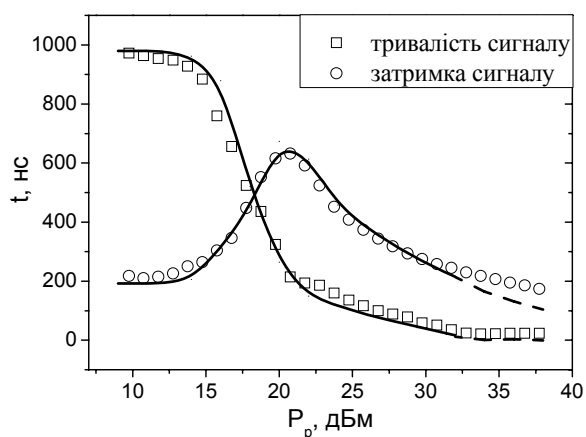


Рис. 3 Залежність тривалості та затримки підсиленого сигналу від потужності накачки. Точки експеримент, суцільна лінія – теорія. Область, де теоретична модель стає несправедливою, відмічено штрихованою

Нетривіальною є залежність часу затримки сигналу від потужності накачки. Як бачимо, на початковому етапі затримка сигналу зростає, досягає певного максимуму, а потім починає зменшуватися. Це пояснюється наступним чином: на початковому етапі підсилення відбувається спотворення форми сигналу від прямокутного до трикутного профілю, зі зсувом його максимуму в бік більших затримок (див. рис. 1). З подальшим ростом потужності накачки вплив домінантної групи підсилюється, сигнал переходить до трикутного профілю і починає звужуватися, при цьому зменшується і його затримка. Варто зазначити, що схожі залежності вже фіксувалися в роботі [6], однак там не було проаналізовано поведінку сигналу при невеликих потужностях накачки і для розрахунку використовувалась проста модель, з просторово-однорідним розподілом амплітуди, тобто процесами поширення сигналу по плівці було знехтувано.

Нарешті, варто звернути увагу на межі

застосовності розробленої теоретичної моделі. Взагалі кажучи, з ростом потужності накачки в багатомодовій параметричній системі починають проявлятися нелінійності вищих порядків, які в поточній моделі не враховані. Окрім того, оскільки має місце значне «звуження» сигналу в часі, стає необхідним врахування дисперсійних доданків в рівнянні (1). У випадку якщо буде виконано критерій Лайтхілла в системі можуть формуватися т.з. солітони огинаючої [7]. Перехід в солітонний режим виходить за рамки нашого дослідження і лежить далеко поза межами розробленої теоретичної моделі. Саме тому область, в якій тривалість вихідного імпульсу стає порядку 15-20 нс, була відмічена на рис. 3, як така, де теоретична модель перестає бути справедливою.

Висновки

В статті розглянуто роботу параметричного підсилювача спінових хвиль в багатомодовому режимі та вплив багатомодовості на форму вихідного сигналу. Експериментально та за допомогою числових методів досліджено еволюцію форми вихідного підсиленого сигналу в багатомодовій параметричній системі на базі тонких плівок залізо-ітрієвого гранату, в якій одночасно відбувається параметричне підсилення сигнальної хвилі та параметричне збудження домінантної групи хвиль. Встановлено, що такий режим роботи параметричних підсилювачів призводить до спотворення часових профілів вихідних сигналів. Показано, що залежність затримки сигналу від потужності накачки є немонотонною, а при зменшенні його тривалості виникають особливості, що за критерієм Лайтхілла можна пояснити переходом в солітонний режим. Було досягнуто мінімальне значення тривалості підсиленого сигналу порядку 20 нс при тривалості імпульсу накачки 1 мкс і середній за період потужності накачки не більше 1 Вт. Теоретичний аналіз такої параметричної системи проводився в межах S-теорії, числові розрахунки якісно та кількісно узгоджуються з експериментальними даними.

Отримані результати дозволяють оптимізувати роботу параметричних підсилювачів на основі тонких плівок ЗІГ з метою збереження ступеня нелінійних спотворень вихідного сигналу в прийнятних межах.

Автори висловлюють глибоку подяку Мелкову Г.А. за підтримку та консультації при виконанні досліджень.

Список використаних джерел

References

1. Эткин В. С. Параметрические системы СВЧ на полупроводниковых диодах / В. С. Эткин, Е. М. Гершензон. – М.: Сов. радио, 1964. – 351 с.
2. Богданович Б. М. Нелинейные искажения в приемно-усилительных устройствах / Б. М. Богданович. – М.: Связь, 1980. – 279 с.
3. Сухоруков А.П. Нелинейные волновые взаимодействия в оптике и радиофизике / А.П. Сухоруков – Москва: Наука, 1985. – 247 с.
4. Беспалов В.И. Нелинейная оптика и адаптивные лазерные системы / В.И. Беспалов. – Москва: Наука, 1986. – 133 с.
5. Львов В.С. Нелинейные спиновые волны / В.С.Львов. – Москва: Наука, 1987. –258 с.
6. Chumak A.V. Parametrically-stimulated recovery of a microwave signal using standing spin-wave modes of a magnetic film / A.V.Chumak, A.A.Serga, B.Hillebrands, G.A.Melkov // *Phys.Rev. B*. – 2009. – **79**. – P. 014405-1 – 014405-10.
7. Мелков Г.А. Параметричне підсилення та обернення хвильового фронту пакетів магнітостатичних хвиль / Г.А. Мелков, О.О. Серга, В.С. Тиберкевич, О. М. Олійник // Вісник Київського університету. Серія: фізико-математичні науки. – 1998. - № 3. – С. 327-333.
8. Vasyuchka V.I. Non-resonant wave front reversal of spin waves used for microwave signal processing / V.I. Vasyuchka, G.A. Melkov, A.N. Slavin, A.V. Chumak, V.A. Moiseienko, B. Hillebrands // *Journal of Physics D: Applied Physics* – 2010. – **43**. – P. 325001-1 - 325001-4.
9. Kalinikos B.A. Parametric amplification of spin wave envelope solitons in ferromagnetic films by parallel pumping / B.A. Kalinikos, M.P. Kostylev // *IEEE Trans. on Magn.* – 1997. – **33**. – No. 5. – P. 3445-3447.
1. ETKIN, V. and GERSHENZON, E. (1964) *Parametricheskie sistemyi SVCh na poluprovodnikovyyih diodah*. Moskva: Sov. radio.
2. BOGDANOVICH, B. (1980) *Nelineynnye iskazheniya v priemno-usilitelnyih ustroystvakh*. Moskva: Svyaz.
3. SUHORUKOV, A. (1985) *Nelineynnye volnovyye vzaimodeystviya v optike i radtofizike*. Moskva: Nauka.
4. BESPALOV, V. (1986) *Nelineynnaya optika i adaptivne lazernyye sistemyi*. Moskva: Nauka.
5. LVOV, V. (1987) *Nelineynnye spinovyye volny*. Moskva: Nauka.
6. CHUMAK, A., SERGA, A., HILLEBRANDS, B. and MELKOV, G. (2009) Parametrically-stimulated recovery of a microwave signal using standing spin-wave modes of a magnetic film. *Phys. Rev. B*, 79(1). p. 014405-1 – 014405-10.
7. MELKOV, G.A., SERGA, O.O., TIBERKEVICH, V.S. and OLIYNIK, O. M. (1998) Parametrichne pidsilennya ta obrernennya hvilovogo frontu paketiv magnitostatichnih hvil. *Bulletin of Taras Shevchenko National University of Kyiv. Series Physics & Mathematics*. 3. p. 327-333.
8. VASYUCHKA, V.I., MELKOV, G.A., SLAVIN, A.N., CHUMAK, A.V., MOISEIENKO, V.A., and HILLEBRANDS, B. (2010) Non-resonant wave front reversal of spin waves used for microwave signal processing. *Journal of Physics D: Applied Physics*. 43(32). p. 325001-1 - 325001-4.
9. KALINIKOS, B. and KOSTYLEV, M. (1997) Parametric amplification of spin wave envelope solitons in ferromagnetic films by parallel pumping. *IEEE Trans. on Magn.* 33(5). p. 3445-3447.

Надійшла до редколегії 19.06.2015

СУЧАСНА ФІЗИКА

УДК 536, 538.9,

Гаврюшенко Д. А.¹, д.ф.-м.н., проф,
Дворецька В. І.², студ.

Осмотичний тиск бінарного розчину поблизу критичної точки розшарування

^{1,2}Київський національний університет імені
Тараса Шевченка, 01601, м. Київ,
вул. Володимирська 64/13,
e-mail: ¹dg@univ.kiev.ua, ²ms.dvorecka@mail.ru

D. A. Gavryushenko¹, Dr. Sci., Prof.,
V. I. Dvoretska², stud.

The osmotic pressure of binary solution in the vicinity of the critical point of stratification

^{1,2}Taras Shevchenko National University of Kyiv,
01601, Kyiv, Volodymyrska st. 64/13,
e-mail: ¹dg@univ.kiev.ua, ²ms.dvorecka@mail.ru

Робота присвячена дослідженню особливостей поведінки флюїдів в обмежених системах. Показано, що як хімічний потенціал бінарного розчину, так і осмотичний тиск в плоскопаралельній порі залежать від радіуса кореляції поблизу критичної точки розшарування. Побудовано вирази для розрахунку хімічного потенціалу та осмотичного тиску бінарного розчину, що дозволяють проводити обчислення цих величин в широкому інтервалі зміни термодинамічних параметрів, включаючи окіл критичної точки розшарування. Встановлено, що хімічний потенціал та осмотичний тиск бінарної суміші, викликаний наявністю зовнішнього поля, в обмеженій системі суттєвим чином залежить від геометрії та розмірів системи. Показано, що з наближенням до критичної точки розшарування в плоскопаралельній порі хімічний потенціал та осмотичний тиск суттєво змінюються.

Ключові слова: хімічний потенціал, осмотичний тиск, критична точка розшарування, радіус кореляції, бінарний розчин.

Approaching to the critical point, parameters that characterize the system (isothermal compressibility, specific heat, thermal conductivity, concentration etc) behave abnormally. This work is devoted to investigation of the behavior of fluids in confined geometry systems. The dependence of the chemical potential and osmotic pressure on the correlation radius within the flat-parallel pore have been demonstrated. The expressions for calculation of the chemical potential and osmotic pressure of binary solution, allowing to calculate this values within a wide range of thermodynamic parameters variation including the vicinity of the critical point of stratification have been obtained. The significant dependence of the chemical potential and osmotic pressure of a binary solution under an external field on the geometry and size of system as well as the correlation radius has been demonstrated. It was shown that in the vicinity of the critical point of stratification in the flat-parallel pore chemical potential and osmotic pressure change significantly.

Key Words: chemical potential, osmotic pressure, the critical point of stratification, correlation radius, binary solution.

Статтю представив академік НАН України, д.ф.-м.н., проф. Булавін Л.А.

Результати досліджень властивостей речовин поблизу критичної точки як випаровування, так і розшарування важливі для практичного застосування. Унікальні властивості речовини в критичній області мають важливе значення для використання у новітніх сучасних технологіях. Існуючі на сьогоднішній день роботи, в яких враховано неоднорідність обмежених систем і проведено розрахунок просторового розподілу густини системи, або присвячені однокомпонентним рідинам, або базуються на

певних припущеннях, що різко звужують область застосування отриманих результатів.

В даній роботі розглядається залежність хімічного потенціалу та осмотичного тиску [7] бінарного розчину поблизу критичної точки розшарування [1] від термодинамічних змінних. Дослідження проведено для ідеальних та неідеальних пор з різними пристінковими потенціалами.

Якщо бінарний розчин відділений від розчинника напівпроникною мембраною (рис. 1), то можлива лише одностороння дифузія. У цьому

випадку осмотичний тиск стає доступною для прямого вимірювання величиною, яка дорівнює надлишковому тиску, який виникає з боку розчину для виконання умови термодинамічної рівноваги. Тенденція системи вирівнювати хімічні потенціали у всіх частинах свого об'єму та переходити до стану з мінімальним значенням відповідного термодинамічного потенціалу викликає осмотичний (дифузійний) перенос речовин.

Розглянемо випадок, коли система складається з двох підсистем, розділених напівпроникною мембраною, тобто вони можуть обмінюватись частинками лише одна з одною, при заданих (T, p_0, N) . Поведінку кожної з підсистем необхідно розглядати в рамках великого канонічного ансамблю із заданими значеннями відповідних хімічних потенціалів. Зліва від цієї мембрани знаходиться розчин компоненти b з молярною концентрацією x_b в розчиннику a з концентрацією x_a , що міститься також справа від стінки. Мембрана проникна лише для молекул розчинника. Очевидно, що $x_a + x_b = 1$. Загальна кількість молекул розчиненої речовини, які знаходяться по різні сторони мембрани становить $N = N_a + N_b$.

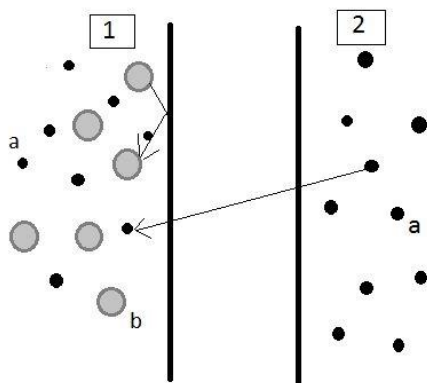


Рис. 1. Механізм виникнення осмотичного тиску при наявності напівпроникної мембрани: молекули розчинника (a) та молекули розчиненої речовини (b)

Необхідною умовою термодинамічної рівноваги між двома підсистемами в цьому випадку є рівність хімічних потенціалів розчинника зліва і справа від мембрани [4]:

$$\mu_{1a}(T, p_1, x_a) = \mu_{2a}(T, p_2), \quad (1)$$

де p_1 – тиск компонент a та b в бінарному розчині на стінку мембрани з боку підсистеми 1, p_2 – тиск компоненти a на стінку з боку підсистеми 2. Умову (1) можна задовольнити лише у випадку, коли $p_1 \neq p_2$, що і призводить до появи осмотичного тиску $\Delta p = p_2 - p_1 \neq 0$.

Розглянемо випадок, коли рідина знаходиться поблизу критичної точки розшарування. Вважаючи, що величина Δp_1 мала, виконаємо розклад $\mu_{1a}(T, p_1, x_a)$ за Δp_1 :

$$\Delta \mu_{1a} = \left(\frac{\partial \mu_{1a}}{\partial p_1} \right)_T \Delta p_1 + \frac{1}{2!} \left(\frac{\partial^2 \mu_{1a}}{\partial p_1^2} \right)_T (\Delta p_1)^2 + \dots \quad (2)$$

Для конкретизації залежності хімічного потенціалу $\mu_{1a}(\Delta x_1(z))$ розглянемо випадок малих відхилень концентрації неоднорідної системи від концентрації однорідної системи. В цьому випадку можна застосувати наближення плавної неоднорідності, запропоноване [6]:

$$\mu_{1a}(\Delta x_1(z)) - \mu_0 \approx \Delta x_1(z) \left(\frac{\partial \mu_{1a}}{\partial x_1} \right)_T^{-1}, \quad (3)$$

де μ_0 – хімічний потенціал за відсутності зовнішнього поля.

Не враховуючи квадратичні члени розкладу, з виразу (2) отримаємо наближений вираз для осмотичного тиску Δp_1 :

$$\Delta p_1 = \mu_{2a}(p_2) - \mu_{1a}(p_1) \left(\frac{\partial \mu_{1a}}{\partial p_1} \right)_T^{-1}, \quad (4)$$

де $\left(\frac{\partial \mu_1}{\partial p_1} \right)_T = v$ – молекулярний об'єм чистого

розчинника. Таким чином вираз для хімічного потенціалу відповідно можна записати:

$$\Delta \mu_{1a} = \frac{\Delta x_a(z)}{A n_0}. \quad (5)$$

Розглянемо плоскопаралельний шар. Введемо тривимірну систему координат так, щоб вісь z була перпендикулярна до площини системи, яка вздовж осей x та y прямує до нескінченності. Система обмежена вздовж осі z , тобто, z може змінюватись в межах від $-L$ до L .

Очевидно, що за відсутності поля, що діє в площині xOy концентрація розчину вздовж осей x та y не змінюється.

Поблизу стінок, що є напівпроникними, діють симетричні сили притягання (відштовхування), потенціал яких промодельємо наступними виразами [2]:

$$\begin{aligned} u_a(z) &= 2B_a e^{-kL} \operatorname{ch} kz, \\ u_b(z) &= 2B_b e^{-kL} \operatorname{ch} kz. \end{aligned} \quad (6)$$

В цьому випадку диференціальне рівняння відносно $\Delta x_a(z)$ має вигляд [3]:

$$\frac{d^2 \Delta x_a}{dz^2} - \kappa^2 \Delta x_a = A \kappa^2 (\lambda + 2B_{ba} e^{-kL} \operatorname{ch} kz) \quad (7)$$

де $\kappa = \frac{1}{\sqrt{2\xi}}$, ξ – радіус кореляції, $A = A(\xi)$ – стала, що залежить від параметрів критичної точки даного розчину, $B_{ba} = (B_b - B_a)N_l$, N_l – кількість частинок (вузлів) у шарі товщиною l . B_a та B_b – відповідні амплітуди ($B > 0$ відповідає силам відштовхування, $B < 0$ – притягання), k^{-1} – ефективний радіус дії цих сил, який вважаємо однаковим для частинок обох сортів, λ – стала, пов'язана з визначенням хімічного потенціалу обмеженої системи, значення якої може бути визначено з умові ізопериметризму. Рівняння (7) необхідно розв'язувати за умов трансверсальності та ізопериметризму [5]:

$$\vec{n}_s \cdot \vec{\nabla} \Delta x(\vec{r}) \Big|_s = 0 \quad (8)$$

$$\int_V d\vec{r}_1 \Delta x(\vec{r}_1) = 0 \quad (9)$$

Розв'язок рівняння (7) з врахуванням умов (8) та (9) в цьому випадку має вигляд:

$$\begin{aligned} \Delta x_a(z) &= 2B_{ba} e^{-kL} A \kappa^2 \times \\ &\times \left[-\frac{k}{\kappa} \frac{1}{k^2 - \kappa^2} \cdot \frac{\operatorname{sh} kL}{\operatorname{sh} \kappa L} \operatorname{ch} \kappa z + \right. \\ &\left. \frac{1}{k^2 - \kappa^2} \operatorname{ch} kz + \frac{1}{k \kappa^2 L} \operatorname{sh} \kappa L \right]. \end{aligned} \quad (10)$$

Одержаний просторовий розподіл концентрації бінарного розчину (10) дозволяє отримати залежність хімічного потенціалу μ_{1a}

від радіусу кореляції, що змінюється в межах від 10^{-7} м. до $0,999 \times 10^{-5}$ м.

$$\begin{aligned} \mu_{1a}(\xi) &= \frac{2B_{ba} e^{-kL}}{x_0} \frac{1}{k^2 \xi^2 - 1} \times \\ &\times \left[k \xi \left(\operatorname{sh} \frac{L}{\xi} \right)^{-1} \left(\operatorname{ch} \frac{L}{\xi} - \left(\frac{L}{\xi} \right)^{-1} \operatorname{sh} \frac{L}{\xi} \right) \operatorname{sh} kL + \frac{\operatorname{sh} kL}{kL} - \operatorname{ch} kL \right]. \end{aligned} \quad (11)$$

На рис. 2 наведено залежність хімічного потенціалу від радіусу кореляції модельного бінарного розчину для різних значень амплітуди B_i . Аналіз наведених на рис. 2 даних вказує на зменшення відносного хімічного потенціалу

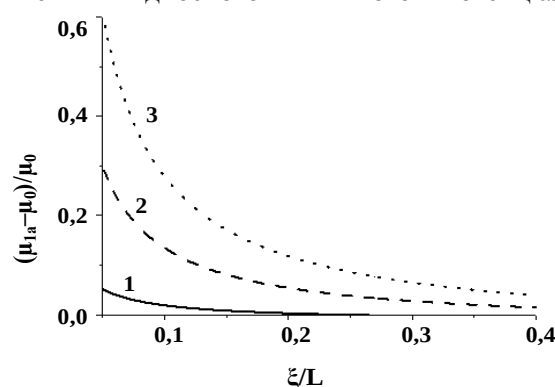


Рис. 2. Залежність відносного осмотичного тиску $(\mu_{1a} - \mu_0) / \mu_0$ бінарного розчину в плоскій порі з симетричним пристінковим потенціалом від радіусу кореляції ξ для системи з $k^{-1} = 3 \times 10^{-10}$ м, $L = 10^{-5}$ м та різних значень B : $B = 0,1 \times 10^{-24}$ Дж (1), $B = 5 \times 10^{-24}$ Дж (2), $B = 1 \times 10^{-24}$ Дж (3).

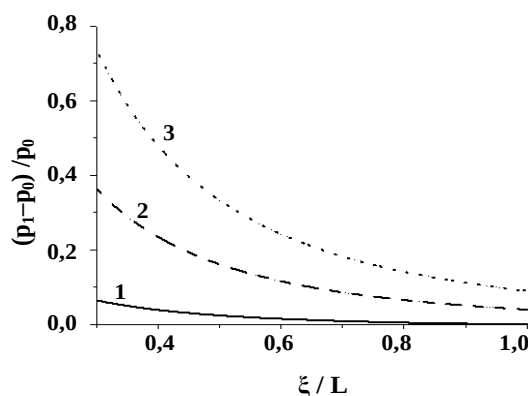


Рис. 3. Залежність відносного осмотичного тиску $(p - p_0) / p_0$ бінарного розчину в плоскій порі з симетричним пристінковим потенціалом від радіусу кореляції ξ для системи з $k^{-1} = 3 \times 10^{-10}$ м, $L = 10^{-5}$ м та різних значень B : $B = 0,1 \times 10^{-24}$ Дж (1), $B = 5 \times 10^{-24}$ Дж (2), $B = 1 \times 10^{-24}$ Дж (3).

$\frac{\mu_{1a} - \mu_0}{\mu_0}$ на 20 – 60% при зростанні радіусу кореляції ξ .

Одержана просторова залежність хімічного потенціалу (11) за допомогою формули (3) дозволяє отримати вираз залежності осмотичного тиску $p(\xi)$ від радіусу кореляції ξ :

$$p_1(\xi) = \frac{2B_{BA}e^{-kL}}{Vx_0} \frac{1}{k^2\xi^2 - 1} \times \left[k\xi \left(sh \frac{L}{\xi} \right)^{-1} \left(ch \frac{L}{\xi} - \left(\frac{L}{\xi} \right)^{-1} sh \frac{L}{\xi} \right) sh kL + \frac{sh kL}{kL} - ch kL \right] \quad (12)$$

На рис. 3 наведено залежність відносного осмотичного тиску від радіусу кореляції для систем з різними значеннями сталих B . Аналіз

наведених на рис. 3 даних вказує на те, що зі збільшенням радіусу кореляції ξ відносний осмотичний тиск для даної системи зменшується на величину порядку 30 – 50%.

Результати та висновки

Побудовано вирази для обчислення осмотичного тиску бінарного розчину, що дозволяють проводити обчислення в широкому інтервалі зміни термодинамічних параметрів, включаючи окіл критичної точки розшарування.

Встановлено, що з наближенням до критичної точки розшарування бінарного розчину відносна зміна хімічного потенціалу та осмотичного тиску досягає величини 20 – 50 %.

Список використаних джерел

1. Булавін Л.А. Фізика фазових переходів / Л.А. Булавін, В.М. Сисоєв. – К.: ВПЦ КНУ, 2010. – 400 с.
2. Булавін Л.А. Просторовий розподіл концентрації бінарного розчину в полоскому шарі поблизу критичної точки рідина-пара / Л.А. Булавін, Д.А. Гаврюшенко, В.М. Сисоєв, О.О. Старовойтов. // УФЖ. – 2003. –Р. 48, №10. – С. 1072–1074..
3. Булавін Л. А.. Концентрація рідини в плоскопаралельній порі з неідеальною стінкою в гравітаційному полі / Булавін Л.А., Гаврюшенко Д.А., Сисоєв В.М // Журн. Фіз. Хімії. – 2004. – Р. 78, № 11. – С. 1814–1817.
4. Иванов Д.Ю. Критические явления в чистых жидкостях / Д.Ю. Иванов. – М: Физматлит, 2003. – 248с.
5. Лаврентев М.А. Курс вариационного исчисления / М.А. Лаврентев, Л. А Люстерник. – М.–Л.: ГИТТЛ, 1950. – 133 с
6. Лакоза Е.Л. Критические флуктуации во внешних полях / Е.Л. Лакоза, В.М. Сысоев, А. В. Чалый // Журн. эксперим. и теор. физики. – 1973. – Т. 65, №2(8). – С. 605–616.
7. Семюхин И.А. Физическая химия. Учебник / И. А. Семюхин. – Москва: МГУ, 2001. –.272с.

References

1. BULAVIN, L., SYSOEV, V. (2010) *Physics of phase transitions*. K: VPC KNU.
2. BULAVIN, L., GAVRYUSHENKO, D., SYSOEV, V., STAROVOITOV, O. (2003) Spatial distribution of components of a binary mixture in a bounded System. *Ukr. Journ. of Phys.* 48(10). p.1072–1074.
3. BULAVIN, L., GAVRYUSHENKO, D., SYSOEV, V. (2004) Density profile of a fluid in a plane-parallel pore with nonideal walls in a gravitational field. *Journ. Phys. Chem.* 78(11) p.1814–1817.
4. IVANOV, D. (2003) *Critical phenomena in pure liquids*. M:Phismatlit.
5. LAVRENTEV, M., LYUSTERNIK, L. (1950) *The course of the calculus of variations*. M.–L.: HITTL.
6. LAKOZA, E., SYSOEV, V., CHALYI, A., (1973) Critical fluctuations in external fields. *Journ. Exper. and Theor. of Phys.* 2(8). p.605–616.
7. SEMIOHIN, I. (2001) *Physical chemistry*. Moskva: MHU.

Надійшла до редколегії 28.05.15

УДК 538.915

Горкавенко Т. В.¹, к.ф.-м.н., м.н.с.,
Плющай І. В.², к.ф.-м.н., доц.,
Макара В. А.³, д.ф.-м.н., проф.

T. V. Gorkavenko¹, PhD,
I. V. Plyushchay², PhD
V. A. Makara³, Dr. Sci., Prof.,

Першопринципний розрахунок рівноважного положення домішки кисню в кремнії

Ab initio calculation of equilibrium position of oxygen impurity in silicon

^{1,2,3}Київський національний університет імені
Тараса Шевченка, 01601, м. Київ, вул.
Володимирська 64/13, e-mail: tvgorka@gmail.com

^{1,2,3}Taras Shevchenko National University of Kyiv,
01601, Kyiv, Volodymyrska st. 64/13,
e-mail: tvgorka@gmail.com

Проведено *ab initio* розрахунок рівноважного положення домішкового атома кисню в надкомірці з 64 атомів кремнію. Розрахунок проведено методом функціоналу густини в узагальненому градієнтному наближенні за допомогою пакета програм ABINIT. Розрахований кут квазімолекули Si-O-Si дорівнює $\sim 136^\circ$. Представлені та проаналізовані електронні спектри надкомірки з 64 атомів кремнію, що містить домішковий атом кисню в тетраедричній порі та у рівноважному положенні. Обговорюється можливість формування магнітних моментів на дефектах. Показано, що частково заповнена домішкова підзона, яка спостерігається для випадку кисню в центрі тетраедричної пори, може призводити до формування магнітного моменту.

Ключові слова: електронна структура, кремній, домішка кисню.

Ab initio calculation of the electronic spectra of a supercell composed of 64 Si atoms with point defect O are presented. The oxygen impurity is considered in the interstitial position. The concentration of oxygen is $\sim 1.5\%$ (10^{20} cm^{-3}). The density functional theory in the general gradient approximation using the software package ABINIT has been used for numerical calculation. The equilibrium position of oxygen in silicon has been investigated. We obtained that the angle of Si-O-Si quasimolecules is $\sim 136^\circ$. The electronic spectra of a supercell composed of 64 Si atoms with oxygen impurity in the tetrahedral pore and in the equilibrium position are presented and analyzed. In general, the obtained spectra corresponds to the spectra of pure Si, the main different is formation narrow impurity peak nearly Fermi level. Analysis of local spectra allows to associate a narrow peak in the vicinity of the Fermi level with the electronic states of impurity atoms O. The possible formation of magnetic moments on impurity atoms are discussed. It is shown that partially filled impurity subzone observed in the case of oxygen in the tetrahedral pore can lead to the formation of the magnetic moment.

Key Words: electronic structure, silicon, oxygen impurity.

Статтю представив член-кор. НАН України, д.ф.-м.н., проф. Макара В.А.

Вступ

Технологічною основою сучасної мікроелектроніки та приладобудування є монокристалічний кремній. Електрофізичні, механічні та технологічні властивості кристалічного кремнію значною мірою зумовлені наявністю домішок, зокрема кисню – домінуючої домішки в монокристалах кремнію. Саме тому вивчення впливу домішки кисню на властивості кремнію залишається актуальною проблемою та привертає увагу багатьох дослідників.

На даний час існують дані про незвичні магнітні (навіть феромагнітні) властивості монокристалів кремнія, які пов'язують з

кисневмісними дефектами. Наприклад, у роботі [1] отримано нелінійні польові залежності магнітної сприйнятливості $\chi(H)$ монокристалів кремнію при $T = 80 \text{ K}$, які корелюють з концентрацією цих дефектів. Згідно [1], таку феромагнітну поведінку магнітної сприйнятливості можна пояснити прямою обмінною взаємодією між кисневмісними дефектами в областях так званих мікрофлуктуацій концентрації кисню з концентрацією тих самих дефектів – 10^{18} – 10^{21} cm^{-3} .

У роботі [2] зафіксовано феромагнітний сигнал від зразків поруватого кремнію (аж до температури 570 K) і вперше, наскільки нам відомо, обговорювалася можливість виникнення

магнетизму в системах дефектів високої густини в кремнії. За допомогою електронного парамагнітного резонансу була зафіксована наявність парамагнітних центрів, пов'язаних з обірваними зв'язками.

Питання наявності або відсутності магнітних моментів на домішках кисню також цікаве з точки зору нещодавно відкритого магнітопластичного ефекту на монокристалах Si [3]. Послідовного пояснення цього ефекту досі не існує, але деякі дослідники пов'язують вплив магнітного поля на механічні характеристики кристалів Si саме через відповідні реакції дефектів, що мають магнітний момент.

Попередньо проведені розрахунки [4] показали можливість формування магнітних моментів на домішках кисню, але в цих роботах не було проведено дослідження рівноважного положення кисню.

З літературних джерел відомо, що атоми кисню займають в кристалічній ґратці кремнію міжвузлове положення, утворюючи з найближчими атомами кремнію ланцюг Si-O-Si [5]. Квазімолекула Si-O-Si паралельна напрямку [111] і є дещо вигнутою, утворюючи кут приблизно 150° (рис. 1). Беручи до уваги симетрію кристалічної ґратки, атом кисню може займати шість еквівалентних положень відносно двох найближчих атомів кремнію. Причому енергетичний бар'єр переорієнтації між цими положеннями настільки малий, що навіть при кімнатній температурі атом кисню досить швидко змінює своє положення. Зміна положення атома кисню не пов'язана з його дифузійним стрибком, що вимагає розриву зв'язку Si-O-Si, як це відбувається при дифузії.

З огляду на вищезазначене метою даної роботи було встановити рівноважне положення кисню в монокристалах кремнію, а також

особливості електронних станів домішок кисню та проаналізувати можливість виникнення на них локалізованих магнітних моментів. Для цього було проведено *ab initio* розрахунок електронних спектрів надкомірки кремнію з 64 атомів, кисень розглядався у міжвузловому положенні. Концентрація кисню $\sim 1,5\%$ (10^{20} см^{-3}). Розрахунок проводився методом функціоналу густини [6] в узагальненому градієнтному наближенні [7] за допомогою пакета програм ABINIT [8].

Результати та обговорення

Спочатку атом кисню ми розташували в центрі тетраедричного міжвузля в середині надкомірки з 64 атомів кремнію. В даному положенні атом кисню мав координати (000). Після цього ми провели чисельний експеримент по розрахунку рівноважного положення домішкового атома кисню у надкомірці кремнію з 64 атомів, рухаючи його з кроком 0,01 ат.од. до найближчих двох атомів кремнію, перпендикулярно напрямку [111].

Залежність повної енергії (E_{tot}) надкомірки з 64 атомів кремнію, що містить кисень у міжвузловому положенні, від зміщення кисню показано на рис. 2. Зміщення атома кисню ми відраховували від центра тетраедричної пори, тобто від початкового положення кисню. Мінімум цієї кривої знаходиться на відстані 0,05 ат.од. від центра тетраедричного міжвузля, або на відстані 0,066 ат.од. від двох найближчих атомів кремнію. Згідно з результатами розрахунку, повна енергія атома O у рівноважному положенні на $\sim 0,05 \text{ eV}$ менша за енергію атома O у тетраедричній порі. Останнє підтверджує, що атоми кисню в кристалічній ґратці кремнію займають міжвузлове положення, утворюючи з сусідніми атомами кремнію ланцюг

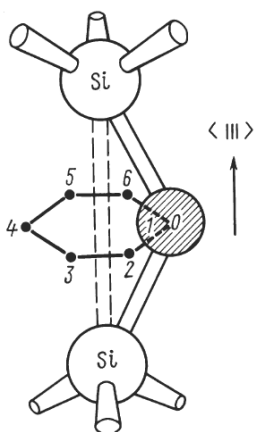


Рис.1. Структурна модель квазімолекули Si-O-Si.

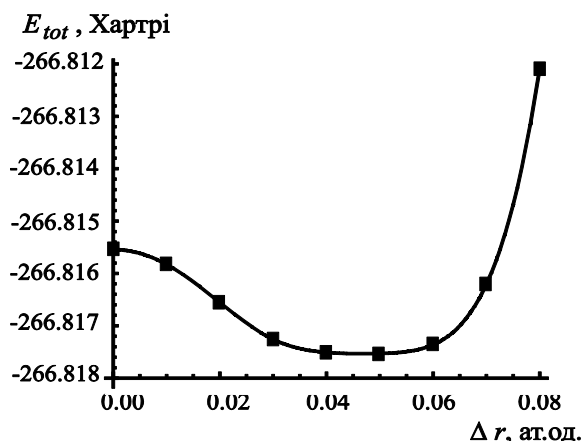


Рис. 2. Залежність повної енергії від положення атома кисню.

Si-O-Si. На основі першопринципних розрахунків ми встановили, що кут квазімолекули Si-O-Si дорівнює $\sim 136^\circ$, отримане значення є дещо меншим від раніше передбаченого $\alpha(\text{Si-O-Si}) = 150^\circ$ [5] на основі модельних уявлень.

На рис. 3 представлено розраховану нами енергетичну залежність густини електронних станів чистого кремнію (а), надкомірки з 64 атомів кремнію, що містить один міжвузловий атом кисню в центрі тетраедричної пори (б), та у рівноважному положенні (в). Положення рівня Фермі показано вертикальною лінією. У цілому, отримані спектри відповідають спектру чистого Si, що свідчить про адекватність розрахунку – 1,5% домішок атомів О не змінюють значно електронну структуру кристалів Si. Єдиною якісною відмінністю є формування вузького додаткового піка в забороненій зоні безпосередньо над валентною зоною. Напівширина цього піка дорівнює $\sim 0,11$ eV для положення атома кисню в центрі тетраедричної пори та $\sim 0,07$ eV для рівноважного положення кисню. Висота домішкового піку для випадку розташування кисню в рівноважному положенні у $\sim 1,6$ рази менше, ніж для випадку кисню у

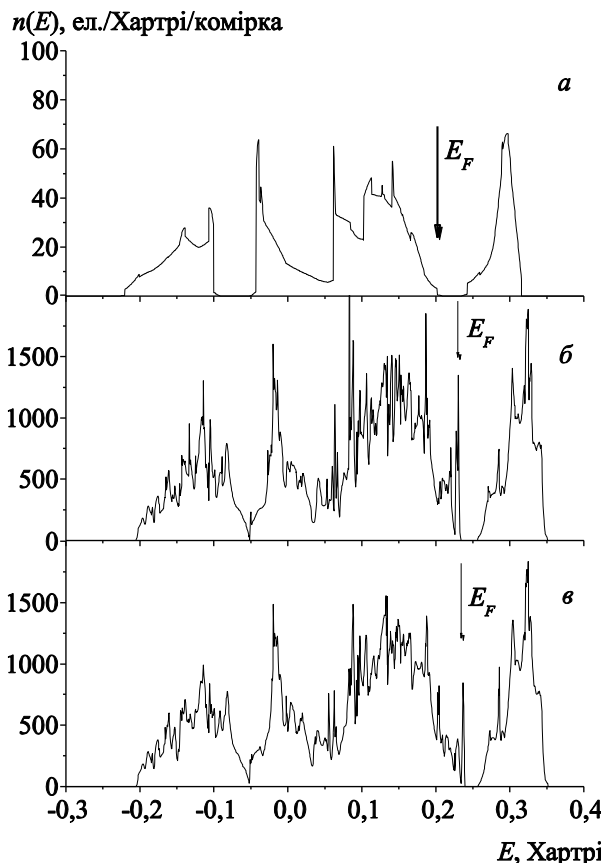


Рис. 3. Енергетична залежність густини електронних станів: а – кристалічний кремній; б – надкомірки з 64 атомів Si та атома кисню в тетраедричній порі; в – надкомірки з 64 атомів Si та атома кисню в рівноважному положенні.

тетраедричній порі. Рівень Фермі по мірі досягнення рівноважного положення дещо підвищується.

На рис. 4 наведено локальні електронні спектри домішкового атома О в центрі тетраедричної пори (а) надкомірки з 64 атомів Si та у рівноважному положенні (б), а також найближчого сусіднього атома Si у квазімолекулі Si-O-Si (в). Аналіз локальних спектрів дозволяє зв'язати вузький пік в околі рівня Фермі саме зі станами домішкових атомів О. Вказаний гострий пік в локальних спектрах усіх інших атомів розглянутої надкомірки виявляється у десятки разів меншим. Отже, частково заповнений гострий пік, який спостерігається в електронному спектрі надкомірки з 64 атомів Si з точковим дефектом О, пов'язаний з домішковими атомами О.

Проаналізуємо результати наших розрахунків з точки зору можливості виникнення зонного магнетизму. Утворення домішкової підзони в околі забороненої зони є типовим для напівпровідників. Згідно з критерієм Стонера, для виникнення магнітного впорядкування необхідна наявність вузького напівзаповненого

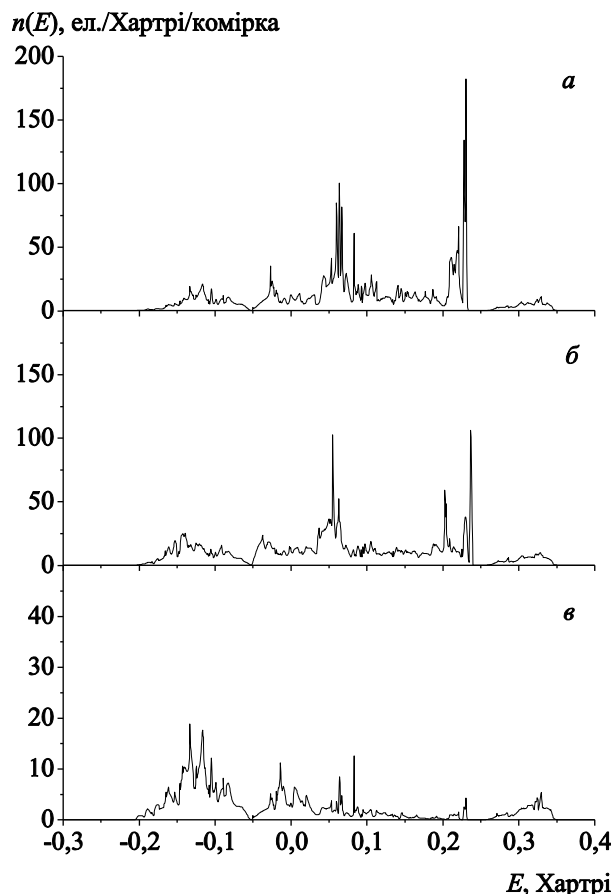


Рис. 4. Енергетична залежність густини електронних станів: а – атома кисню в центрі тетраедричної пори; б – атома кисню у рівноважному положенні; в – атома Si, що є першим сусідом домішки кисню.

піка поблизу рівня Фермі. Як раз таким і є домішковий субпік (див. б, в на рис. 3), пов'язаний з наявністю домішкових атомів О у центрі тетраедричної пори (б) та у рівноважному положенні (в). В обох випадках домішкова підзона електронних станів кисню у міжвузловому стані дуже вузька (плоска) і для випадку О у тетраедричній порі частково заповнена, а у випадку кисню в рівноважному положенні майже незаповнена, тобто рівень Фермі потягнута в область початку цього домішкового піку. Частково заповнена вузька домішкова підзона у випадку розміщення атома кисню у тетраедричній порі, згідно з [4], може призвести до виникнення зонного магнетизму, як це відбувається в графітових структурах. Як показали проведені нами розрахунки, у випадку

коли домішковий кисень знаходиться точно у рівноважному положенні вищезазначений ефект стає менш виразним.

Висновки

Проведений нами першопринципний розрахунок рівноважного положення домішкового атома кисню в кремнії показав, що кут квазімолекули Si-O-Si становить $\sim 136^\circ$. Аналіз електронних спектрів надкомірки кремнію з домішкою кисню в центрі тетраедричної пори і в рівноважному положенні виявив можливість формування магнітних моментів на домішковому кисні.

Список використаних джерел

1. Неймаш В.Б. Процеси трансформації станів домішки кисню в монокристалах кремнію при високоенергетичному опроміненні та термообробках : дис. д-ра фіз.-мат. наук / В.Б. Неймаш – Київ, 2007. – 325 с.
2. Laiho R. Magnetic Properties of Light-Emitting Porous Silicon / R. Laiho, E. L'ahderanta, L. Vlasenko et al. // J. of Luminescence. – 1993. – V. 57. – P. 197–201.
3. Makara V.A. Effect of weak magnetic field on structural arrangement of extrinsic oxygen atoms and mechanical properties of silicon monocrystals / V.A. Makara, L.P. Steblenko, Yu.L. Kolchenko et al. // Semicond. Phys., Quantum Electron. and Optoelectron. – 2006. – V. 9, No 2. – P. 1–3.
4. Plyushchay I.V. Ab initio calculation of magnetic interaction between edge dislocation and oxygen impurity in silicon / I.V. Plyushchay, O.I. Plyushchay, V.A. Makara // Metallofizika i Noveishie Tekhnologii. – 2014. – V. 36, No 5, P. 589–596.
5. Бабич В.М. Кислород в монокристаллах кремнія / В.М. Бабич, Н.И. Блецкан, Е.Ф. Венгер. – Киев: Интерпрес ЛТД, 1997. – 240 с.
6. Gonzea X. ABINIT: First-principles approach of materials and nanosystem properties / X. Gonzea, B. Amadond, P.-M. Anglade et al. // Comput. Phys. Com. – 2009. – 180. – P. 2582–2615.
7. Perdew J. P. Generalized gradient approximation made simple / J. P. Perdew, K. Burke, M. Ernzerhof // Phys. Rev.Lett. – 1996. – 77. – P. 3865–3868.
8. Abinit (2014) [Online] Application. Available from: <http://www.abinit.org> [Accessed: 2004-2015].

References

1. NEYMASH, V. B. (2007) Processes of oxygen impurity state transformation in the silicon single crystals under high energy irradiation and heat treatments (Doctoral dissertation).
2. LAIHO, R. et al. (1993) Magnetic Properties of Light-Emitting Porous Silicon. *Journal of Luminescence*. 57. p.197-201.
3. MAKARA, V. A. et al. (2006) Effect of weak magnetic field on structural arrangement of extrinsic oxygen atoms and mechanical properties of silicon monocrystals. *Semiconductor Physics, Quantum Electronics and Optoelectronics*. 9 (2). p.1-3.
4. PLYUSHCHAY, I. V., PLYUSHCHAY, O. I. and MAKARA, V. A. (2014) Ab initio calculation of magnetic interaction between edge dislocation and oxygen impurity in silicon. *Physics of Metals and Advanced Technologies*. 36 (5), p. 589-596.
5. BABYICH, V. M., BLETSKAN, N. I. and VENGER, E. F. (1997) Oxygen in the silicon single crystals. Kiev: Interpress LTD.
6. GONZEA, X. et al. (2009) ABINIT: First-principles approach of materials and nanosystem properties. *Computer Physics Communications*. 180. p.2582-2615.
7. PERDEW, J. P., BURKE, K. and ERNZERHOF, M. (1996) Generalized gradient approximation made simple. *Physical Review Letters*. 77. p.3865-3868.
8. Abinit (2014) [Online] Application. Available from: <http://www.abinit.org> [Accessed: 2004-2015].

Надійшла до редколегії 20.04.15

УДК 535.324, 539.2

Гоцульський В.Я.^{1,2}, к.ф.-м.н., доц.,
Свечнікова О.С.¹, к.ф.-м.н.,
Стиранець М.В.², аспірант

Тонка структура низькоконцентраційного піку розсіювання світла водними розчинами етанолу

¹Київський національний університет імені Тараса Шевченка, 01601, м. Київ, вул. Володимирська, 64/13,
e-mail: svechnikova@univ.kiev.ua

²Одеський національний університет імені І. І. Мечникова, фізичний факультет, вул. Пастера 42, Одеса, Україна e-mail: ²vygot@onu.edu.ua

V.Ya.Gotsulskiy^{1,2}, PhD, Asoc. Prof.,
O.S. Svechnikova¹ PhD,
M.V. Stiranets², CPhD stud.

The fine structure of low concentration peak light scattering aqueous solutions of ethanol

¹Taras Shevchenko National University of Kyiv, 01601, Kyiv, Volodymyrska str., 64/13,
e-mail: svechnikova@univ.kiev.ua

²I.I.Mechnikov National University of Odessa, Faculty of physics, 42, Pastera Str., Odesa 65001, Ukraine
e-mail: ²vygot@onu.edu.ua

Робота присвячена експериментальному дослідженню концентраційної залежності низькоконцентраційного піку інтенсивності розсіяного світла у водних розчинах етанолу в околі особливої точки $0.05 < x < 0.1$ (x - мольна частка етанолу у розчині). Увага приділяється околу особливої точки, яка визначається як точка перетину кривих концентраційних залежностей контракції при різних температурах або як точка, в якій спостерігається аномальний максимум інтенсивності молекулярного розсіювання світла. Показано, що рівноважні властивості розчинів встановлюються на протязі тривалого часу (тиждень та більше), при цьому аномальний максимум розпадається на два. Положення цих піків корелює з даними аналізу контракції водних розчинів етанолу таким чином, що вони розташовуються симетричне відносно границі області де суттєві квазікристалічні властивості води. Локальна в'язкість розчинів з максимальним розсіюванням визначалась динамічним розсіюванням світла на завісі у розчинах латексних частинок розміром 60 нм. Поведінка концентраційної залежності локальної в'язкості суттєво відрізняється для піків тонкої структури аномального розсіювання світла.

Ключові слова: водні розчини етанолу, низькоконцентраційне розсіювання світла, динамічне розсіювання світла, особлива точка.

This paper is concerned with experimental study of the concentration dependence of the scattered light's low-concentration intensity peak in water-ethanol solutions in the neighborhood of the peculiar point $0.05 < x < 0.1$ (x is a mole fraction of ethanol in solution). We pay attention to the neighborhood of the peculiar point defined as intersection point of the contraction's concentration dependence curves taken at different temperatures. Also peculiar point can be defined as the point of an anomalous maximum intensity of molecular light scattering. It is shown that the equilibrium properties of solutions are established for a long time (a week or more), while the maximum anomalous splits into two. The position of these peaks correlated with the contraction to the analysis of aqueous solutions of ethanol. They are arranged symmetrically with respect to border areas where significant quasicrystal properties of water. Local viscosity of solutions with maximum scattering defined by dynamic light scattering on suspension of latex particles 60 nm differs markedly for these peaks.

Key Words: aqueous solutions of ethanol, low concentration of light scattering, dynamic light scattering, a singular point.

Статтю представив академік НАН України, д.ф.-м.н., проф. Булавін Л. А.

Молекулярне розсіювання світла стало одним з основних засобів вивчення властивостей розчинів ще в 50-ті роки минулого сторіччя і залишалося таким протягом усіх наступних 60-ти

років. Це пов'язано з тим, що розсіювання світла визначається мікроскопічними оптичними неоднорідностями різного масштабу, які можуть вимірюватися шляхом зміни кутів розсіювання, а

характер їх часової залежності - шляхом вимірювання частоти розсіяного світла [1-5].

В [1,6] було показано, що у водно-спиртових розчинах залежність повної інтенсивності молекулярного розсіяння світла від концентрації має два максимуми: один з них - основний при концентрації спирту $x \sim 0,2 \div 0,5$ і другий, більш різкий максимум, при концентрації спиртів $x \sim 0,01 \div 0,1$ (рис.1).

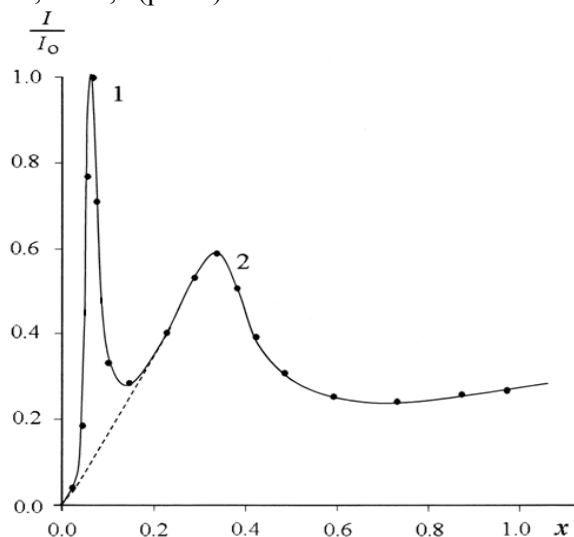


Рис. 1. Характерна концентраційна залежність інтегральної інтенсивності молекулярного розсіяння світла у водних розчинах спиртів [1-4, 6]: 1 – низькоконцентраційний максимум, 2 – основний максимум.

Основний максимум зумовлений розсіюванням світла на флуктуаціях концентрації. Його положення і амплітуда чітко корелюють з розрахунками, проведеними в рамках термодинамічної теорії флуктуацій, а поведінка низькоконцентраційного максимуму є аномальною: він зменшується зі збільшенням температури і зміщується при цьому в область великих концентрацій. При цьому для водних розчинів етанолу і метанолу основний максимум концентраційного розсіювання світла практично відсутній, оскільки показники заломлення води і спирту близькі [7]. Експерименти по дослідженню концентраційної залежності теплоємності водно-спиртових розчинів показали наявність максимуму там же, де знаходиться додатковий максимум молекулярного розсіяння світла. Вважається, що це свідчить про наявність структурних змін в розглянутих системах [1, 6].

Підкреслимо, що для отримання відтворюваних результатів, необхідно тривале витримання розчину після його приготування [8,9].

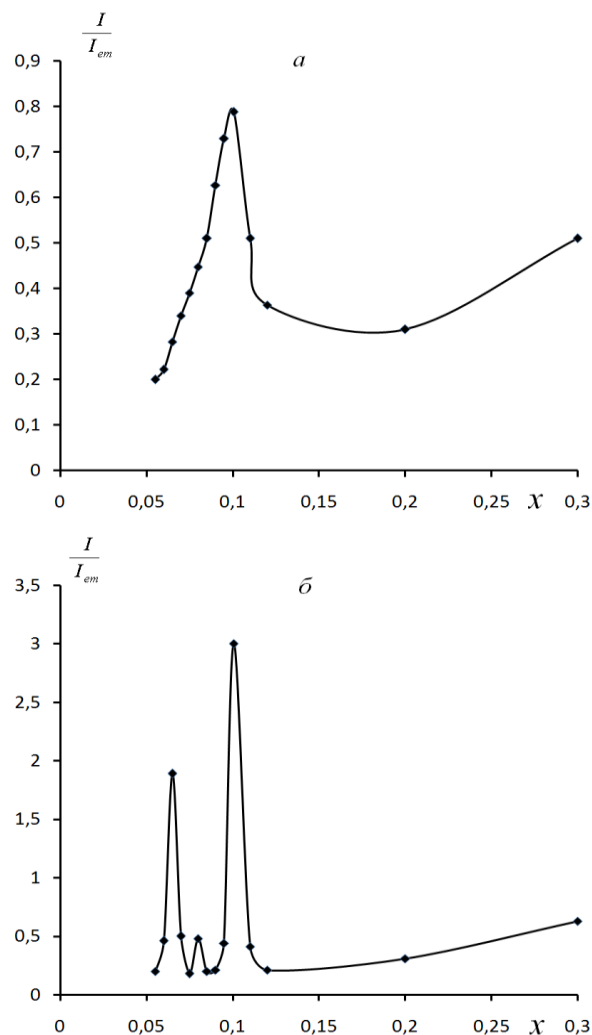


Рис. 2. Залежність мутності (turbidity) I/I_{em} водних розчинів етанолу від його мольної частки, яка отримана відносним методом: а) в першу добу після змішання розчину, б) через тиждень після приготування розчину. Суцільна лінія – результат згладжування. I, I_{em} – інтенсивності розсіяного світла на розчинах вода-етанол і на еталону завісі латекса відповідно.

У цих роботах було показано, що рівноважний стан у водно-спиртових розчинах досягається протягом тривалого часу (від доби до декількох тижнів) після їх приготування. Ця обставина також вказує на те, що перерозподіл молекул спирту в розчині зумовлений не простим дифузійним механізмом, а має більш складний характер, який пов'язаний з виникненням і руйнуванням зародків метастабільної фази.

Нами була поставлена задача по дослідженню зміни характеристик аномального піку розсіяння світла у водних розчинах етанолу з часом. Для цього забезпечувалось одночасне

приготування у однакових умовах 20 зразків у концентраційному діапазоні $0,05 < x < 0,1$ мольної частки етанолу у водному розчині. Як правило, при дослідженні аномального максимуму розсіяння світла автори обмежувались констатуванням його наявності і на його концентраційний діапазон приходилось не більше 5 точок.

Навіть візуальне спостереження всього ряду зразків при сталій температурі 22°C дозволило виявити те, що протягом тижня лише у двох зразках з концентраціями $x = 0,065$ і $x = 0,095$ мольної частки утворювався стан з розсіянням світла по інтенсивності близький до опалесценції. Результати нефелометрії, які наведені на рис. 2, показують, що концентраційна залежність мутності розчину (turbidity), що визначена в початковий момент часу (рис. 2 а) і має один максимум, через два тижні трансформується у сукупність піків (рис. 2 б). Їх положення повністю узгоджується з тим, що при цій температурі гранична концентрація області, де суттєві квазікристалічні властивості води у розчині, приходить на концентрацію $x = 0,08$, яка була отримана аналізом залежностей контракції [10,11].

На рис. 3 представлено результати вимірювань часів кореляції світла, розсіяного на латексних частинках з паспортним розміром 60 нм, які були зважені в тих же водних розчинах етанолу, що були використані при попередніх експериментах.

Автокореляційна функція інтенсивності розсіяного світла $I(t)$ на рухомих неоднорідностях має вигляд [3-5]

$$G^2(\tau) = \langle I(t)I(t+\tau) \rangle = a + b \cdot e^{-\frac{2\tau}{\tau_c}} \quad (1)$$

де τ_c – характерний час флуктуацій інтенсивності розсіяного світла, a , b – константи. У разі, якщо неоднорідності в розчині є броунівськими частинками, використовуючи формулу Ейнштейна, можна пов'язати характерний час τ_c з коефіцієнтом дифузії D та хвильовим вектором розсіяння q виразами

$$\tau_c = \frac{1}{Dq^2}, \quad q = \frac{4\pi n}{\lambda} \sin(\Theta/2) \quad (2)$$

n – показник заломлення, λ – довжина хвилі зонduючого випромінювання, Θ – кут розсіяння.

На рис. 3 в концентраційній залежності часів кореляції видно присутність двох піків. Їх концентраційне положення збігається з піками інтенсивності розсіяння (рис. 2).

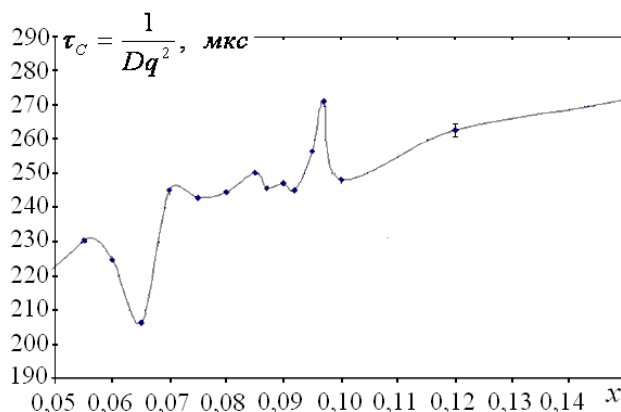


Рис. 3. Залежність від мольної частки етанолу часів кореляції інтенсивності світла, що розсіяне на зависі частинок латексу розміром 60 нм у розчинах вода-етанол. Температура 22°C .

Особливість їх в тому, що в околі одного при $x = 0,065$ помітно зменшення часу кореляції, тобто збільшення рухливості частинок латексу, а при $x = 0,095$ мольної частки – зменшення рухливості. Це можна трактувати як зменшення або збільшення локальної в'язкості на наномасштабах, які визначають рух броунівський частинки розміром десятки нанометрів. Раніше розмір неоднорідностей в області аномального розсіювання світла оцінювався як мезомасштабний: 100-1000 нм [9,12].

Висновки

1. Рівноважний стан у водних розчинах етанолу у околі особливої точки $0,05 < x < 0,1$ мольної частки досягається на протязі тижня.
2. При утворенні рівноважного стану аномальний низькоконцентраційний пік розсіяння світла приймає складну форму – сукупність піків.
3. Положення піків тонкої структури аномального піку розсіяння світла корелює з результатами аналізу контракції водних розчинів етанолу та наявністю області, де суттєві квазікристалічні властивості води.
4. Динамічне розсіювання на частинках латексу у розчинах, що відповідають максимумам тонкої структури піку розсіяння, показує суттєву відмінність локальної в'язкості цих розчинів.

Список використаних джерел

1. Вукс М.Ф. Рассеяние света в газах, жидкостях и растворах / М.Ф. Вукс. – Ленинград: ЛГУ, 1977. – 320с.
2. Фабелинский И.Л. Молекулярное рассеяние света / И.Л. Фабелинский.- Москва: Высшая школа, 1965. – 512с.
3. Pecora R. (ed) *Dynamic Light Scattering. Applications of Photon Correlation Spectroscopy* / Ed. R. Pecora. - New York: Plenum Press, 1985. 420p.
4. *Photon Correlation and Light Beating Spectroscopy* / Eds H.Z.Cummins, E.R. Pike. [autor H.Z.Cummins]. - New York: Plenum Press, 1974. - 285-330.
5. Pike E.R. *Light Scattering and Photon Correlation Spectroscopy* // E.R. Pike, J.B. Abbiss. - New York: Kluwer Academic Publishers, 1997. – 470 p.
6. Анисимов М.А. Критические явления в жидкостях и жидких кристаллах / М.А. Анисимов. – Москва: Наука, 1987. – 271с.
7. Вукс М.Ф. Рассеяние света растворами пропиловых спиртов в воде и тяжелой воде / М.Ф. Вукс, Л.В. Шурупова // Журнал структурной химии. – 1971. Т. 12. – С. 712-713.
8. Булавін Л.А. Особливості встановлення рівноважного стану в розбавлених водних розчинах гліцерину / Л.А. Булавін, В.Я. Гоцульський, В.Є. Чечко // УФЖ. – 2014. – Т. 59. – №7. – С. 689-695.
9. Chechko V.E. On the nature of relaxation processes in dilute water-glycerol solutions /, V.Ya. Gotsulskiy, V.G. Zaremba // *Journal of Molecular Liquids*. - 2003. - V. 105/2-3.- P. 211-214.
10. Gotsulskiy V.Ya. Features of the Temperature and Concentration Dependences of the Contraction of Aqueous Solutions of Ethanol / V.Ya. Gotsulskiy, N.P. Malomuzh, V.E. Chechko // *Russian Journal of Physical Chemistry A*. – 2013. – V. 87. – No. 10. – P. 1638-1644.
11. Gotsulskiy V.Ya. Particular Points of Water-Alcohol Solutions / V.Ya. Gotsulskiy, N.P. Malomuzh, V.E. Chechko // *Russian Journal of Physical Chemistry A*. – 2015. - V. 89. - No. 2. - P. 207–213.
12. Subramanian D. Resolving the Mystery of Aqueous Solutions of Tertiary Butyl Alcohol / D. Subramanian, M.A. Anisimov // *Journal of Physical Chemistry B*. – 2011. – V. 115. – P. 9179-9183.

References

1. VUKS, M. (1977) *Light scattering in gases, liquids and solutions*. Leningrad: LSU.
2. FABELINSKII, I. (1965) *Molecular scattering of light*. Moskva: Vusshaya Shkola.
3. PECORA, R. (ed.) (1985) *Dynamic Light Scattering. Applications of Photon Correlation Spectroscopy*. New York: Plenum Press.
4. CUMMINS, H. (1974) *Photon Correlation and Light Beating Spectroscopy*. New York: Plenum Press.
5. PIKE, E., ABBISS, J. (1997) *Light Scattering and Photon Correlation Spectroscopy* New York: Kluwer Academic Publishers.
6. ANISIMOV, M. (1987) *Critical phenomena in liquids and liquid crystals*. Moskva: Nauka.
7. VUKS, M., SHURUPOVA, L. (1971) Light scattering propyl alcohol solutions in water and heavy water *Zhurnal structurnoy Khimii*. 12 (4). p.712-713.
8. BULAVIN, L., GOTSULSKIY V., CHECHKO V. (2014) Peculiarities in the establishment of equilibrium state in diluted aqueous *Ukrainian Journal of Physics*. 59 (7). p.689-695.
9. CHECHKO, V., GOTSULSKIY, V., ZAREMBA, V. (2003) On the nature of relaxation processes in dilute water-glycerol solutions. *Journal of Molecular Liquids* 105/2-3. p.211-214.
10. GOTSUL'SKII, V., MALOMUZH, N., CHECHKO V. (2013) Features of the Temperature and Concentration Dependences of the Contraction of Aqueous Solutions of Ethanol. *Russian Journal of Physical Chemistry A*. 87 (10). p.1638-1644.
11. GOTSULSKIY, V., MALOMUZH, N., CHECHKO, V. (2015) Particular Points of Water-Alcohol Solutions. *Russian Journal of Physical Chemistry A*. 89 (2). p.207–213.
12. SUBRAMANIAN, D., ANISIMOV, M. (2011) Resolving the Mystery of Aqueous Solutions of Tertiary Butyl Alcohol. *Journal of Physical Chemistry B*. 115. p.9179-9183.

Надійшла до редколегії 26.04.15

УДК 539.213

Лисов В.І.¹, д.ф.м. н., проф.,
Цареградська Т.Л.², к.ф.м.н., доц.,
Саєнко Г.В.³, інж.,
Турков О.В.⁴, інж

V.I. Lysov¹, Dr.Sc., Prof.,
T.L. Tsaregradskaya², PhD.,
G.V. Saenko³, eng.,
O.V. Turkov⁴, eng.

Підвищення термічної стабільності аморфних сплавів за рахунок процесу ініційованої висхідної дифузії

The increasing of thermal stability amorphous alloys is due to initiate of ascending diffusion process

^{1,2,3,4} Київський національний університет імені
Тараса Шевченка, 01601, м. Київ,
вул. Володимирська 64/13,
e-mail: ²tsar_grd@ukr.net

^{1,2,3,4} Taras Shevchenko National University of Kyiv,
01601, Kyiv, Volodymyrska st. 64/13,
e-mail: ²tsar_grd@ukr.net

За допомогою термодинамічних розрахунків проведена оцінка можливості протікання процесу висхідної дифузії в аморфних сплавах системи Fe-B. Показано, що при ізотермічному відпалі аморфних сплавів системи Fe-B в області температур $T = 0,6T_{\text{ев}}$ існує можливість розсмоктування вмерзених центрів кристалізації за рахунок висхідної дифузії, що у свою чергу може підвищити термічну стабільність аморфного сплаву. Даний висновок підтверджено експериментально за допомогою дилатометричної методики.

Ключові слова: аморфний сплав, вмерзнені центри кристалізації, висхідна дифузія

The temperature stability of amorphous alloys is determined by their composition, the receipt technology, and also by the temperature interval of phase transformations' activation. The study of kinetics and the mechanisms of amorphous alloys' structural transformations allows to give concrete recommendations as for their thermal stability increase and the changes of physical properties in the necessary direction. It has been conducted the estimation of ascending diffusion process possibility in the amorphous alloys of Fe-B system through thermodynamics calculations. It has been also theoretically grounded and experimentally implemented the process of amorphous matrix cleaning from the frozen crystallization centers by the isothermal annealing of amorphous alloy at the temperature $T = 0,6T_{\text{ev}}$, due to the initiated ascending diffusion caused by the difference between the chemical potentials of i-component in the amorphous and crystalline phases. The dilatometric method has revealed that the isothermal annealing at temperatures $T = 0,6T_{\text{ev}}$ results in the expansion of the temperature stability interval of amorphous alloys on (20-60) K. The increase of temperature of intensive crystallization beginning, which is explained by dissolution of the frozen crystallization centers, also testifies the previous issue.

Key Words: amorphous alloy, freezed crystallization centers, ascending diffusion

Статтю представив член-кор. НАН України, д.ф.-м.н., проф. Макара В.А.

Вступ

Сучасні термодинамічно нерівноважні матеріали з аморфною структурою характеризуються комплексом унікальних фізико-хімічних властивостей. Проте більшість аморфних сплавів мають суттєвий недолік – невисоку термічну стабільність. При нагріванні до певної температури аморфні металеві сплави переходять в більш стабільний кристалічний стан, втрачаючи при цьому свої характерні властивості. Високотемпературну стабільність аморфних сплавів часто пов'язують із повною

відсутністю зародків кристалічної фази, але численні електронномікроскопічні дослідження показують, що в аморфних сплавах існують центри кристалізації, які є вмерзеними в аморфну фазу [1, 2]. Згідно теорії гомогенного зародкоутворення при швидкому гартуванні розплаву зародки кристалічної фази утворюються в області температур $T \sim 0,7 T_p$, при цьому переохолоджена рідина переходить в тверду аморфну фазу. Таким чином, утворюється гетерогенна система: вмерзнені центри кристалізації – аморфна фаза, в якій виникають пружні напруження. Термодинамічна рівновага в

такій системі визначається різницею хімічних потенціалів компонентів сплаву в аморфній та кристалічній фазах $\Delta\mu_i$. На процес зародження центрів кристалізації впливає різниця хімічних потенціалів між аморфною та кристалічною фазами $\Delta\mu_i$, при цьому зменшення $\Delta\mu_i$ сприяє підвищенню термічної стабільності сплавів. В роботі [3] отримано термодинамічні умови існування в заморожених центрів кристалізації в аморфній фазі. Врахування додаткового тиску на поверхні розділу кристалічний зародок – аморфна фаза призводить до значного зменшення величини $\Delta\mu_i$ в порівнянні зі значенням $\Delta\mu_i$ в тій області, де пружні напруження відсутні. За умови $\Delta\mu < 0$ розмір заморожених центрів кристалізації може зменшуватись та відбуватися їх розчинення в аморфній матриці. В даному випадку має підвищуватись термічна стабільність аморфного сплаву.

Термодинамічний аналіз можливості протікання процесу висхідної дифузії в аморфних сплавах

Причиною переміщення частинки i -го компонента у розчині є різниця між хімічними потенціалами $\Delta\mu_i$ в α (аморфній) і β (кристалічній) фазах, яка призводить до виникнення рушійної сили

$$\vec{F}_i = -\nabla\mu_i. \quad (1)$$

Термодинамічний аналіз показує, що рушійна сила $\vec{F}_i$, яка викликає дифузійний перенос речовини i -го компонента в багатокомпонентній системі визначається градієнтом хімічного потенціалу i -го компонента, а не градієнтом концентрації, як це слідує із першого рівняння Фіка.

Під дією сили $\vec{F}_i$ атоми здійснюють спрямований рух із середньою швидкістю

$$\langle v_i \rangle_F = u_i F_i, \quad (2)$$

де u_i – рухливість атомів i -го компонента.

В результаті виникає дифузійний потік атомів i -го компонента, густина якого дорівнює

$$J_i = c_i \langle v_i \rangle = c_i u_i F_i = -c_i u_i \nabla\mu_i. \quad (3)$$

Хімічний потенціал i -го компонента в реальному розчині визначається так:

$$\mu_i = RT \ln a_i + \mu_{0i}, \quad (4)$$

де a_i – активність i -го компонента, μ_{0i} – хімічний потенціал чистого i -го компонента в стандартному стані.

Враховуємо, що активність визначається рівнянням

$$a_i = \gamma_i c_i, \quad (5)$$

де γ_i – коефіцієнт активності.

Отримаємо вираз для хімічного потенціалу i -го компонента

$$\mu_i = RT \ln(\gamma_i c_i) + \mu_{0i}. \quad (6)$$

Таким чином, для густини дифузійного потоку атомів i -го компонента маємо вираз

$$J_i = -u_i RT \frac{\partial \ln a_i}{\partial \ln c_i} \nabla c_i, \quad (7)$$

або

$$J_i = -u_i RT \left(1 + \frac{\partial \ln \gamma_i}{\partial \ln c_i} \right) \nabla c_i. \quad (8)$$

Порівнюючи (8) із першим законом Фіка отримаємо вираз

$$D_i = u_i RT \left(1 + \frac{\partial \ln \gamma_i}{\partial \ln c_i} \right). \quad (9)$$

Вираз (9) показує можливість виникнення явища дифузії проти градієнту концентрації за умови

$$\frac{\partial \ln \gamma_i}{\partial \ln c_i} < -1. \quad (10)$$

Дана умова описує явище висхідної дифузії, при якій потік компонента, що дифундує, спрямований в бік збільшення градієнта концентрації; при цьому кристалічний зародок буде розчинятись в аморфній фазі.

З урахуванням рівняння (9), перше рівняння Фіка для реальних розчинів можна записати так

$$J_i = -D_i \nabla c_i + c_i \langle v_i \rangle F_i. \quad (11)$$

Порівнюючи (8) та (11) приходимо до висновку, що в реальному розчині з'являється додаткова сила

$$F_i = -kT \nabla \ln \gamma_i. \quad (12)$$

Ця сила характеризує вплив градієнта власного потенціального поля кристала, в якому рухається атом, що дифундує.

Рушійна сила дифузії $(-kT \nabla \ln \gamma_i)$ відрізняється від концентраційної ∇c_i , яка відповідає ентропії змішування.

Розглянемо можливість теоретичного розрахунку коефіцієнта активності γ_i . Коефіцієнт активності i -компонента визначається рівнянням

$$\gamma_i = \exp\left(\frac{\Delta H_i}{RT}\right), \quad (13)$$

де ΔH_i – відносна парціальна ентальпія перетворення кристал – аморфна фаза для i -го компонента.

Відносні парціальні ентальпії перетворення кристал – аморфна фаза для i -го компонента знайдемо через відносну інтегральну ентальпію сплаву

$$\Delta H_i = \Delta H + (1 - c_i) \frac{d\Delta H}{dc_i}, \quad (14)$$

де ΔH – відносна інтегральна ентальпія перетворення кристал – аморфна фаза, яку можна визначити з експерименту.

Для бінарного розчину можна записати

$$\begin{aligned} \Delta H_1 &= \Delta H + (1 - c_1) \frac{d\Delta H}{dc_1}, \\ \Delta H_2 &= \Delta H + (1 - c_2) \frac{d\Delta H}{dc_2}. \end{aligned} \quad (15)$$

Оцінимо можливість виконання критерію (10) протікання процесу висхідної дифузії в бінарних аморфних сплавах системи Fe-B.

Скористаємось експериментальними залежностями відносної інтегральної ентальпії перетворення $\Delta H(T)$ кристалічного сплаву $\text{Fe}_{82}\text{B}_{18}$ в переохолоджений (метастабільний) розплав при температурі 850 К [4] та напівемпіричним виразом для бінарного сплаву, отриманим в роботі [5]

$$\Delta H = \frac{\beta c_1 (1 - c_1)}{\alpha^2 + (1 - 2\alpha)(1 - c_1)}, \quad (16)$$

де α – положення максимуму функції $\Delta H(c_1)$,

β – значення функції $\Delta H(c_1)$ в точці максимуму.

Розрахунок параметра $\frac{\partial \ln \gamma_{Fe}}{\partial \ln c_{Fe}}$ згідно формул

(10), (13), (14) для кристалічного зародка в сплаві $\text{Fe}_{80}\text{B}_{20}$ при $T = 0,6 T_{eg} = 850$ К та $\Delta H = 8,0$ кДж/моль, показав, що умова можливості протікання процесу висхідної дифузії виконується:

$$\frac{\partial \ln \gamma_{Fe}}{\partial \ln c_{Fe}} = -1,35.$$

Таким чином, при ізотермічному відпалі аморфного сплаву $\text{Fe}_{80}\text{B}_{20}$ в області температур $T = 0,6 T_{eg}$ можливе розчинення в заморожених центрів кристалізації за рахунок висхідної дифузії, що в свою чергу може підвищити термічну стабільність аморфного сплаву.

Результати експериментальних досліджень та їх обговорення

Даний теоретичний висновок було підтверджено експериментально. Визначення температури початку інтенсивної кристалізації проводилось за допомогою дилатометричного методу, суть якого полягає в наступному [6]. Молярний об'єм більшості сплавів в аморфному та кристалічному станах відрізняється на (1-3) %. При нагріванні аморфного сплаву з постійною швидкістю його об'єм монотонно зростає, коли температура досягає певного значення (температури початку інтенсивної кристалізації) відбувається перехід аморфного сплаву в кристалічний стан. Цей процес супроводжується різким зменшенням об'єму. Процес охолодження сплаву характеризується лінійною залежністю відносної зміни об'єму від температури.

На матеріал з відомим коефіцієнтом термічного розширення закріплюється відрізок аморфної фольги та задається прогин. Вимірювання довжини досліджуваного зразка в процесі нагрівання безперервно фіксується за зміною стріли прогину, що викликана розширенням фольги відносно станка. Таким чином, можна дослідити процес кристалізації аморфного сплаву, фіксуючи зміну довжини зразка та перераховуючи її в об'ємні зміни. З отриманої температурної залежності відносної зміни об'єму визначають температуру початку інтенсивної кристалізації.

Для вихідного аморфного сплаву $\text{Fe}_{83}\text{B}_{17}$ температура початку інтенсивної кристалізації складає (690 ± 5) К, після ізотермічного відпалу при температурі $T = 0,6 T_{eg}$ досягає (730 ± 5) К, отже інтервал термічної стабільності розширюється на 40 К. Аналогічні дослідження було проведено для низки аморфних сплавів системи Fe-B: $\text{Fe}_{80}\text{B}_{20}$, $\text{Fe}_{70}\text{Cr}_{15}\text{B}_{15}$, $\text{Fe}_{80}\text{B}_{14}\text{Si}_6$, $\text{Fe}_{70}\text{Mo}_{10}\text{Si}_6\text{B}_{14}$, $\text{Fe}_{76,2}\text{B}_{14}\text{Si}_6\text{Ni}_{3,8}$, $\text{Fe}_{77,5}\text{B}_{16}\text{Si}_2\text{Ni}_{3,5}\text{Mo}_1$. Були визначені температури початку інтенсивної кристалізації для вихідних аморфних сплавів та для сплавів, що пройшли попередній ізотермічний відпал при відповідних температурах, нижчих за температуру початку інтенсивної кристалізації.

Аналіз отриманих експериментальних результатів (табл. 1). показав, що ізотермічний відпал призводить до розширення інтервалу термічної стабільності аморфних сплавів на $(20 \div 60)$ К, про що свідчить збільшення температури початку інтенсивної кристалізації.

Таблиця 1

Температури початку інтенсивної кристалізації для вихідних аморфних сплавів та сплавів після попереднього ізотермічного відпалу

Склад аморфного сплаву	$(T_k \pm 5) \text{ K}$	
	вихідний	після відпалу
$\text{Fe}_{80}\text{B}_{20}$	650	690
$\text{Fe}_{83}\text{B}_{17}$	695	730
$\text{Fe}_{70}\text{Cr}_{15}\text{B}_{15}$	750	810
$\text{Fe}_{80}\text{B}_{14}\text{Si}_6$	770	800
$\text{Fe}_{76}\text{Ni}_4\text{Si}_6\text{B}_{14}$	770	790
$\text{Fe}_{77,5}\text{B}_{16}\text{Si}_2\text{Ni}_{3,5}\text{Mo}_1$	750	790

Підвищення термічної стабільності аморфних сплавів пояснюється розчиненням в заморожених центрів кристалізації за рахунок висхідної дифузії.

Висновки

• Теоретично обґрунтовано і експериментально реалізовано процес очищення аморфної матриці від заморожених центрів кристалізації шляхом ізотермічного відпалу аморфного сплаву в області температур $T = 0,6 T_{\text{ев}}$, за рахунок ініційованої висхідної дифузії, викликані різницею хімічних потенціалів i -го компоненту в аморфній та кристалічній фазах.

• За допомогою дилатометричної методики показано, що ізотермічний відпал в області температур $T = 0,6 T_{\text{ев}}$ призводить до розширення інтервалу температурно-часової стабільності аморфних сплавів на $(20 \div 60) \text{ K}$, про що свідчить збільшення температури початку інтенсивної кристалізації, що пояснюється розчиненням заморожених центрів кристалізації.

Список використаних джерел

1. Шпак А.П. Кристалізація і аморфізація металевих систем. / А.П. Шпак, В.І. Лисов, Ю.А. Куницький, Т.Л. Цареградська – Київ.: Академперіодика, 2002. – 208 с.
2. Лисов В.І. Дослідження процесів стабілізації структури та властивостей металевих стекол під дією зовнішніх впливів / В.І. Лисов, Т.Л. Цареградська, О.В. Турков, Г.В. Саєнко // Журнал фізичних досліджень. – 2013. – т.7. – № 2. – С.2701.1-2701.5.
3. Харьков Е.И. Термодинамическая теория высокотемпературной устойчивости аморфных сплавов / Е.И. Харьков, В.И. Лысов // Расплавы. – 1987. – т.1, – №1. – С. 30-37.
4. Зайцев А.И. Термодинамические свойства расплава и фазовые равновесия в системе железо-бор. Превращение жидких сплавов Fe-B в аморфное состояние / А.И. Зайцев, Н.Е. Зайцева // Журнал физической химии. – 2002. – т.76. – №1. – С.33-44.
5. Лысов В.И. Фазовое расслоение в металлических неупорядоченных системах / В.И. Лысов, Т.Л. Цареградская, О.В. Турков, Г.В. Саєнко, В.В. Ярыш // Журнал физической химии. – 2007. – т.81. – №10. – С. 1765-1769.
6. Новиков В.Н. Методика дилатометрических исследований металлических стекол / В.Н. Новиков, Б.Е. Рябчиков // Заводская лаборатория. – 1989. – т.55. – №37. – С.49-53.

References

1. SHPAK, A.P. *et al.* (2002) *Crystallisation and amorphisation of metallic systems*. Kyiv: Akademperiodika.
2. LYSOV, V.I., TSAREGRADSKAYA, T.L., TURKOV, O.V., SAENKO G.V. (2013) The studies of processes stabilizing structure and properties of metallic glasses under external influences. *Journal of Physical Studies*. 7 (2). p. 2701.1-2701.5.
3. KHAR'KOV, E.I., LYSOV, V.I. (1987) Thermodynamic theory of high-temperature stability of amorphous alloys. *Fusions*. 1(1). p. 30-37.
4. ZAYTSEV, A.I., ZAYTSEVA, N.E. (2002) Thermodynamics properties of fusion and phase equilibrium in iron-boron system. Transformation liquid alloys Fe-B in amorphous state. *Journal of Physical Chemistry* 76 (1). p.33-44.
5. LYSOV, V.I., TSAREGRADSKAYA, T.L., TURKOV, O.V., SAENKO, G.V., YARYSH, V.V. (2007) Phase stratification is in the metallic unregulated systems. *Russian Journal of Physical Chemistry*, 81 (10), p. 1765-1769.
6. NOVIKOV, V.N., RYABCHIKOV, B.E. (1989) The method of dilatometrical researches of metallic glasses. *Factory laboratory*. 55(37). p. 49-53.

Надійшла до редколегії 29.05.2015

УДК 537.6

Ніколаєнко Т. Ю.¹, к.ф.-м.н.
Пилипчук Є. В.², к.х.н.

Визначення фізико-хімічних властивостей рідинних наносистем на основі вимірювання магнітної проникності

¹Київський національний університет імені
Тараса Шевченка, 01601, м. Київ, вул.
Володимирська 64/13,
e-mail: tim_mail@ukr.net

²Інститут хімії поверхні ім. О.О. Чуйка НАН
України, 03164 м. Київ, вул. Генерала
Наумова, 17, e-mail: chemind@ukr.net

T. Yu. Nikolaienko¹, Ph.D.,
Ie. V. Pylypchuk², Ph.D.,

Determination of physico-chemical properties of liquid nanosystems on the basis of magnetic permeability measurements

¹Taras Shevchenko National University of Kyiv,
01601, Kyiv, Volodymyrska st. 64/13,
e-mail: tim_mail@ukr.net,

²Chuiko Institute of Surface Chemistry, National
Academy of Sciences of Ukraine, 03164 Kyiv,
General Naumov str., 17,
e-mail: chemind@ukr.net

Описано установку для вимірювання магнітної проникності рідинних наносистем і показано, що за вимірними значеннями магнітної проникності можливо визначати концентрацію наночастинок магнетиту в зразках (як рідинах, так і порошках). Проаналізовано джерела похибок вимірювань та вказано можливі шляхи подальшого збільшення точності вимірювань.

Ключові слова: магнітні наночастинки, нанорідини, магнітна сприйнятливість.

A measurement system for the determination of magnetic permeability of liquid nanosystems is presented and it is shown that with the proposed system it is possible to determine the concentration of magnetic nanoparticle in liquids and powder samples in the range of concentration from 0.1 to 500 mg/ml with accuracy not worse than 20%. Possible sources of measurement errors are analyzed possible ways to further increase the accuracy of measurements are sketched. An accuracy in magnetic permeability measurement is estimated to be better than $1 \cdot 10^{-5}$ and linear relationship between the measured value of magnetic permeability and magnetite contents has been shown to work well for both liquid systems and dry mixtures. Test measurements of magnetic permeability were made for liquid system containing magnetite nanoparticle covered with sodium oleate in distilled water and the same particles in the form of aggregates mixed with silicon oxide. The proposed system can be used in medicinal physics, medical biophysics, biology, medicine to control the number of magnetic nanoparticles in biological samples.

Key Words: magnetic nanoparticles, nanoliquids, magnetic permeability.

Статтю представив академік НАН України, д.ф.-м.н., проф. Булавін Л.А.

Вступ

Задача визначення концентрації магнітних наночастинок у зразках, зокрема – біологічного походження, виникає при проведенні біомедичних досліджень на піддослідних тваринах з метою експериментального вивчення впливу зовнішнього магнітного поля на перерозподіл магнітних наночастинок-носіїв лікарських засобів в їх організмах і розробці магнітокерованих методів доставки молекул лікарських засобів, з'єднаних з магнітними наночастинами, у задану область організму [1, 2].

Серед відомих методів визначення концентрації магнітних наночастинок у зразках слід згадати магнітну силову мікроскопію [3], рентгенівську фотоелектронну спектроскопію [4], метод люмінесцентних міток [5], а також методи електронного (ЕПР) та ядерного магнітного резонансу (ЯМР) [6] і аналогічні їм [7]. Однак, всі ці методи або потребують використання доволі складного обладнання, яке є порівняно дорогим і потребує особливого налаштування та обслуговування, або потребують хімічної або ізотопної модифікації наночастинок, що не завжди є припустимим.

Значно простішим у реалізації, але водночас доволі ефективним способом визначення кількісного вмісту магнітних наночастинок в зразку є метод, заснований на вимірюванні магнітних властивостей речовин, а саме – магнітної сприйнятливості, або однозначно пов'язаної з нею магнітної проникності. Зазначимо, що оскільки мова йде про систему із великої кількості незв'язаних однодомених магнітних частинок сферичної форми, магнетик, що їх містить, будемо вважати ізотропним, а його магнітну сприйнятливості – скалярною величиною.

Метою даної роботи є встановити можливість визначення вмісту магнітних наночастинок в зразку біологічного походження на основі вимірювання його магнітної проникності та реалізувати цю можливість за допомогою конкретної конструкції вимірювального приладу.

Матеріали і методи

Досліджувані зразки являли собою розчини наночастинок магнетиту (Fe_3O_4), вкритих бішаром олеату натрію, у дистильованій воді. Концентрація магнетиту в розчинах складала від 0,05 до 50 мг/мл. Крім того, досліджували зразки, що являли собою суміш порошків магнетиту і оксиду кремнію (SiO_2) Kieselgel-40 з розміром частинок 0,063 – 0,200 мкм. Середній діаметр наночастинок магнетиту складав 10 нм.

Для вимірювання магнітної проникності досліджуваних зразків було розроблено прилад, принцип роботи якого пояснюється блок-схемою, наведеною на рис. 1.

Вимірювальний прилад побудовано на базі генератора синусоїдальних коливань 1 на основі операційного підсилювача зі струмовим

зворотним зв'язком LT1228 (вир. Linear Technologies), частото задаючим елементом якого є коливальний контур з паралельно з'єднаних між собою конденсатора ємністю 6800 пФ та котушки індуктивності L (рис. 1), виконаної таким чином, що всередину неї може бути вміщений досліджуваний зразок S . Індуктивність котушки без зразка складає 486 мкГн, тож робоча частота створеного генератора (частота власних коливань коливального контура) складає близько 88 кГц. При вміщенні у котушку зразка, що має відмінну від одиниці магнітну проникність μ (ненульову магнітну сприйнятливості χ , $\chi = \mu - 1$) індуктивність котушки змінюватиметься пропорційно до μ ($L = \mu L_0$, де L_0 – індуктивність котушки без зразка), а період коливань генератора складатиме

$$T = \sqrt{LC} = \sqrt{\mu L_0 C} = T_0 \sqrt{\mu} \quad (1)$$

де T_0 – період коливань генератора за умови, що до котушки не вміщено зразок.

До виходу генератора приєднано компаратор 2 типу LT1715 (вир. Linear Technologies), який перетворює сигнал синусоїдальної форми у сигнал прямокутної форми, який далі подається на один з цифрових входів мікроконтролера 5 (тип PIC18F14K50). Програма мікроконтролера визначає час 100000 повних періодів вхідного сигналу від генератора, використовуючи при цьому для вимірювання часу опорний сигнал з частотою 12 МГц від вбудованого кварцового резонатора 4. Блок живлення 3 забезпечує подачу необхідної стабілізованої напруги на вищезазначені елементи схеми. Приладом індикації слугує персональний комп'ютер, з'єднаний з мікроконтролером 5 по інтерфейсу USB.

Для забезпечення якомога вищої точності вимірювань при конструюванні приладу особливу увагу було приділено часовій стабільності роботи генератора 1 та генератора опорного сигналу на основі кварцового резонатора 4. У частото задаючому коливальному контурі було використано конденсатор КСО-1Г з посрібленим слюдяним діелектриком, що має температурний коефіцієнт ємності не вищий від $50 \cdot 10^{-6}$ 1/К, а у кварцовому генераторі було використано кварц HC-49S серії 9В (вир. ТХС), який характеризується відносною стабільністю частоти на рівні $\pm 10 \cdot 10^{-6}$ в діапазоні температур від -20 до +70°C та гладкістю залежності резонансної частоти від температури. В результаті за умови виходу приладу на штатний режим роботи було досягнуто дрейф вимірюваного значення магнітної проникності не

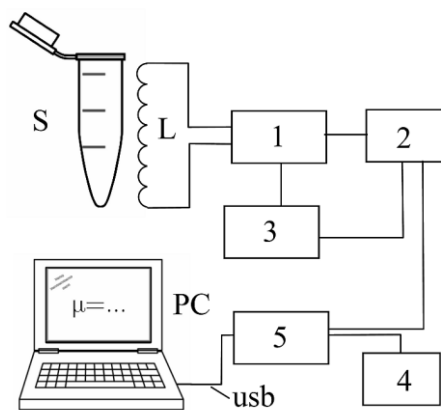


Рис. 1. Блок-схема комп'ютеризованого приладу для вимірювання магнітної проникності наносистем.

вище $0,2 \cdot 10^{-7}$ 1/с. Таким чином, при тривалості одного виміру порядку 5 с похибка магнітної проникності, пов'язана із несталістю параметрів компонентів приладу не перевищувала $1 \cdot 10^{-6}$.

Результати та обговорення

На рис. 2 представлено результати вимірювання магнітної проникності μ водних рідинних систем на основі наночастинок магнетиту, вкритих бішаром олеїнової кислоти, проведені при різних концентраціях c розчинів, а також – два можливі варіанти апроксимації залежності $\mu - 1 = f(c)$ лінійними функціями. Відмітимо, що для представлення даних на графіку використано двічі логарифмічний масштаб.

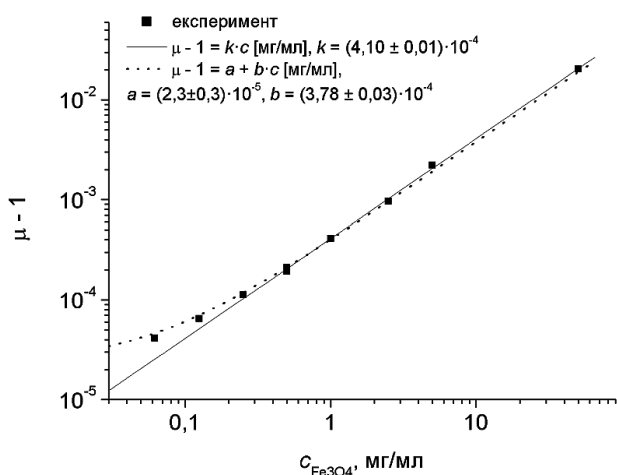


Рис. 2. Залежність магнітної проникності рідинних систем «магнетит + вода» від концентрації магнетиту та її апроксимації лінійними функціями.

Можна зробити висновок, що лінійна функція досить добре описує залежність магнітної проникності від концентрації наночастинок магнетиту в усьому діапазоні досліджених концентрацій (середньоквадратична похибка апроксимації виду $\mu - 1 = a + b \cdot c$ складає $7 \cdot 10^{-6}$), вказуючи тим самим на можливість розв'язку і оберненої задачі – визначення концентрації наночастинок магнетиту у рідинній системі за виміряними значеннями магнітної проникності.

Аналогічну залежність $\mu - 1 = f(c)$ для випадку зразків, що являли собою сухі суміші високодисперсного магнетиту (агрегати розміром до 400 нм) із оксидом кремнію представлено на рис. 3. В цьому випадку за величину, що відкладено по осі абсцис, було вибрано загальну масу магнетиту як величину, що характеризує

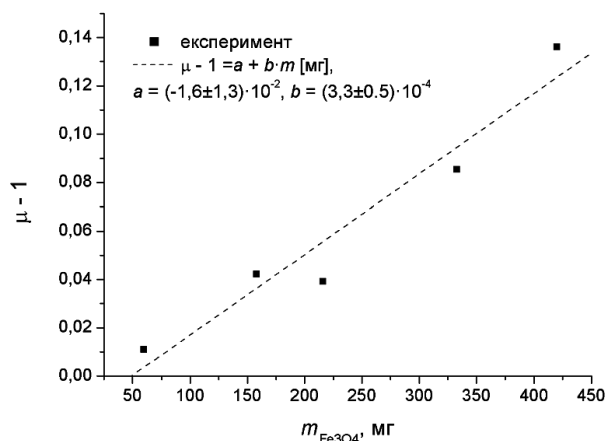


Рис. 3. Залежність магнітної проникності сухих сумішей «магнетит + SiO_2 » від концентрації магнетиту та її апроксимація лінійною функцією.

загальну кількість частинок, що мають магнітний момент і можуть орієнтуватися під впливом зовнішнього магнітного поля.

Слід відзначити, що якщо у випадку водних рідинних систем з магнетитом об'єм усіх зразків складав 1 мл (а маса рідинної системи 1,0 г), то у другому випадку загальна маса лежала в межах від 0,9 до 1,5 г залежно від маси оксиду кремнію, що використовувався для доведення об'єму суміші до 1,5 мл, необхідного, для заповнення вимірювальної комірки.

Утім, незважаючи на згадані додаткові джерела похибок, лінійна функція $\mu - 1 = a + b \cdot c$ досить добре апроксимує одержану залежність $\mu - 1 = f(c)$ (середньоквадратична похибка апроксимації складає 0,014), причому значення коефіцієнта b виявляється слабко чутливим до агрегатного стану зразка ($3,8 \cdot 10^{-4}$ 1/мг у випадку рідинної системи та $3,3 \cdot 10^{-4}$ 1/мг у випадку сухої суміші). При цьому слід відзначити, що загалом в проведених експериментах було досліджено діапазон масового вмісту магнетиту від 0,1 до 500 мг, який сягає 4 порядків величини.

Окремої уваги заслуговує питання про відмінність коефіцієнта апроксимації a від нуля ($2 \cdot 10^{-5}$) у випадку рідинних наносистем та його рівність нулеві в межах похибки апроксимації ($-1,6 \pm 1,3$) у випадку сухої суміші. Причина такої відмінності пов'язана із наявністю магнітних властивостей у води та її значною діелектричною проникністю, яка могла підвищити ефективну ємність котушки індуктивності вимірювальної комірки. Цю особливість необхідно враховувати при використанні запропонованої методики визначення вмісту наночастинок магнетиту у рідинних системах на основі води.

Висновки

Запропоновано вимірювальну систему для визначення вмісту магнітних наночастинок у рідинних системах та сухих сумішах. Система дозволяє визначати ефективну магнітну проникність середовищ із точністю не гірше $1 \cdot 10^{-5}$ та вміст наночастинок магнетиту в одиниці об'єму зразка – не гірше 20% у діапазоні від 0,1 до 500 мг/мл. Показано, що магнітна проникність досліджених водних рідинних систем на основі наночастинок магнетиту пов'язана з концентрацією c (мг/мл) магнетиту лінійним

співвідношенням $\mu = 2,3 \cdot 10^{-5} + 3,8 \cdot 10^{-4} \cdot c$. Розроблена система може бути застосована у медичній біофізиці, біології, медицині для здійснення контролю кількості магнітних наночастинок у зразках біологічного походження.

Роботу виконано в рамках теми молодих вчених Київського національного університету імені Тараса Шевченка №15БФ051-01.

Список використаних джерел

1. Синтез и свойства магнитных жидкостей на основе наноразмерного Fe_3O_4 / П.П. Горбик, Н.В. Абрамов, А.Л.Петрановская [и др.] // Поверхность. – 2011. – Вып. 3(18). – С. 245 – 257.
2. Пилипчук Е.В. Синтез и свойства нанокомпозитов на основе магнетита, модифицированного диэтилентриазинпентауксусной кислотой / Е.В. Пилипчук, А.Л. Петрановская, П.П. Горбик // Наноструктурное материаловедение. – 2012. – N 3. – С. 47.
3. Sarid D. Scanning force microscopy: with applications to electric, magnetic, and atomic forces / D. Sarid. – New York: Oxford University Press, 1994. – 263 p.
4. Нефедов В.И. Рентгеноэлектронная спектроскопия химических соединений / В.И. Нефедов. – Москва: Химия. – 1984. – 256 с.
5. Aqueous ferrofluid of magnetite nanoparticles: fluorescence labeling and magnetophoretic control. / Y. Sahoo, A. Goodarzi, M. T. Swihart, [et al] // J. Phys. Chem. B. – 2005. – 109(9). – P. 3879-3885.
6. Kempe S. Application of electron paramagnetic resonance (EPR) spectroscopy and imaging in drug delivery research—chances and challenges. / S. Kempe, H. Metz, K. Mäder // Eur J Pharm Biopharm. – 2010. – V. 74(1). – P. 55-66.
7. Knopp T. Magnetic particle imaging: an introduction to imaging principles and scanner instrumentation / T. Knopp, T. M. Buzug. – Berlin: Springer-Verlag, 2012. – 204 p.

References

1. GORBYK, P.P., ABRAMOV, N.V., PETRANOVSKAYA, A.L., TURELYK, M.P., PILIPCHUK, E.V., ORANSKAYA, E.I., KONCHITS, A.A. & SHEVCHENKO, YU.B. (2011) Synthesis and properties of magnetic fluids based on nanosized Fe_3O_4 . *Poverhnost (in Ukrainian)* 3(18). p. 245–257.
2. PILIPCHUK, E.V., PETRANOVSKAYA, A.L., & GORBYK, P.P. (2012) Synthesis and properties of nanocomposites based on magnetite modified with DTPA. *Nanostrukturnoe Materialovedenie (in Russian)* 3. p. 47.
3. SARID, D. (1994) *Scanning force microscopy: with applications to electric, magnetic, and atomic forces*. New York: Oxford University Press.
4. NEFEDOV, V.I. *X-ray photoelectron spectroscopy of chemical compounds (in Russian)*, Moscow: Khimiya.
5. SAHOO, Y., GOODARZI, A., SWIHART, M. T., OHULCHANSKY, T. Y., KAUR, N., FURLANI, E. P., & PRASAD, P. N. (2005). Aqueous ferrofluid of magnetite nanoparticles: fluorescence labeling and magnetophoretic control. *J. Phys. Chem. B.*, 109(9), p. 3879-3885.
6. KEMPE, S., METZ, H., & MÄDER, K. (2010). Application of electron paramagnetic resonance (EPR) spectroscopy and imaging in drug delivery research—chances and challenges. *Eur J Pharm Biopharm*, 74(1), p. 55-66.
7. KNOPP, T. and BUZUG, T. M. (2012) *Magnetic particle imaging: an introduction to imaging principles and scanner instrumentation*. Berlin: Springer-Verlag.

Надійшла до редколегії 03.06.15

Підписано до друку _____. Формат _____. Друк. офс.
Папір офс. № _____. Друк. арк. _____. Тираж _____. Зам. _____.
Друкарня _____.